

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/316789928>

Higher Order Dynamic Mode Decomposition

Article in *SIAM Journal on Applied Dynamical Systems* · January 2017

DOI: 10.1137/15M1054924

CITATIONS

2

READS

141

2 authors:



[Soledad Le Clainche Martínez](#)

Universidad Politécnica de Madrid

18 PUBLICATIONS 82 CITATIONS

[SEE PROFILE](#)



[José M. Vega](#)

Universidad Politécnica de Madrid

159 PUBLICATIONS 1,519 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Higher order dynamic mode decomposition [View project](#)

Higher Order Dynamic Mode Decomposition*

Soledad Le Clainche[†] and José M. Vega[†]

Abstract. This paper deals with an extension of *dynamic mode decomposition* (DMD), which is appropriate to treat general periodic and quasi-periodic dynamics, and transients decaying to periodic and quasi-periodic attractors, including cases (not accessible to standard DMD) that show limited spatial complexity but a very large number of involved frequencies. The extension, labeled as *higher order dynamic mode decomposition*, uses time-lagged snapshots and can be seen as superimposed DMD in a sliding window. The new method is illustrated and clarified using some toy model dynamics, the Stuart–Landau equation, and the Lorenz system. In addition, the new method is applied to (and its robustness is tested in) some permanent and transient dynamics resulting from the complex Ginzburg–Landau equation (a paradigm of pattern forming systems), for which standard DMD is seen to only uncover trivial dynamics, and the thermal convection in a rotating spherical shell subject to a radial gravity field.

Key words. dynamic mode decomposition, Koopman operator, delayed snapshots, nonlinear dynamical systems, transient dynamics, quasi-periodic attractors, complex Ginzburg–Landau equation, thermal convection in spherical shells

AMS subject classifications. 35B10, 35B40, 35B41, 35C99, 37L15, 37L30, 37M99

DOI. 10.1137/15M1054924

1. Introduction. *Dynamic mode decomposition* (DMD) is closely related to (and inspired by) Koopman-operator analysis [36, 43], initiated by Koopman [31] in 1931. DMD itself was developed by Schmid [48], and has become a useful tool for postprocessing massive spatio-temporal data in numerical and experimental fluid mechanics [49, 50, 51, 52, 32] and other fields [28, 40]. The ability of DMD to extract *relevant patterns* makes this method potentially useful in identifying *nonlinear dynamics* in many physical systems, but this application requires some care since the standard DMD may give completely spurious results. The main goal of this paper is to analyze the application of DMD to general (both low-dimensional and infinite-dimensional) dynamical systems, which will require a nontrivial extension of the method.

In order to fix ideas, DMD applies to spatio-temporal data organized in K equispaced J -dimensional *snapshots* as

$$(1.1) \quad \mathbf{v}_k \equiv \mathbf{v}(t_k) \in \mathbb{R}^J \quad \text{for } t_k = t_1 + (k-1)\Delta t \quad \text{with } k = 1, \dots, K.$$

Thus, the *spatial and temporal dimensions* of the given data are J and K , respectively. Stan-

*Received by the editors January 4, 2016; accepted for publication (in revised form) by T. Sauer December 20, 2016; published electronically DATE.

<http://www.siam.org/journals/siads/x-x/M105492.html>

Funding: This research was partially supported by the Spanish Ministry of Economy and Competitiveness, under grant TRA2013-45808-R.

[†]E.T.S.I. Aeronáutica y del Espacio, Universidad Politécnica de Madrid, 28040 Madrid, Spain (soledad.leclainche@upm.es, josemanuel.vega@upm.es).

standard DMD relies on the following *Koopman assumption*:

$$(1.2) \quad \mathbf{v}_{k+1} = \mathbf{R} \mathbf{v}_k \quad \text{for } k = 1, \dots, K-1$$

for a certain $J \times J$ -matrix \mathbf{R} , which can be called the *Koopman matrix* (or operator). For simplicity in the exposition, this equation is assumed at the moment to be satisfied exactly, though it will only hold approximately in most applications (see below). The assumption (1.2) implies that the snapshots lie in the *invariant subspace* under the action of the *Koopman group* [58, 11], generated by \mathbf{R} . This assumption leads to the following DMD representation of the snapshots:

$$(1.3) \quad \mathbf{v}(t_k) = \sum_{m=1}^M a_m \mathbf{u}_m e^{(\delta_m + i\omega_m)(k-1)\Delta t} \quad \text{for } k = 1, \dots, K,$$

where the exponents appear in complex conjugate pairs when the snapshots are real. This is a Fourier-like expansion that involves, not only the frequencies ω_m , but also the growth rates δ_m .

The *linear time invariant* (LTI) [60] system (1.2) is not assumed to be the actual *physical model* that has produced the snapshots. In other words, the role of the linear system (1.2) is only *instrumental* in this context, as a means to compute the modes \mathbf{u}_m , amplitudes a_m , growth rates δ_m , and frequencies ω_m appearing in the Fourier-like expansion (1.3). Instead, the underlying physical model may well be nonlinear, and the snapshots (1.1) satisfying (1.2) may be particular outcomes of the system for a periodic (if $\delta_m = 0$ and the frequencies ω_m are commensurable) or quasi-periodic (if $\delta_m = 0$ and the frequencies ω_m are incommensurable) attractor of the system, or for a transient approaching a periodic or quasi-periodic attractor (if $\delta_m \leq 0$).

Since any terms in (1.3) with the same exponents may be collected into a single term by summing their coefficients, we assume (without loss of generality) that the various exponents in (1.3) are distinct. In this case, the number of involved modes, M , can be called the *spectral complexity*, while the dimension of the subspace generated by the M DMD-modes, namely

$$(1.4) \quad N = \dim(\text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_M\}) \leq \min\{M, J\},$$

is the *spatial complexity*. The spatial complexity can be elucidated via either truncated *proper orthogonal decomposition* (POD) [16] or truncated *singular value decomposition* (SVD) [25], which allows reducing the dimension of the snapshots set and, combined with the pseudo-inverse, also permits computing the Koopman matrix \mathbf{R} . Once \mathbf{R} has been calculated, the growth rates and frequencies appearing in (1.2) are related to the nonzero eigenvalues of \mathbf{R} , μ_m , as

$$\delta_m + i\omega_m = \frac{1}{\Delta t} \log \mu_m,$$

while the modes \mathbf{u}_m are the associated eigenvectors; the amplitudes a_m can be computed by various means (see below). At the moment, it is convenient to note that, if (1.2) and (1.3) are consistent with each other, then

- The DMD modes appearing in (1.3) are linearly independent, as eigenvectors of a matrix associated with different eigenvalues.
- The number of terms appearing in (1.3), namely the spectral complexity M , cannot be larger than the spatial complexity N , which, invoking (1.4), means that $M = N$.

In fact, if $M = N$ and conditions (1.2) and (1.3) are exact, then these conditions are equivalent. This is a fundamental limitation of standard DMD [48]. More general expansions (1.3), with $M > N$, cannot be obtained via standard DMD, which gives spurious results, even in the simplest case in which the snapshots are identically of the form (1.3) (see the examples in section 3.1). Instead, according to Theorem D in Appendix A, the expansion (1.3) is consistent (always, for general M and N) with the following *higher order Koopman assumption*:

$$(1.5) \quad \mathbf{v}_{k+d} = \mathbf{R}_1 \mathbf{v}_k + \mathbf{R}_2 \mathbf{v}_{k+1} + \dots + \mathbf{R}_d \mathbf{v}_{k+d-1} \quad \text{for } k = 1, \dots, K-d,$$

where $d \geq 1$ is tunable. Summarizing, although the assumption (1.2) leads to the DMD expansion (1.3), the converse is not true, namely not all expansions of the form (1.3) can be calculated from the assumption (1.2). The more general assumption (1.5), instead, is equivalent to (1.3).

It must be noted that the spectral complexity M is larger than the spatial complexity N in dynamical systems of scientific and industrial interest. For instance, $N = 3$ in the Lorenz system (considered in section 3.2), which exhibits chaotic dynamics (infinite spectral complexity) and also periodic attractors with very large spectral complexity M . On the other hand, M may be larger than N in dissipative, infinite-dimensional systems modeled by partial differential equations. For example,

- The weakly nonlinear theory near a Hopf bifurcation in these systems is described by a Stuart–Landau equation [26]. As further explained in section 3.3, the weakly nonlinear theory predicts transient dynamics involving a large number M of decaying modes that are approximately contained in the two-dimensional *center manifold* built around the steady state, meaning that $N = 3 < M$.
- Similarly, weakly nonlinear descriptions of quasi-periodic phenomena in these systems (see [2] and references therein) may involve few spatial modes.
- Fully nonlinear dynamics can behave similarly in these systems since, for large t , the solution of many infinite-dimensional systems converges to a nonlinear, finite-dimensional *inertial manifold* [22], contained in a generally larger (but finite-dimensional, according to the Whitney embedding theorem [27]) N -dimensional linear manifold. The dynamics in this manifold may well be chaotic (e.g., transitional flows) and also exhibit attractors with large M ($> N$). A very clear example with finite-dimensional inertial manifold where fairly complex dynamics occur is the complex Ginzburg–Landau equation, which will be considered in section 4. This is a very convenient pattern forming system to illustrate the methods in this paper because (i) it exhibits fairly complex dynamics, (ii) it is a “normal form” that applies to a variety of oscillatory bifurcations in infinite-dimensional systems [19, 26], and (iii) it is simple enough as to allow for fast numerical simulation. Also, this application suggests that *higher order dynamic mode decomposition* (HODMD) may be very useful to analyze periodic and quasi-periodic phenomena in pattern forming systems. Let us note here that

identifying and computing quasi-periodic attractors in infinite-dimensional systems is a subtle matter [45, 46].

The counterpart of standard DMD using the assumption (1.5) will be labeled as HODMD, and will be seen in this paper to be a fairly precise and robust means to compute the expansion (1.3) for sets of snapshots resulting from general periodic and quasi-periodic dynamics, namely consistent with (1.3), with arbitrary finite values of the spectral and spatial complexities. In any event, irrespective of whether the correct DMD expansion is calculated using standard DMD (which can be seen as a particular case of HODMD) or HODMD, the continuous extension of (1.3),

$$(1.6) \quad \mathbf{v}(t) = \sum_{m=1}^M \mathbf{u}_m e^{(\delta_m + i\omega_m)(t-t_1)} \quad \text{for } t_1 \leq t \leq t_K,$$

may be used to reconstruct (very general, transient, or permanent) *nonlinear dynamics*. For permanent dynamics (when all $\delta_m = 0$), (1.6) can be seen as *interpolation* in the time variable, while for transient dynamics (with $\delta_m \leq 0$), skipping in (1.6) those terms with $\delta_m < 0$ permits anticipating the final attractor for $t \gg 1$ from transient behavior, which involves *extrapolation* and, as noticed elsewhere [3], may be used to reduce the CPU time required to approach the final attractors using numerical simulation; this CPU time may be huge near bifurcation points.

The good performance of HODMD can be attributed to the fact that, according to Theorem D in Appendix A, the assumption (1.5) is general enough as to allow for general spatio-temporal behaviors of the form (1.3). Another explanation of the improved performance of HODMD follows by interpreting (1.5) as a modified form of (1.2) containing also time-lagged snapshots. Namely, (1.5) can also be written as

$$(1.7) \quad \tilde{\mathbf{v}}_{k+1} = \tilde{\mathbf{R}} \tilde{\mathbf{v}}_k,$$

where the *modified snapshots* $\tilde{\mathbf{v}}_k$ and the *modified Koopman matrix* $\tilde{\mathbf{R}}$ are defined as

$$(1.8) \quad \tilde{\mathbf{v}}_k \equiv \begin{bmatrix} \mathbf{v}_k \\ \mathbf{v}_{k+1} \\ \dots \\ \mathbf{v}_{k+d-2} \\ \mathbf{v}_{k+d-1} \end{bmatrix}, \quad \tilde{\mathbf{R}} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \\ \mathbf{R}_1 & \mathbf{R}_2 & \mathbf{R}_3 & \dots & \mathbf{R}_{d-1} & \mathbf{R}_d \end{bmatrix},$$

with \mathbf{I} = the $J \times J$ unit matrix. In fact, the algorithm presented in this paper to perform HODMD, which will be called the *DMD- d algorithm*, roughly consists of applying standard DMD to the enlarged snapshots defined in (1.8). The good performance of the computations based on (1.7) could be seen as consistent with the use of sliding windows in the improvement of the fast Fourier transform (FFT) known as *power spectral density* (PSD) [38]. A more relevant explanation follows noting that if the spatial complexity of the original snapshots is smaller than the spectral complexity, including the time-lagged snapshots in (1.8) recovers the missing degrees of freedom and increases the spatial complexity in the enlarged snapshots as to make it equal to the spectral complexity. This is consistent with the *method of delay-coordinate embedding* for state space reconstruction of an attractor using time-series of a

limited number of observables (even scalar observables). In this method, the number of observables is increased by using time-lagged observables (see also [58, 59] for alternative ways of extending the space of observables); delayed information has also been recently used in model identification [12]. The method of delay-coordinate embedding relies on seminal ideas by Packard et al. [37], which were formalized by Takens [54] in his *delay-embedding theorem* (see also [44] for extensions of this theorem), and further pursued by Broomhead and King [10]. In this sense, HODMD can be seen as a synergic combination of standard DMD and Takens' delay embedding and could be useful for the delay-embedding community, which is very interested in forecasting. However, most of the current forecasting methods are restricted to *stationary time-series*, which are analyzed using classical techniques such as the method of analogs [34] and local linear forecasting [21, 13]. These methods have been recently revisited, combined with diffusion maps [62, 9] and the Koopman operator [24], and generalized to a probabilistic framework for stochastic systems [42, 7, 8]. In contrast to these *local linear methods*, HODMD is a *global linear method* that provides all involved frequencies and growth rates simultaneously. Moreover, by its own nature, the outcomes of HODMD may involve nonzero decaying rates in (1.6), which permits *extrapolation* (and thus dealing with nonstationary time series) while the delay-embedding methods mentioned above are all based on interpolation on stationary time series.

Although, for simplicity in the presentation, (1.2), (1.3), and (1.5) were assumed above to be exact, these equations are usually only approximate in typical applications, due to either the presence of noise or the fact that only a limited accuracy is sought. In the latter case, the spatial complexity M and the spectral complexity N depend on the required accuracy. In fact, the accuracy may be an issue [61, 29] when calculating the DMD expansion (1.3). Concerning noise, this is filtered out by the preliminary application of SVD (whose ability to filter out noise is well known [53]), which is performed as the previous step to dimension-reducing the snapshots set and reinforced through the use of delayed snapshots (whose insensitivity to noise is well documented [6]). In this context, it is necessary to robustly capture the spatio-temporal redundancies, which requires having enough data in both space and time. Concerning the selected snapshots, the sampling frequency must be somewhat large compared to the largest frequency appearing in (1.3) and the sampled timespan be somewhat large compared to $2\pi/\omega_{\min}$, where ω_{\min} is the smallest frequency in (1.3). These conditions imply, in particular, that K should be such that

$$(1.9) \quad K \gg M.$$

The good performance of the standard DMD and HODMD methods requires testing the accuracy of the reconstructed snapshots via the approximation (1.3) obtained by these methods. However, the aim is not using DMD-like methods as a data-processing tool, just to obtain good data-reconstructions, but to uncover the underlying dynamics. The difference between both goals will be made clear in this paper (see section 3.1) noting that high-order (namely, exhibiting very small amplitudes a_m) DMD modes in (1.3) may be useful to obtain good data reconstructions, but still be associated with “errors” resulting from the finite sampled timespan, the finite sampling frequencies, and/or noise. In other words, these small amplitude modes could be dynamically spurious. Thus, accuracy is not enough, but robust-

ness (against changes in the sampling frequency, the sampled time interval, and the index d) is also necessary to identify the dynamically relevant DMD modes.

On the other hand, HODMD may be used to recover the relevant frequencies in periodic and quasi-periodic dynamics using a very limited amount of noisy data. This is because, as already mentioned, HODMD takes advantage, not only of the spatial redundancies (via the truncated SVD step that is performed at the outset), but also of the temporal redundancies (through the time-lagged snapshots). This is of interest to, e.g., minimizing the number of sensors in experimental tests (e.g., of accelerometers in aeroelastic wind tunnel and flight tests [30]). In fact, standard DMD is somewhat similar to the so-called autoregressive moving average (ARMA) method [35], which is already used in aeroelasticity. It would be interesting to know how the time-lagged snapshots considered in the HODMD method improves ARMA, though this application is well beyond the scope of this paper.

The examples and applications below are performed using MATLAB. Sufficient details are given in all cases, allowing the reader to reconstruct and check results.

The remainder of this paper is organized as follows. The HODMD method is introduced in section 2, and the new method is developed in the context of the standard DMD method. The new ideas and the performance of the method are illustrated in section 3, where several examples are considered that include both simple toy models (in section 3.1 and low-dimensional systems (in sections 3.2 and 3.3). In particular, for the Lorenz system considered in section 3.2, HODMD will be used to recover the final attractors from transient dynamics. The more complex periodic and quasi-periodic attractors appearing in the complex Ginzburg–Landau equation (CGLE) will be considered in section 4, where the ability of HODMD to recover the relevant frequencies in complex dynamics using a limited amount of noisy data will also be tested and compared with FFT and PSD. The HODMD method will be applied in section 5 to the three-dimensional thermal convection in a rotating spherical shell subject to a radial gravity field, a problem that is of fundamental interest in geophysical and astrophysical fluid dynamics [17]. This paper ends with some concluding remarks in section 6.

2. Higher order dynamic mode decomposition. For convenience, the standard DMD method is first revisited and reformulated in the spirit of this paper. Then the standard DMD is extended using delayed snapshots, which gives the HODMD method. Below, we consider several *snapshot matrices* (whose columns are snapshots), denoted as $V_{k_1}^{k_2} = [\mathbf{v}_{k_1}, \mathbf{v}_{k_1+1}, \dots, \mathbf{v}_{k_2}]$. For instance, the full snapshot matrix is

$$V_1^K = [\mathbf{v}_1, \dots, \mathbf{v}_K].$$

The aim is to calculate the DMD expansion

$$(2.1) \quad \mathbf{v}_k \simeq \mathbf{v}_k^{\text{DMD}} \equiv \sum_{m=1}^M a_m \mathbf{u}_m e^{(\delta_m + i\omega_m)(k-1)\Delta t} \quad \text{for } k = 1, \dots, K,$$

in a robust and precise way.

2.1. Standard DMD revisited and reformulated: The DMD-1 algorithm. Standard DMD relies on the assumption (1.2), namely

$$(2.2) \quad \mathbf{v}_{k+1} \simeq \mathbf{R} \mathbf{v}_k \quad \text{for } k = 1, \dots, K-1,$$

which can be written in matrix form (in terms of snapshot matrices) as

$$(2.3) \quad \mathbf{V}_2^K \simeq \mathbf{R} \mathbf{V}_1^{K-1}.$$

These approximate equations, and many other appearing below, could be written as exact equations with a residual. For instance, (2.4) could be written as $\mathbf{V}_1^K = \mathbf{U} \mathbf{\Sigma} \mathbf{T}^\top + \mathbf{R}$, where the residual depends on the neglected singular values, $\sigma_{N+1}^2 + \dots$, and associated modes, and is readily seen to exhibit a Frobenius norm $\|\mathbf{R}\|_{\text{Fro}} = \sqrt{\sigma_{N+1}^2 + \dots}$. Likewise, the norm of the residual in the remaining approximate equations could be estimated noting that these equations are obtained upon direct application of truncated SVD (for which well known error estimates are available [25]) or the pseudoinverse (which also relies on SVD). Moreover, these estimates could be used to obtain an a priori error estimate for the snapshots reconstruction via the DMD expansion (2.1). However, none of these will be done below to avoid both a too involved analysis and fairly messy equations, which would divert from the main focus of the paper. In any event, standard DMD and HODMD are both postprocessing methods, and the actual reconstruction error can well be calculated a posteriori.

The (somewhat straightforward and fairly similar to the method proposed by Schmid [48]) method to calculate the standard DMD expansion is now described. The method proceeds in three steps, considered in the following subsections.

2.1.1. Step 1: Dimension reduction. This step is performed by applying *truncated SVD* (implemented in the MATLAB command “svd”, option “econ”) to the full snapshot matrix, as

$$(2.4) \quad \mathbf{V}_1^K \simeq \mathbf{U} \mathbf{\Sigma} \mathbf{T}^\top, \quad \text{with } \mathbf{U}^\top \mathbf{U} = \mathbf{T}^\top \mathbf{T} = \text{the } N \times N \text{ unit matrix},$$

where $\mathbf{\Sigma}$ is the diagonal matrix containing the retained SVD singular values sorted in decreasing order, $\sigma_1, \sigma_2, \dots, \sigma_N$, and the (orthonormal) columns of the $J \times N$ matrix \mathbf{U} and the $K \times N$ matrix \mathbf{T} are the *spatial and temporal SVD-modes*, respectively. Invoking well known SVD formulae [25], the number N of retained modes is selected in terms of the rank of the snapshots matrix, $R \leq \min\{J, K\}$, to be such that

$$(2.5) \quad \text{EE}(N) \equiv \frac{\sigma_{N+1}^2 + \dots + \sigma_R^2}{\sigma_1^2 + \dots + \sigma_R^2} \leq \varepsilon_1,$$

where $\text{EE}(N)$ is the relative root mean square (RMS) error of the approximation (2.4). The threshold ε_1 is tunable and can be selected in view of the singular values distribution, using the error estimate (2.5). The default value could be $\varepsilon_1 = 10^{-8}$.

Now, (2.4) can also be written as

$$(2.6) \quad \mathbf{V}_1^K \simeq \mathbf{U} \hat{\mathbf{V}}_1^K,$$

where the *reduced snapshot matrix* is defined as

$$(2.7) \quad \hat{\mathbf{V}}_1^K = \mathbf{\Sigma} \mathbf{T}^\top (\equiv \mathbf{U}^\top \mathbf{V}_1^K).$$

The rows of the $N \times K$ -matrix $\hat{\mathbf{V}}_1^K$ are proportional to the SVD temporal modes and can thus be seen as *rescaled temporal modes*. Rescaling is important to get consistent results in the dimension-reduced formulation; see below. The columns of $\hat{\mathbf{V}}_1^K$ will be called the *reduced snapshots* and exhibit a much smaller dimension than the original snapshot matrix if $N \ll J$ (the usual case when simulating infinite-dimensional systems). Thus, it is this reduced snapshot matrix that will be used below in all computations. In particular, (2.6) implies that the snapshots and reduced snapshots are such that

$$(2.8) \quad \mathbf{v}_k \simeq \mathbf{U} \hat{\mathbf{v}}_k.$$

Similarly, the counterpart of the DMD expansion (2.1) for the reduced snapshots is

$$(2.9) \quad \hat{\mathbf{v}}_k \simeq \hat{\mathbf{v}}_k^{\text{DMD}} \equiv \sum_{m=1}^M a_m \hat{\mathbf{u}}_m e^{(\delta_m + i\omega_m)(k-1)\Delta t} \quad \text{for } k = 1, \dots, K,$$

with

$$\mathbf{u}_m = \mathbf{U} \hat{\mathbf{u}}_m.$$

2.1.2. Step 2: Computation of the reduced Koopman matrix and the DMD modes.

Premultiplying (2.2) and (2.3) by \mathbf{U}^T and invoking (2.8) yields

$$(2.10) \quad \hat{\mathbf{v}}_{k+1} \simeq \hat{\mathbf{R}} \hat{\mathbf{v}}_k \quad \text{or} \quad \hat{\mathbf{V}}_2^K \simeq \hat{\mathbf{R}} \hat{\mathbf{V}}_1^{K-1} \quad \text{with } \hat{\mathbf{R}} = \mathbf{U}^\top \mathbf{R} \mathbf{U}.$$

The $N \times N$ -matrix $\hat{\mathbf{R}}$ will be called the *reduced Koopman matrix* and is calculated using the pseudoinverse, as follows. Standard (no truncation) SVD applied to the matrix $\hat{\mathbf{V}}_1^{K-1}$ leads to

$$(2.11) \quad \hat{\mathbf{V}}_1^{K-1} = \hat{\mathbf{U}} \hat{\mathbf{\Sigma}} \hat{\mathbf{T}}^\top,$$

where, assuming that $N \leq K-1$ (see (1.9)), the $N \times N$ matrix $\hat{\mathbf{U}}$ and the $(K-1) \times N$ matrix $\hat{\mathbf{T}}$ are such that $\hat{\mathbf{U}}^\top \hat{\mathbf{U}} = \hat{\mathbf{U}} \hat{\mathbf{U}}^\top = \text{unit matrix}$, $\hat{\mathbf{T}}^\top \hat{\mathbf{T}} = \text{unit matrix}$, and the $N \times N$ diagonal matrix $\hat{\mathbf{\Sigma}}$ is nonsingular and (because of the rescaling (2.7)) is very close to its counterpart in (2.4), $\mathbf{\Sigma}$. Substituting (2.11) into (2.10) and postmultiplying the resulting equation by $\hat{\mathbf{T}} \hat{\mathbf{\Sigma}}^{-1} \hat{\mathbf{U}}^\top$ leads to

$$\hat{\mathbf{R}} = \hat{\mathbf{V}}_2^K \hat{\mathbf{T}} \hat{\mathbf{\Sigma}}^{-1} \hat{\mathbf{U}}^\top.$$

Now, once the reduced snapshot matrix $\hat{\mathbf{R}}$ has been calculated, its eigenvectors \mathbf{q}_m , rescaled as

$$(2.12) \quad \|\mathbf{q}_m\|/\sqrt{J} = 1,$$

and eigenvalues μ_m readily yield the modes, growth rates, and frequencies appearing in the DMD expansions (2.1) and (2.9) as

$$(2.13) \quad \hat{\mathbf{u}}_m = \mathbf{q}_m, \quad \mathbf{u}_m = \mathbf{U} \mathbf{q}_m, \quad \delta_m + i\omega_m = \frac{1}{\Delta t} \log \mu_m.$$

The rescaling (2.12) means that the RMS of the modes is 1, which makes the mode amplitudes a_m as independent as possible from J . The amplitudes a_m are left aside at the moment.

Now, this calculation is readily seen to differ from that in [48, p. 9] only in one point. The projection matrix \mathbf{U} is calculated here in (2.4) by applying SVD to the whole snapshot matrix \mathbf{V}_1^K , while Schmid [48] calculates \mathbf{U} by applying SVD to the smaller snapshot matrix \mathbf{V}_1^{K-1} (ignoring the last snapshot). If the number of snapshots $K \gg 1$ (the usual case), then both calculations give very similar results. In fact, the method in [48, p. 9] gives almost identical results as DMD-1 in all calculations performed in this paper and will not be further considered.

2.1.3. Step 3: Computation of the DMD mode amplitudes. The DMD amplitudes are now calculated using the reduced DMD expansion (2.9), which, invoking (2.13), can also be written as

$$(2.14) \quad \hat{\mathbf{v}}_k \simeq \sum_{m=1}^M a_m \mathbf{q}_m \mu_m^{k-1} \quad \text{for } k = 1, \dots, K.$$

Since the dimension of the (known) reduced snapshots $\hat{\mathbf{v}}_k$ is N ($= M$ in the present case) and the eigenvectors \mathbf{q}_m are linearly independent, each of these equations uniquely determines the amplitudes a_m . The whole system of equations, instead, is highly overdetermined if $K \gg 1$ (the usual case), but can be solved using the pseudoinverse, which represents a minimization of the least-squares-error in the approximation (2.14) and is essentially equivalent to the so-called optimized DMD method [18]. This step is performed by rewriting (2.14) in matrix form as

$$(2.15) \quad \mathbf{L} \mathbf{a} = \mathbf{b},$$

where the $(NK \times M)$ -matrix \mathbf{L} , the unknown amplitudes vector \mathbf{a} , and the forcing term \mathbf{b} are defined as

$$(2.16) \quad \mathbf{L} = \begin{bmatrix} \mathbf{Q} \\ \mathbf{Q} \mathbf{M} \\ \vdots \\ \mathbf{Q} \mathbf{M}^{K-1} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \hat{\mathbf{v}}_1 \\ \hat{\mathbf{v}}_2 \\ \vdots \\ \hat{\mathbf{v}}_K \end{bmatrix}.$$

Here, \mathbf{L} is exactly the *observability matrix* for a linear system with observation given by the $M \times M$ -matrix $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_M]$ (formed by the eigenvectors) and dynamics given by the $M \times M$ diagonal matrix \mathbf{M} (formed by the associated eigenvalues, μ_1, \dots, μ_M). The pseudoinverse is calculated by applying standard SVD (no truncation) to the matrix \mathbf{L} as

$$(2.17) \quad \mathbf{L} = \tilde{\mathbf{U}}_1 \tilde{\Sigma} \tilde{\mathbf{U}}_2^\top,$$

with the $(NK \times M)$ -matrix $\tilde{\mathbf{U}}_1$ and the $M \times M$ -matrix $\tilde{\mathbf{U}}_2$ such that

$$(2.18) \quad \tilde{\mathbf{U}}_1^\top \tilde{\mathbf{U}}_1 = \tilde{\mathbf{U}}_2^\top \tilde{\mathbf{U}}_2 = \tilde{\mathbf{U}}_2 \tilde{\mathbf{U}}_2^\top = \text{the } M \times M \text{ unit matrix.}$$

Substituting (2.17) into (2.15), premultiplying the resulting equation by $\tilde{\mathbf{U}}_2 \tilde{\Sigma}^{-1} \tilde{\mathbf{U}}_1^\top$, and invoking (2.18) yields

$$(2.19) \quad \mathbf{a} = \tilde{\mathbf{U}}_2 \tilde{\Sigma}^{-1} \tilde{\mathbf{U}}_1^\top \mathbf{b}.$$

As a final step, the mode amplitudes are all set ≥ 0 by changing sign in both a_m and \mathbf{q}_m when needed in the expansion (2.14). This completes the calculation of the amplitudes, which is usually the most computationally expensive step if $K \gg 1$ (the usual case).

Finally, the DMD modes can be reordered sorting the mode amplitudes in decreasing order, and a smaller number of modes, M , may be selected imposing that

$$(2.20) \quad a_{M+1}/a_1 < \varepsilon,$$

for some tunable (small) threshold ε .

2.1.4. Summary of the DMD-1 method. Summarizing, the following algorithm has been developed to apply standard DMD to the snapshots (1.1). As a first step, truncated SVD is applied to the snapshot matrix (see (2.4)), with the number of retained terms as defined in (2.5), for some tunable threshold ε_1 . Then, the reduced snapshot matrix $\hat{\mathbf{V}}_1^K$ is calculated using (2.6)–(2.7). The M eigenvectors \mathbf{q}_m and eigenvalues μ_m of $\hat{\mathbf{R}}$ and the pseudoinverse solution of (2.16) (with the additional truncation of the DMD modes performed using (2.20)) yield the ingredients of the DMD-expansion (2.1) using (2.13) and (2.19).

For convenience, this algorithm will be labeled as the *DMD-1 algorithm*. Since this algorithm relies on the assumption (2.10) (which is typically not known a priori), consistency of the results requires that

$$(2.21) \quad \left\| \hat{\mathbf{V}}_2^K - \hat{\mathbf{R}} \hat{\mathbf{V}}_1^{K-1} \right\|_2 / \left\| \hat{\mathbf{V}}_2^K \right\|_2 \ll 1.$$

If this condition does not hold, then standard DMD gives a spurious approximation and the more general HODMD method, considered in the next subsection, should be used.

2.2. The higher order dynamic mode decomposition: DMD- d algorithm with $d > 1$.

As anticipated in the introduction, HODMD relies on the higher order Koopman condition (1.5), namely

$$(2.22) \quad \mathbf{v}_{k+d} \simeq \mathbf{R}_1 \mathbf{v}_k + \mathbf{R}_2 \mathbf{v}_{k+1} + \cdots + \mathbf{R}_d \mathbf{v}_{k+d-1} \quad \text{for } k = 1, \dots, K-d.$$

This more general condition is now treated in a similar way as we did with the assumption (2.2) in section 2.1, in three steps described in the following subsections.

2.2.1. Step 1: Dimension reduction. To begin with, we perform exactly the same dimension reduction developed in step 1 of the DMD-1 method (see section 2.1.1), using (2.4)–(2.7) with a convenient threshold ε_1 , which yields exactly the same projection equation, namely (2.8). This is rewritten here for convenience as

$$(2.23) \quad \mathbf{v}_k \simeq \mathbf{U} \hat{\mathbf{v}}_k.$$

Using this, the counterpart of (2.22) in the reduced linear manifold is

$$(2.24) \quad \hat{\mathbf{v}}_{k+d} \simeq \hat{\mathbf{R}}_1 \hat{\mathbf{v}}_k + \hat{\mathbf{R}}_2 \hat{\mathbf{v}}_{k+1} + \cdots + \hat{\mathbf{R}}_d \hat{\mathbf{v}}_{k+d-1} \quad \text{for } k = 1, \dots, K-d.$$

As anticipated in the introduction, this equation can be seen as a modified Koopman equation and written as

$$\tilde{\mathbf{v}}_{k+1} \simeq \tilde{\mathbf{R}} \tilde{\mathbf{v}}_k,$$

where the *modified snapshots* $\tilde{\mathbf{v}}_k$ and the *modified Koopman matrix* $\tilde{\mathbf{R}}$ are

$$(2.25) \quad \tilde{\mathbf{v}}_k \equiv \begin{bmatrix} \hat{\mathbf{v}}_k \\ \hat{\mathbf{v}}_{k+1} \\ \dots \\ \hat{\mathbf{v}}_{k+d-2} \\ \hat{\mathbf{v}}_{k+d-1} \end{bmatrix}, \quad \tilde{\mathbf{R}} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \\ \hat{\mathbf{R}}_1 & \hat{\mathbf{R}}_2 & \hat{\mathbf{R}}_3 & \dots & \hat{\mathbf{R}}_{d-1} & \hat{\mathbf{R}}_d \end{bmatrix}.$$

Here, \mathbf{I} and $\mathbf{0}$ are the $N \times N$ unit and zero matrices, respectively. Note that the modified snapshots $\tilde{\mathbf{v}}_k$ are of dimension dN and the modified Koopman matrix $\tilde{\mathbf{R}}$ is a $(dN \times dN)$ -matrix, which may be fairly large but still reasonable. In fact, this matrix will not be calculated below because the set of snapshots $\tilde{\mathbf{v}}_k$ will be further dimension-reduced in the next step. Because d and K are usually fairly large, this step is usually fairly computationally expensive. Nonetheless, the spatial dimension reduction performed above (see (2.23)) has been essential because if this were not performed, the dimension of the counterpart of the modified snapshot matrix (see (2.26)) would be a $(Jd \times Jd) \times (K-d+1)$ -matrix. For typical applications in spatially three-dimensional fluid flows, $J \sim 10^7$. Since $K \sim 1000$ and $d \sim 100$ for typical quasi-periodic dynamics in infinite-dimensional systems (see below), the method would be absolutely impractical without the spatial dimension reduction. In addition, the spatial dimension reduction will not permanently remove information since Takens' delay-embedding theorem guarantees that the full state can be reconstructed from any generic observable. So, introducing the lagged snapshots in the next section will reconstruct the variables necessary to reconstruct the state.

2.2.2. Step 2: Computation of the DMD modes. The *DMD- d method* proceeds by applying steps 1 and 2 in the DMD-1 method (see sections 2.1.1 and 2.1.2, respectively) to the modified snapshot matrix (with $\tilde{\mathbf{v}}_k$ as defined in (2.25))

$$(2.26) \quad \tilde{\mathbf{V}}_1^{K-d+1} = [\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_{K-d+1}],$$

which is of order $(dN) \times (K-d+1)$. Note that when step 1 in the DMD-1 method is applied to this snapshot matrix, then a dimension reduction results, which is an additional dimension reduction to that performed in step 1 above and is made using the counterpart of (2.5), namely

$$(2.27) \quad \text{EE}(N) \equiv \frac{\sigma_{N+1}^2 + \cdots + \sigma_R^2}{\sigma_1^2 + \cdots + \sigma_R^2} \leq \varepsilon_1 \quad \text{with } R = \min\{dN, k-d+1\}.$$

The threshold ε_1 may coincide with its counterpart in the first dimension reduction performed in step 1 above (see section 2.2.1) or not. The outcome of DMD-1, step 2 (using (2.13)) is a set

of DMD modes, $\tilde{\mathbf{q}}_m$, growth rates, δ_m , and frequencies ω_m . Consistently with the definition (2.25) of the modified snapshots to whom DMD-1 has been applied, the N first components of $\tilde{\mathbf{q}}_m$ (rescaled as indicated in (2.12)) give the reduced DMD-modes \mathbf{q}_m for the reduced DMD expansion

$$(2.28) \quad \hat{\mathbf{v}}_k \simeq \hat{\mathbf{v}}_k^{\text{DMD}} \equiv \sum_{m=1}^M a_m \mathbf{q}_m e^{(\delta_m + i\omega_m)(k-1)\Delta t} \quad \text{for } k = 1, \dots, K,$$

which is the counterpart of (2.1). The growth rates and frequencies are precisely those computed in the application of the DMD method to the modified snapshot matrix (2.26).

2.2.3. Step 3: Computation of the DMD mode amplitudes. Once the reduced DMD modes, \mathbf{q}_m , growth rates δ_m and frequencies ω_m have been calculated, the DMD amplitudes a_m are calculated precisely as we did for the DMD-1 algorithm in section 2.1.1. Namely, the amplitudes in (2.28) are calculated by rewriting (2.28) in the form (2.14) and calculating the amplitudes vector via the pseudoinverse of the matrix \mathbf{L} . A further truncation using (2.20) is also performed. This completes the derivation of the reduced DMD expansion (2.28). As in the DMD-1 method, this is usually the most computationally expensive step.

2.2.4. Summary of the DMD- d method, with $d > 1$. Summarizing the above, given a set of snapshots \mathbf{v}_k and an index $d > 1$, the snapshots are dimension reduced as explained in step 1 (see section 2.2.1), which gives the reduced snapshots $\hat{\mathbf{v}}_k$ defined in (2.23). The reduced modes \mathbf{q}_m , growth rates δ_m , and frequencies ω_m appearing in the reduced DMD expansion (2.28) are calculated as explained in section 2.2.2 and the mode amplitudes a_m , as explained in section 2.2.3. Premultiplying the reduced expansion (2.28) by the projection matrix \mathbf{U} appearing in (2.23) yields the DMD expansion (2.1), with $\mathbf{u}_m = \mathbf{U} \mathbf{q}_m$.

2.3. Practical implementation and calibration of the DMD- d algorithm. The DMD-1 and DMD- d algorithms developed above are straightforwardly implemented in MATLAB. In fact, as developed, the methods (and all formulae) are applicable to both real and complex data, as automatically performed in MATLAB by using the command $'$ for the transpose $^\top$. As anticipated, these algorithms can be used, in principle, either (i) to reconstruct the set of given snapshots (using the method as a data processing tool) or (ii) to uncover the underlying dynamics, which is more subtle. The goal (ii) requires identifying the right damping rates and frequencies and permits extrapolating the identified dynamics for large t .

Whatever the goal, these methods exhibit various tunable parameters, namely, the sampled time interval, the number of snapshots K , the threshold ε_1 used in SVD truncation (see (2.5)), the threshold ε defining the retained modes (see (2.20)), and the index d for the DMD- d method, which should be chosen in each particular application. Note that condition (2.21) is a very useful means to elucidate whether the standard method is appropriate or not. When DMD-1 is not appropriate, the calibration of the index $d > 1$ deserves some attention. Obviously, d must not be too close to 1. On the other hand, when applying the DMD- d method, the modified snapshots matrix (2.26) contains $K - d$ modified snapshots (of size dN), meaning that $d > 1$ must not be too close to K , to avoid that only a small number of modified snapshots are really used by the method. As a consequence, d must be chosen in an appropriate interval that, fortunately, is usually fairly wide, as it will be repeatedly seen

below. The appropriate values of d obviously depend on the involved dynamics, the sampled time interval, and the sampling frequency. For a given dynamics and sampled interval, d scales with the sampling frequency, namely with the number of snapshots K for a given sampled timespan. In other words, if K is doubled, d should be doubled, too. This conclusion is further justified in Appendix A.

If the method is to be used as a purely data processing tool, the reconstruction error should be made as small as possible. The accuracy of the reconstruction will be measured below in terms of the relative RMS error, defined as

$$(2.29) \quad \text{RRMSE} \equiv \left[\frac{\sum_{k=1}^K \|\mathbf{v}_k - \mathbf{v}_k^{\text{DMD}}\|_2^2}{\sum_{k=1}^K \|\mathbf{v}_k\|_2^2} \right]^{1/2}.$$

If, instead, the method is used to capture the relevant dynamics, then it must be ensured that the approximation captures well the dynamically relevant modes, which could, of course, be identified beforehand by other means (i.e., physical relevance). However, in the absence of a priori information, the relevant modes can be identified elucidating the consistency of the results when the parameters of the method are varied. In particular, consistency requires that the retained frequencies and growth rates be stable when the sampled time interval is shifted or enlarged and the index d is varied. In other words, the DMD-d parameters can be tuned using standard statistical cross-validation methods. This will be the means to distinguish between the relevant and spurious modes below. It must be kept in mind that spurious modes are to be expected due to the unavoidable errors, which in practice come from two sources:

- Numerical and experimental errors that are already present in the snapshots. If the RMS level of these errors is known beforehand, then they can be filtered out by an appropriate selection of the threshold ε_1 appearing in (2.5), taking advantage of the well known [57] error-filtering ability of POD/SVD. If, instead, the error level is not known, but the errors are uncorrelated with the physically meaningful data, then they will promote a change of tendency in the usual semilogarithmic plot of the singular values of the primary SVD performed in (2.5) versus the retained number of modes. This change of tendency can help to identify the error level; see Figure 1.
- Errors that are promoted by the method itself, such as round-off or truncation errors, finite-sampled-timespan errors, finite-sampling-frequency errors, and errors resulting from the various pseudoinverse calculations above. Because of these errors, the appropriate number of retained DMD modes in the DMD expansions will not generally coincide with its maximum possible number. It is precisely because of this that the last truncation (2.20) was performed. The threshold ε in this truncation should be selected attending to the robustness of the results when the various parameters of the method are varied, as explained above.

Once the dynamically relevant modes have been identified, they naturally provide the structure of the attractor in permanent dynamics. Also, in transient dynamics, the relevant DMD-expansion (1.3) can be used to *extrapolate to* $t \gg 1$, which gives an approximation of the final attractor. The extrapolation is performed by just ignoring in (1.3) those DMD modes with negative δ_m , which generally gives a smaller number of *asymptotic modes*, M_∞ . It must be noted that, because of errors, some relevant asymptotic modes may exhibit a small-but-

nonzero $|\delta_m|$. Thus, a threshold $\varepsilon_2 \ll 1$ is defined and the growth rate of those modes with $|\delta_m| < \varepsilon_2$ is set to zero before performing extrapolation. When comparing two extrapolations, a relative time-shift t_0 may be present, which must be dealt with as explained in the next paragraph.

Consistency between DMD- d reconstructions of attractors using different snapshots sets is problematic because of the possible time-shifts that naturally appear. In other words, two DMD reconstructions can only be required to be close to each other up to a time shift. For periodic dynamics, the shift must be smaller than the period and it can be calculated upon (nonlinear) least squares fit. For quasi-periodic attractors, instead, the time shift can be extremely large and the least square fit be impractical. Thus, in these more involved cases, the comparison will be made below only in terms of the DMD amplitudes and frequencies.

3. Illustration of the HODMD method. Let us now illustrate the HODMD method in both several toy model dynamics and two low-dimensional systems: the Lorenz system and the Stuart–Landau equation.

3.1. Some toy model dynamics. A simple but quite illustrative application of the methods proceeds as follows: Select randomly N orthonormal complex vectors of dimension $J > N$, $\mathbf{w}_1, \dots, \mathbf{w}_N$, and construct $M > N$ linear combinations,

$$(3.1) \quad \mathbf{u}_m = \sum_{j=1}^N \alpha_j \mathbf{w}_j,$$

with randomly selected coefficients α_j , such that the Euclidean norm of the vectors \mathbf{u}_m is \sqrt{J} (see (2.12)). Here, we are assimilating the various components of \mathbf{u} as “spatial coordinates.”

With these, construct the following toy-model dynamics:

$$\mathbf{u}(t) = \sum_{m=1}^M a_m \mathbf{u}_m e^{i\omega_m t},$$

with the amplitudes a_m and the frequencies ω_m selected below. Because of the random character of this example (which has been included to avoid any bias in the example), the performance of the methods will slightly vary from one run to another, but the results will be essentially consistent. Now, $K > M$ equispaced snapshots are selected in the interval $0 \leq t \leq 1$. Note that these $J \times K$ data are a “simple” instance of the DMD expansion (2.1) that, by construction, exhibits spatial and spectral complexities exactly equal to N and M , respectively. Once the various random selections above have been fixed, several interesting cases to illustrate the methods will be considered by appropriately varying the integers J , K , M , and N .

To begin with, we set $J = 100$, $K = 1000$, $M = N = 10$ (i.e., equal spectral and spatial complexities), and select the amplitudes and frequencies as

$$(3.2) \quad a_m = 10^{-m/2}, \quad \omega_m = 10m \quad \text{for } m = 1, \dots, M.$$

Note that the sampling frequency is $K = 1000$, namely ten times larger than the largest frequency, $\omega_1 = 100$. Since $M = N$, the spectral and spatial complexities coincide and the

standard DMD-1 algorithm should do a good job. In fact, setting

$$(3.3) \quad \varepsilon_1 = 10^{-6}, \quad \varepsilon = 10^{-6},$$

the DMD-1 algorithm calculates the frequencies with a relative error 10^{-13} and reconstructs the snapshots with relative RMS error, as defined in (2.29), $\text{RRMSE} \sim 10^{-12}$. The algorithm DMD- d provides slightly better results for d in the range $1 < d < 400$.

Now, we maintain the parameters of the method, (3.3), the spectral complexity $M = 10$, and the amplitudes and frequencies defined in (3.2), but decrease the spatial complexity to $N = 5$. Since $N = 5 < M = 10$, the algorithm DMD-1 is only able to recognize five modes, with frequencies

$$\omega_1 = 55.85, \quad \omega_2 = 48.28, \quad \omega_3 = 27.40, \quad \omega_4 = 93.03, \quad \omega_5 = 66.69,$$

which compared with (3.2) are seen to be spurious; the reconstruction relative error RRMSE is $\sim 10^{-2}$. Applying the HODMD algorithm DMD- d , with $d = 50$ (which is somewhat optimal for this case), instead, recovers both the $M = 10$ frequencies defined in (3.2) with a relative error $\sim 10^{-13}$ and the snapshots within a relative error $\text{RRMSE} \sim 10^{-13}$. Slightly increasing and decreasing d (to, e.g., 80 and 20, respectively) the very small reconstruction errors are essentially maintained.

Let us now add some noise to the data. Specifically, we add uniformly distributed positive noise (using the MATLAB command *rand*) with size $\varepsilon_{\text{noise}} = 5 \cdot 10^{-7}$. Note that the noise is positive and thus exhibits a nonzero mean. Repeating the two cases considered above to this noisy data, with $\varepsilon_1 = \varepsilon = 10^{-6}$ (both two times larger than the noise), the following results are obtained:

- a. For the case $M = N = 10$, the relative RMS reconstruction (comparing with the clean solution) is $\text{RRMSE} \sim 10^{-4}$ and $\sim 10^{-7}$ using the algorithms DMD-1 and DMD-300 (note that now the convenient value of d is larger than in the clean case), respectively. Thus, both the standard DMD and HODMD are able to filter errors out (despite the fact that the errors are positive), though HODMD gives much better results.
- b. For the case $M = 10$, $N = 5$, DMD-1 fails, as above, since it only recognizes five modes and reconstructs the solution with $\text{RRMSE} \sim 10^{-2}$. DMD-300 (note that d is larger now than for clean data), instead, recovers the ten frequencies with a relative error $\sim 10^{-9}$ and reconstructs the solution with $\text{RRMSE} \sim 10^{-7}$ (which is fairly smaller than the added noise!). Again, HODMD is quite robust in connection with varying the index d , and the results improve if the sampled interval and/or the number of snapshots is increased.

As expected, the noise reduces the performance of both methods, but the new HODMD method gives far more robust estimates and reconstructions. This is because, as anticipated, HODMD takes full advantage of both the spatial and temporal redundancies in the given data. Note that the threshold $\varepsilon = 10^{-6}$ has been chosen as somewhat larger than the error level, $5 \cdot 10^{-7}$. If the latter were not known in advance, then it could be guessed from the singular value distribution in the SVD that is performed in the method. For instance, in the two cases (a and b) considered above, using DMD-1 and DMD-300, the singular value distributions are plotted in Figure 1. As can be seen, a change of tendency occurs in a region below 10^{-6} ,

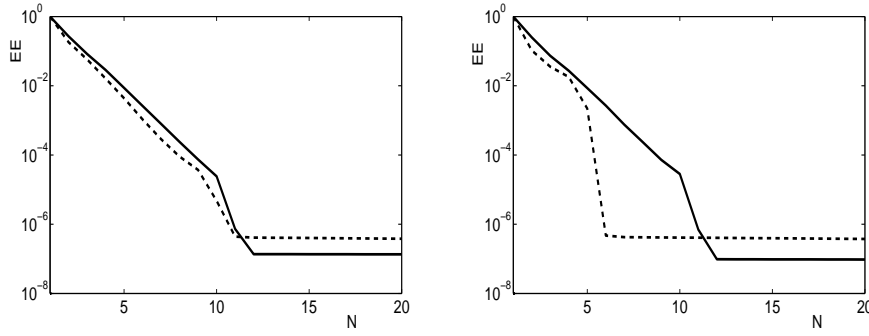


Figure 1. The error estimate $EE(N)$ defined in (2.5) and (2.27) versus the retained number of modes N when applying DMD-1 (dashed lines) and DMD-300 (solid lines), respectively, to the noisy databases considered in items a, with $M = N$ (left) and b, with $M < N$ (right) above.

meaning that the selected threshold was a reasonable choice.

The dynamics considered in (3.2) are periodic, but the conclusions above stand for quasi-periodic dynamics, involving incommensurable frequencies and defining a linear flow in a torus. For instance, replacing the frequencies and amplitudes defined in (3.2) by

$$a_m = 10^{-\omega_m/30}, \quad \omega_m = 30(m_1 - 1) + 7\sqrt{3}(2m_2 - 1),$$

with $m_1, m_2 = 1, 2, \dots$ chosen such that $\omega_1 < \omega_2 < \omega_3 < \dots$. Except for using these new frequencies, the application is identical to the above application for the periodic flow and the results are completely similar, even quantitatively. In particular, setting $J = 100$, $K = 1000$, and $M = N = 10$, and selecting the thresholds (3.3), both the standard algorithm DMD-1 and DMD- d , with $1 < d \leq 400$ recognize the ten involved frequencies with a relative error 10^{-13} and reconstruct the snapshots with a relative RMS error $RRMSE \sim 10^{-12}$. If, instead, we set $M = 10$ and $N = 5$, DMD-1 gives five frequencies that are spurious, but DMD-50 still recovers the ten involved frequencies and reconstructs the snapshots with $RRMSE \sim 10^{-12}$. Also, adding noise of the same size as above, for the case $M = N = 10$, the algorithms DMD-1 and DMD-300 reconstruct the snapshots with $RRMSE \sim 10^{-6}$ and $\sim 10^{-7}$, respectively, while for the noisy snapshots with $M = 10$ and $N = 5$, the reconstruction errors using DMD-1 and DMD-300 is $RRMSE \sim 10^{-1}$ and $\sim 10^{-6}$, respectively. In summary, quasi-periodic dynamics is not more demanding than periodic dynamics for the DMD- d algorithm with appropriate d .

Decaying/growing dynamics are treated similarly. As it happens with standard DMD (when the method works because the spatial complexity equals the spectral complexity), HODMD works equally well in the decaying case for arbitrary spatial and spectral complexities.

The dynamics outcome (3.1) somewhat mimics linear dynamics. Let us now turn to the fully nonlinear case, considering one-dimensional data, namely with $J = 1$, which is not accessible to standard DMD. Fully nonlinear driving very easily leads to a large number of frequencies, which is now illustrated in the following cases, in which HODMD can be seen as a very advantageously alternative to FFT. The advantages are that the method determines growth rates (in addition to frequencies) and that the sampled timespan can be comparable to

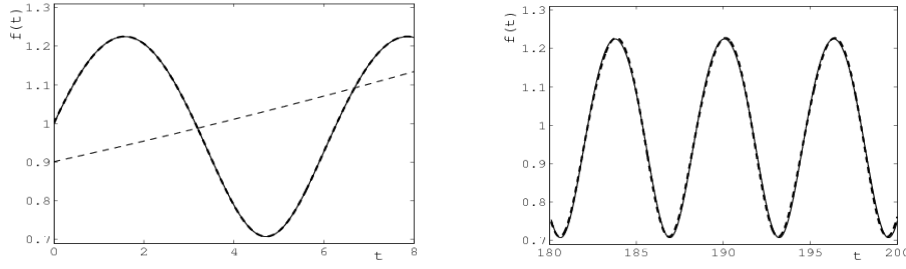


Figure 2. The function defined in (3.4) (thin solid lines) in the indicated time intervals. The reconstruction using DMD-270 in $0 \leq t \leq 8$ (left) and extrapolation (right) are plotted thick dashed lines, and the spurious reconstruction using DMD-1, with thin dashed line.

the largest period (which does not need to be known in advance) without introducing sideband artifacts. These examples are also used to get some insights on the calibration of the methods.

We consider the fully nonlinear 2π -periodic function

$$(3.4) \quad f(t) = \sqrt{1 + \frac{1}{2} \sin t}$$

and take $K = 1000$ snapshots in the sampled time interval $0 \leq t \leq 8$ (slightly larger than one period). Setting $\varepsilon_1 = 10^{-10}$ and $\varepsilon = 10^{-3}$, the HODMD algorithm DMD-270 identifies seven frequencies with a relative error $\sim 10^{-4}$ and reconstructs the data with $\text{RRMSE} = 2.5 \cdot 10^{-4}$. Thus, the DMD-270 reconstruction is plot-indistinguishable from its exact counterpart, as seen in Figure 2-left, where the DMD-1 reconstruction is also plotted for reference. It is also interesting to note that the extrapolation of the DMD-270 expansion to the interval $180 \leq t \leq 200$ (Figure 2-right) is also plot-indistinguishable from its exact counterpart, since the RRMSE is $6.19 \cdot 10^{-3}$. Concerning robustness, it turns out that decreasing ε increases the number of identified frequencies and the accuracy of the reconstruction, as expected. The index $d = 270$ is somewhat optimal though decreasing it by 50 does not affect the outcome of the method much. On the other hand, doubling the number of snapshots to $K = 2000$ also doubles the optimal value of d , but does not produce a significant benefit. Similarly, dividing K by 2, the optimal value of d is also divided by two and does not improve the results either. Still, doubling the sampled timespan (to $0 \leq t \leq 16$), the accuracy of the seven frequencies that are identified improves significantly, although the reconstruction only slightly improves (recall that only seven modes are retained). Shifting the sampled time interval produces essentially the same results, which was to be expected for consistency. On the other hand, dividing the sampled interval by 2, to $0 \leq t \leq 4$ (smaller than a period), the method only identifies five modes, which are spurious since they exhibit the following frequencies and damping rates:

$$\begin{aligned} \delta_1 + i\omega_1 &= 0.18, & \delta_{2,3} + i\omega_{2,3} &= -0.039 \pm i0.53, \\ \delta_4 + i\omega_4 &= -1.60, & \delta_5 + i\omega_5 &= 1.36, \end{aligned}$$

but the reconstruction is fairly good, with $\text{RRMSE} = 1.8 \cdot 10^{-4}$. The fact that these frequencies are spurious may be ensured by shifting and/or enlarging the sampled time interval and getting

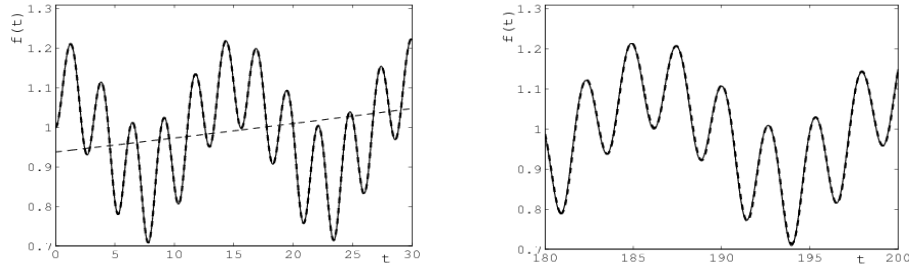


Figure 3. Counterpart of Figure 2 for the quasi-periodic function defined in (3.5).

completely different frequencies. This ill-functioning of the method is very interesting and illustrates the capacity of DMD-like methods (both DMD and HODMD) to fit given data in a given timespan, without a good approximation of the “true dynamics,” which is what allows for safely extrapolating the fitted approximation. This is why consistency and robustness is emphasized in all examples and applications below.

Let us consider the fully nonlinear, quasi-periodic function

$$(3.5) \quad f(t) = \sqrt{1 + \frac{1}{2} \sin t \sin(\sqrt{2}t)},$$

which exhibits the incommensurable frequencies $\omega = 1$ and $\omega = \sqrt{2}$, thus defining a nonlinear flow in a torus. Now, we consider a larger timespan, $0 \leq t \leq 32$ and extract $K = 2000$ equispaced snapshots. These values of the parameters of the method are somewhat optimal. We now set $\varepsilon_1 = 10^{-10}$, $\varepsilon = 10^{-4}$, and $d = 800$ (four times larger than in the previous case). The method identifies 27 correct (not spurious) frequencies and reconstructs the snapshots with $\text{RRMSE} = 2.5 \cdot 10^{-4}$. As in the former example, the reconstruction (Figure 3-left) is plot indistinguishable from its exact counterpart, as is the extrapolation (Figure 3-right) to the interval $180 \leq t \leq 200$, whose RRMSE is $= 2.8 \cdot 10^{-3}$. Note that, for a similar precision, the sampled timespan and the index d both need to be larger than for the periodic function (3.4), as could be expected, and the number of required frequencies is much larger. If the tunable parameters of the method are varied, the method behaves quite as it did in the periodic case. In particular, shifting and/or enlarging the sampled time interval produces essentially the same results (in particular, the same frequencies), as required for consistency.

As a third example, we consider the decaying toy model dynamics (cf (3.5))

$$f(t) = \sqrt{1 + \frac{\exp(-t/32)}{2} \sin t \sin(\sqrt{2}t)},$$

for which, using exactly the same values for the parameters of the method as in the previous case, 29 modes with negative damping rates are identified, giving a $\text{RRMSE} = 1.3 \cdot 10^{-4}$.

3.2. Periodic attractors in the Lorenz system. The necessity of DMD- d , with $d > 1$, and the essence of the method are clearly seen considering the well known third-order system (spatial dimension: $N = 3$)

$$(3.6) \quad v'_1 = \sigma(v_2 - v_1), \quad v'_2 = v_1(r - v_3) - v_2, \quad v'_3 = v_1 v_2 - b v_3,$$

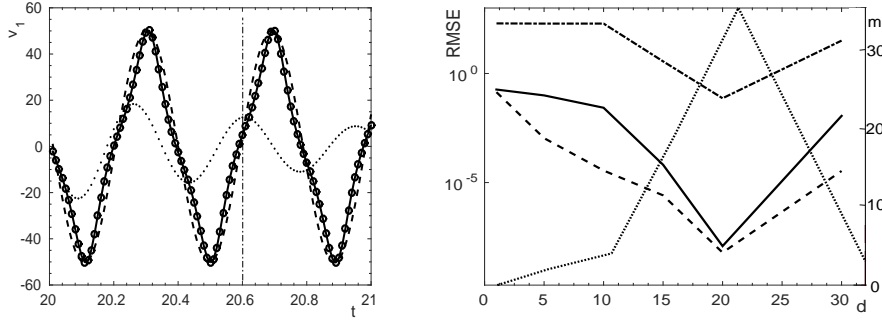


Figure 4. The Lorenz system. Left: v_1 versus t for the considered periodic orbit (solid), its monochromatic counterpart (dashed), and the reconstructions using DMD-1 (dotted) and DMD-20 (circles); the vertical line at $t = 20.6$ indicates the upper limit of the sampled interval. Right: RRMSE (as defined in (2.29)) of the DMD- d reconstruction (solid) and the DMD- d extrapolation (dot-dashed), error in the frequency ω (dashed), and the number of identified harmonics (dotted line) m versus d . In the right plot, the various errors correspond to the left scale and the number of identified modes, to the right scale.

first derived by Lorenz [33] as a rough approximation of atmospheric terrestrial convection. As a paradigm of chaos, this system exhibits fairly complex periodic orbits [20], such as that for $\sigma = 10$, $r = 350$, and $b = 8/3$, whose precise representation requires a large number of harmonics ($M \gg N = 3$); see Figure 4-left, where to appreciate the nonmonochromatic character of the orbit, its monochromatic counterpart, $a \sin(\omega t + \delta)$, with appropriate a and δ , is also plotted.

To apply DMD, we integrate (3.6) with initial condition $(v_1, v_2, v_3) = (0, 0.1, 0)$, using MATLAB “ode45” (with relative and absolute tolerances both equal to 10^{-8}) and consider $K = 60$ equispaced snapshots in the interval $20 \leq t \leq 20.6$, which is comparable to 1.5 times the period of the orbit. It is to be noted that the solution in this interval corresponds to a transient behavior, with the dominant mode exhibiting a damping rate $\delta \sim 10^{-2}$. Thus, identifying the attractor from these snapshots involves extrapolation.

Standard DMD, using the algorithm DMD-1 produces just three modes, which give a poor approximation (plotted with dotted line in Figure 4-left), while the DMD- d produces a quite good reconstruction (plotted with circles). After some calibration, the tunable parameters of the DMD- d method as set as $\varepsilon_1 = \varepsilon = 10^{-8}$ for all d , which means that the same threshold, ε_1 , is used in the SVDs that are needed at steps 1 and 2 of the DMD- d method. The resulting RMS errors of the reconstructions/extrapolations and the error in the frequency are as plotted versus d in Figure 4-right; the extrapolation errors are calculated comparing with the numerical solution in the interval $19990 \leq t \leq 20000$. Note that DMD- d is fairly robust when increasing d . At $d = 20$, the method identifies $m = 35$ DMD harmonics and the relative RMS error reaches its minimum at $\text{RRMSE} \sim 10^{-7}$, which is comparable to both the tolerance of the numerics and the SVD-threshold chosen above ($\sim 10^{-8}$), meaning that the decomposition is somewhat optimal. Beyond that value of d , the outcomes worsen, namely the RRMSE increases and the number of identified frequencies decreases. This is consistent with the expected behavior as d increases, anticipated in section 2.3.

Additional calculations, not shown here, demonstrate that (as could be expected) main-

taining the timespan $20 \leq t \leq 20.6$ but increasing K by a factor (of, say, 10), d must be consistently increased by the same factor to obtain similar results. And, obviously, increasing the sampled interval (maintaining both the sampling frequency and d), the method improves.

3.3. On the application to the Stuart–Landau equation. Let us now consider the Stuart–Landau equation [26], which conveniently rescaled is written in terms of a complex amplitude A as

$$(3.7) \quad A' = \mu(1 + i\nu)A - (1 + i\beta)|A|^2 A,$$

where the bifurcation parameter μ measures departure from the bifurcation point and ν is the rescaled detuning. This equation applies in the vicinity of Hopf bifurcation [26] in autonomous dynamical systems. The state vector is approximated as

$$(3.8) \quad \mathbf{v}(t) \simeq \mathbf{v}_0 + [\mathbf{V}A(t)e^{i\omega_0 t} + c.c.],$$

where higher order harmonics are ignored, \mathbf{V} is an eigenvector of the linearized problem at threshold, and $i\omega_0$ is the associated eigenvalue. Thus, this equation is relevant at the onset of vortex shedding in the two-dimensional cylinder wake, considered in this context by Bagheri [5].

Equation (3.7) is solved in closed-form setting

$$(3.9) \quad A = Re^{i\theta},$$

which yields

$$(3.10) \quad R(t) \simeq \frac{\sqrt{\mu} R(0)}{\sqrt{R(0)^2 + [\mu - R(0)^2]e^{-2\mu t}}}, \quad \theta'(t) = \mu\nu - \beta R(t)^2.$$

It is interesting to note that, even though the Stuart–Landau equation itself is associated with weakly nonlinear dynamics in the underlying physical problem, (3.10) represents fully nonlinear dynamics of the Stuart–Landau equation, namely A is not slowly varying. For $t > 1/(2\mu)$, $R(t)$ can be expanded in powers of $e^{-2\mu t} \ll 1$, which requires a large number of terms if $e^{-2\mu t}$ is only moderately small. Substituting this expansion into the second expression in (3.10), integrating in t , and substituting the resulting expansions for R and θ into (3.9) and the resulting expression for $A(t)$ into (3.8) yields a DMD expansion of the form (1.3), where the number of required terms M (namely, the spectral complexity) may be quite large. The spatial complexity, instead, is $N = 3$ to the approximation (3.8) (namely, sufficiently close to the bifurcation point). In other words, standard DMD may not give good descriptions of transient dynamics near the Hopf bifurcation, and HODMD should be safely used instead. On the other hand, even though the Stuart–Landau equation may also apply beyond the Hopf bifurcation (not close to threshold) in, e.g., the von Karman instability [41], departure from the bifurcation point may increase the spatial dimension, which could make the difference between standard DMD and HODMD less dramatic.

4. Application to the CGLE. Let us consider the CGLE for the complex independent variable u ,

$$\partial_t u = (1 + i\alpha)\partial_{xx}u + \mu u - (1 + i\beta)|u|^2 u, \quad \text{with } \partial_x u = 0 \text{ at } x = 0, 1,$$

which depends on the real parameters α and β , which account for dispersion and nonlinear detuning, respectively, and μ , which measures departure from marginal instability and is usually taken as a bifurcation parameter. This equation is invariant under the $D_1 \times SO(2)$ group generated by the actions $x \rightarrow 1 - x$ and $u \rightarrow ue^{ic}$, where c is a constant. Because of the stabilizing cubic term, the solutions to this equation are globally bounded, namely both u and the spatial derivatives of u are bounded for all t . The CGLE is a well known paradigm of pattern forming systems that can be considered as a normal form for extended dissipative systems modeled by partial differential equations near the onset of oscillatory instabilities. The equation itself is a simple nonlinear equation that exhibits intrinsically complex dynamics [4], due to the modulational instability if $\alpha\beta < -1$ (Newell's condition) and μ exceeds a threshold value. It must be noted that as μ increases, the complexity of the attractors does not necessarily increase, as seen in the bifurcations diagram given in Figure 5. Also note that if $\mu \gg 1$, then the solution must show fast spatio-temporal oscillations because of the spatial and temporal steepness scale with

$$1/\sqrt{\mu} \quad \text{and} \quad 1/\mu,$$

respectively. For varying μ , the attractors can be periodic, quasi-periodic, and chaotic. In particular, the equation exhibits attractors with very large spectral complexity, which are not accessible to standard DMD. The state variable can be written as

$$u(x, t) = u_0(x, t) e^{i\gamma t},$$

where γ can be seen as a (real) frequency shift, which allows for classifying the attractors [55] as follows:

- *Type I*: Monochromatic, spatially uniform, periodic solutions if u_0 is constant.
- *Type II*: Monochromatic, spatially nonuniform, periodic solutions if $u_0 = u_0(x)$.
- *Type III*: Quasi-periodic solutions if $u_0 = u_0(x, t)$ is time periodic, with a frequency ω_0 incommensurable with γ (which is generically expected).
- *Type IV*: Quasi-periodic solutions, with three involved frequencies if $u_0 = u_0(x, t)$ is quasi-periodic, with basic frequencies ω_0^1 and ω_0^2 , both incommensurable with γ .

Thus, type I and II attractors are monochromatic and exhibit just one spatial mode. In type III attractors, the relevant frequencies are of the form $\omega = m\omega_0 + \gamma$, for appropriate integers m , which correspond to the retained frequencies of the periodic function u_0 shifted by the second frequency γ . In type IV attractors, the relevant frequencies are of the form $\omega = m_1\omega_0^1 + m_2\omega_0^2 + \gamma$, for appropriate integers m_1 and m_2 . Note that in these four cases $|u| = |u_0|$ is (I) steady and spatially uniform, (II) steady and spatially nonuniform, (III) time periodic, and (IV) quasi-periodic, respectively.

The attractors of type IV were not considered in [55], where no inexpensive method to identify these attractors was available. Here, instead, HODMD gives us a very efficient

means to ascertain the nature of these attractors. On the other hand, the nature of these attractors will be guessed (not ascertained) below by plotting $|u(3/4, t)|$ versus $|u(1/4, t)|$. For a given accuracy, these attractors exhibit finite spectral and spatial complexities and, thus, they can be identified using DMD and HODMD. The CGLE also exhibits chaotic attractors, with finite spatial complexity (for a given accuracy) but arbitrarily large spectral complexity. These attractors cannot be approximated using DMD-like approximations, but they can be identified as chaotic using these methods, see below.

The snapshots for the various applications of DMD and HODMD below will be numerically calculated by using a standard Crank–Nicolson plus Adams–Bashforth scheme [14], using a time step $\Delta t = 10^{-5}$ and discretizing spatial derivatives by centered finite differences in a uniform grid of 1000 points. The initial condition will be

$$u = \sqrt{\mu}(1 + i).$$

After discarding the transient behavior in $0 \leq t < t_0$, the snapshots are calculated in the interval $t_0 \leq t \leq t_1$, meaning that we have $(t_1 - t_0)/\Delta t$ snapshots at our disposal, but we shall consider a smaller number of snapshots below. Concerning the J spatial points, these will all be used in the first set of applications of the DMD-like methods considered in the next subsection. However, the performance of the HODMD method using a limited number of spatial points will also be addressed in section 4.2.

4.1. Analysis of the attractors using all spatial data. The snapshots sets will consist of K -equispaced snapshots in an interval $t_0 < t < t_1$, and denoted as

$$S_{[t_0, t_1]}^K.$$

On the other hand, in order to get comparable results, the thresholds

$$(4.1) \quad \varepsilon_1 = 10^{-8} \quad \text{and} \quad \varepsilon = 10^{-6}$$

will be taken in all applications below for the various truncations mentioned in sections 2.1.1–2.1.3 for DMD-1 and in sections 2.2.1–2.2.3 for DMD- d , with $d > 1$.

Taking the above into account, we now proceed with several representative attractors of the CGLE. To begin with, we consider the case

$$\alpha = -10, \quad \beta = 10$$

for various values of μ . The bifurcation diagram for varying μ is given in Figure 5. As further explained in [56], this bifurcation diagram is obtained by plotting $|u(3/4, t)|$ for the intersections of an orbit in the attractor with the Poincaré hypersurface $\int_0^1 [(1 + i\alpha)\partial_{xx}u + \mu u - (1 + i\beta)|u|^2u]u \, dx = 0$. The considered values of α and β mean, since $|\alpha| = 10$ and $|\beta| = 10$ are both large, that the equation is almost conservative (dispersion and nonlinear detuning somewhat large compared to diffusion and nonlinear damping, respectively), namely the equation is somewhat close to the conservative cubic Schrodinger equation [1], which is fairly demanding from the computational point of view. This case was considered in [56] to

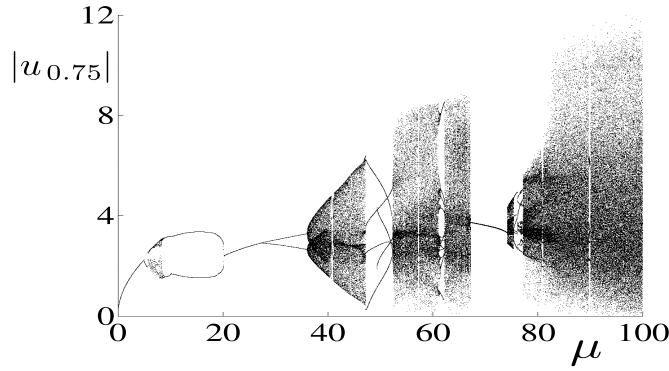


Figure 5. The bifurcation diagram for the CGLE for $\alpha = -10$, $\beta = 10$, and $0 \leq \mu \leq 100$. Courtesy of Dr. Filippo Terragni.

illustrate the construction of bifurcation diagrams for varying μ using an adaptive reduced order model.

For $\mu = 4$ and 25 , the attractors are of types I and II, respectively. The snapshots set is taken as $S_{[20,20.125]}^{125}$. In this case, the spatial and spectral complexities are both equal to 1, meaning that standard DMD solves the problem. Thus, DMD-1 gives a unique frequency, which is $\omega = -40$ and -45.29 for $\mu = 4$ and 25 , respectively. The attractors are reconstructed with a RRMSE $\sim 10^{-12}$ and damping rate $\sim 10^{-13}$ in both cases. HODMD does not improve this reconstruction. The remaining attractors considered below exhibit a spectral complexity larger than the spatial complexity and thus require using HODMD.

For $\mu = 15$ and 35 (Figure 6), the attractors are both type III, with $\gamma \simeq -127$ and 56 , respectively, exhibiting spectral complexities that are larger than the spatial complexities, meaning that the standard DMD yields spurious approximations: using DMD-1 in the snapshots set $S_{[20,20.125]}^{125}$, the RRMSE defined in (2.29) errors are $O(1)$ and $O(0.1)$ for $\mu = 15$ and 35 , respectively, and the (10 and 14 for $\mu = 15$ and 35 , respectively) selected frequencies and amplitudes plotted in red in Figure 6 are spurious. HODMD, instead, gives very accurate, consistent, and robust results, since using the thresholds (4.1), the snapshots sets and the values of the index d (chosen after some calibration) indicated in the caption of Figure 6, all give RRMSE $\sim 10^{-6}$ using the same number of almost identical frequencies and amplitudes, which are 25 for $\mu = 15$ and 21 for $\mu = 35$. These are the spectral complexities within this approximation, while the spatial complexities coincide with their counterparts using DMD-1 and are smaller, namely 10 and 14 for $\mu = 15$ and 35 , respectively; the numbers of SVD modes (see section 2.2.2) retained by DMD-30 and DMD-150 for $\mu = 15$ are both equal to 33, while their counterparts using DMD-50 and DMD-250 for $\mu = 35$ are both equal to 28. Note that, to the considered approximations, the retained amplitudes decay spectrally. The damping rates vary, in all cases, from $\sim 10^{-10}$ for the smallest frequencies to $\sim 10^{-4}$ for the largest frequencies, meaning that the considered snapshots are well within the attractor. The results above are fairly insensitive to the considered value of d , since essentially the same results are obtained for $\mu = 15$ in the ranges $20 \leq d \leq 40$ and $80 \leq d \leq 200$ using 125 and 625 snapshots, respectively, and for $\mu = 35$ in the ranges $5 \leq d \leq 100$ and $20 \leq d \leq 500$

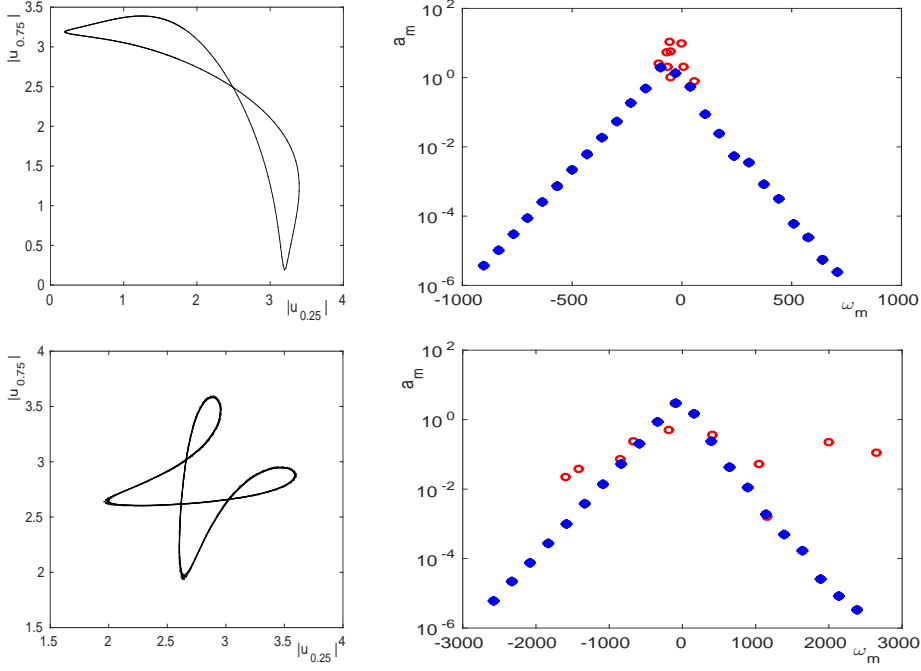


Figure 6. The CGLE for $\alpha = -10$, $\beta = 10$, and $\mu = 15$ (top) and 35 (bottom). Left: $|u(3/4, t)|$ versus $|u(1/4, t)|$. Right: The DMD modes amplitudes a_m (as calculated in section 2.2.3) versus the associated frequencies ω_m , as obtained using standard DMD for the whole snapshots set $S_{[20,20,125]}^{125}$ (red circles) and using DMD-30 and DMD-50 for $\mu = 15$ and 35, respectively, considering the snapshots sets $S_{[20,20,125]}^{125}$ (black circles), $S_{[20,875,21]}^{125}$ (black crosses), and DMD-150 and DMD-250 for $\mu = 15$ and 35, respectively, using $S_{[20,20,125]}^{625}$ (blue circles) and $S_{[20,875,21]}^{625}$ (blue crosses). Black and blue symbols are plot indistinguishable.

using 125 and 625 snapshots, respectively. Decreasing or increasing d outside these ranges the accuracy decreases and, in fact, increasing d too much (i.e., $d \geq 500$ when 625 snapshots are considered), the method diverges.

For $\mu = 50$, the attractor is also type III, with $\gamma \simeq -150$, but more complex than in the former cases. The counterpart of Figure 6 is given in Figure 7, where it can be seen that, again, standard DMD (red circles), which selects 21 DMD modes, gives completely spurious results; as above the RRMSE is $O(1)$. The snapshots sets are indicated in the caption; the sampled timespans are comparable to the period of $u_0(x, t)$. Also, to the intended accuracy:

- The lowest sampling frequency (black symbols) is not enough since it only gives consistently a part of the relevant frequencies, namely those 65 frequencies with mode amplitudes larger than 10^{-4} . The spatial complexity is 21, the total number of SVD modes selected by DMD-70 is 141, and the number of retained frequencies is 104. Retaining these 104 modes, the reconstruction shows a RRMSE $\sim 10^{-6}$ and plot-indistinguishable amplitudes and frequencies, both using the snapshots sets $S_{[20,20,25]}^{250}$ $S_{[20,75,21]}^{250}$ (black circles and crosses, respectively) and the associated values of d indicated in the caption. However, only the first 65 DMD modes correspond to a type III solution since the remaining $104 - 65 = 39$ frequencies would suggest that the solution

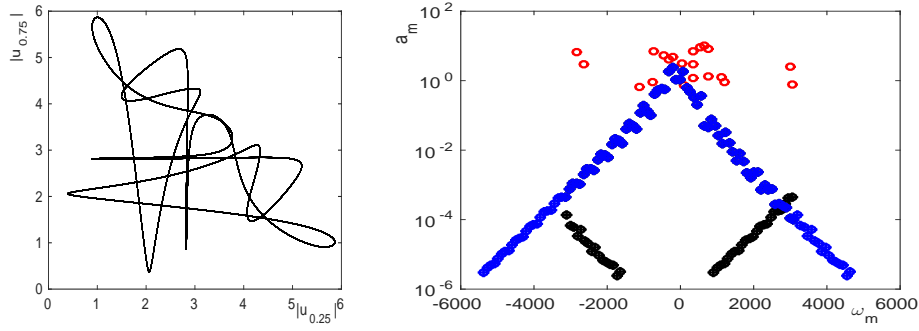


Figure 7. Counterpart of Figure 6 for $\mu = 50$, doubling the intervals where the snapshots are calculated and maintaining the sampling frequency. Thus, in the right plot, DMD-1 is applied to the snapshots set $S_{[20,20.25]}^{250}$ (red circles), DMD-70 is applied to the snapshots sets $S_{[20,20.25]}^{250}$ (black circles), $S_{[20.75,21]}^{250}$ (black crosses), and DMD-400 applied to $S_{[20,20.25]}^{1250}$ (blue circles), and $S_{[20.75,21]}^{1250}$ (blue crosses). Blue symbols are plot indistinguishable

is type IV and, furthermore, they do not coincide with their counterparts calculated with the larger sampling frequency considered in the next item. In other words, the 39 spurious frequencies are not associated with the actual dynamics of the system, but with the inappropriate, sampling frequency, which is too small for the intended accuracy.

- With the largest sampling frequency (blue symbols) using DMD-400 instead, the plotted 104 frequencies are given consistently, since results for the snapshots sets and the values of d indicated in the caption coincide. Moreover, these 104 frequencies also coincide with their counterparts (not plotted in Figure 7) calculated with an even larger sampling frequency, using, e.g., DMD-600 or DMD-800 in the snapshots sets $S_{[20,20.5]}^{2500}$ and $S_{[20.5,21]}^{2500}$. The spatial complexity is again 21 (as both, in DMD-1 and in the results using the smaller sampling frequency), the number of SVD modes selected by DMD-400 is 144 and, retaining the available 104 modes, which are all correct, the reconstruction shows a RRMSE $\sim 10^{-6}$. It is interesting to note that this value of RRMSE essentially coincides with its counterpart (retaining 104 modes also) calculated with the smallest sampling frequency considered in the previous item. However, the number of considered and reconstructed snapshots is four times larger in the present case, meaning that the new snapshots promote the identification of the correct frequencies ω_m with $65 < m \leq 104$. Note that the amplitudes of the relevant 104 modes decay spectrally. The damping rates vary from $\sim 10^{-10}$ for the lowest frequencies to 10^{-4} for the largest frequencies.

As in the previously considered cases, the results are quite insensitive to the index d , since essentially the same results as in Figure 7 are obtained in the ranges $40 \leq d \leq 100$ and $150 \leq d \leq 800$ using 250 and 1250 snapshots, respectively.

Let us now turn into the quasi-periodic attractor for $\mu = 40$, which is type IV with $\gamma \simeq 67$. Since the approach to the attractor is slower than in the former cases considered above, we discard a larger transient stage, $0 \leq t \leq 40$, before calculating the snapshots, which are taken in the time interval $40 \leq t \leq 42$. Also, the sampled time intervals and the

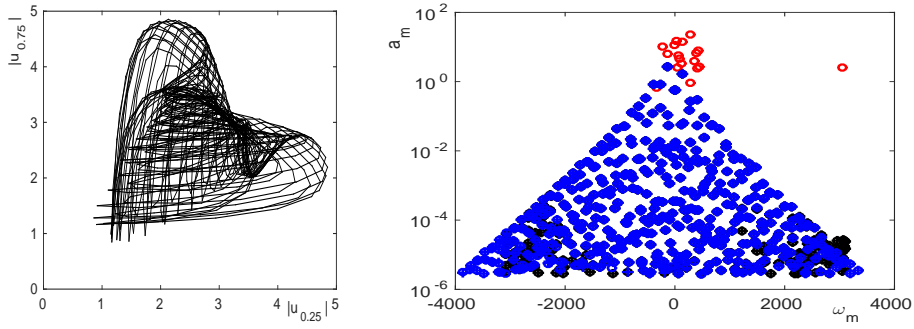


Figure 8. Counterpart of Figure 7 for $\mu = 40$, with new snapshots sets: in the right plot, DMD-1 is applied to the snapshots set $S_{[40,41]}^{1000}$ (red circles), DMD-200 is applied to the snapshots sets $S_{[40,41]}^{1000}$ (black circles), $S_{[41,42]}^{1000}$ (black crosses), and DMD-1000 applied to $S_{[40,41]}^{5000}$ (blue circles), and $S_{[41,42]}^{5000}$ (blue crosses).

sampling frequencies are both larger. As a consequence, for the lowest sampling frequency, the considered snapshots sets $S_{[40,41]}^{1000}$ and $S_{[41,42]}^{1000}$ are treated using DMD-200, while $S_{[40,41]}^{5000}$ and $S_{[41,42]}^{5000}$ require using DMD-1000. For the intended accuracy (see (4.1)), the spatial complexity is 18. The counterpart of Figure 7 is given in Figure 8. The results are similar to those obtained for the previous cases. As expected, DMD-1 gives a completely spurious solution. Using the smallest sampling frequency (black symbols) does not solve the problem to the given accuracy. Specifically, the number of SVD modes retained by DMD-200 is 626 and the total number of retained frequencies, 481; the reconstruction error retaining all modes is $\text{RRMSE} \sim 5 \cdot 10^{-6}$. However, only the first 157 frequencies (namely, those showing modes amplitudes $> 10^{-4}$) are dynamically meaningful, as seen comparing with the results obtained from the snapshots set using the largest sampling frequency (blue symbols). In this case, the number of SVD modes retained by DMD-1000 (larger than with the smaller sampling frequency) is 705 and the total number of retained frequencies is 440 (smaller than with the smaller sampling frequency). The reconstruction error retaining all modes is $\text{RRMSE} \sim 5 \cdot 10^{-6}$, comparable to its counterpart using the smaller sampling frequency, but now the 440 modes are dynamically relevant, which has been tested comparing with the results (omitted here) obtained by multiplying both the sampling frequency and the index d by 2. The damping rates vary from $\sim 10^{-7}$ for the lowest frequencies to 10^{-2} for the largest frequencies. Note that, as anticipated, for the same required accuracy, the number of relevant frequencies is much larger here than in the periodic cases considered above, as seen comparing the number of blue symbols in Figure 8 with its counterparts in Figures 6 and 7. As in the formerly considered cases, the results are quite insensitive to the index d , since essentially the same results as in Figure 8 are obtained in the ranges $180 \leq d \leq 300$ and $900 \leq d \leq 1300$ using 1000 and 5000 snapshots, respectively.

Finally, for $\mu = 7$, the attractor is chaotic, which means that the Fourier expansion of the orbits are broadband and, moreover, each orbit is unstable. This means that no method can be able to produce consistent and robust (i.e., coinciding in shifted sampled intervals) DMD expansions with a finite number of modes, as those obtained above for periodic and quasi-periodic orbits. Namely, the obtained amplitudes and frequencies strongly depend on the sampled interval because the orbit is chaotic and visits different regions of the attractor

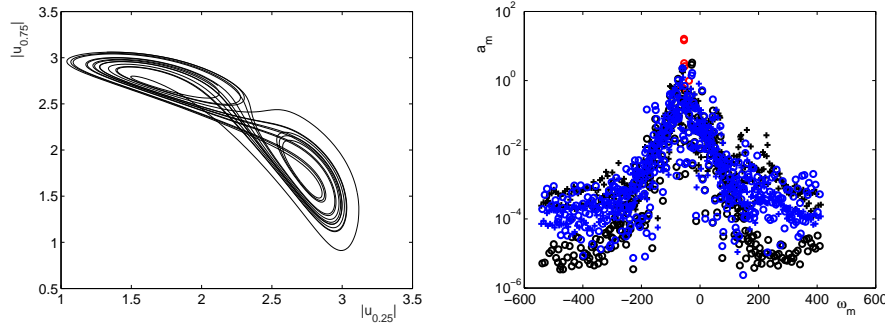


Figure 9. Counterpart of Figure 8 for $\mu = 7$, considering the snapshots sets using DMD-600: $S_{[40,44]}^{4000}$, $S_{[46,50]}^{4000}$ and using DMD-4000: $S_{[40,44]}^{20000}$, and $S_{[46,50]}^{20000}$.

in different timespans. However, DMD- d , with appropriate d is able to give approximations in individual sampled intervals, while DMD-1 is not. Since the dynamics are slower (because μ is smaller than in the former case), the sampled time intervals are chosen larger than in the previous cases, as indicated in the caption of Figure 9. The various HODMD applications (namely, those plotted in Figure 9 and many others, in different sampled time intervals, with even larger sampling frequencies and sampled timespans) give different results, which is a good indication that the dynamics in this case are more complex than quasi-periodic. In spite of the lack of robustness, DMD-600 is able to reconstruct the chaotic orbit in each of the sampled intervals considered in Figure 9 reasonably well: the RRMSE of the reconstruction is $\text{RRMSE} \sim 10^{-2}$ for both sampled intervals, while DMD-1 produces completely spurious results. In other words, for chaotic dynamics, DMD- d , with appropriate d is able to fit data in given timespans, but (contrary to what happens for periodic and quasi-periodic dynamics), the results are not consistent when comparing one sampling interval to another, as expected.

Additional calculations, not given here for the sake of brevity, show that for chaotic dynamics at, e.g., $\mu = 1000$ (see [55]), the performance of the methods is similar to that illustrated in Figure 9. Namely, DMD- d gives (slightly accurate but) inconsistent results, whatever the snapshots sets and the value of the index d . Similarly, the HODMD method has been tested for the bifurcation diagrams that are encountered for other values of α and β . For instance, for $\alpha = -10$ and $\beta = 1$ the system is still highly dispersive though no longer almost conservative, but the system still exhibits fairly complex periodic and quasi-periodic attractors (see [55]). In these additional computations, not shown here for the sake of brevity, performance of the HOSVD method is as good as in the various cases considered above. The standard DMD, instead, provides accurate and consistent results for the simplest periodic attractors only.

Summarizing the above, in contrast to standard DMD, the new HODMD method, using the developed DMD- d algorithm for appropriate snapshots sets in reasonable sampled time intervals (comparable to the $2\pi/\omega$, where ω is the smallest involved frequency), is a very useful means to identify general periodic and quasi-periodic attractors in complex dynamical systems. The method calculates well the involved frequencies/amplitudes (in fact, a quite large number of them for quasi-periodic attractors) and reconstructs well the snapshots. Moreover,

the dynamically relevant modes are identified by comparing the obtained results using various snapshot sets. It must be noted that, since we had a very large number of snapshots at our disposal (from the numerical computations) no attempt has been made to minimize the snapshots number, K . Instead, we have taken K sufficiently large to reconstruct the solution with $\text{RRMSE} \sim 10^{-6}$, which has required identifying fairly large frequencies. If, instead, only a $\text{RRMS} \sim 10^{-3}$ were sought, then the required numbers of snapshots in Figure 6 would have been 85 with $d = 5$ (instead of 125 with $d = 30$) and 25 with $d = 5$ (instead of 125 with $d = 50$) for $\mu = 15$ and 35, respectively. Similarly, the numbers of snapshots in Figures 7 and 8 would have been 140 with $d = 10$ (instead of 250 with $d = 70$) and 700 with $d = 100$ (instead of 1000 with $d = 200$), respectively. On the other hand, chaotic attractors can be identified from the ill-functioning of the method. As done for the Lorenz system in section 3.2, decaying dynamics may be used to identify the large time attractors via extrapolation, but this application has not been illustrated in this section for the sake of brevity.

4.2. Recovering the relevant modes amplitudes and frequencies using a limited amount of noisy data. As anticipated in the introduction and illustrated in some toy-model examples in section 3.1, the HODMD algorithm DMD- d , with appropriate d , takes full advantage of both the spatial and temporal redundancies in periodic and quasi-periodic dynamics of the type

$$(4.2) \quad \mathbf{v} \simeq \sum_{m=1}^M a_m \mathbf{u}_m e^{i\omega_m t},$$

which has a potential interest to do the following:

- Recover the relevant mode amplitudes and frequencies using a limited amount of spatial data.
- Filter out small errors.

On the other hand, as DMD-1 applied to time periodic and quasi-periodic flows is related to FFT [18], DMD- d can be compared to PSD [39], where the number of segments used in PSD is the counterpart of the index d in the DMD- d algorithm. All these are the object of the present section. For the sake of brevity, to avoid defining a new dynamical system, this illustration relies in the CGLE considered in this section.

Before proceeding further, some remarks are in order:

- When comparing results obtained using different numbers of spatial points, comparison of the relevant frequencies is straightforward, but comparison of the mode amplitudes is more tricky. Scaling the modes \mathbf{u}_m appearing in (4.2) as we have done in this paper (see (2.12)), namely such that $\|\mathbf{u}_m\| = \sqrt{J}$, where J is the spatial dimension of the given data, makes the mode amplitudes a_m as independent as possible from J . However, when comparing data from two quite different values of J (e.g., $J = 1000$ and 1), the two sets of amplitudes will only coincide if the data is spatially uniform. Otherwise, only the frequencies can be safely compared.
- For $J = 1$, the scaling (2.29) yields $\|\mathbf{u}_m\| = 1$, namely the mode amplitudes exactly coincide with their counterparts in PSD and FFT calculations.
- After some calibration for the CGLE, it has been found that HODMD also recovers the mode amplitudes if the considered number of spatial data is comparable to the

spatial complexity.

- The following advantages of HODMD can be anticipated in periodic dynamics:
 - As illustrated in the last section, HODMD gives good results in arbitrary sampled time intervals whose length is comparable to the period. FFT and PSD, instead, produce artifacts unless the sampled timespan either is sufficiently large or coincides with the period.
 - For slightly decaying dynamics, HODMD identifies both the frequencies and the (small) damping rates, which in principle gives good reconstructions without any additional treatment. Moreover, setting to zero the damping rates produces a good approximation of the final attractors. Instead, the nonzero damping rates produce undesirable sideband artifacts when using FFT and PSD.

These two observations are even more evident when treating quasi-periodic dynamics.

In what follows, several cases will be considered in which the following calculations will be compared:

- *DMD-d-full*: DMD- d calculations using all spatial data (as already done in the last section). These will be taken as reference.
- *DMD-d-N*: DMD- d calculations using a number of spatial points equal to the spatial complexity.
- *DMD-d-1*: DMD- d calculations using just one spatial point, namely $x = 1/4$.
- *PSD-d*: PSD calculations using d segments (and just one spatial point).
- *FFT*: FFT calculations (for just one spatial point).

It must be noted that the mode amplitudes appearing in the DMD-expansions involving different spatial points need not coincide because the spatial modes are generally different. In the application of DMD- d to the various cases below, the index d will be taken as that calibrated in the last section. The HODMD thresholds will also be taken identical to those used in the last section, namely as defined in (4.1). The same threshold $\varepsilon = 10^{-6}$ will also be used for PSD and FFT truncation.

To begin with, we address the case $\alpha = -10$, $\beta = 10$, and $\mu = 15$, already considered in Figure 6. The results applying the five methods indicated above are depicted in Figure 10. As can be seen in this figure, DMD-150-full, DMD-150-10 (note that 10 is the spatial complexity), and DMD-500-1 all give very similar results, since the identified frequencies coincide though, as anticipated there are some differences in the associated mode amplitudes. In fact, both, DMD-150-10 and DMD-500-1 (with the sampled timespan doubled) correctly identify the 25 and 24 relevant (to the given accuracy) frequencies, respectively. On the other hand, the results are very insensitive to the index d since, e.g., DMD-d-1, with $450 \leq d \leq 800$ yields essentially the same result as that plotted in Figure 10. Instead, both PSD and FFT only approximately identify the first 20 frequencies, which, moreover, do not show the largest amplitudes. This is in spite of the fact that the sampled time interval used by DMD-500-1 has been multiplied by 240 and 720 for PSD and FFT, respectively (because using the same sampled interval as when using DMD-500-1, the results are completely spurious).

The ill-functioning of both FFT and PSD is due to the artifacts introduced by these methods when the sampled interval is not commensurable with the period of the solution. By its own nature, these artifacts are not present in HODMD (and neither in standard DMD when

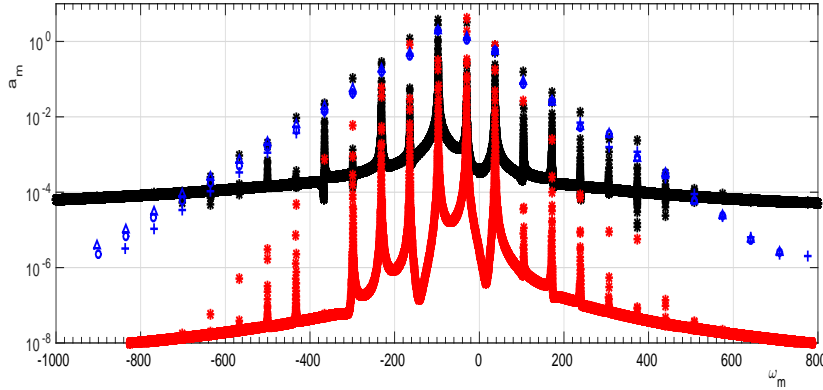


Figure 10. Counterpart of Figure 6-top-right constructed by applying DMD-150-full (blue triangles) and DMD-150-10 (blue circles) in the snapshots set $S_{[20,20,125]}^{625}$, DMD-500-1 in the snapshots set $S_{[20,20,250]}^{1250}$ (blue crosses), applying PSD-100 (red asterisks) in the snapshots set $S_{[20,80]}^{3 \cdot 10^5}$, and FFT (black asterisks) in the snapshots set $S_{[20,200]}^{1.8 \cdot 10^6}$.

this method is effective because the spatial complexity is larger than the spectral complexity). Thus, this example illustrates well both the power of HODMD to identify the frequencies using just one spatial point and the advantages of this method compared to Fourier-like methods. The comparison is even more advantageous to HODMD in the remaining applications below, but this is already clear from the present example and, thus, it will only be HODMD that will be used below. Moreover, in order to illustrate how robust the HODMD method is against noise (even using a limited number of spatial data), the snapshots are contaminated in the remaining applications with a uniformly distributed, positive random error (as given by the MATLAB command “rand”) of size 10^{-7} . Note that this noise level is somewhat smaller than the parameter $\varepsilon = 10^{-6}$ that is being used to truncate the mode amplitudes, according to (4.1).

The simple type III attractor considered in Figure 10 and the more complex type III attractor in Figure 7 are addressed in Figure 11. As can be seen comparing with the blue symbols corresponding to DMD- d -full applied to the clean snapshots, which are taken for reference, the performance of DMD- d -1, DMD- d - N , and DMD- d -full are consistent, precise, and robust, since the identified frequencies very approximately coincide though, as anticipated, the associated amplitudes show differences when different spatial data are considered. In particular, the results with noise coincide with their counterparts for the clean snapshots, meaning that the methods efficiently clean the errors out. DMD- d -1 correctly identifies 24 of the 25 frequencies in the left plot and 87 of the 104 relevant frequencies in the right plot. The remaining frequencies are not identified because they exhibit amplitudes that are smaller than the threshold $\varepsilon = 10^{-6}$; decreasing this threshold helps to identify some of the lost frequencies but also introduces spurious frequencies associated with the noise. As in the previous cases, the method is fairly insensitive to the index d (for, e.g., $400 \leq d \leq 900$ in DMD- d -1).

The counterpart of the plots in Figure 11 for the more complex type IV attractor already considered in Figure 8 would be very confusing because of the extremely large number of

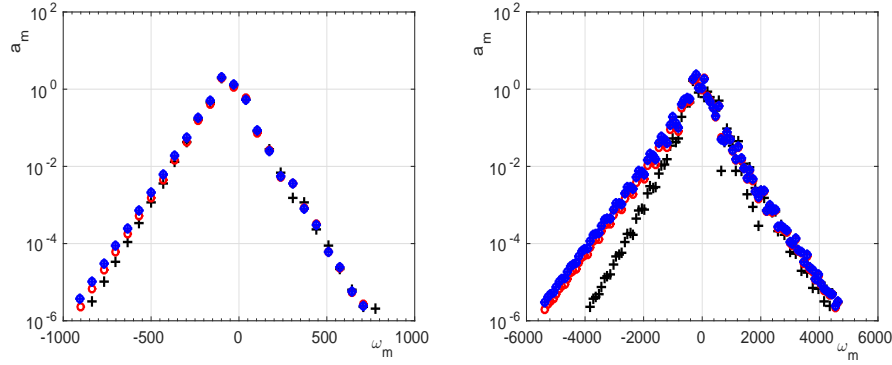


Figure 11. Left: Counterpart of Figure 6-top-right ($\mu = 15$) constructed by applying DMD-150-full without noise (blue circles), and with noise by applying DMD-150-full (blue crosses) and DMD-150-10 (red circles) in the snapshots set $S_{[20,20.125]}^{625}$, and DMD-500-1 (black crosses) in the snapshots set $S_{[20,20.250]}^{1250}$. Right: Counterpart of Figure 7-right ($\mu = 50$) constructed by applying DMD-400-full without noise (blue circles), and with noise by applying DMD-400-full (blue crosses), DMD-700-21 (red circles) in the snapshots set $S_{[20,20.250]}^{1250}$, and DMD-700-1 (black crosses) in the snapshots set $S_{[20,20.250]}^{2500}$.

involved frequencies. However, the method also performs very well since the 157 amplitudes and frequencies in Figure 8 are recovered using the either DMD-1000-18 method in the snapshots set $S_{[40,41]}^{5000}$, and the 157 frequencies are recovered using DMD-4000-1 in the snapshots set $S_{[40,44]}^{20000}$.

Summarizing the above, DMD- d , with appropriate d , using appropriate snapshots sets (defined in short sampled timespans), is able to capture very well the relevant frequencies in periodic and quasi-periodic attractors, even using a very limited number of spatial data (just one point in the applications to the CGLE). Moreover, even when the data are affected by errors, the method still gives very good results.

5. The three-dimensional thermal convection in a rotating spherical shell. Let us now consider a more computationally demanding three-dimensional problem, which has been selected by requiring that it is physically relevant and involves periodic and quasi-periodic dynamics that are easily illustrated for the reader. More complex dynamics (e.g., transitional flows) are not easily illustrated in a few pages, which would penalize clarity in the exposition. The selected test problem is the thermal convection in a rotating spherical fluid shell subject to a radial gravity field, which is a basic problem in geophysical and astrophysical fluid dynamics [17]. In particular, the study of this configuration is useful to estimate the transport of energy in the interior of planets and stars, and to determine their internal structure and its influence on the patterns observed in the upper atmospheres of giant planets. All of these provide models to fit to experimental observations of the drifting of these internal patterns, which in turn is needed to, e.g., understand the generation of magnetic fields if coupled with the induction equation. Identification of periodic and quasi-periodic azimuthal waves, which appear for sufficiently large Rayleigh number, are of particular interest.

The problem is formulated in terms of the nondimensional velocity vector \mathbf{v} , pressure Π , and the nondimensional perturbation of the temperature with respect to its conduction state

profile, Θ . The units of length, time, temperature, and pressure for nondimensionalization are d , d^2/ν , $\nu^2/(\alpha\gamma d^4)$, and $\rho_0\nu^2/d$, respectively, where $d = r_o^* - r_i^*$ is the difference between the (dimensional) outer and inner radii, ν is the kinematic viscosity, γr_i^* is the imposed radial gravity at the inner radius, α is the thermal expansion coefficient, and ρ_0 is the density at the reference temperature. Using a rotating Cartesian coordinate frame, with origin at the common center of the inner and outer spheres, with the z axis along the rotation axis, the nondimensional continuity, momentum, and energy equations governing the dynamics of the fluid in the Boussinesq approximation are

$$\begin{aligned}\nabla \cdot \mathbf{v} &= 0, \\ \partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} + 2E^{-1} \mathbf{k} \times \mathbf{v} &= -\nabla \Pi + \nabla^2 \mathbf{v} + \Theta \mathbf{r}, \\ \sigma(\partial_t \Theta + \mathbf{v} \cdot \nabla \Theta) &= \nabla^2 \Theta + \mathbf{R} \eta (1 - \eta)^{-2} \mathbf{r}^{-3} \mathbf{r} \cdot \mathbf{v},\end{aligned}$$

where \mathbf{k} is the vertical unit vector, $\mathbf{r} = (\mathbf{x}, \mathbf{y}, \mathbf{z})$, $\eta = r_i^*/r_o^*$ is the inner to outer radii ratio, $\sigma = \nu/\kappa$ is the Prandtl number (with κ = thermal conductivity), $E = \nu/(\omega d^2)$ is the Eckman number, and $R = \gamma \alpha \Delta T d^4/(\nu \kappa)$ is the Rayleigh number. The boundary conditions are homogeneous Dirichlet at both the inner and outer spheres. For nonzero rotation, the resulting problem is invariant under the symmetry group $SO(2) \times \mathbb{Z}_2$, generated by the rotations about the vertical z axis and the up-down reflection with respect to the equatorial plane. Invariance under rotation is essential to obtain rotating waves, most of which are also (instantaneously) invariant under the up/down symmetry.

For convenience, the pressure and the continuity equation are both eliminated from the formulation by taking the curl of the momentum equation and writing the (solenoidal) velocity field as

$$\mathbf{v} = \nabla \times (\Psi \mathbf{r}) + \nabla \times \nabla \times (\Phi \mathbf{r}),$$

where Ψ and Φ are the toroidal and poloidal scalar potentials, respectively, first introduced by Chandrasekhar [15]. The resulting problem is written in spherical coordinates (r, θ, ϕ) , where θ and ϕ are the colatitude and longitude, respectively, in the rotating frame of reference linked to the spherical shell. The non-dimensional temperature and the scalar potentials are expanded in spherical harmonics, and spatially discretized using Gauss–Lobatto collocation (with $25 \times 64 \times 128$ collocation points in (r, θ, ϕ)). Several fairly subtle ingredients are used for dealiasing and, especially, for continuation along the periodic and quasi-periodic branches; see [23, 47] and references therein for further details.

The results considered below are concerned with the case $\sigma = 0.1$ and $E = 10^{-4}$, which somehow corresponds to the outer atmosphere of some planets such as Jupiter. It turns out that, for the outer/inner radii ratio η in the range $0.32 \leq \eta \leq 0.35$, the system exhibits a fairly complex bifurcation diagram, which is quite efficiently calculated in [46] using continuation, not only of periodic orbits but also of invariant tori (a fairly demanding problem). Generically, the purely *conductive state* (nonuniform temperature distribution, with no convection) loses stability in a Hopf bifurcation, where a new branch of periodic solutions is born, which is a *pure rotating wave* (namely, a steady pattern in reference frame which is rotating in the already considered rotating frame linked to the spherical shell). An example is given in Figure 12 that

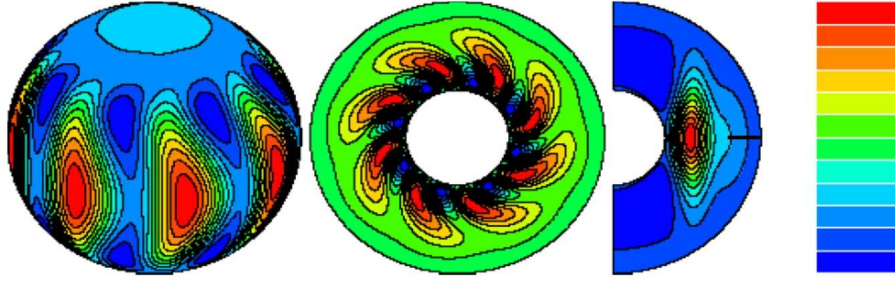


Figure 12. Sketch of a snapshot showing the contour levels of Θ in an intermediate spherical surface (left), the equatorial plane (middle), and a meridian plane (right); courtesy of Professor Marta Net. The plotted snapshot is rigidly rotating in the azimuthal direction. The horizontal segment plotted with thick solid line in the right plot corresponds to the 15% outer part of the radial equatorial segment, which will be used in Figure 13.

shows an azimuthal wavenumber $m = 8$ and is rigidly rotating in the azimuthal direction with a velocity c , meaning that the frequency of the associated periodic pattern is $\omega = c/8$. In other words, concentrating for simplicity in the temperature field, the spatio-temporal structure of the pattern is given by

$$(5.1) \quad \Theta(r, \theta, \phi, t) = \sum_{n=1}^N \Theta_n(r, \theta) e^{in(\omega t - 8\phi)} + c.c.,$$

which can be seen as the relevant DMD-expansion.

There are several branches of rotating waves, exhibiting different azimuthal wave-numbers m in the range $3 \leq m \leq 8$, which bifurcate from the conduction states at different critical Rayleigh numbers, $R \sim 10^5$. Some of these solution branches are stable but lose stability via Neimark–Sacker bifurcations (complex Floquet multipliers; see [26]) that produce invariant tori (quasi-periodic solutions). On the other hand, for some critical values of η , two of the periodic branches coalesce, namely the bifurcation points associated with two different values of m occur at the same critical Rayleigh number, giving a (codimension-two) double Hopf bifurcation that also produces invariant tori. For instance, at $\eta_c = 0.33079$ and $R_c = 2.0069 \cdot 10^5$, a double Hopf bifurcation [26] takes place that involves the $m = 5$ and 6 modes; see [46] for further details.

The cases analyzed in this article are two representative stable flows, one periodic and another quasi-periodic. The spatio-temporal data that will be used for both have been provided to us by Professors Joan Sanchez and Marta Net, who obtained them using the algorithm outlined above. Both flows are symmetric with respect to the meridional plane. Consequently, DMD-d has been applied only to half of the three-dimensional domain (positive colatitude $\theta \geq 0$). For simplicity in the presentation, only the thermal field, $\Theta(r, \theta, \phi, t)$ will be considered below, both in the performed DMD computations and in all illustrations.

The considered periodic flow is the purely rotating wave illustrated in Figure 12, which corresponds to $R = 8 \cdot 10^5$ and shows that such flow is steady in a rotating (with the phase velocity) reference frame. Such steady flow, in turn, is periodic in the azimuthal direction, with a fundamental wavenumber equal to 8. Therefore, this is of the form (5.1), where N

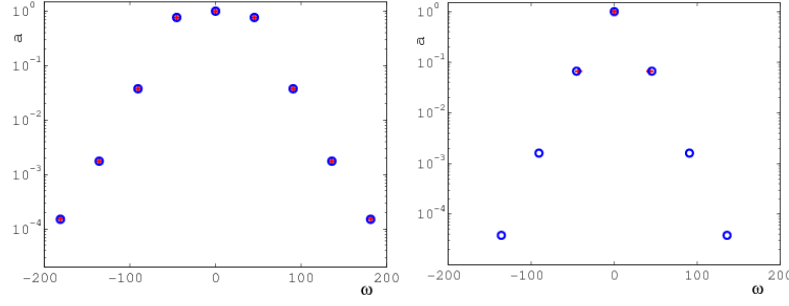


Figure 13. Mode amplitude (right) versus the frequency for the modes obtained by applying DMD-1 (red crosses) and DMD-40 (blue circles) to a set of 100 equispaced snapshots, with a nondimensional time shift between snapshots equal to $2.766 \cdot 10^{-3}$, considering the meridian plane $\phi = 0$ (left) and the 85% outer part of the radial line $\phi = \varphi = 0$ (right).

is the number of modes to be retained and $\omega = 45.3354$ is the fundamental frequency. The temporal structure of this three-dimensional flow is fairly simple and, moreover, the spatial and temporal complexities of the associated DMD expansion are both equal to the number of retained frequencies, N . This is because, due to the factor $e^{-8in\phi}$, the spatial modes $\Theta_n(r, \theta)e^{-8in\phi}$ are obviously linearly independent (in fact, they are pairwise orthogonal with the usual L_2 inner product). Consequently, as emphasized in the paper, the advantages of HODMD with respect to the standard DMD method are very weak in the whole three-dimensional flow. In fact, retaining the mean flow and the four most energetic oscillatory modes, both DMD-40 and DMD-1 give relative RMS errors $\text{RRMSE} \sim 10^{-5}$, comparable to the truncation errors in the numerical scheme. Restricting to the meridian plane $\theta = 0$, the restricted DMD expansion is (cf. (5.1))

$$(5.2) \quad \Theta(r, \theta, 0, t) = \sum_{n=1}^N \Theta_n(r, \theta) e^{in\omega t} + c.c., \quad \text{with } r_{\text{in}} \leq r \leq r_{\text{out}}, 0 \leq \theta \leq \pi/2.$$

Using the restricted snapshots, the plots of the amplitudes versus the frequencies computed by DMD-1 and DMD-40 are both very good and still plot indistinguishable, as seen in Figure 13-left, though DMD-40 slightly outperforms DMD-1, since the relative errors of the frequency of the fourth harmonic (comparing with the three-dimensional computations) are $\sim 10^{-6}$ and $\sim 10^{-3}$, respectively. The advantages of using delayed snapshots are appreciated by drastically reducing the number and quality of the observables. This is seen by considering only the last six Gauss–Lobatto collocation points in the radial line $\theta = \phi = 0$, which corresponds to the outer 15% segment of this radial (plotted with a thick solid line in the right plot in Figure 12) and yields the restricted DMD expansion (cf. (5.1)–(5.2))

$$(5.3) \quad \Theta(r, 0, 0, t) = \sum_{n=1}^N \Theta_n(r) e^{in\omega t} + c.c., \quad \text{with } 0.85 \leq r - r_{\text{in}} \leq r_{\text{out}} - r_{\text{in}} (= 1).$$

Note in Figure 12-right that the thermal activity in this small segment of the radial line is very small. In spite of this, DMD-40 (see Figure 13-right) is still able to identify the mean flow and

Table 1

Wavenumbers n_1 and n_2 and frequencies, as defined in (5.5), for the 12 most energetic modes in the quasi-periodic attractor considering the whole amount of three-dimensional data (3D), the meridian plane $\theta = 0$ (2D), and the line $\phi = \theta = 0$ (1D), as calculated by applying DMD-200 and DMD-1, as indicated.

		3D	2D		1D	
		DMD-200	DMD-200	DMD-1	DMD-200	DMD-1
n_1	n_2	ω				
0	0	0.0000	0.0000	0.0000	0.0000	0.0000
1	0	47.5890	47.5890	47.9306	47.5891	55.6587
0	1	239.8411	239.8412	240.8474	239.8408	—
−1	1	192.2520	192.2535	192.9154	192.2524	187.5543
2	0	95.1783	95.1744	96.0682	95.1749	—
1	1	287.4301	287.4285	287.7189	287.4288	280.4798
2	1	335.0192	335.0139	344.1997	335.0210	—
0	2	479.6822	479.6717	—	479.6807	—
−2	1	144.6444	144.1315	—	143.8763	—
−1	2	432.0838	432.0914	—	432.0938	—
1	2	527.2712	527.1714	—	527.2698	—
3	0	142.7519	142.4148	124.5031	144.9524	—
3	1	382.7911	382.4096	—	382.9238	351.8692

three oscillating modes and reconstruct the restricted snapshots with a relative RMS error $\text{RRMSE} \sim 10^{-5}$; also, DMD-40 the relative errors in the three identified frequencies is $\sim 10^{-4}$. The fourth oscillating mode exhibits a too small amplitude in this restricted segment, smaller than truncation errors in the numerical approximation, and cannot be identified. The very good identification of the frequencies and the very small reconstruction error clearly illustrate the advantages of using the delayed snapshots. DMD-1, instead, only recognizes the mean flow and the first harmonic.

Let us now consider the quasi-periodic attractor that is obtained for $R = 1.1 \cdot 10^6$. This attractor can be seen as resulting from the nonlinear interaction of two purely rotating waves, with wavenumbers $m = 6$ and $m = 8$. Thus, the counterpart of (5.1) is

$$(5.4) \quad \Theta(r, \theta, \phi, t) = \sum_{n_1, n_2} \Theta_{n_1 n_2}(r, \theta) e^{i[n_1(\omega_1 t - 8\phi) + n_2(\omega_2 t - 6\phi)]} + c.c.,$$

for some (incommensurable) fundamental frequencies ω_1 and ω_2 ; the sum is extended to a number of retained pairs of values of the integers n_1 and n_2 . As in the periodic case, the factor $e^{-i(8n_1 + 6n_2)\phi}$ makes the spatial modes fairly uncorrelated, which means that using the whole amount of three-dimensional data, retaining a reasonable number of modes, HODMD is not expected to dramatically outperform standard DMD. However, it is interesting to note that both DMD-1 and DMD-200 (applied to a set of 500 equispaced snapshots, with a nondimensional time-shift between snapshots equal to $5 \cdot 10^{-4}$) are able to identify quite well the fundamental frequencies and the harmonics, defined as

$$(5.5) \quad \omega = n_1 \omega_1 + n_2 \omega_2,$$

which are displayed in Table 1 (third column) and will be used for reference below. As can

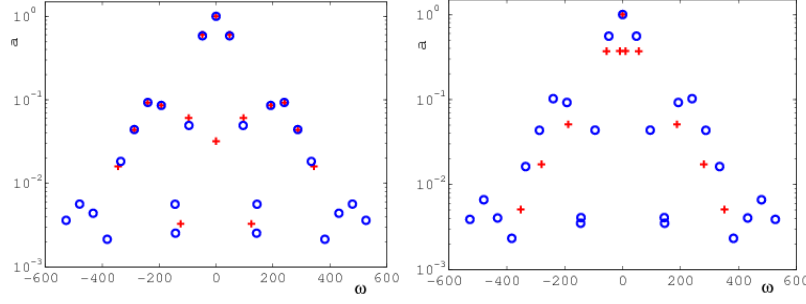


Figure 14. Counterpart of Figure 13 for the quasi-periodic attractor.

be checked by the reader, these frequencies satisfy quite well the commensurability relation (5.5).

As in the periodic case, restricting to the meridian plane $\theta = 0$ and the radial line $\theta = \phi = 0$, namely restricting (5.4) to

$$\Theta(r, \theta, 0, t) = \sum_{n_1, n_2} \Theta_{n_1 n_2}(r, \theta) e^{i(n_1 \omega_1 t + n_2 \omega_2) t} + c.c.,$$

in $r_{\text{in}} \leq r \leq r_{\text{out}}, 0 \leq \theta \leq \pi/2$ and

$$\Theta(r, 0, 0, t) = \sum_{n_1, n_2} \Theta_{n_1 n_2}(r) e^{i(n_1 \omega_1 + n_2 \omega_2) t} + c.c. \quad \text{with } r_{\text{in}} \leq r \leq r_{\text{out}},$$

respectively, decreases the spatial complexity. In particular, restricting to the meridian plane $\theta = 0$, Table 1 and Figure 14 show that while DMD-200 identifies all frequencies and reconstructs the snapshots with great precision (RRMSE $\sim 10^{-3}$); DMD-1 gives much worse results, since the reconstruction involves a larger error (RRMSE $\sim 10^{-2}$) and the frequencies are identified with an error that increases as the amplitude of the mode decreases. Moreover, restricting to the radial line $\phi = \theta = 0$, DMD-1 gives completely spurious results in connection with both the identified frequencies (see Table 1) and the reconstruction (RRMSE ~ 0.4), but DMD-200 approximates the considered modes with a reasonable accuracy, since RRMSE $\sim 10^{-3}$ and the frequencies are well identified (see Table 1). Thus, the advantages of using delayed snapshots are clearly seen, for a given spectral complexity, as the number of observables (and thus the spatial complexity) is decreased.

Finally, the mean flow and the spatial modes associated with the two fundamental frequencies are considered in Figure 15. As can be seen:

- The thermal activity mainly concentrates (for both the mean flow and the fundamental modes) near the equator and also somewhat near the inner boundary of the spherical shell for the 10-mode and 20-mode, but in more interior regions for both the mean flow and the remaining two oscillatory modes.
- The real and imaginary parts of the oscillatory modes are not proportional to each other, meaning that these modes are not standing but progressive, namely the associated activity oscillates back and forth in the meridian plane.

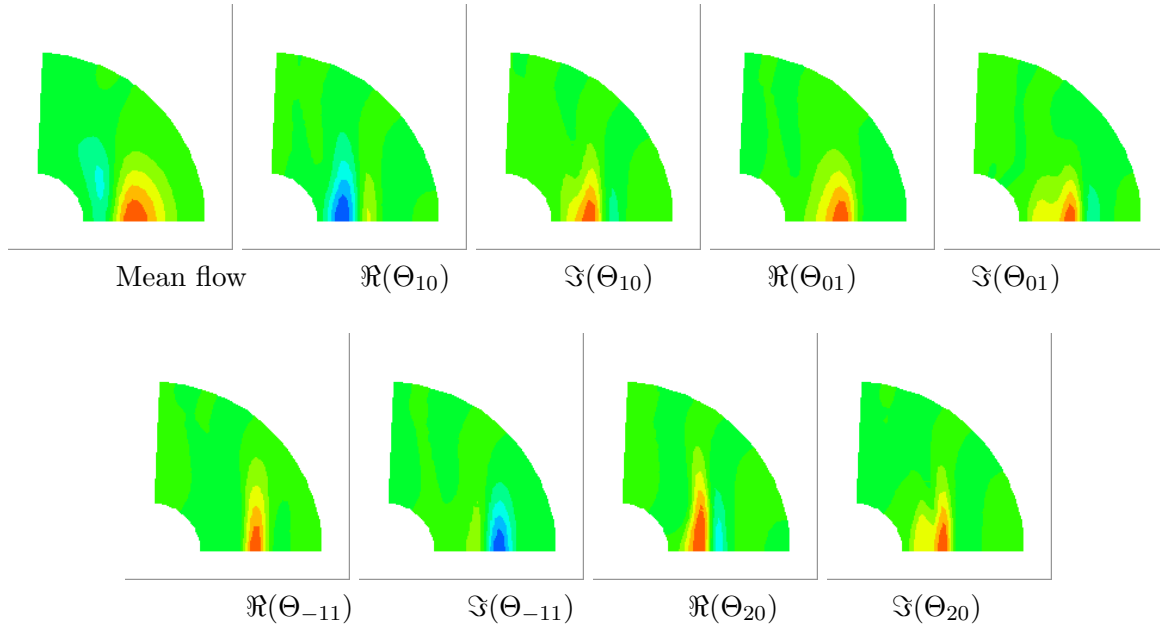


Figure 15. The five most energetic modes obtained for the quasi-periodic attractor in the meridian plane $\theta = 0$, via DMD-200, considering the mean flow, and the real and imaginary parts of the oscillatory modes associated with the two fundamental frequencies.

Summarizing, as anticipated, the considered dynamics is (on purpose) somewhat simple, even in the quasi-periodic case, where the various DMD-modes are associated with purely rotating waves, whose three-dimensional spatial structures are fairly uncorrelated (see (5.1) and (5.4)). Thus, the advantages of using delayed snapshots are not appreciated when the whole amount of three-dimensional data is used. However, as the data used is restricted to the meridian plane and a radial line in the equatorial plane, the advantages of using HODMD become clear. This is relevant keeping in mind the identification of the spatio-temporal structures using a limited number of observables in the applications outlined at the beginning of this section. Obviously, as it happened with the CGLE considered in section 4, delayed snapshots are unavoidable in the analysis of fluid flows involving a larger temporal complexity (and noise, when dealing with experimental data), but the analysis and description of these flows is more involved and well beyond the scope of this paper.

6. Concluding remarks. A new HODMD method has been developed that generalizes standard DMD by using the standard method with time-lagged snapshots. An algorithm labeled as DMD- d , where $d \geq 1$ indicates the number of involved time-delayed snapshots, has been derived that permits calculating the new decomposition in a robust and precise way.

The simplest algorithm, DMD-1, is very similar to the original DMD method developed by Schmid [48]. There is a *conceptual difference* between both, however, since DMD-1 separates calculations in two steps: the first SVD application to the whole snapshot matrix involves the *spatial truncation*, which defines from the outset the spatial modes, the spatial complexity N , and the reduced snapshots. The latter are used in the second step to calculate a DMD de-

composition for the reduced snapshots, which is readily translated into a DMD decomposition of the original snapshots by using the spatial modes.

For $d > 1$, the DMD- d algorithm is conceptually new and able to cope with cases (not accessible to standard DMD) in which the DMD-modes associated with different frequencies/growth rates are linearly dependent. The DMD- d algorithm turns out to be quite robust in connection with varying d . Since, after identification of the temporal modes, the application of DMD- d with large d is computationally inexpensive, a good strategy is to apply DMD-1 and DMD- d for several sample values of d and compare results to guess the appropriate value of d .

The new method has been illustrated and tested in several toy models, the Stuart–Landau equation, the Lorenz system, and the CGLE. In particular, in the applications to noisy toy models, it has been proven that the new HODMD method is much more precise and robust than standard DMD in the presence of noise.

For a periodic attractor in the Lorenz system involving a large number of harmonics, the new HODMD method quite efficiently identifies the correct basic frequency of the periodic solution and the amplitude of the harmonics. In addition, using the method in transient dynamics approaching the periodic attractor, the method is able to efficiently extrapolate to the final attractor.

Concerning the CGLE considered in section 4, a battery of representative periodic and quasi-periodic attractors has been considered. Except for trivial cases (namely, the attractors labeled as type I and II in section 4), the standard DMD gives spurious results but the new method correctly identifies the relevant frequencies and amplitudes to the precision that is being sought. In addition, the method has been used to recover the relevant frequencies considering a limited number of grid points, both for the clean databases and after adding random noise. Again, the method performs quite well since, in particular, applying the method with just one grid point, the results are much better than using FFT and PSD.

Concerning the more computationally demanding problem describing thermal convection in a rotating spherical shell considered in section 5, the advantages of using the delayed snapshots in the HODMD method have been illustrated in two representative attractors of the system, namely one periodic consisting in a purely rotating wave, and another quasi-periodic involving the nonlinear interaction between two rotating waves. Even though no intention was made to deal with actual observational data (which bears its own difficulties due to the unavoidable errors, whose treatment is well beyond the scope of this paper), the various tests have kept in mind one important application of numerical simulations on this problem, namely fitting observations and numerics. In this case, observations cannot cover a large amount of three-dimensional data, but a limited amount concentrated in some regions. When the amount of data is decreased too much, the analysis in section 5 illustrates that standard DMD may give completely spurious results, while the HODMD algorithm DMD- d , with appropriate d , always gives consistent results for general periodic and quasi-periodic dynamics, according to Theorem B in Appendix A, with obvious limitations in connection with accuracy, the sampling frequency (which must be somewhat large compared to the largest involved frequency), and the sampled timespan (which must be larger than the largest involved period). Obviously, when increasingly complex quasi-periodic phenomena is considered, the spatial complexity rapidly increases (in, e.g., transitional flows) and the standard DMD method may give spurious results

even using all available spatial data, as it happened with the CGLE. However, transitional flows exhibit their own difficulties (already in connection with numerical simulation), which once more are beyond the scope of this paper.

On the other hand, the DMD- d algorithm derived in the paper is also useful for other tasks, which are relevant from both the purely scientific and industrial points of view. Namely:

- As already mentioned, obtaining the growth rates and frequencies (and thus elucidating the dynamics type) using a limited amount of spatial data. In other words, the method provides a means to substitute spatial data by temporal data in this context. This is very useful to, e.g., calculating the natural frequencies from (wind tunnel or flight tests) aeroelastic or aeroacoustic data using a limited amount of accelerators (which decreases both cost and the experimental complexity). However, once more, this application is well beyond the scope of this paper.
- Elucidating the nature (periodic or quasi-periodic) of a given attractor by calculating the dimension of the attractor. It must be noted that, in practice, periodicity/quasi-periodicity cannot be ascertained by identifying the fundamental frequencies and elucidating whether they are incommensurable or not. As anticipated in the introduction, the computation of quasi-periodic attractors in large scale systems is a highly nontrivial issue [45, 46].

These are open problems which invite further analysis.

In any event, we hope that the results in this paper are a step further to increase robustness and flexibility of DMD in the many applications that can be envisaged.

Appendix A. Exact DMD-1 and DMD- d decompositions. Let us rewrite (1.3) as

$$(A.1) \quad \mathbf{v}_k = \sum_{m=1}^M a_m \mathbf{u}_m \mu_m^{k-1} \quad \text{for } k = 1, \dots, K,$$

with

$$(A.2) \quad \mu_m = e^{(\delta_m + i\omega_m)\Delta t} \quad \text{for } m = 1, \dots, M.$$

Moreover, we assume that (i) the complex scalars μ_1, \dots, μ_M are nonzero and distinct from each other (otherwise, those summands with the same μ_m are collected together), (ii) the complex amplitudes a_m are nonzero, and (iii) $K \geq M$ (in fact, the case of interest is $K \gg M$). Note that no assumption is made in connection with the dimension of the snapshots, J , which may well be smaller than M .

Now, the calculation of the DMD expansion (A.1) depends on whether the DMD modes are linearly independent or not. The former case is considered in the following well-known theorem, which is included here for completeness.

Theorem B. *Under the assumptions above, the linearly independent complex snapshots $\mathbf{v}_1, \dots, \mathbf{v}_K$ satisfy (A.1) if and only if there is a diagonalizable matrix \mathbf{R} , whose nonzero eigenvalues and associated eigenvectors are μ_m and \mathbf{u}_m (for $m = 1, \dots, M$), respectively, such that*

$$(B.1) \quad \mathbf{v}_{k+1} = \mathbf{R} \mathbf{v}_k \quad \text{for } k = 1, \dots, K.$$

Proof. If (A.1) holds, then all snapshots belong to the vector space $V = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_M\}$. If the vectors $a_1\mathbf{u}_1, \dots, a_M\mathbf{u}_M$ are taken as a basis in V , then the coordinates of the snapshots in this basis satisfy (B.1), with \mathbf{R} = diagonal matrix whose elements are μ_1, \dots, μ_M . This proves the direct statement. And conversely, if (B.1) holds, then we consider the M nonzero eigenvalues, μ_1, \dots, μ_M and the M linearly independent associated eigenvectors, $\mathbf{u}_1, \dots, \mathbf{u}_M$, which span a M -dimensional vector space, V . Equation (B.1) readily implies that the snapshots all belong to V and, taking $\mathbf{u}_1, \dots, \mathbf{u}_M$ as a basis for V , the matrix \mathbf{R} is the diagonal matrix whose diagonal elements are μ_1, \dots, μ_M , which readily yields (A.1), with $a_1 = \dots = a_M = 1$. ■

If the DMD modes are linearly dependent, then for each $d > 1$, we consider the Jd -dimensional *modified snapshots* and *modified modes* defined as

$$(B.2) \quad \tilde{\mathbf{v}}_k^d = \begin{bmatrix} \mathbf{v}_k \\ \mathbf{v}_{k+1} \\ \vdots \\ \mathbf{v}_{k+d-1} \end{bmatrix}, \quad \tilde{\mathbf{u}}_m^d = \begin{bmatrix} \mathbf{u}_m \\ \mu_m \mathbf{u}_m \\ \vdots \\ \mu_m^{d-1} \mathbf{u}_m \end{bmatrix}$$

for $k = 1, \dots, K - d + 1$ and $m = 1, \dots, M$. These allow for rewriting (A.1) as

$$(B.3) \quad \tilde{\mathbf{v}}_k^d = \sum_{m=1}^M \mu_m^{k-1} \tilde{\mathbf{u}}_m^d \quad \text{for } k = 1, \dots, K - d + 1.$$

The counterpart of (B.1) would be

$$(B.4) \quad \tilde{\mathbf{v}}_{k+1} = \tilde{\mathbf{R}} \tilde{\mathbf{v}}_k \quad \text{for } k = 1, \dots, K - d + 1.$$

Now, the idea is that the modified modes defined in (B.2) are linearly independent if d is sufficiently large, as proven in the following lemma.

Lemma C. *Under the assumptions above, there is a minimum value of d , $d_0 \leq M$, such that for each $d \geq d_0$, the modified modes defined in (B.2) are linearly independent.*

Proof. The aim is to prove that there is a minimum value of $d \leq M$ such that the $Jd \times M$ matrix

$$(C.1) \quad \tilde{\mathbf{U}}^d = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_M \\ \mu_1 \mathbf{u}_1 & \mu_2 \mathbf{u}_2 & \dots & \mu_M \mathbf{u}_M \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1^{d-1} \mathbf{u}_1 & \mu_2^{d-1} \mathbf{u}_2 & \dots & \mu_M^{d-1} \mathbf{u}_M \end{bmatrix}$$

exhibits maximum rank, namely its rank is M . This is proven in two steps. First, if $\tilde{\mathbf{U}}^d$ exhibits a maximum rank for some d , then for larger d , $\tilde{\mathbf{U}}^d$ obviously exhibits a maximum rank, too. Second, we prove that $\tilde{\mathbf{U}}^M$ exhibits maximum rank as follows. We note that under the stated assumptions, the J -dimensional modes $\mathbf{u}_1, \dots, \mathbf{u}_M$ span a vector space whose dimension $M_0 \leq M \leq J$. Then, there is a unit vector, \mathbf{w}_1 , such that the orthogonal projections of $\mathbf{u}_1, \dots, \mathbf{u}_M$ on \mathbf{w}_1 are all nonzero. Completing this vector with another $J - 1$ unit vectors to construct an orthonormal basis of \mathbb{C}^J , and referring the modes \mathbf{u}_m to this new basis,

we consider the $M \times M$ -submatrix of \tilde{U}^M formed by the M rows associated with the first components of \mathbf{u}_m . Namely, this submatrix is formed by the first, $(J+1)$ th, $(2J+1)$ th, \dots , $((d-1)J+1)$ th rows, and reads

$$(C.2) \quad \begin{bmatrix} u_1^1 & u_2^1 & \dots & u_M^1 \\ \mu_1 u_1^1 & \mu_2 u_2^1 & \dots & \mu_M u_M^1 \\ \dots & \dots & \dots & \dots \\ \mu_1^{M-1} u_1^1 & \mu_2^{M-1} u_2^1 & \dots & \mu_M^{M-1} u_M^1 \end{bmatrix},$$

where, by construction, u_m^1 is the nonzero orthogonal projection of \mathbf{u}_m on the first element of the basis, \mathbf{w}_1 . Now, dividing the m th column of this matrix by u_m^1 , for $m = 1, \dots, M$, the resulting matrix is a Vandermonde matrix, which is nonsingular because the scalars μ_1, \dots, μ_M are all nonzero and different from each other, as assumed. Thus, the rank of the submatrix (C.2) is M and so is the rank of the larger matrix \tilde{U}^M (defined as in (C.1) with $d = M$), which completes the proof. ■

Using this, the case of linearly dependent modes is considered in the following theorem.

Theorem D. *If (A.1) holds with the vectors $\{\mathbf{u}_m\}_{m=1}^M$ not necessarily linearly independent and the complex scalars μ_1, \dots, μ_M different among each other, then there exists a smallest value of d , with $1 < d_0 \leq M$, such that if $d \geq d_0$, then*

$$\mathbf{v}_{k+d} = \mathbf{R}_1 \mathbf{v}_{k+d-1} + \dots + \mathbf{R}_d \mathbf{v}_k \quad \text{for } k = 1, \dots, K - d + 1$$

for appropriate matrices, $\mathbf{R}_1, \dots, \mathbf{R}_d$.

Proof. Let d_0 be as in Lemma C. Applying Theorem B, (B.3) holds if and only if there is a matrix $\tilde{\mathbf{R}}$ such that (B.4) holds. Thus, we need only consider the J first rows in (B.3) and (B.4) to complete the proof. ■

Note that the matrices appearing in (C.1) are not unique, especially when $d > d_0$. A trivial counterexample of this occurs when the modes are linearly independent, in which the snapshots satisfy, e.g., $\mathbf{v}_{k+1} = \mathbf{R}_1 \mathbf{v}_k$, $\mathbf{v}_{k+1} = \mathbf{R}_1^2 \mathbf{v}_{k-1}$, and $\mathbf{v}_{k+1} = (\mathbf{R}_1 \mathbf{v}_k + \mathbf{R}_1^2 \mathbf{v}_{k-1})/2$.

The results above could be seen as theoretical foundations of the standard DMD and the HODMD methods. However, this theorem applies to the case in which (A.1) is exact and calculations are performed with infinite precision, which is not possible in practice. Thus, some consequences from the above can be misleading in practical situations.

The proof of Lemma C suggests that, for a given index d , the complex scalars μ_m must be sufficiently different among each other for a good performance of the HODMD methods. Invoking (A.2), this condition implies that Δt should not be too small, namely the sampling frequency should not be too large for a given sampled timespan. On the other hand, the sampling frequency should be sufficiently large. As further explained in this paper, the sampling frequency must be somewhat larger than the largest frequency that is present in the DMD-expansion. Solving these contradictory conclusions will require taking a sufficiently large value of the index d , which will somehow scale with the sampling frequency.

Acknowledgments. The authors are indebted to Dr. Esteban Ferrer, Professors Marta Net, Joan Sanchez, and Fernando Varas, and also to two anonymous referees for useful comments on an earlier version of the paper.

REFERENCES

- [1] M.J. ABLOWITZ, *Nonlinear Dispersive Waves. Asymptotic Analysis and Solitons*, Cambridge University Press, New York, 2011.
- [2] T. ALBRECHT, H.M. BLACKBURN, J.M. LOPEZ, R. MANASSEH, AND P. MEUNIER, *Triadic resonances in precessing rapidly rotating cylinder flows*, J. Fluid Mech., 778 (2015), R1, <https://doi.org/10.1017/jfm.2015.377>.
- [3] N. ANDERSSON AND L.-E. ERIKSSON, *A novel solver acceleration technique based on dynamic mode decomposition*, in Proceedings of the 6th European Conference on Computational Fluid Dynamics (ECFD VI), Barcelona, Spain, 2014, pp. 4832–4851.
- [4] I.S. ARANSON AND L. KRAMER, *The world of the complex Ginzburg-Landau equation*, Rev. Modern Phys., 74 (2002), pp. 99–143.
- [5] S. BAGHERI *Koopman-mode decomposition of the cylinder wake*, J. Fluid Mech., 726 (2013), pp. 596–623.
- [6] T. BERRY, R. CRESSMAN, Z. GREGURIC-FERENCSEK, AND T. SAUER, *Time-scale separation from diffusion-mapped delay coordinates*, SIAM J. Appl. Dyn. Syst., 12 (2013), pp. 618–649, <https://doi.org/10.1137/12088183X>.
- [7] T. BERRY, D. GIANNAKIS, AND J. HARLIM, *Nonparametric forecasting of low-dimensional dynamical systems*, Phys. Rev. E, 91 (2015), 032915, <https://doi.org/10.1103/PhysRevE.91.032915>.
- [8] T. BERRY AND J. HARLIM, *Forecasting turbulent modes with nonparametric diffusion models: Learning from noisy data*, Phys. D, 329 (2016), pp. 57–76.
- [9] E. BRADLEY AND H. KANTZ, *Nonlinear time-series analysis revisited*, Chaos, 25 (2015), 097610.
- [10] D.S. BROOMHEAD AND G.P. KING, *Extracting qualitative dynamics from experimental data*, Phys. D, 20 (1986), pp. 217–236.
- [11] S.L. BRUNTON, B.W. BRUNTON, J.L. PROCTOR, AND J.N. KUTZ, *Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control*, PLoS ONE, 11 (2016), e0150171, <https://doi.org/10.1371/journal.pone.0150171>.
- [12] S.L. BRUNTON, J.L. PROCTOR, AND J.N. KUTZ, *Discovering governing equations from data: Sparse identification of nonlinear dynamical systems*, Proc. Natl. Acad. Sci. USA, 113 (2016), pp. 3932–3937.
- [13] M. CASDAGLI, *Nonlinear prediction of chaotic time series*, Phys. D, 35 (1989), pp. 335–356.
- [14] T. CEBECI, *Convective Heat Transfer*, Springer-Verlag, Berlin, 2002.
- [15] S. CHANDRASEKHAR, *Hydrodynamic and Hydromagnetic Stability*, Oxford University Press, Oxford, UK, 1961.
- [16] A. CHATTERJEE, *An introduction to the proper orthogonal decomposition*, Current Science, 78 (2000), pp. 808–817.
- [17] P. CHARBONNEAU, *Solar and Stellar Dynamos*, Springer-Verlag, Berlin, Heidelberg, 2013.
- [18] 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 (2012), pp. 8871–8875.
- [19] M. CROSS AND H. GREENSIDE, *Pattern Formation and Dynamics in Nonequilibrium Systems*, Cambridge University Press, Cambridge, UK, 2009.
- [20] B. ECKHARDT AND G. OTT, *Periodic orbit analysis of the Lorenz attractor*, Z. Phys. B, 93 (1994), pp. 259–266.
- [21] J.D. FARMER AND J.J. SIDOROWICH, *Predicting chaotic time series*, Phys. Rev. Lett., 59 (1987), pp. 845–848.
- [22] C. FOIAS, G.R. SELL, AND R. TEMAM, *Inertial manifolds for nonlinear evolution equations*, J. Differential Equations, 73 (1988), pp. 309–353, [https://doi.org/10.1016/0022-0396\(88\)90110-6](https://doi.org/10.1016/0022-0396(88)90110-6).
- [23] F. GARCIA, M. NET, B. GARCIA-ARCHILLA, AND J. SANCHEZ, *A comparison of high-order time integrators for thermal convection in rotating spherical shells*, J. Comput. Phys., 229 (2010), pp. 7997–8010, <https://doi.org/10.1016/j.jcp.2010.07.004>.
- [24] D. GIANNAKIS, J. SLAWINSKA, AND Z. ZHAO, *Spatiotemporal feature extraction with data-driven Koopman operators*, JMLR: Workshop and Conference Proceedings, 44 (2015), pp. 103–115.
- [25] G.H. GOLUB AND G.T. VAN LOAN, *Matrix Computations*, John Hopkins University Press, Baltimore, MD, 1996.

- [26] M. HARAGUS AND G. IOOSS, *Local Bifurcations, Center Manifolds, and Normal Forms in Infinite Dimensional Dynamical Systems*, Springer-Verlag, London, 2011.
- [27] M.W. HIRSCH, *Differential Topology*, Springer-Verlag, New York, Heidelberg, 1976.
- [28] G. JOURDAIN, L.-E. ERIKSSON, S. H. KIM, AND C. H. SOHN, *Application of dynamic mode decomposition to acoustic-modes identification and damping in a 3-dimensional chamber with baffled injectors*, J. Sound Vibration, 332 (2013), pp. 4308–4323, <https://doi.org/10.1016/j.jsv.2013.02.041>.
- [29] M.R. JOVANOVIĆ, P.J. SCHMID, AND J.W. NICHOLS, *Sparsity-promoting dynamic mode decomposition*, Phys. Fluids, 26 (2014), 024103.
- [30] F. VON KNOBLAUCH, R. MORENO, P.F. TAYLOR, AND J. NEWSOM, *Analytical correlation of a flexible empennage wind tunnel flutter test at high transonic Mach number*, in Proceedings of the 55th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, National Harbor, Maryland, AIAA2014-0676, 2014, pp. 1–20.
- [31] B. KOOPMAN, *Hamiltonian systems and transformations in Hilbert space*, Proc. Natl Acad. Sci. USA, 17 (1931), pp. 315–318.
- [32] S. LE CLAINCHE, D. RODRÍGUEZ, V. THEOFILIS, AND J. SORIA *Flow around a hemisphere-cylinder at high angle of attack and low Reynolds number. Part II: POD and DMD applied to reduced domains*, Aerosp. Sci. Technol., 44 (2015), pp. 88–100.
- [33] E.N. LORENZ, *Deterministic nonperiodic flow*, J. Atmos. Sci., 20 (1963), pp. 130–141.
- [34] E.N. LORENZ, *Atmospheric predictability as revealed by naturally occurring analogues*, J. Atmos. Sci., 26 (1969), pp. 636–646, [https://doi.org/10.1175/1520-0469\(1969\)26\(636:APARBN\)2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)26(636:APARBN)2.0.CO;2).
- [35] J.J. MCNAMARA AND P.P. FRIEDMAN, *Flutter boundary identification for time-domain computational aeroelasticity*, AIAA J., 45 (2007), pp. 1546–1555, <https://doi.org/10.2514/1.26706>.
- [36] I. MEZIĆ, *Analysis of fluid flows via spectral properties of the Koopman operator*, Ann. Rev. Fluid Mech., 45 (2013), pp. 357–378.
- [37] N.H. PACKARD, J.P. CRUTCHFIELD, J.D. FARMER, AND R.S. SHAW, *Geometry from a time series*, Phys. Rev. Lett., 45 (1980), pp. 712–716.
- [38] W.H. PRESS, S.A. TEUKOLSKY, W.T. VETTERLING, AND B.P. FLANNERY, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, 1988.
- [39] W.H. PRESS, S.A. TEUKOLSKY, W.T. VETTERLING, AND B.P. FLANNERY, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, New York, 1992.
- [40] J.L. PROCTOR AND P.A. ECKHOFF, *Discovering dynamic patterns from infectious disease data using dynamic mode decomposition*, Int. Health, 7 (2015), pp. 139–145.
- [41] M. PROVANASAL, C. MATHIS, AND L. BOYER, *Bénard-von Kármán instability: transient and forcing regimes*, J. Fluid Mech., 182 (1987), pp. 1–22.
- [42] M. RAGWITZ AND H. KANTZ, *Markov models from data by simple nonlinear time series predictors in delay embedding spaces*, Phys. Rev. E, 65 (2002), 056201.
- [43] C.W. ROWLEY, I. MEZIĆ, S. BAGHERI, P. SCHLATTER, AND D.S. HENNINGSON, *Spectral analysis of nonlinear flows*, J. Fluid Mech., 641 (2009), pp. 115–127.
- [44] T. SAUER, J.A. YORKE, AND M. CASDAGLI, *Embedology*, J. Statist. Phys., 65 (1991), pp. 579–616.
- [45] J. SANCHEZ, M. NET, AND C. SIMO, *Computation of invariant tori by Newton-Krylov methods in large-scale dissipative systems*, Phys. D, 239 (2010), pp. 123–133.
- [46] J. SANCHEZ AND M. NET, *A parallel algorithm for the computation of invariant tori in large-scale dissipative systems*, Phys. D, 252 (2013), pp. 22–33.
- [47] J. SANCHEZ, F. GARCIA, AND M. NET, *Computation of azimuthal waves and their stability in thermal convection in rotating spherical shells with application to the study of a double-Hopf bifurcation*, Phys. Rev. E, 87 (2013), 033014.
- [48] P.J. SCHMID, *Dynamic mode decomposition of numerical and experimental data*, J. Fluid Mech., 656 (2010), pp. 5–28.
- [49] P.J. SCHMID, *Application of the dynamic mode decomposition to experimental data*, Exp. Fluids, 50 (2011), pp. 1123–1130, <https://doi.org/10.1007/s00348-010-0911-3>.
- [50] P.J. SCHMID, L. LI, M.P. JUNIPER, AND O. PUST, *Applications of the dynamic mode decomposition*, Theor. Comp. Fluid Dyn., 25 (2011), pp. 249–259, <https://doi.org/10.1007/s00162-010-0203-9>.
- [51] P.J. SCHMID, D. VIOLATO, AND F. SCARANO, *Decomposition of time-resolved tomographic PIV*, Exp. Fluids, 52 (2012), pp. 1567–1779, <https://doi.org/10.1007/s00348-012-1266-8>.

- [52] P. STEGEMAN, A. OOI, AND J. SORIA, *Proper orthogonal decomposition and dynamic mode decomposition of under-expanded free-jets with varying nozzle pressure ratios*, Fluid Mech. Appl. 107, Springer, Cham, 2015, pp. 85–90, https://doi.org/10.1007/978-3-319-06260-0_12.
- [53] G.W. STEWART, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Rev., 15 (1973) 727–764, <https://doi.org/10.1137/1015095>.
- [54] F. TAKENS, *Detecting strange attractors in turbulence*, Lecture Notes in Math. 898, D.A. Rand and L.-S. Young, eds., Springer, Berlin, New York, 1981, pp. 366–381.
- [55] F. TERRAGNI AND J.M. VEGA, *On the use of POD-based ROMs to analyze bifurcations in some dissipative systems*, Phys. D, 241 (2012), pp. 1393–1405, <https://doi.org/10.1016/j.physd.2012.04.009>.
- [56] F. TERRAGNI AND J.M. VEGA, *Construction of bifurcation diagrams using POD on the fly*, SIAM J. Appl. Dyn. Syst., 13 (2014), pp. 339–365, <https://doi.org/10.1137/130927267>.
- [57] M.E. WALL, A. RECHTSTEINER, AND L.M. ROCHA, *Singular value decomposition and principal component analysis*, in A Practical Approach to Microarray Data Analysis, D.P. Berrar, W. Dubitzky, and M. Granzowr, eds., Kluwer, Norwell, MA, 2003, pp. 91–109.
- [58] 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 (2015), pp. 1307–1346.
- [59] M.O WILLIAMS, C.W. ROWLEY, AND I.G. KEVREKIDIS, *A kernel-based method for data-driven Koopman spectral analysis*, J. Comput. Dyn., 2 (2015), pp. 247–265.
- [60] J.C. WILLEMS, *From time series to linear system. Part I. Finite dimensional linear time invariant systems*, Automatica, 22 (1986), pp. 561–580, [https://doi.org/10.1016/0005-1098\(86\)90066-X](https://doi.org/10.1016/0005-1098(86)90066-X).
- [61] A. WYNN, D.S. PEARSON, B. GANAPATHISUBRAMANIA, AND P.J. GOULARTA, *Optimal mode decomposition for unsteady flows*, J. Fluid Mech., 733 (2013), pp. 473–503.
- [62] Z. ZHAO AND D. GIANNAKIS, *Analog forecasting with dynamics-adapted kernels*, Nonlinearity, 29 (2016), pp. 2888–2939.