

# Data-driven modeling and prediction of non-linearizable dynamics via spectral submanifolds

Mattia Cenedese <sup>1</sup>, Joar Axås <sup>1</sup>, Bastian Bäuerlein <sup>2,3</sup>, Kerstin Avila <sup>2,3</sup> & George Haller <sup>1</sup>✉

We develop a methodology to construct low-dimensional predictive models from data sets representing essentially nonlinear (or *non-linearizable*) dynamical systems with a hyperbolic linear part that are subject to external forcing with finitely many frequencies. Our data-driven, sparse, nonlinear models are obtained as extended normal forms of the reduced dynamics on low-dimensional, attracting spectral submanifolds (SSMs) of the dynamical system. We illustrate the power of data-driven SSM reduction on high-dimensional numerical data sets and experimental measurements involving beam oscillations, vortex shedding and sloshing in a water tank. We find that SSM reduction trained on unforced data also predicts nonlinear response accurately under additional external forcing.

<sup>1</sup>Institute for Mechanical Systems, ETH Zürich, Leonhardstrasse 21, 8092 Zürich, Switzerland. <sup>2</sup>University of Bremen, Faculty of Production Engineering, Badgasteiner Strasse 1, 28359 Bremen, Germany. <sup>3</sup>Leibniz Institute for Materials Engineering IWT, Badgasteiner Strasse 3, 28359 Bremen, Germany.  
✉email: [georgehaller@ethz.ch](mailto:georgehaller@ethz.ch)

Low-dimensional reduced models of high-dimensional nonlinear dynamical systems are critically needed in various branches of applied science and engineering. Such simplified models would significantly reduce computational costs and enable physical interpretability, design optimization and efficient controllability. As of yet, however, no generally applicable procedure has emerged for the reliable and robust identification of nonlinear reduced models.

Instead, the most broadly used approach to reducing nonlinear dynamical systems has been a fundamentally linear technique, the proper orthogonal decomposition (POD), followed by a Galerkin projection<sup>1–3</sup>. Projecting the full dynamics to the most energetic linear modes, POD requires the knowledge of the governing equations of the system and hence is inapplicable when only data is available. As purely data-based alternatives, machine learning methods are broadly considered and tested in various fields<sup>4–7</sup>. While the black-box approach of machine learning might often seem preferable to a detailed nonlinear analysis, the resulting neural network models require extensive tuning, lack physical interpretability, generally perform poorly outside their training range and tend to be unnecessarily complex<sup>8</sup>. This has inspired a number of approaches that seek a blend of machine learning with a priori information about the underlying physics<sup>9,10</sup>. Still within the realm of machine learning, sparse regression has also shown promise in approximating the right-hand sides of low-dimensional, simple dynamical systems with functions taken from a preselected library<sup>4</sup>. Another recent approach is cluster-based network modeling, which uses the toolkit of network science and statistical physics for modeling nonlinear dynamics<sup>11</sup>.

A popular alternative to POD and machine learning is the dynamic mode decomposition (DMD)<sup>12</sup>, which approximates directly the observed system dynamics. The original DMD and its later variants fit a linear dynamical system to temporally evolving data, possibly including further functions of the original data, over a given finite time interval<sup>13</sup>. DMD provides an appealingly simple yet powerful algorithm to infer a local model near steady states where the nonlinear dynamics is always approximately linear. This linear model is also more globally valid if constructed over observables lying in a span of some eigenfunctions of the Koopman operator, which maps observables evaluated over initial states into their evaluations over current states<sup>14–16</sup>. This relationship between DMD and the Koopman operator has motivated an effort to machine-learn Koopman eigenfunctions from data in order to linearize nonlinear dynamical systems globally on the space of their observables<sup>17–19</sup>.

Finding physically realizable observables that fall in a Koopman eigenspace is, however, often described as challenging or difficult<sup>20</sup>. A more precise assessment would be that such a find is highly unlikely, given that the probability of any countable set of a priori selected observables falling in any Koopman eigenspace is zero. In addition, those eigenspaces can only be determined explicitly in simple, low-dimensional systems. In practice, therefore, DMD can only provide a justifiable model near an attracting fixed point of a dynamical system. While Koopman modes still have the potential to linearize the observer dynamics on larger domains, those domains cannot include more than one attracting or repelling fixed point<sup>19–21</sup>. Indeed, DMD and Koopman mode expansions fail to converge outside neighborhoods of fixed points even in the simplest, one-dimensional nonlinear systems with two fixed points<sup>20,22</sup>. In summary, while these data-driven model reduction methods are powerful and continue to inspire ongoing research, their applicability is limited to locally linearized systems and globally linearizable nonlinear systems, such as the Burgers equation<sup>23</sup>.

The focus of this paper is the development of data-driven, simple and predictive reduced-order models for essentially

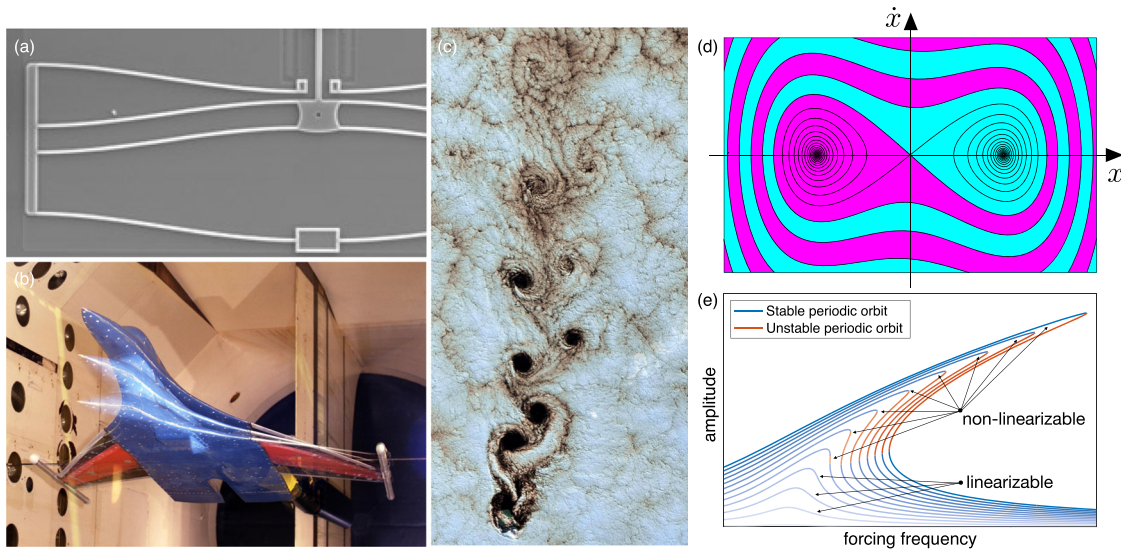
nonlinear dynamical systems, i.e., nonlinearizable systems. Determining exact linearizability conclusively from data is beyond reach. In contrast, establishing that a dynamical system is nonlinearizable in a domain of interest is substantially simpler: one just needs to find an indication of coexisting isolated stationary states in the data. By an isolated stationary state, we mean here a compact and connected invariant set with an open neighborhood that contains no other compact and connected invariant set. Examples of such stationary states include hyperbolic fixed points, periodic orbits, invariant spheres and quasi-periodic tori; closures of homoclinic orbits and heteroclinic cycles; and chaotic attractors and repellers. If a data set indicates the coexistence of any two sets from the above list, then the system is conclusively non-linearizable in the range of the available data. Specifically, there will be no homeomorphism (continuous transformation with a continuous inverse) that transforms the orbits of the underlying dynamical system into those of a linear dynamical system. While this is a priori clear from dynamical systems theory, several studies have specifically confirmed a lack of convergence of Koopman-mode expansions already for the simplest case of two coexisting fixed points, even over subsets of their domain of attraction or repulsion<sup>20,22</sup>.

Non-linearizable systems are ubiquitous in science, technology and nature. Beyond the well-known examples of chaotic dynamical systems and turbulent fluid flows<sup>1</sup>, any bifurcation phenomenon, by definition, involves coexisting steady states and hence is automatically non-linearizable. Indeed, aerodynamic flutter<sup>24</sup>, buckling of beams and shells<sup>25</sup>, bistable microelectromechanical systems<sup>26</sup>, traffic jams<sup>27</sup> or even tipping points in climate change<sup>28</sup> are all fundamentally non-linearizable, just to name a few. Figure 1 shows some examples of non-linearizable systems emerging in technology, nature and scientific modeling.

We will show here that a collection of classic and recent mathematical results from nonlinear dynamical systems theory enables surprisingly accurate and predictive low-dimensional modeling from data for a number of non-linearizable phenomena. Our construct relies on the recent theory of spectral submanifolds (SSMs), the smoothest invariant manifolds that act as nonlinear continuations of non-resonant eigenspaces from the linearization of a system at a stationary state (fixed point, periodic orbit or quasiperiodic orbit<sup>29</sup>). Using appropriate SSM embeddings<sup>30–32</sup> and an extended form of the classic normal form theory<sup>33</sup>, we obtain sparse dynamical systems describing the reduced dynamics on the slowest SSMs of the system, which are normally hyperbolic and hence robust under perturbations<sup>34</sup>.

We construct the extended normal form within the slowest SSM as if the eigenvalues of the linearized dynamics within the SSM had zero real parts, although that is not the case. As a result, our normalization procedure will not render the simplest possible (linear) normal form for the SSM dynamics, valid only near the underlying isolated stationary state. Instead, our procedure yields a sparsified nonlinear, polynomial normal form on a larger domain of the SSM that can also capture nearby coexisting stationary states. This fully data-driven normalization algorithm learns the normal form transformation and the coefficients of the normal form simultaneously by minimizing an appropriately defined conjugacy error between the unnormalized and normalized SSM dynamics.

For a generic observable of an oscillatory dynamical system without an internal resonance, a two-dimensional data-driven model calculated on the slowest SSM of the system turns out to capture the correct asymptotic dynamics. Such an SSM-reduced model is valid on domains in which the nonlinearity and any possible external forcing are strong enough to create non-linearizable dynamics, yet are still moderate enough to render the eigenspace of the linear system relevant. More generally,



**Fig. 1** Examples of non-linearizable systems. **a** Snap-through instability of a microelectro-mechanical (MEMS) device with three coexisting equilibria (Sandia National Laboratories). **b** Wind-tunnel flutter of an airplane prototype, involving a fixed point and coexisting limit cycles (NASA Langley Research Center). **c** Swirling clouds behind an island in the Pacific ocean, forming a vortex street with coexisting isolated hyperbolic and elliptic trajectories for the dynamical system describing fluid particle motion (USGS/NASA). **d** Phase portrait of the damped, double-well Duffing oscillator  $\ddot{x} + \dot{x} - x + \beta x^3 = 0$  with  $\beta > 0$ , the most broadly used model for nonlinear systems with coexisting domains of attraction (colored), such as the MEMS device in plot (a). **e** Nonlinear response amplitude ( $|x(t)|_{\max}$ ) in the forced-damped, single-well Duffing oscillator,  $\ddot{x} + \dot{x} + x + \beta x^3 = f \cos \omega t$  with  $\beta > 0$ , under variations of the forcing frequency  $\omega$  and forcing amplitude  $f$ . Coexisting stable and unstable periodic responses show non-linearizable dynamics conclusively for this classic model.

oscillatory systems with  $m$  independent internal resonances in their spectrum can be described by reduced models on  $2(m + 1)$ -dimensional SSMs. In both the resonant and the non-resonant cases, the models can be refined by increasing the degree of their nonlinearity rather than by increasing their dimension. As we show in examples, the resulting SSM-based models are explicit, deterministic and even have the potential to predict system behavior outside the range of the training data away from bifurcations. Most importantly, we find that the models also accurately predict forced response, even though they are only trained on data collected from unforced systems.

We illustrate the power of data-driven SSM-reduced models on high-dimensional numerically generated data sets and on experimental data. These and further examples are also available as MATLAB® live scripts, which are part of a general open-source package, SSMLearn, that performs this type of model reduction and prediction for arbitrary data sets.

**Results**

**Spectral submanifolds and their reduced dynamics.** A recent result in dynamical systems is that all eigenspaces (or spectral subspaces) of linearized systems admit unique nonlinear continuations under well-defined mathematical conditions. Specifically, spectral submanifolds (SSMs), as defined by<sup>29</sup>, are the unique smoothest invariant manifolds that serve as nonlinear extensions of spectral subspaces under the addition of nonlinearities to a linear system. The SSM formulation and terminology we use here is due to<sup>29</sup>; the Methods section “Existence of SSMs” discusses the history of these results and further technical details.

We consider  $n$ -dimensional dynamical systems of the form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{f}_0(\mathbf{x}) + \epsilon \mathbf{f}_1(\mathbf{x}, \boldsymbol{\Omega}t; \epsilon), \quad \mathbf{f}_0(\mathbf{x}) = \mathcal{O}(|\mathbf{x}|^2), \quad 0 \leq \epsilon \ll 1, \tag{1}$$

with a constant matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , and with class  $C^r$  functions  $\mathbf{f}_0 : \mathcal{U} \rightarrow \mathbb{R}^n$  and  $\mathbf{f}_1 : \mathcal{U} \times \mathbb{T}^\ell \rightarrow \mathbb{R}^n$ , where  $\mathbb{T}^\ell = S^1 \times \dots \times S^1$  is the  $\ell$ -dimensional torus. The elements of the frequency vector

$\boldsymbol{\Omega} \in \mathbb{R}^\ell$  are rationally independent, and hence the function  $\mathbf{f}_1$  is quasiperiodic in time. The assumed degree of smoothness for the right-hand side of (1) is  $r \in \mathbb{N}^+ \cup \{\infty, a\}$ , with  $a$  referring to analytic. The small parameter  $\epsilon$  signals that the forcing in system (1) is moderate so that the structure of the autonomous part is still relevant for the full system dynamics. Rigorous mathematical results on SSMs are proven for small enough  $\epsilon$ , but continue to hold in practice for larger values of  $\epsilon$  as well, as we will see in examples. Note that eq. (1) describes equations of motions of physical oscillatory systems. It does not cover phenomenological models of phase oscillators, such as the Kuramoto model<sup>35</sup>.

The eigenvalues  $\lambda_j = \alpha_j + i\omega_j \in \mathbb{C}$  of  $\mathbf{A}$ , with multiplicities counted, are ordered based on their real parts,  $\text{Re}\lambda_j$ , as

$$\text{Re}\lambda_n \leq \text{Re}\lambda_{n-1} \leq \dots \leq \text{Re}\lambda_1. \tag{2}$$

Their corresponding real modal subspaces (or eigenspaces),  $E_j \subset \mathbb{R}^n$ , are spanned by the imaginary and real parts of the corresponding eigenvectors and generalized eigenvectors of  $\mathbf{A}$ . To analyze typical systems, we assume that  $\text{Re}\lambda_j = \alpha_j \neq 0$  holds for all eigenvalues, i.e.,  $\mathbf{x} = \mathbf{0}$  is a hyperbolic fixed point for  $\epsilon = 0$ .

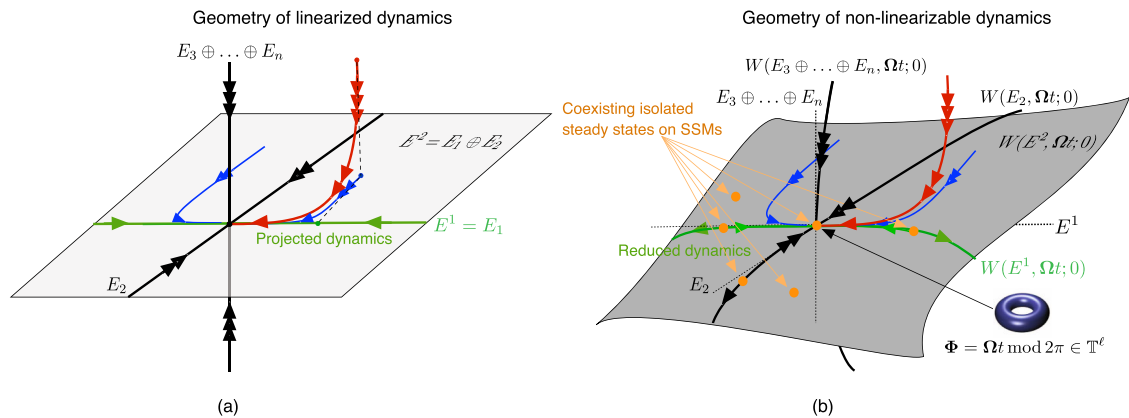
A spectral subspace  $E_{j_1, \dots, j_q}$  is a direct sum

$$E_{j_1, \dots, j_q} = E_{j_1} \oplus E_{j_2} \oplus \dots \oplus E_{j_q} \tag{3}$$

of an arbitrary collection of modal subspaces, which is always an invariant subspace for the linear part of the dynamics in (1). Classic examples of spectral subspaces are the stable and unstable subspaces, comprising all modal subspaces with  $\text{Re}\lambda_k < 0$  and  $\text{Re}\lambda_k > 0$ , respectively. Projections of the linearized system onto the nested hierarchy of slow spectral subspaces,

$$E^1 \subset E^2 \subset E^3 \subset \dots, \quad E^k := E_{1, \dots, k}, \quad k = 1, \dots, n, \tag{4}$$

provide exact reduced-order models for the linearized dynamics over an increasing number of time scales under increasing  $k$ , as sketched in panel (a) of Fig. 2. This is why a Galerkin projection onto  $E^k$  is an exact model reduction procedure for linear systems, whose accuracy can be increased by increasing  $k$ . A fundamental question is whether nonlinear analogues of spectral subspaces



**Fig. 2 Linear vs. nonlinear model reduction.** **a** Reduction of linear dynamics via Galerkin projection. The slowest spectral subspace,  $E^1 = E_1$  (green), and the modal subspace,  $E_2$  (black), span together the second slowest spectral subspace,  $E^2 = E_1 \oplus E_2$ . The full dynamics (red curve) can be projected onto  $E^1$  to yield a reduced slow model without transients. Projection of the full dynamics onto  $E^2$  (blue curve) yields a reduced model that also captures the slowest decaying transient. Further, faster-decaying transients can be captured by projections onto slow spectral subspaces,  $E^k$ , with  $k > 1$ . **b** Reduction of nonlinearizable dynamics via restriction to spectral submanifolds (SSMs) in the  $\epsilon = 0$  limit of nonlinear, non-autonomous systems forced with  $\ell$  frequencies. An SSM,  $W(E, \Omega t; 0)$ , is the unique, smoothest, nonlinear continuation of a nonresonant spectral subspace  $E$ . Specifically, the slowest SSM,  $W(E^k, \Omega t; 0)$  (green), is the unique, smoothest, nonlinear continuation of the slowest spectral subspace,  $E^k$ . Nonlinearizability of the full dynamics follows if isolated stationary states coexist on at least one of the SSMs. The time-quasiperiodic SSMs for  $\epsilon > 0$ , denoted  $W(E, \Omega t; \epsilon)$ , are not shown here but they are  $C^\ell$ -close to the structures shown, as discussed by<sup>29</sup>.

continue to organize the dynamics under the addition of nonlinear and time-dependent terms in the full system (1).

Let us fix a specific spectral subspace  $E = E_{j_1, \dots, j_d}$  within either the stable or the unstable subspace. If  $E$  is non-resonant (i.e., no nonnegative, low-order, integer linear combination of the spectrum of  $\mathbf{A}|_E$  is contained in the spectrum of  $\mathbf{A}$  outside  $E$ ), then  $E$  has infinitely many nonlinear continuations in the system (1) for  $\epsilon$  small enough<sup>29</sup>. These invariant manifolds are of smoothness class  $C^{\Sigma(E)}$ , with the spectral quotient  $\Sigma(E)$  measuring the ration of the fastest decay exponent outside  $E$  to the slowest decay exponent inside  $E$  (see eq. (13) of the Methods section “Existence of SSMs”). All such manifolds are tangent to  $E$  for  $\epsilon = 0$ , have the same quasiperiodic time dependence as  $\mathbf{f}_1$  does and have a dimension equal to that of  $E$ .

Of these infinitely many invariant manifolds, however, there will be a unique smoothest one, the *spectral submanifold* (SSM) of  $E$ , denoted  $W(E, \Omega t; \epsilon)$ . This manifold is  $C^r$  smooth if  $r > \Sigma(E)$  and can therefore be approximated more accurately than the other infinitely many nonlinear continuations of  $E$ . In particular, SSMs have convergent Taylor expansions if the dynamical system (1) is analytic ( $r = a$ ). Then the reduced dynamics on a slow SSM,  $E^k$ , can be approximated with arbitrarily high accuracy using arbitrarily high-order Taylor expansions, without ever increasing the dimension of  $E^k$ , see panel (b) of Fig. 2. Such an approximation for dynamical systems with known governing equations is now available for any required order of accuracy via the open-source MATLAB<sup>®</sup> package `SSMToolbox`<sup>36</sup>. In contrast, reduced models obtained from projection-based procedures can only be improved by increasing their dimensions.

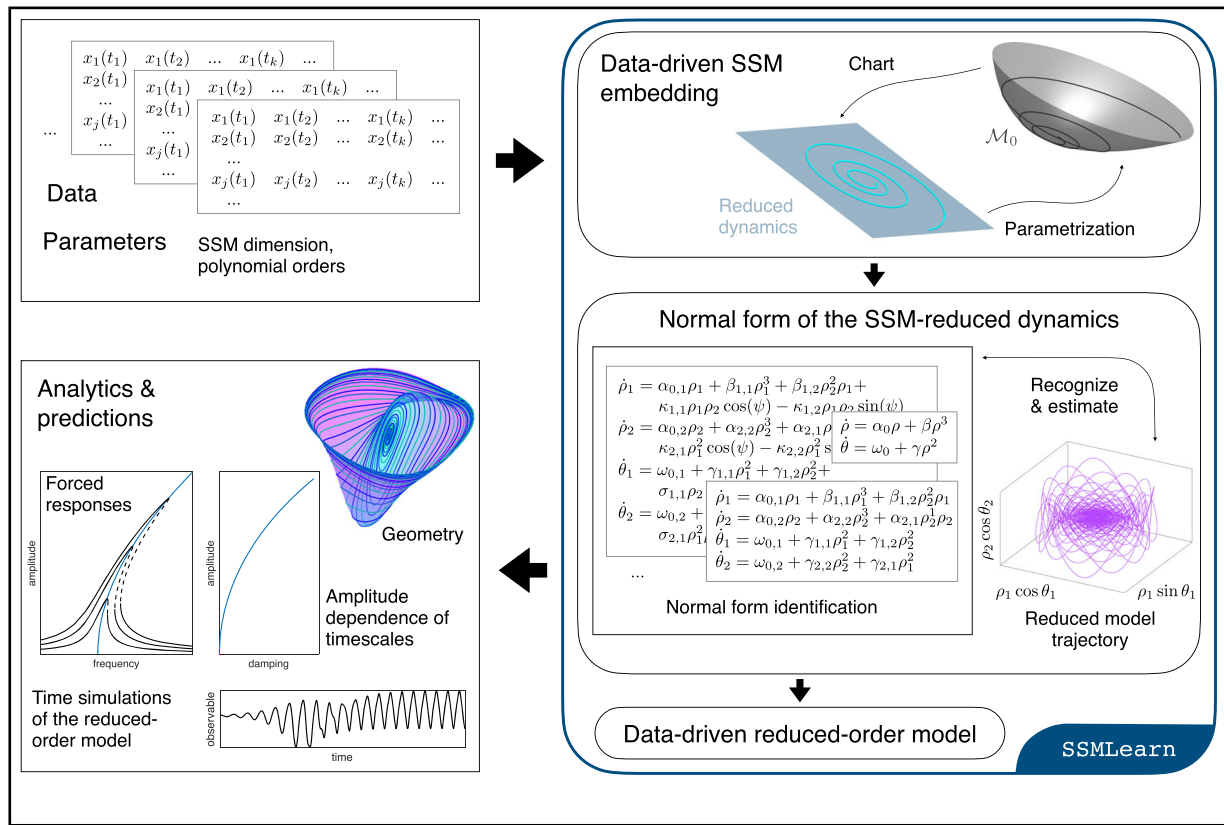
The nearby coexisting stationary states in Fig. 2 happen to be contained in the SSM. In specific examples, however, these states may also be off the SSM, contained instead in one of the infinitely many additional nonlinear continuations,  $\tilde{W}(E, \Omega t; \epsilon)$ , of the spectral subspace  $E$ . The Taylor expansion of the dynamics on  $\tilde{W}(E, \Omega t; \epsilon)$  and  $W(E, \Omega t; \epsilon)$  are, however, identical up to order  $\Sigma(E)$ . Therefore, the reduced models we will compute on the SSM  $W(E, \Omega t; \epsilon)$  also correctly capture the nearby stationary states on  $\tilde{W}(E, \Omega t; \epsilon)$ , as long as the polynomial order of the model stays below  $\Sigma(E)$ . In large physical systems, this represents no limitation, given that  $\Sigma(E) \gg 1$ .

**Embedding SSMs via generic observables.** If at least some of the real parts of the eigenvalues in (2) are negative, then longer-term trajectory data for system (1) will be close to an attracting SSM, as illustrated in panel (b) of Fig. 2. This is certainly the case for data from experiments that are run until a nontrivial, attracting steady state emerges, see, e.g., in panel (e) of Fig. 1. Measurements of trajectories in the full phase space, however, are seldom available from such experiments. Hence, if data about system (1) is only available from observables, the construction of SSMs and their reduced dynamics has to be carried out in the space of those observables.

An extended version of Whitney’s embedding theorem guarantees that almost all (in the sense of prevalence) smooth observable vectors  $\mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), \dots, y_p(\mathbf{x})) \in \mathbb{R}^p$  provide an embedding of a compact subset  $C \subset W(E, \Omega t; \epsilon)$  of a  $d$ -dimensional SSM,  $W(E, \Omega t; \epsilon)$ , into the observable space  $\mathbb{R}^p$  for high enough  $p$ . Specifically, if we have  $p > 2(d + \ell)$  simultaneous and independent continuous measurements,  $\mathbf{y}(\mathbf{x})$ , of the  $p$  observables, then almost all maps  $\mathbf{y} : C \rightarrow \mathbb{R}^p$  are embeddings of  $C$ <sup>37</sup>, and hence the top right plot of Fig. 3 is applicable with probability one.

In practice, we may not have access to  $p > 2(d + \ell)$  independent observables and hence cannot invoke Whitney’s theorem. In that case, we invoke the Takens delay embedding theorem<sup>38</sup>, which covers observable vectors built from  $p$  uniformly sampled, consecutive measured instances of a single observable. More precisely, if  $s(t)$  is a generic scalar quantity measured at times  $\Delta t$  apart, then the observable vector for delay-embedding is formed as  $\mathbf{y}(t) = (s(t), s(t + \Delta t), \dots, s(t + (p - 1)\Delta t)) \in \mathbb{R}^p$ . We discuss the embedding,  $\mathcal{M}_0 \subset \mathbb{R}^p$ , of an autonomous SSM,  $W(E, \Omega t_0; 0)$ , in the observable space  $\mathbb{R}^p$  in more detail in the Methods section “Embedding the SSM in the observable space”.

**Data-driven extended normal forms on SSMs.** Once the embedded SSM,  $\mathcal{M}_0$ , is identified in the observable space, we seek to learn the reduced dynamics on  $\mathcal{M}_0$ . An emerging requirement for learning nonlinear models from data has been model sparsity<sup>4</sup>, without which the learning process would be highly sensitive. The dynamics on  $\mathcal{M}_0$ , however, is inherently non-sparse, which suggests that we learn its Poincaré normal form<sup>39</sup>



**Fig. 3 Schematics of SSMLearn.** First, the data-driven, SSM-based model reduction algorithm implemented in SSMLearn diagnoses and approximates the dominant SSM from the input data. Next, it constructs a data-driven reduced-order model as an extended normal form on the SSM. Finally, the algorithm uses this model to predict individual unforced trajectories and the response of the system under additional forcing.

instead. This classic normal form is the simplest polynomial form to which the dynamics can be brought via successive, near-identity polynomial transformations of increasing order.

Near the origin on a slow SSM, however, this simplest polynomial form is just the restriction of the linear part of system (1) to  $\mathcal{M}_0$ , as long as infinitely many nonresonance conditions are satisfied for the operator  $\mathbf{A}^{40}$ . The Poincaré normal form on  $\mathcal{M}_0$  would, therefore, only capture the low-amplitude, linearized part of the slow SSM dynamics.

To construct an SSM-reduced model for non-linearizable dynamics, we use *extended normal forms*. This idea is motivated by normal forms used in the study of bifurcations of equilibria on center manifolds depending on parameters<sup>33,41</sup>. In that setting, the normal form transformation is constructed at the bifurcation point where the system is non-linearizable by definition. The same transformation is then used away from bifurcations, even though the normal form of the system would be linear there. One, therefore, gives up the maximal possible simplicity of the normal form but gains a larger domain on which the normal form transformation is invertible and hence captures truly nonlinear dynamics. In our setting, there is no bifurcation at  $\mathbf{x} = \mathbf{0}$ , but we nevertheless construct our normal form transformation as if the eigenvalues corresponding to the slow subspace  $E$  were purely imaginary. This procedure leaves additional, near-resonant terms in the SSM-reduced normal form, enhancing the domain on which the transformation is invertible and hence the normal form is valid.

We determine the normal form coefficients directly from data via the minimization of a conjugacy error (see the Methods section). This least-square minimization procedure renders simultaneously the best-fitting normal form coefficients and the best fitting normal form transformation. As we will find in a

specific example, this data-driven procedure can yield accurate reduced models even beyond the formal domain of convergence of equation-driven normal forms.

The simplest extended normal form on a slow SSM of an oscillatory system arises when the underlying spectral subspace  $E$  corresponds to a pair of complex conjugate eigenvalues. Writing in polar coordinates and truncating at cubic order,<sup>42</sup> finds this normal form on the corresponding two-dimensional, autonomous SSM,  $\mathcal{M}_0$ , to be

$$\begin{aligned} \dot{\rho} &= \alpha_0 \rho + \beta \rho^3, \\ \dot{\theta} &= \omega_0 + \gamma \rho^2. \end{aligned} \tag{5}$$

This equation is also known as the Stuart–Landau equation arising in the unfolding of a Hopf bifurcation<sup>43–45</sup>.

The dynamics of (5) is characteristically non-linearizable when  $\alpha_0 \beta < 0$ , given that a limit cycle coexists with the  $\rho = 0$  fixed point in that case. Further coexisting steady states will arise when forcing is added to the system, as we discuss in the next section. We note that the cubic normal form on two-dimensional SSMs has also been approximated from data in<sup>46</sup>. That non-sparse procedure fits the full observer dynamics to a low-dimensional, discrete polynomial dynamical system, then performs an analytic SSM reduction and a classic normal form transformation on the SSM.

For higher accuracy, the extended normal form on an oscillatory SSM of dimension  $2m$  is of the form

$$\begin{aligned} \dot{\rho}_j &= \alpha_j(\boldsymbol{\rho}, \boldsymbol{\theta}) \rho_j, \\ \dot{\theta}_j &= \omega_j(\boldsymbol{\rho}, \boldsymbol{\theta}), \end{aligned} \quad j = 1, 2, \dots, m, \quad \boldsymbol{\rho} \in \mathbb{R}_+^m, \quad \boldsymbol{\theta} \in \mathbb{T}^m. \tag{6}$$

If the linearized frequencies are nonresonant, then the functions  $\alpha_j$  and  $\omega_j$  only depend on  $\boldsymbol{\rho}$ <sup>42</sup>. Our numerical procedure

determines these functions up to the necessary order that ensures a required accuracy for the reduced-order model on the SSM. This is illustrated schematically for a four-dimensional slow SSM ( $m = 2$ ) in the bottom right plot of Fig. 3.

**Predicting forced dynamics from unforced data.** With the normalized reduced dynamics (6) on the embedded SSM,  $\mathcal{M}_0$ , at hand, we can also make predictions for the dynamics of the embedded quasiperiodic SSM,  $\mathcal{M}_\epsilon(\Omega t)$ , of the full system (1). This forced SSM is guaranteed to be an  $\mathcal{O}(\epsilon)C^r$ -close perturbation of  $\mathcal{M}_0$  for moderate external forcing amplitudes. A strict proof of this fact is available for small enough  $\epsilon > 0$ <sup>29</sup>, but as our examples will illustrate, the smooth persistence of the SSM,  $\mathcal{M}_\epsilon(\Omega t)$ , generally holds for all moderate  $\epsilon$  values in practice. Such moderate forcing is highly relevant in a number of technological settings, including system identification in structural dynamics and fluid-structure interactions, where the forcing must be moderate to preserve the integrity of the structure.

We discuss the general extended normal form on  $\mathcal{M}_\epsilon(\Omega t)$  in the Methods section “SSM dynamics via extended normal forms”. In the simplest and most frequent special case, the external forcing is periodic ( $\ell = 1$ ) and  $\mathcal{M}_\epsilon(\Omega t)$  is the embedding of the slowest, two-dimensional SSM corresponding to a pair of complex conjugate eigenvalues. Using the modal forcing amplitude  $f_{1,1}$  and modal phase shift  $\phi_{1,1}$  in the general normal form (25)<sup>47</sup>, introduces the new phase coordinate  $\psi = \theta - \Omega t - \phi_{1,1}$  and lets  $f = f_{1,1}$ ,  $\alpha = \alpha_1$ ,  $\omega = \omega_1$  to obtain the planar, autonomous, extended normal form on  $\mathcal{M}_\epsilon(\Omega t)$  as

$$\begin{aligned}\dot{\rho} &= \alpha(\rho)\rho + f \sin \psi, \\ \dot{\psi} &= \omega(\rho) - \Omega + \frac{f}{\rho} \cos \psi\end{aligned}\quad (7)$$

at leading order in  $\epsilon$ . All stable and unstable periodic responses on the SSM are fixed points of system (7), with their amplitudes  $\rho_0$  and phases  $\psi_0$  satisfying the equations

$$\Omega = \omega(\rho_0) \pm \sqrt{\frac{f^2}{\rho_0^2} - \alpha^2(\rho_0)}, \quad \psi_0 = \tan^{-1} \left[ \frac{\alpha(\rho_0)}{\omega(\rho_0) - \Omega} \right]. \quad (8)$$

The first analytic formula in (8) predicts the *forced response curve* (FRC) of system (1), i.e., the relationship between response amplitude, forcing amplitude and forcing frequency, from the terms  $\alpha(\rho)$  and  $\omega(\rho)$  of the extended normal form of the autonomous SSM,  $\mathcal{M}_0$ . These terms are constructed from trajectories of the unforced system, thus eq. (8) predicts the behavior of a nonlinearizable dynamical system under forcing based solely on unforced training data. The stability of the predicted periodic response follows from a simple linear analysis at the corresponding fixed point of the ODE (7). The first formula in (8) also contains another frequently used notion of nonlinear vibration analysis, the *dissipative backbone curve*  $\omega(\rho)$ , which describes the instantaneous amplitude-frequency relation along freely decaying vibrations within the SSM.

As we will also show in examples, our unforced model-based predictions for forced periodic response (see the Methods section “Prediction of forced response from unforced training data”) are confirmed by numerical simulation or dedicated laboratory experiments on forced systems.

**Examples.** We now illustrate data-driven, SSM-based modeling and prediction on several numerical and experimental data sets describing non-linearizable physical systems. Further applications are described in<sup>48</sup>. Both the numerical and the experimental data sets were initialized without knowledge of the exact SSM. All our computations have been carried out by the publicly available

MATLAB<sup>®</sup> package, *SSMLearn*, whose repository also contains further examples not discussed here. The main algorithm behind *SSMLearn* is illustrated in Fig. 3, with more detail given in the Methods section “Summary of the algorithm”.

To quantify the errors of an SSM-based reduced model, we use the *normalized mean-trajectory-error* (NMTE). For  $P$  observations of the observable vector  $\mathbf{y}_j$  and their model-based reconstructions,  $\hat{\mathbf{y}}_j$ , this modeling error is defined as

$$\text{NMTE} = \frac{1}{\|\underline{\mathbf{y}}\|} \frac{1}{P} \sum_{j=1}^P \|\mathbf{y}_j - \hat{\mathbf{y}}_j\|. \quad (9)$$

Here  $\underline{\mathbf{y}}$  is a relevant normalization vector, such as the data point with the largest norm. When validating the reduced dynamics for a given testing trajectory, we run the reduced model from the same initial condition for the comparison. Increasing the order of the truncated normal form polynomials in eq. (6) generally reduces the NMTE error to any required level but excessively small errors can lead to overfitting. In our examples, we will be allowing model errors in the order of 1%–4% to avoid overfitting.

As a first example, we consider a finite-element discretization of a von Kármán beam with clamped-clamped boundary conditions<sup>49</sup>, shown in panel (a) of Fig. 4. In contrast to the classic Euler-Bernoulli beam, the von Kármán model captures moderate deformations by including a nonlinear, quadratic term in the kinematics. We first construct a 33 degree-of-freedom, damped, unforced finite element model (i.e.,  $n = 66$  and  $\epsilon = 0$  in eq. (1)) for an aluminum beam of length 1 [m], width 5 [cm], thickness 2 [cm] and material damping modulus  $10^6$  [] (see the Supplementary Information for more detail).

Our objective is to learn from numerically generated trajectory data the reduced dynamics on the slowest, two-dimensional SSM,  $W(E_1)$ , of the system, defined over the slowest two-dimensional ( $d = 2$ ) eigenspace  $E_1$  of the linear part. To do so, we generate two trajectories starting from initial beam deflections caused by static loading of 12 [kN] and 14 [kN] at the midpoint, as shown in panel (a) of Fig. 4. The latter trajectory, shown in panel (b) of Fig. 4, is used for training, the other for testing. Along the trajectories, we select our single observable  $s(t)$  to be the midpoint displacement of the beam.

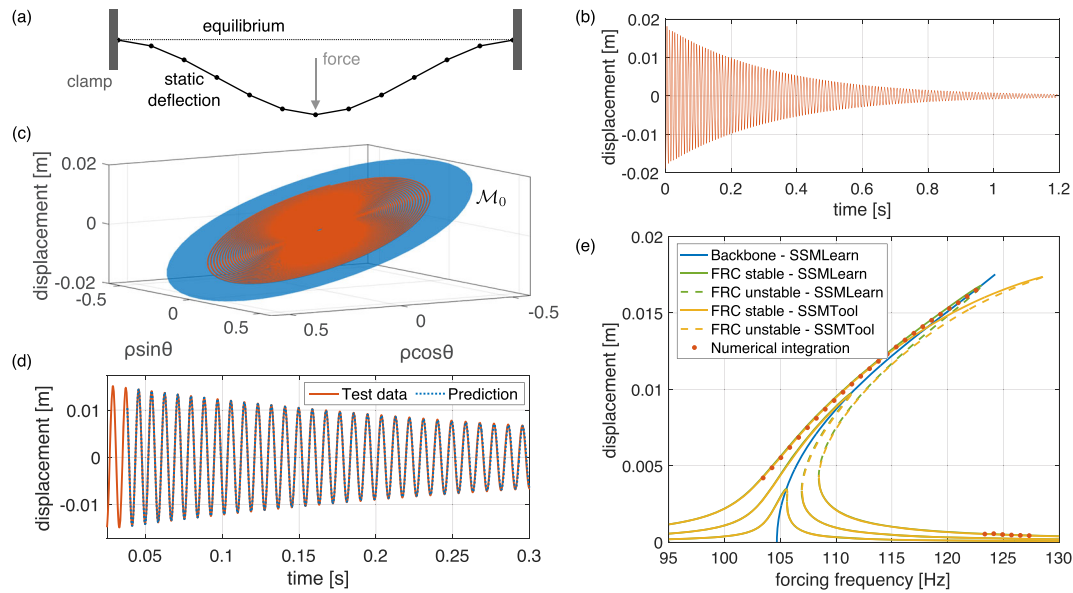
The beam equations are analytic ( $r = a$ ), and hence the SSM,  $W(E_1)$ , admits a convergent Taylor expansion near the origin. The minimal embedding dimension for the two-dimensional,  $W(E_1)$ , as required by Whitney’s theorem, is  $p = 5$ , which is not satisfied by our single scalar observable  $s(t)$ . We therefore employ delay-embedding using  $\mathbf{y}(t) = (s(t), s(t + \Delta t), \dots, s(t + 4\Delta t))$  with  $\Delta t = 0.0955$  [ms]. By Takens’s theorem, this delayed observable embeds the SSM in  $\mathbb{R}^5$  with probability one.

A projection of the embedded SSM,  $\mathcal{M}_0 \in \mathbb{R}^5$ , onto three coordinates is shown in panel (c) of Fig. 4. On  $\mathcal{M}_0$ , *SSMLearn* returns the 7th-order extended normal form

$$\begin{aligned}\dot{\rho} &= \alpha(\rho)\rho, & \alpha(\rho) &= -3.02 - 5.79\rho^2 + 57.5\rho^4 - 191\rho^6, \\ \dot{\theta} &= \omega(\rho), & \omega(\rho) &= 658 + 577\rho^2 - 347\rho^4 - 387\rho^6,\end{aligned}\quad (10)$$

to achieve our preset reconstruction error bar of 3% on the test trajectory (NMTE = 0.027), shown in panel (d) of Fig. 4.

We now use the model (10), trained on a single decaying trajectory, to predict the forced response of the beam for various forcing amplitudes and frequencies in closed form. We will then compare these predictions with analytic forced response computations for the forced SSM,  $\mathcal{M}_\epsilon(\Omega t)$ , obtained from *SSMTool*<sup>36</sup> and with numerical simulations of the damped-forced beam. The periodic forcing is applied at the midpoint node; the Taylor expansion order in *SSMTool* for the analytically computed dynamics on  $\mathcal{M}_\epsilon(\Omega t)$  is set to 7, as in (10). Panel (e) of Fig. 4



**Fig. 4 Construction of a data-driven nonlinear reduced-order model on the slowest SSM of a von Kármán beam.** (a) System setup and the initial condition for the decaying training trajectory shown in (b) in terms of the midpoint displacement. (c) The SSM,  $\mathcal{M}_0$ , in the delay embedding space, shown along with the reconstructed test trajectory in extended normal form coordinates. (d) Zoom of the prediction of the reduced order model for the test trajectory not used in learning  $\mathcal{M}_0$ . (e) Closed-form backbone curve and forced response curve (FRC) predictions ( $\epsilon > 0$ ,  $\ell = 1$ ) by `SSMLearn` are compared with analytic FRC calculations performed by `SSMTool`<sup>36</sup> and with results from numerical integration of the forced-damped beam.

shows the FRCs (green) and the backbone curve (blue) predicted by `SSMLearn` based on formula (8) from the single unforced trajectory in panel (b) of Fig. 4. To obtain the relevant forcing amplitudes  $f$  in the delay-observable space, we have followed the calibration procedure described in the Methods section “Prediction of forced response from unforced training data” for the forcing values  $|ef_1| = 15, 45, 95$  [N] at the single forcing frequency  $\Omega = 103.5$  [Hz]. Recall that coexisting stable (solid lines) and unstable (dashed lines) periodic orbits along the same FRC are hallmarks of non-linearizable dynamics and hence cannot be captured by the model reduction techniques we reviewed in the Introduction for linearizable systems.

The data-based prediction for the FRCs agrees with the analytic FRCs for low forcing amplitudes but departs from it for higher amplitudes. Remarkably, as the numerical simulations (red) confirm, the data-based FRC is the correct one. The discrepancy between the two FRCs for large amplitudes only starts decreasing under substantially higher-order Taylor series approximations used in `SSMTool` (see the Supplementary Information). This suggests the use of the data-based approach for this class of problems even if the exact equations of motion are available.

As a second example, we consider the classic problem of vortex shedding behind a cylinder<sup>8</sup>. Our input data for SSM-based reduced modeling are the velocity and pressure fields over a planar, open fluid domain with a hole representing the cylinder section, as shown in panel (a) of Fig. 5. The boundary conditions are no-slip on the circular inner boundary, standard outflow on the outer boundary at the right side, and fixed horizontal velocity on the three remaining sides<sup>50</sup>. The Reynolds number for this problem is the ratio between the cylinder diameter times the inflow velocity and the kinematic viscosity of the fluid.

Available studies<sup>8,50,51</sup> report that, at low Reynolds number, the two-dimensional unstable manifold,  $W^u(SS)$ , of the wake-type steady solution,  $SS$ , in panel (b) of Fig. 5 connects  $SS$  to the limit cycle shown in panel (c) of Fig. 5. Here we evaluate the performance of `SSMLearn` on learning this unstable manifold as an SSM, along with its reduced dynamics, from trajectory data at Reynolds number equal to 70. For this SSM, we again have  $d = 2$

and  $r = a$ , as in our previous example. There is no external forcing in this problem, and hence we have  $\epsilon = 0$  in eq. (1). In contrast to prior studies that often consider a limited number of observables<sup>8,51,52</sup>, here we select the full phase space of the discretized Navier-Stokes simulation to be the observable space for illustration, which yields  $n = p = 76,876$  in eq. (1). We generate nine trajectories numerically, eight of which will be used for training and one for testing the SSM-based model.

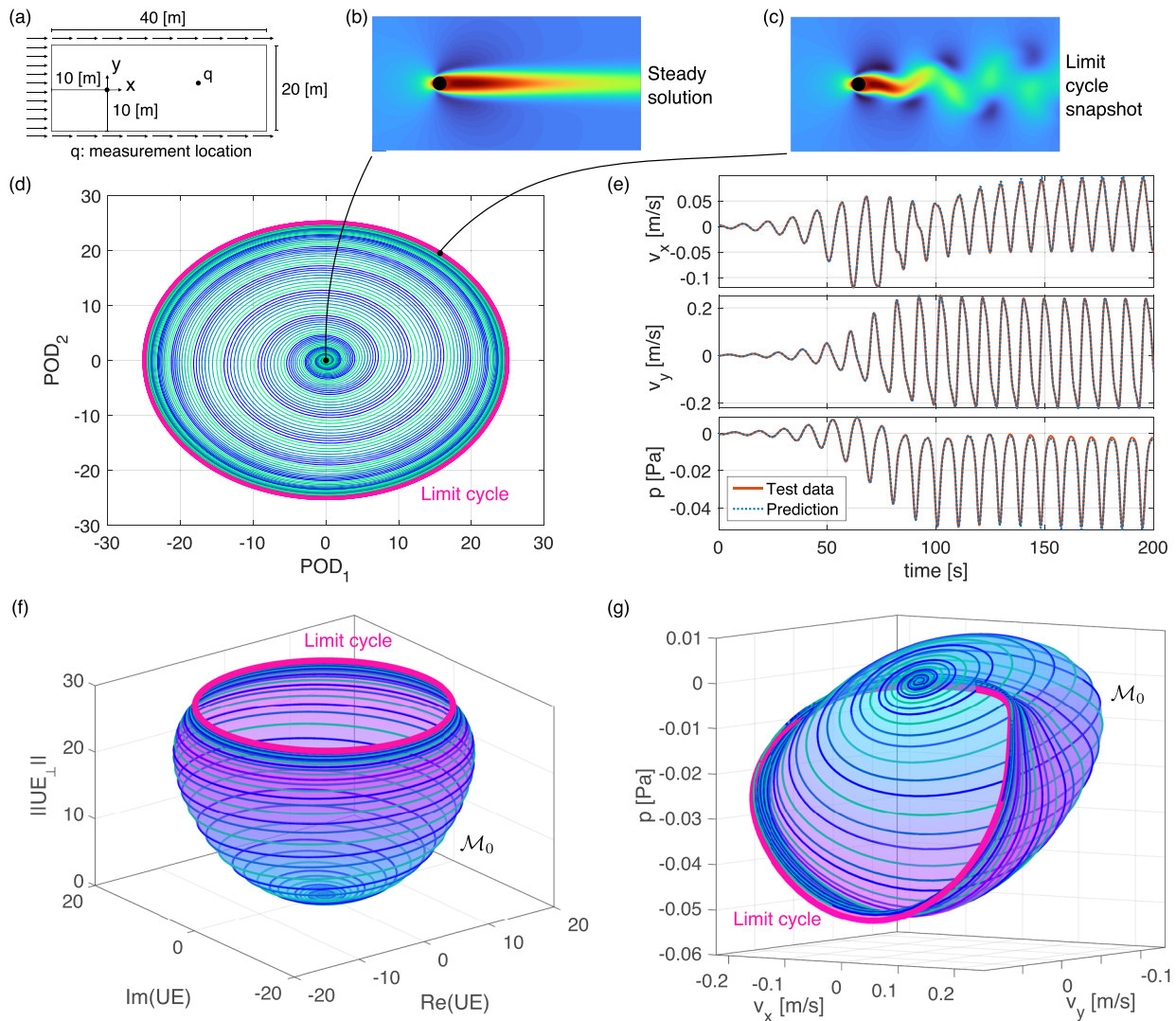
The nine initial conditions of our input trajectory data are small perturbations from the wake-type steady solution along its unstable directions, equally spaced on a small amplitude circle on this unstable plane. All nine trajectories quickly converge to the unstable manifold and then to the limit cycle representing periodic vortex shedding.

We choose to parametrize the SSM,  $\mathcal{M}_0 = W^u(SS)$ , with two leading POD modes of the limit cycle, which have been used in earlier studies for this problem. The training trajectories projected onto these two POD modes are shown in panel (d) of Fig. 5. To limit the modeling error (9) to less than  $NMTE = 1\%$ , `SSMLearn` requires a polynomial order of 18 in the SSM computations. For this order, our approach can accommodate the strong mode deformation observed for this problem<sup>51</sup>, manifested by a fold of the SSM over the unstable eigenspace in panel (f) of Fig. 5. Panel (g) of Fig. 5 shows the strongly nonlinear geometry of  $\mathcal{M}_0$  projected to the observable subspace formed by the velocities and the pressure of a probe point in the wake.

To capture the SSM-reduced dynamics with acceptable accuracy, we need to compute the extended normal form up to order 11, obtaining

$$\begin{aligned} \dot{\rho} &= \alpha(\rho)\rho = 0.0584\rho - 0.479\rho^3 + 1.27\rho^5 + 6.80\rho^7 - 58.9\rho^9 + 108\rho^{11}, \\ \dot{\theta} &= \omega(\rho) = 0.553 + 0.441\rho^2 - 3.38\rho^4 + 55.5\rho^6 - 321\rho^8 + 626\rho^{10}. \end{aligned} \tag{11}$$

To describe a transition qualitatively from a fixed point to a limit cycle, the reduced-order dynamical model should be at least of cubic order<sup>51</sup>. Capturing the qualitative behavior (i.e., the unstable fixed point and the stable limit cycle), however, does not imply a low  $NMTE$  error for the model. Indeed, the data-



**Fig. 5** Data-driven nonlinear SSM-reduced model on the unstable manifold of the steady solution of the flow past a cylinder. **a** Problem setup. **b, c** Snapshots of the steady solution and the time-periodic vortex-shedding solution (limit cycle, in magenta). **d** Trajectories projected on the 2-dim. subspace spanned by the two-leading POD modes of the limit cycle. **e** Model-based reconstruction of the test trajectory (not used in learning the SSM) in terms of velocities and pressures measured at a location  $q$  shown in plot **a**. **f** The SSM formed by the unstable manifold of the origin, along with some reduced trajectories, plotted over the unstable eigenspace  $UE \equiv E_i$ ;  $\|UE_{\perp}\|$  denotes the normed projection onto the orthogonal complement  $UE_{\perp}$ . **g** Same but projected over velocity and pressure coordinates.

driven cubic normal form for this example gives a reconstruction error of  $NMTE = 117\%$  normalized over the limit cycle amplitude, mainly arising from an out-of-phase convergence to the limit cycle along the testing trajectory. In contrast, the  $\mathcal{O}(11)$  normal form in eq. (11) reduced this error drastically to  $NMTE = 3.86\%$  on the testing trajectory, as shown in panel (e) of Fig. 5.

We show in Section 1.2.3 of the Supplementary Information that for comparable accuracy, the Sparse Identification of Nonlinear Dynamics (SINDy) approach of<sup>4</sup> returns non-sparse nonlinear models for this example. Similarly, while the DMD<sup>13</sup> can achieve highly accurate curve-fitting on the available training trajectories with a high-dimensional linear model, that model only captures linearizable dynamics near the origin. As a consequence, its trajectories grow without bound over longer integration times and hence fail to capture the limit cycle.

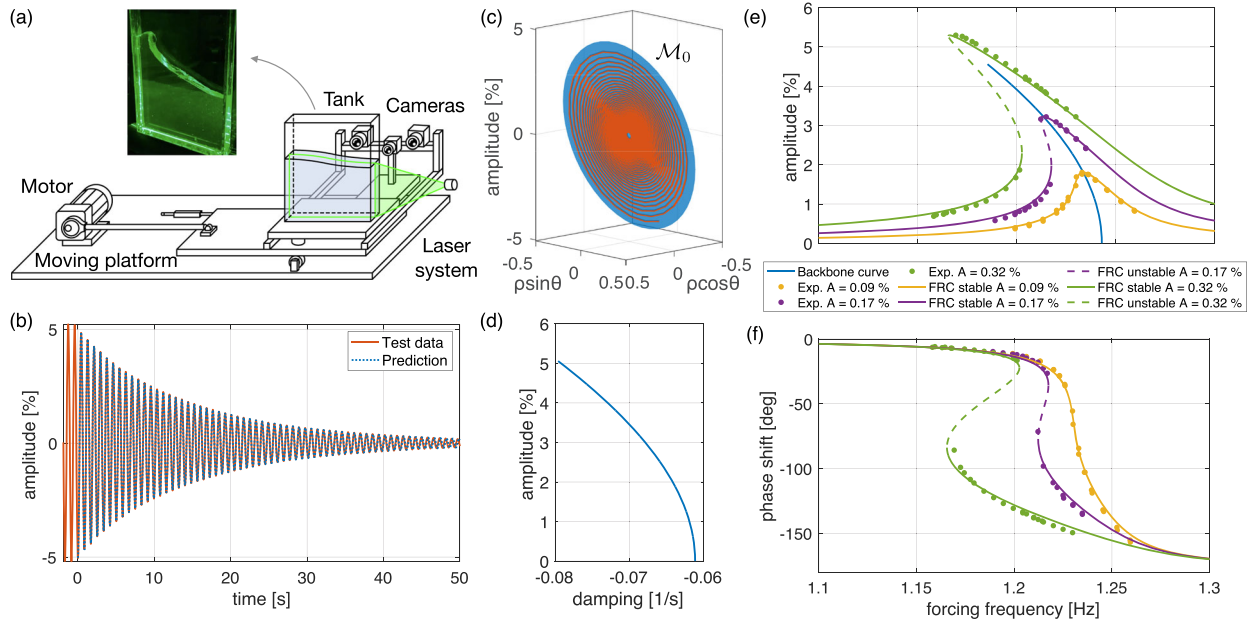
As a third example, we consider fluid oscillations in a tank, which exhibit highly nonlinear characteristics<sup>53</sup>. To describe such

non-linearizable softening effects observed in the sloshing motion of surface waves, Duffing-type models have been proposed<sup>54</sup>. While amplitude variations observed in forced experiments can be fitted to forced softening Duffing equations, nonlinear damping remains a challenge to identify<sup>55</sup>.

The experiments we use to construct an SSM-reduced nonlinear model for sloshing were performed in a rectangular tank of width 500 [mm] and depth 50 [mm], partially filled with water up to a height of 400 [mm], as shown in panel (a) of Fig. 6. The tank was mounted on a platform excited harmonically by a motor. The surface level was detected via image processing from a monochrome camera. As an observable  $s(t)$  we used the horizontal position of the computed center of mass of the water at each time instant, normalized by the tank width. This physically meaningful scalar is robust with respect to image evaluation errors<sup>55</sup>.

We identify the unforced nonlinear behavior of the system from data obtained in resonance decay experiments<sup>56</sup>. In those





**Fig. 6** Data-driven nonlinear reduced-order model on the slowest SSM of fluid sloshing in a tank. **a** Setup for the sloshing experiment<sup>55</sup>. **b** Decaying model-testing trajectory and its reconstruction from an unforced, SSM-based model **c** The geometry of the embedded SSM **d** Nonlinear damping  $\alpha(\rho)$  from the SSM-reduced dynamics **e, f** Closed form, SSM-based predictions of the FRCs and the response phases  $\psi_0$  for three different forcing amplitudes (solid lines), with their experimental confirmation superimposed (dots).

experiments (as in Fig. panel (a) of 6, but with a shaker instead of a motor), once a periodic steady state is reached under periodic horizontal shaking of the tank, the shaker is turned off and the decaying sloshing is recorded. We show such a decaying observable trajectory (orange line) in panel (b) of Fig. 6, with the shaker switched off slightly before zero time. This damped oscillation is close, by construction, to the two-dimensional, slowest SSM of the system. We use three such decaying observer trajectories (two for training and one for model testing) for the construction of a two-dimensional ( $d = 2$ ), autonomous, SSM-based reduced-order model for  $s(t)$ . For delay embedding dimension, we again pick  $p = 5$ , the minimal value guaranteed to be generically correct for embedding the SSM by Takens’s theorem. The delay used in sampling  $s(t)$  is  $\Delta t = 0.033$  [s]. For this input and for a maximal reconstruction error of 2%, SSMLearn identifies a nearly flat SSM in the delayed observable space—see panel (c) of Fig. 6—with a cubic extended normal form

$$\dot{\rho} = -0.063179\rho - 0.041214\rho^3, \quad \dot{\theta} = 7.8144 - 1.5506\rho^2. \tag{12}$$

This lowest-order, Stuart–Landau-type normal form, cf. (5), already constitutes an accurate reduced-order model with  $NMTE = 1.88\%$  on the testing data set, see panel (b) of Fig. 6. The amplitude-dependent nonlinear damping,  $\alpha(\rho)$ , provided by this model is plotted in panel (d) of Fig. 6 with respect to the physical amplitude.

In another set of experiments with the setup of panel (a) Fig. 6, steady states of periodically forced sloshing were measured in sweeps over a range of forcing frequencies under three different shaker amplitudes. As in the previous beam example, we identify the corresponding forcing amplitude,  $f$ , in (7) at the maximal amplitude response of each frequency sweep. Shown in panels (e, f) of Fig. 6, the closed-form predictions for FRCs from eq. (8) (solid lines) match closely the experimental FRCs (dots). Given the strong nonlinearity of the FRC, any prediction of this curve

from a DMD-based model is bound to be vastly inaccurate, as we indeed show in Section 1.3 of the Supplementary information.

The phase  $\psi_0$  of the forced response relative to the forcing has been found difficult to fit to forced Duffing-type models<sup>55</sup>, but the present modeling methodology also predicts this phase accurately using the second expression in (8). The blue curve in panel (e) of Fig. 6 shows the backbone curve of decaying vibrations, which terminates at the highest amplitude occurring in the training data set. This plot therefore shows that the closed-form FRC predictions obtained from the SSM-based reduced model are also effective for response amplitudes outside the training range of the reduced model.

### Discussion

We have described a data-driven model reduction procedure for non-linearizable dynamical systems with coexisting isolated stationary states. Our approach is based on the recent theory of spectral submanifolds (SSMs), which are the smoothest nonlinear continuations of spectral subspaces of the linearized dynamics. Slow SSMs form a nested hierarchy of attractors and hence the dynamics on them provide a hierarchy of reduced-order models with which generic trajectories synchronize exponentially fast. These SSMs and their reduced models smoothly persist under moderate external forcing, yielding low-dimensional, mathematically exact reduced-order models for forced versions of the same dynamical system. The normal hyperbolicity of SSMs also ensures their robustness under small noise.

All these results have been implemented in the open-source MATLAB® package, SSMLearn, which we have illustrated on data sets arising from forced nonlinear beam oscillations, vortex shedding behind a cylinder and water sloshing in a vibrating tank. For all three examples, we have found that two-dimensional data-driven extended normal forms on the slowest SSMs provide sparse yet accurate models of non-linearizable dynamics in the space of the chosen observables. Beyond matching training and testing data, SSM-reduced models prove their intrinsic,

qualitative meaning by predicting non-linearizable, forced steady states purely from decaying, unforced data.

In this brief report, examples of higher-dimensional SSMs and multi-harmonic forcing have not been considered, even though SSMLearn is equipped to handle them. Higher-dimensional SSMs are required in the presence of internal resonances or in non-resonant problems in which initial transients also need to be captured more accurately. A limitation of our approach for non-autonomous systems is the assumption of quasiperiodic external forcing. Note, however, that even specific realizations of stochastic forcing signals can be approximated arbitrarily closely with quasiperiodic functions over any finite time interval of interest. A further limitation in our work is the assumption of smooth system dynamics. For data from non-smooth systems, SSMLearn will nevertheless return an equivalent smooth reduced-order model whose accuracy is a priori known from the available mean-squared error of the SSM fitting and conjugacy error of the normal form construction. We are addressing these challenges in ongoing work to be reported elsewhere. Further applications of SSMLearn to physical problems including higher-dimensional coexisting steady states (see, e.g.,<sup>57</sup>) are also underway.

**Methods**

**Existence of SSMs.** In the context of rigid body dynamics, invariant manifolds providing generalizations of invariant spectral subspaces to nonlinear systems were first envisioned and formally constructed as nonlinear normal modes by<sup>58</sup> (see<sup>59</sup> for a recent review of related work). Later studies, however, pointed out the nonuniqueness of nonlinear normal modes in specific examples<sup>(60,61)</sup>.

In the mathematics literature,<sup>62</sup> obtained general results on the existence, smoothness and degree of uniqueness of such invariant manifolds for mappings on Banach spaces. These results use a special parameterization method to construct the manifolds even in evolutionary partial differential equations that admit a well-posed flow map in both time directions (see<sup>63</sup> for a mechanics application). The results have been extended to a form applicable to dynamical systems with quasiperiodic time dependence<sup>64</sup>. An extensive account of the numerical implementation of the parametrization method with a focus on computing invariant tori and their whiskers in Hamiltonian systems is also available<sup>65,29</sup>. Discussed the existence of the SSM,  $W(E, \Omega; \epsilon)$ , depending on its absolute spectral quotient,

$$\Sigma(E) = \text{Int} \left[ \frac{\max_{\lambda \in \text{Spect}(\mathbf{A}|_S)} |\text{Re}\lambda|}{\min_{\lambda_e \in \text{Spect}(\mathbf{A}|_E)} |\text{Re}\lambda_e|} \right], \tag{13}$$

where  $\text{Spect}(\mathbf{A}|_S)$  is the stable (unstable) spectrum of  $\mathbf{A}$  if the SSM is stable (unstable). For a stable SSM,  $\Sigma(E)$  is the integer part of the quotient of the minimal real part in the spectrum of  $\mathbf{A}$  and the maximal real part of the spectrum of  $\mathbf{A}$  restricted to  $E$ .

Based on  $\Sigma(E)$ , we call a  $d$ -dimensional spectral subspace  $E$  non-resonant if for any set  $(m_1, \dots, m_d)$  of nonnegative integers satisfying  $2 \leq \sum_{j=1}^d m_j \leq \Sigma(E)$ , the eigenvalues,  $\lambda_k$ , of  $\mathbf{A}$  satisfy

$$\sum_{j=1}^d m_j \text{Re}\lambda_j \neq \text{Re}\lambda_k, \quad \lambda_k \in \text{Spect}(\mathbf{A}) - \text{Spect}(\mathbf{A}|_E). \tag{14}$$

This condition only needs to be verified for resonance orders between 2 and  $\Sigma(E)$ <sup>64</sup>. In particular, a 1:1 resonance between  $E_1$  and  $E_2$  is allowed if  $\dim E_1 = \dim E_2 = 1$ , in which case each strongly resonant spectral subspace gives rise to a unique nearby spectral submanifold.

If  $E$  violates the nonresonance condition (14), then  $E$  can be enlarged to a higher-dimensional spectral subspace until the nonresonance relationship (14) is satisfied. In the absence of external forcing ( $\epsilon = 0$ ), the nonresonance condition (14) can also be relaxed with the help of the relative spectral quotient,

$$\sigma(E) = \text{Int} \left[ \frac{\max_{\lambda \in \text{Spect}(\mathbf{A}|_S) - \text{Spect}(\mathbf{A}|_E)} |\text{Re}\lambda|}{\min_{\lambda_e \in \text{Spect}(\mathbf{A}|_E)} |\text{Re}\lambda_e|} \right], \tag{15}$$

to the form

$$\sum_{j=1}^d m_j \lambda_j \neq \lambda_k, \quad \lambda_k \in \text{Spect}(\mathbf{A}) - \text{Spect}(\mathbf{A}|_E), \quad 2 \leq \sum_{j=1}^d m_j \leq \sigma(E). \tag{16}$$

This is indeed a relaxation because condition (16) is only violated if both the real and the imaginary parts of eigenvalues involved are in the exact same resonance with each other. In contrast, (14) is already violated when the real parts are in resonance with each other.

If  $\text{Re}\lambda_k < 0$  in eq. (2) and all  $E^k$  subspaces are nonresonant, then the nested set of slow spectral submanifolds,

$$W(E^1, \Omega; \epsilon) \subset W(E^2, \Omega; \epsilon) \subset W(E^3, \Omega; \epsilon) \subset \dots,$$

gives a hierarchy of local attractors. All solutions in a vicinity of  $\mathbf{x} = \mathbf{0}$  approach the reduced dynamics on one of these attractors exponentially fast, as sketched in panel (b) of Fig. 2 for the  $\epsilon = 0$  limit. As we will see, non-linearizable dynamics tend to emerge on  $W(E^k, \Omega; \epsilon)$  due to near-resonance between the linearized frequencies within  $E^k$  and the forcing frequencies  $\Omega$ . The specific location of nontrivial steady states in  $W(E^k, \Omega; \epsilon)$  is then determined by a balance between the nonlinearities, damping and forcing.

A resonant  $E^k$  subspace can be enlarged by adding the next  $k'$  modal subspaces to it until  $E^{k+k'}$  in the hierarchy (4) becomes non-resonant and hence admits an SSM,  $W(E^{k+k'}, \Omega; \epsilon)$ . This technical enlargement is also in agreement with the physical expectation that all interacting modes have to be included in an accurate reduced-order model. Finally, we note that SSMs are robust features of dynamical systems: they inherit smooth dependence of the vector field in (1) on parameters<sup>29</sup>.

For discrete-time dynamical systems of the form

$$\mathbf{x}_{k+1} = \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{f}}_0(\mathbf{x}_k) + \epsilon \bar{\mathbf{f}}_1(\mathbf{x}_k, \phi_k; \epsilon), \quad \phi_{k+1} = \phi_k + \bar{\Omega}, \tag{17}$$

the above results on SSMs apply based on the eigenvalues  $\mu_k$  of  $\bar{\mathbf{A}}$ . One simply needs to replace  $\lambda_k$  with  $\log \mu_k$  and  $\text{Re}\lambda_k$  with  $\log |\mu_k|$  in formulas (13)-(16)<sup>29</sup>.

We close by noting that in a neighborhood of an SSM, an invariant family of surfaces resembling the role of coordinate planes in a linear system exists<sup>66</sup>. This invariant spectral foliation (ISF) can, in principle, be used to generate a nonlinear analogue of linear modal superposition in a vicinity of a fixed point. Constructing the ISF from data has shown both initial promise and challenges to be addressed.

**Embedding the SSM in the observable space.** Originally conceived for autonomous systems, the Takens delay embedding theorem<sup>38</sup> has been strengthened and generalized to externally forced dynamics<sup>32</sup>. By these results, the embedding for a  $d$ -dimensional compact SSM subset,  $\mathcal{C} \subset W(E, \Omega; \epsilon)$ , in the delay observable space as  $\mathcal{M}(\Omega t)$  is guaranteed for almost all choices of the observable  $s(t)$  if  $p > 2(d+1)$  and some generic assumptions regarding periodic motions on  $\mathcal{M}(\Omega t)$  are satisfied<sup>37</sup>.

Of highest importance in technological applications is the case of time-periodic forcing ( $\ell = 1$ ), with frequency  $\Omega = \Omega \in \mathbb{R}$  and period  $T = 2\pi/\Omega$ . In this case, the Whitney and Takens embedding theorems can be applied to the associated period- $T$  sampling map (or Poincaré map)  $\mathbf{P}_{t_0} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  of the system based at time  $t_0$ . This map is autonomous and has a time-independent SSM that coincides with the  $d$ -dimensional SSM,  $\mathcal{M}(\Omega t_0)$ , of the full system (1). In this case, by direct application of the embedding theorems to the discrete dynamical system generated by  $\mathbf{P}_{t_0}$ , the typically sufficient embedding dimension estimate is improved to  $p > 2d$  for Whitney's and Takens's theorem.

Technically speaking, the available data will never be exactly on an SSM, as these embedding theorems assume. By the smoothness of the embeddings, however, points close enough to the SSM in the phase space will be close to  $\mathcal{M}(\Omega t)$  in the observable space under the embeddings. Moreover, as slow SSMs attract nearby trajectories exponentially, the distance of observable data from the embedded slow SSM will shrink exponentially fast. Therefore, even under uncorrelated noise in the measurements, mean-squared estimators are suitable for learning slow SSMs from data in the observable space, as we illustrate in the Supplementary Information.

After a possible coordinate shift, the trivial fixed point of the autonomous limit of system (1) will be mapped into the  $\mathbf{y} = \mathbf{0}$  origin of the observable space. To find an embedded,  $d$ -dimensional SSM,  $\mathcal{M}_0 \in \mathbb{R}^p$ , attached to this origin for  $\epsilon = 0$ , we focus on observable domains in which  $\mathcal{M}_0$  is a graph over its tangent space  $T_0\mathcal{M}_0$  at the origin  $\mathbf{y} = \mathbf{0}$ . Such domains always exist and are generally large enough to capture non-linearizable dynamics in most applications (but see below). Note that  $T_0\mathcal{M}_0$  coincides with the image of the spectral subspace  $E$  in the observable space.

To learn such a graph-style parametrization for  $\mathcal{M}_0$  from data, we define a matrix  $\mathbf{U}_1 \in \mathbb{R}^{p \times d}$  with columns that are orthonormal vectors spanning the yet unknown  $T_0\mathcal{M}_0$ . The reduced coordinates  $\boldsymbol{\eta} \in \mathbb{R}^d$  for a point  $\mathbf{y} \in \mathcal{M}_0$  are then defined as the orthogonal projection  $\boldsymbol{\eta} = \mathbf{U}_1^T \mathbf{y}$ . We seek a Taylor-expansion for  $\mathcal{M}_0$  near the  $\boldsymbol{\eta} = \mathbf{0}$  origin, denoting by  $\boldsymbol{\eta}^{2:M}$  the family of all monomials of  $d$  variables from degree 2 to  $M$ . For example, if  $d = 2$  and  $M = 3$ , then  $\boldsymbol{\eta}^{2:3} = (\eta_1^2, \eta_1\eta_2, \eta_2^2, \eta_1^3, \eta_1^2\eta_2, \eta_1\eta_2^2, \eta_2^3)^T$ . As a graph over  $T_0\mathcal{M}_0$ , the manifold  $\mathcal{M}_0$  is approximated as  $\mathbf{y} = \mathbf{V}_1\boldsymbol{\eta} + \mathbf{V}\boldsymbol{\eta}^{2:M}$ , where the matrices  $\mathbf{V}_1$  and  $\mathbf{V}$  contain coefficients for the  $d$ -variate linear and nonlinear monomials, respectively. Learning  $\mathcal{M}_0$  from a data set of  $P$  observations  $\mathbf{y}_1, \dots, \mathbf{y}_P$  then amounts to finding the  $(\mathbf{U}_1^*, \mathbf{V}_1^*, \mathbf{V}^*)$  matrices that minimize the mean-square reconstruction error along the training data:

$$(\mathbf{U}_1^*, \mathbf{V}_1^*, \mathbf{V}^*) = \arg \min_{\mathbf{U}_1, \mathbf{V}_1, \mathbf{V}} \sum_{j=1}^P \|\mathbf{y}_j - \mathbf{V}_1 \mathbf{U}_1^T \mathbf{y}_j - \mathbf{V}(\mathbf{U}_1^T \mathbf{y}_j)^{2:M}\|^2, \tag{18}$$

$$\mathbf{U}_1^T \mathbf{U}_1 = \mathbf{I}.$$

The simplest solution to this problem is  $\mathbf{U}_1 = \mathbf{V}_1$  with the additional constraint  $\mathbf{V}_1^T \mathbf{V} = \mathbf{0}$ , which represents a basic nonlinear extension of the principal component analysis<sup>67</sup>.

The above graph-style parametrization of the SSM breaks down for larger  $\mathbf{y}$  values if  $\mathcal{M}_0$  develops a fold over  $T_0\mathcal{M}_0$ . That creates an issue for model reduction if a nontrivial steady state on  $\mathcal{M}_0$  falls outside the fold, as the limit cycle does in our vortex shedding example. In that case, alternative parametrization methods for  $\mathcal{M}_0$  can be used to enhance the domain of the SSM-reduced model. These methods include selecting the columns of  $\mathbf{U}_1$  to be the leading POD modes of the nontrivial steady state, or enlarging the embedding space with (further) delayed observations. In these cases, the columns of  $\mathbf{V}_1$  are still orthonormal vectors spanning  $T_0\mathcal{M}_0$ .

In both panels (c) of Figs. 4, 6, the SSM,  $\mathcal{M}_0$ , is nearly flat in the delay-embedding space. This turns out to be a universal property of delay embedding for small delays and low embedding dimensions (see the Supplementary Information).

For  $\epsilon > 0$  small (i.e., for moderate forcing), the autonomous SSM,  $\mathcal{M}_0$ , already captures the bulk nonlinear behavior of system (1). Indeed, for this forcing range, the reduced dynamics on the corresponding SSM can simply be computed as an additive perturbation of the autonomous dynamics on  $\mathcal{M}_0$ <sup>47,68,69</sup> (see section “Predicting forced dynamics from unforced data”).

**SSM dynamics via extended normal forms.** For an autonomous SSM  $\mathcal{M}_0$ , the reduced dynamics is governed by a vector field

$$\dot{\boldsymbol{\eta}} = \mathbf{r}(\boldsymbol{\eta}) \tag{19}$$

with a flow map  $\boldsymbol{\varphi}_r^t(\boldsymbol{\eta})$ . We can generically assume that the Jacobian  $D\mathbf{r}(\mathbf{0})$  is semisimple, i.e.,  $D\mathbf{r}(\mathbf{0})\mathbf{B} = \mathbf{B}\boldsymbol{\Lambda}$ , where  $\boldsymbol{\Lambda} \in \mathbb{C}^{d \times d}$  is a diagonal matrix containing the eigenvalues of  $D\mathbf{r}(\mathbf{0})$ . Classic normal form theory would seek to simplify the reduced dynamics (19) in a vicinity of  $\boldsymbol{\eta} = \mathbf{0}$  via a nonlinear change of coordinates,  $\boldsymbol{\eta} = \mathbf{h}(\mathbf{z})$ , so that the transformed vector field  $\dot{\mathbf{z}} = \mathbf{n}(\mathbf{z})$  with flow map  $\boldsymbol{\varphi}_n^t(\mathbf{z})$  has a diagonal linear part and has as few nonlinear terms in its Taylor expansion as possible. In our present setting, the origin is assumed hyperbolic, in which case the classic normal form is simply  $\dot{\mathbf{z}} = \boldsymbol{\Lambda}\mathbf{z}$  under appropriate non-resonance conditions that generically hold<sup>40</sup>. The corresponding normal form transformation  $\mathbf{h}(\mathbf{z})$ , however, is only valid on a small enough domain in which the dynamics is linearizable.

To capture non-linearizable behavior, we employ extended normal forms motivated by those used to unfold bifurcations<sup>33</sup>. In this approach, we construct normal forms that do not remove those polynomial terms from (19) whose removal would result in small denominators in the Taylor coefficients  $\mathbf{h}(\mathbf{z})$  and hence decrease its domain of convergence. Instead, we seek a normal form for (19) of the form

$$\begin{aligned} \mathbf{n}(\mathbf{z}; \mathbf{N}) &= \boldsymbol{\Lambda}\mathbf{z} + \mathbf{N}\mathbf{z}^{2:N}, \\ \mathbf{h}(\mathbf{z}; \mathbf{H}) &= \mathbf{B}(\mathbf{z} + \mathbf{H}\mathbf{z}^{2:N}), \quad \mathbf{h}^{-1}(\boldsymbol{\eta}; \mathbf{H}) = \mathbf{B}^{-1}\boldsymbol{\eta} + \mathbf{H}_*(\mathbf{B}^{-1}\boldsymbol{\eta})^{2:N}, \end{aligned} \tag{20}$$

where the matrices  $\mathbf{N}$ ,  $\mathbf{H}$  and  $\mathbf{H}_*$  contain the coefficients for the appropriate  $d$ -variate monomials. To identify near-resonances, we let  $\mathbf{S}^{2:N}$  be the matrix of integers whose columns are the powers of the  $d$ -variate monomials from order 2 to  $N$ . We then define a matrix  $\boldsymbol{\Delta}^{2:N}$  containing all relevant integer linear combinations of eigenvalues as follows:

$$(\boldsymbol{\Delta}^{2:N})_{j,k} = (\text{Im}\boldsymbol{\Lambda})_{j,j} - \sum_{s=1}^d (\text{Im}\boldsymbol{\Lambda})_{s,s} (\mathbf{S}^{2:N})_{s,k}. \tag{21}$$

Following the approach used in universal unfolding principles<sup>41</sup>, we collect in a set  $S$  the row and column indices of the entries of  $\boldsymbol{\Delta}^{2:N}$  for which near-resonances occur, i.e., for which the corresponding entry of  $\boldsymbol{\Delta}^{2:N}$  is smaller in norm than a small, preselected threshold. (The default threshold is  $10^{-8}$  in SSMLearn.) The entries of  $\mathbf{H}$  and  $\mathbf{H}_*$  with indices contained in  $S$  are then set to zero but the corresponding monomial terms are retained in  $\mathbf{n}(\mathbf{z}; \mathbf{N})$ . Conversely, coefficients of non-near-resonant entries of  $\mathbf{H}$  and  $\mathbf{H}_*$  are selected in a way so that the corresponding non-near-resonant monomials vanish from the normal form  $\mathbf{n}(\mathbf{z}; \mathbf{N})$ . As a result, the matrix  $\mathbf{N}$  is sparse, containing only the coefficients of essential, near-resonant monomials.

For example, if  $d = 2$ ,  $N = 3$  and the eigenvalues of  $D\mathbf{r}(\mathbf{0})$  form a complex pair  $\lambda = \alpha_0 \pm i\omega_0$  with  $\omega_0 = \mathcal{O}(1)$ , then we have

$$\begin{aligned} \mathbf{S}^{2:N} &= [21032100120123], \quad \boldsymbol{\Delta}^{2:N} \\ &= [-\omega_0\omega_03\omega_0 - 2\omega_02\omega_04\omega_0 - 3\omega_0 - \omega_0\omega_0 - 4\omega_0 - 2\omega_02\omega_0]. \end{aligned} \tag{22}$$

Only two elements of  $\boldsymbol{\Delta}^{2:N}$  are (near-) zero, and hence the reduced dynamics in extended normal form will require learning the following coefficients:

$$\begin{aligned} \mathbf{H}_* &= [h_{20}h_{11}h_{02}h_{30}0h_{12}h_{03}\bar{h}_{02}\bar{h}_{11}\bar{h}_{20}\bar{h}_{03}\bar{h}_{12}0\bar{h}_{30}], \\ \mathbf{N} &= [0000h_{21}0000000\bar{h}_{21}0]. \end{aligned} \tag{23}$$

The corresponding cubic polar form (5) is then obtained from the relations  $\mathbf{z} = (\rho e^{i\theta}, \rho e^{-i\theta})$  and  $h_{21} = \beta + i\gamma$ .

For a data-driven construction of the extended normal form (20), we first obtain an estimate for the Jacobian  $D\mathbf{r}(\mathbf{0})$  from linear regression. This determines the matrix  $\mathbf{B}$  and the types of monomials arising in  $\mathbf{h}^{-1}$  and  $\mathbf{n}$ . Next, we note that the flow map  $\boldsymbol{\varphi}_r^t$  of the SSM-reduced dynamics and the flow map  $\boldsymbol{\varphi}_n^t$  of the extended normal form are connected through the conjugacy relationship  $\boldsymbol{\varphi}_n^t = \mathbf{h}^{-1} \circ \boldsymbol{\varphi}_r^t \circ \mathbf{h}$ . We find the nonzero complex coefficients of  $\mathbf{h}^{-1}$  and  $\mathbf{n}$  by minimizing the error in this exact conjugacy over the available  $P$  data points, represented in the  $\boldsymbol{\eta}$  coordinates. Specifically, we determine the nonzero elements

of  $\mathbf{H}_*$  and  $\mathbf{N}$  as

$$\begin{aligned} (\mathbf{H}_*, \mathbf{N}^*) &= \arg \min_{\mathbf{H}, \mathbf{N}} \sum_{j=1}^p \left\| \frac{d}{dt} \mathbf{h}^{-1}(\boldsymbol{\eta}_j; \mathbf{H}_*) - \mathbf{n}(\mathbf{h}^{-1}(\boldsymbol{\eta}_j; \mathbf{H}_*); \mathbf{N}) \right\|^2, \\ (\mathbf{N})_{s,k} &= 0, \quad \forall (s, k) \in S; (\mathbf{H}_*)_{s,k} = 0, \quad \forall (s, k) \notin S. \end{aligned} \tag{24}$$

Once  $\mathbf{h}^{-1}$  is known, we obtain the coefficients  $\mathbf{H}$  of  $\mathbf{h}$  via regression.

As initial condition for the minimization problem (24), we set all unknown coefficients to zero. This initial guess assumes linear dynamics, which the minimization corrects as needed. We can compute the time derivative in (24) reliably using finite differences, provided that the sampling time  $\Delta t$  of the trajectory data is small compared to the fastest timescale of the SSM dynamics. For larger sampling times, one should use the discrete formulation of SSM theory, as discussed in section “Existence of SSMs” and<sup>29</sup>. In that formulation, the conjugacy error must be formulated for the 1-step prediction error of the normal form flow map  $\boldsymbol{\varphi}_n^{\Delta t}(\mathbf{z})$ . The matrix defined in eq. (21) also carries over to the discrete time setting, with  $\boldsymbol{\Lambda}$  defined as the diagonal matrix of the logarithms of the eigenvalues of  $D\boldsymbol{\varphi}_r^{\Delta t}(\mathbf{0})$ .

**Prediction of forced response from unforced training data.** Forced SSMs continue to be embedded in our observable space, provided that we also include the phase of the forcing among our observables<sup>32</sup>. (In the simplest case of periodic forcing, this inclusion is not necessary, as we pointed out Section “Embedding SSMs via generic observables”). The quasiperiodic SSM-reduced normal form of system (1) in the observable space takes the general form

$$\begin{aligned} \dot{\rho}_j &= \alpha_j(\rho, \boldsymbol{\theta})\rho_j - \sum_{\mathbf{k} \in K_j^{\pm}} f_{j,\mathbf{k}} \sin(\langle \mathbf{k}, \boldsymbol{\Omega} \rangle t + \phi_{j,\mathbf{k}} \mp \theta_j), \\ \dot{\theta}_j &= \omega_j(\rho, \boldsymbol{\theta}) + \sum_{\mathbf{k} \in K_j^{\pm}} \frac{f_{j,\mathbf{k}}}{\rho_j} \cos(\langle \mathbf{k}, \boldsymbol{\Omega} \rangle t + \phi_{j,\mathbf{k}} \mp \theta_j), \end{aligned} \quad j = 1, 2, \dots, m, \quad \mathbf{k} \in \mathbb{Z}^d, \quad \boldsymbol{\Omega} \in \mathbb{R}_+^d, \tag{25}$$

where the terms  $f_{j,\mathbf{k}}$  and  $\phi_{j,\mathbf{k}}$  are the forcing amplitudes and phases for each mode of the SSM and for each forcing harmonic  $\langle \mathbf{k}, \boldsymbol{\Omega} \rangle$ , while  $K_j^{\pm}$  are the set containing the indexes  $\mathbf{k}$  of the resonant forcing frequencies for mode  $j$  (see the Supplementary Information). The normal form (25) will capture non-linearizable dynamics arising from resonant interactions between the eigenfrequencies of the spectral subspace  $E$  (which may also contain internal resonances) and the external forcing frequencies in  $\boldsymbol{\Omega}$ . One can use numerical continuation<sup>70</sup> to find nontrivial co-existing steady states (such as periodic orbits and invariant tori) in eq. (25) under varying forcing amplitudes and forcing frequencies.

To predict forced response from the SSM-based model trained on unforced data, the forcing amplitude  $f$  relevant for eq. (7) in the observable space needs to be related to the forcing amplitude  $|\epsilon f_1|$  relevant for system (1) in the physical phase space. This involves (1) employing a single forcing amplitude-frequency pair  $(|\epsilon f_1|, \Omega)$  in the experiment (2) measuring the periodic observable response  $\mathbf{y}(t)$  (3) computing the corresponding normalized reduced and normalized response amplitude  $\rho_0$  (4) substituting  $\rho_0$  into the first formula in (8) and (5) solving for  $f$  in closed form. This  $f$  can then be used to make a prediction for the full FRC and response phase via (8) in the experiment for arbitrary  $\Omega$  forcing frequencies. The predicted FRC may have several connected components, including isolated responses (isolas) that are notoriously difficult to detect by numerical or experimental continuation<sup>68</sup>.

**Summary of the algorithm.** The data-driven model reduction method used in this paper is available in the open-source MATLAB<sup>®</sup> package SSMLearn. User input is the measured trajectory data of the autonomous dynamical system ( $\epsilon = 0$ ), the SSM dimension  $d$ , the polynomial orders or approximation  $(M, N)$  for the SSM and for the extended normal form, as well as the type of the dynamical system (discrete or continuous). If the number of observables is not sufficient for manifold embedding, the data is automatically augmented with delays to reach the minimum embedding dimension  $p = 2d + 1$ . If the manifold learning returns poor results (due to, e.g., insufficient closeness of the data to the SSM), then the starting value of  $p$  can be increased until a good embedding is found. Then, the algorithm learns the SSM geometry in observable space and, after unsupervised detection of the required normal form, identifies the extended normal form of the reduced dynamics. The level of accuracy can be increased with larger polynomial orders, keeping in mind that excessive orders may lead to overfitting.

SSMLearn also offers all the tools we have used in this paper to analyze the reduced dynamics and make predictions for forced response from unforced training data. In particular, it contains the MATLAB<sup>®</sup>-based numerical continuation core COCO<sup>70</sup>, which can compute steady state and help with the design of nonlinear control strategies. In principle, there are no restrictions on the dimensions of the reduced-order model, yet the larger the SSM is, the more computationally expensive the problem becomes.

Qualitative or partial a priori knowledge of the linearized dynamics (e.g., some linearized modes and frequencies) helps in finding good initial conditions for trajectories to be used in SSMLearn. For example, the resonance decay method<sup>56</sup> (which we exploited in our sloshing example), targets a specific 2-dimensional, stable SSM in laboratory experiments. This method consists of empirically isolating

a resonant periodic motion on the SSM based on its locally maximal amplitude response under a forcing frequency sweep. Discontinuing the forcing will then generate transient decay towards the equilibrium in a close proximity of the SSM. For noisy data, filtering or dimensionality reduction can efficiently de-noise the data<sup>67</sup>, provided that the polynomial orders used for the description of the SSM and its reduced dynamics are not excessively large (see the Supplementary Information). For higher-dimensional SSMs, it is desirable to collect diverse trajectories to avoid bias towards specific motions. Good practice requires splitting the data sets into training, testing and validation parts.

#### Algorithm 1.

SSMLearn

**Input parameters:** SSM dimension  $d$ , polynomial approximation orders  $(M, N)$ , selection among discrete or continuous-time dynamics

**Input data:** measured unforced trajectories

**Output:** SSM geometry, extended normal form of reduced dynamics, predictions for forced response.

1 Embed data in a suitable  $p$ -dimensional observable space with  $p > 2d$ .

2 Identify the manifold parameterization in reduced coordinates.

3 Estimate the normalized reduced dynamics after an automated identification of the required type of extended normal form.

4 Run analytics and prediction of forced response on the SSM-reduced and normalized model.

#### Data availability

All data discussed in the results presented here is publicly available in the SSMLearn repository at [github.com/haller-group/SSMLearn](https://github.com/haller-group/SSMLearn).

#### Code availability

The code supporting the results presented here is publicly available in the SSMLearn repository at [github.com/haller-group/SSMLearn](https://github.com/haller-group/SSMLearn).

Received: 24 September 2021; Accepted: 25 January 2022;

Published online: 15 February 2022

#### References

- Holmes, P.J., Lumley, J.L., Berkooz, G., & Rowley, C.W. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* 2nd edn, (Cambridge Monographs on Mechanics. Cambridge University Press, 2012).
- Awrejcewicz, J., Krysko, V.A., & Vakakis, A.F. *Order Reduction by Proper Orthogonal Decomposition (POD) Analysis*, 279–320 (Springer, Berlin, Heidelberg, 2004).
- Lu, K. et al. Review for order reduction based on proper orthogonal decomposition and outlooks of applications in mechanical systems. *Mech. Sys. Signal Proc.* **123**, 264–297 (2019).
- Brunton, S. L., Proctor, J. L. & Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci.* **113**, 3932–3937 (2016).
- Mohamed, K.S. *Machine Learning for Model Order Reduction* (Springer, Cham, 2018).
- Daniel, T., Casenave, F., Akkari, N. & Ryckelynck, D. Model order reduction assisted by deep neural networks (rom-net). *Adv. Model. Simul. Eng. Sci.* **7**, 105786 (2020).
- Calka, M. et al. Machine-learning based model order reduction of a biomechanical model of the human tongue. *Computer Methods Prog. Biomedicine* **198**, 105786 (2021).
- Loiseau, J.-C., Brunton, S.L., & Noack, B.R. *From the POD-Galerkin method to sparse manifold models*, 279–320 (De Gruyter, Berlin, 2020).
- Karniadakis, G. E. et al. Physics-informed machine learning. *Nat. Rev. Phys.* **123**, 422–440 (2021).
- Li, S. & Yang, Y. Data-driven identification of nonlinear normal modes via physics-integrated deep learning. *Nonlinear Dyn.* **106**, 3231–3246 (2021).
- Fernex, D., Noack, B. R. & Semaan, R. Cluster-based network modeling—From snapshots to complex dynamical systems. *Sci. Adv.* **7**, eabf5006 (2021).
- Schmid, P. J. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28 (2010).
- Kutz, J.N., Brunton, S.L., Brunton, B.W., & Proctor, J.L. *Dynamic Mode Decomposition* (SIAM, Philadelphia, PA, 2016).
- Rowley, C. W., Mezić, I., Bagheri, S., Schlachter, P. & Henningson, D. S. Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**, 115–127 (2009).
- Mezić, I. Analysis of fluid flows via spectral properties of the Koopman operator. *Ann. Rev. Fluid Mech.* **45**, 357–378 (2013).
- Mauroy, A., Mezić, I., & Susuki, Y. *The Koopman Operator in Systems and Control Concepts, Methodologies, and Applications: Concepts, Methodologies, and Applications* (Springer, New York, 2020).
- Lusch, B., Kutz, J. N. & Brunton, S. L. Deep learning for universal linear embeddings of nonlinear dynamics. *Nat. Commun.* **9**, 1–10 (2018).
- Otto, S. E. & Rowley, C. W. Linearly recurrent autoencoder networks for learning dynamics. *SIAM J. Appl. Dynamical Syst.* **18**, 558–593 (2019).
- Kaiser, E., Kutz, J. N. & Brunton, S. L. Data-driven discovery of koopman eigenfunctions for control. *Mach. Learn.: Sci. Technol.* **2**, 035023 (2021).
- Page, J. & Kerswell, R. R. Koopman mode expansions between simple invariant solutions. *J. Fluid Mech.* **879**, 1–27 (2019).
- Brunton, S. L., Brunton, B. W., Proctor, J. L. & Kutz, J. N. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE* **11**, 1–19 (2016b).
- Bagheri, S. Koopman-mode decomposition of the cylinder wake. *J. Fluid Mech.* **726**, 596–623 (2013).
- Page, J. & Kerswell, R. R. Koopman analysis of Burgers equation. *Phys. Rev. Fluids* **3**, 071901 (2018).
- Dowell, E. H. Panel flutter - a review of the aeroelastic stability of plates and shells. *AIAA J.* **8**, 385–399 (1970).
- Abramian, A., Virot, E., Lozano, E., Rubinstein, S. M. & Schneider, T. M. Nondestructive prediction of the buckling load of imperfect shells. *Phys. Rev. Lett.* **125**, 225504 (2020).
- Podder, P., Mallick, D., Amann, A. & Roy, S. Influence of combined fundamental potentials in a nonlinear vibration energy harvester. *Sci. Rep.* **6**, 37292 (2016).
- Orosz, G. & Stépán, G. Subcritical hopf bifurcations in a car-following model with reaction-time delay. *Proc. R. Soc. A* **462**, 2643–2670 (2006).
- Ashwin, P., Wiczeorek, S., Vitolo, R. & Cox, P. Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Philos. Trans. R. Soc. A* **370**, 1166–1184 (2012).
- Haller, G. & Ponsioen, S. Nonlinear normal modes and spectral submanifolds: existence, uniqueness and use in model reduction. *Nonlinear Dyn.* **86**, 1493–1534 (2016).
- Whitney, H. The self-intersections of a smooth  $n$ -manifold in  $2n$ -space. *Ann. Math.* **45**, 220–246 (1944).
- Stark, J., Broomhead, D. S., Davies, M. E. & Huke, J. Takens embedding theorems for forced and stochastic systems. *Nonlinear Anal.: Theory, Methods Appl.* **30**, 5303–5314 (1997).
- Stark, J. Delay embeddings for forced systems. I. deterministic forcing. *J. Nonlinear Sci.* **9**, 255–332 (1999).
- Guckenheimer, J. & Holmes, P. *Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields* (Springer, New York, 1983).
- Fenichel, N. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.* **21**, 193–226 (1971).
- Kuramoto, Y. *Chemical Oscillations, Waves and Turbulence* (Springer, Berlin, 1984).
- Jain, S. & Haller, G. How to compute invariant manifolds and their reduced dynamics in high-dimensional finite-element models? (*Nonlinear Dyn.*, 2021).
- Sauer, T., Yorke, J. A. & Casdagli, M. Embedology. *J. Stat. Phys.* **65**, 579–616 (1997).
- Takens, F. Detecting strange attractors in turbulence. In D. Rand and L. Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, 366–381 (Springer Berlin Heidelberg, 1981).
- Poincaré, H. *Les Méthodes Nouvelles de la Mécanique Céleste*. (Gauthier-Villars et Fils, Paris, 1892).
- Sternberg, S. On the structure of local homeomorphisms of euclidean  $n$ -space, II. *Am. J. Math.* **80**, 623–631 (1958).
- Murdock, J. *Normal Forms and Unfoldings for Local Dynamical Systems*. (Springer Monographs in Mathematics. Springer-Verlag New York, 2003).
- Ponsioen, S., Pedergrana, T. & Haller, G. Automated computation of autonomous spectral submanifolds for nonlinear modal analysis. *J. Sound Vib.* **420**, 269–295 (2018).
- Landau, L. D. On the problem of turbulence. *Dokl. Akad. Nauk SSSR* **44**, 339–349 (1944).
- Stuart, J. T. On the non-linear mechanics of wave disturbances in stable and unstable parallel flows. Part 1. The basic behaviour in plane Poiseuille flow. *J. Fluid Mech.* **9**, 353–370 (1960).
- Fujimura, K. Centre manifold reduction and the Stuart-Landau equation for fluid motions. *Proc.: Math., Phys. Eng. Sci.* **453**, 181–203 (1997).
- Szalai, R., Ehrhardt, D. & Haller, G. Nonlinear model identification and spectral submanifolds for multi-degree-of-freedom mechanical vibrations. *Proc. R. Soc. A* **473**, 20160759 (2017).
- Breunung, T. & Haller, G. Explicit backbone curves from spectral submanifolds of forced-damped nonlinear mechanical systems. *Proc. R. Soc. A* **474**, 20180083 (2018).

48. Cenedese, M., Axàs, J., Yang, H., Eriten, M., & Haller, G. Data-driven nonlinear model reduction to spectral submanifolds in mechanical systems. *arXiv:2110.01929*, 2021.
49. Jain, S., Tiso, P. & Haller, G. Exact nonlinear model reduction for a von Kármán beam: slow-fast decomposition and spectral submanifolds. *J. Sound Vib.* **423**, 195–211 (2018).
50. Barkley, D. & Henderson, R. D. Three-dimensional floquet stability analysis of the wake of a circular cylinder. *J. Fluid Mech.* **322**, 215–241 (1996).
51. Noack, B. R., Afanasiev, K., Morzyński, M., Tadmor, G. & Thiele, F. A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.* **497**, 335–363 (2003).
52. Rowley, C. W. & Dawson, S. T. M. Model reduction for flow analysis and control. *Annu. Rev. Fluid Mech.* **49**, 387–417 (2017).
53. Taylor, G. I. An experimental study of standing waves. *Proc. R. Soc. Lond. Ser. A. Math. Phys. Sci.* **218**, 44–59 (1953).
54. Ockendon, J. R. & Ockendon, H. Resonant surface waves. *J. Fluid Mech.* **59**, 397–413 (1973).
55. Bäuerlein, B. & Avila, K. Phase lag predicts nonlinear response maxima in liquid-sloshing experiments. *J. Fluid Mech.* **925**, 2021 (2021).
56. Peeters, M., Kerschen, G. & Golinval, J. C. Dynamic testing of nonlinear vibrating structures using nonlinear normal modes. *J. Sound Vib.* **330**, 486–509 (2011).
57. Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. Low-order model for successive bifurcations of the fluidic pinball. *J. Fluid Mech.* **884**, A37 (2020).
58. Shaw, S. W. & Pierre, C. Normal modes for non-linear vibratory systems. *J. Sound Vib.* **164**, 85–124 (1993).
59. Renson, L., Kerschen, G. & Cochelin, B. Numerical computation of nonlinear normal modes in mechanical engineering. *J. Sound Vib.* **364**, 177–206 (2016).
60. Neild, S. A., Champneys, A. R., Wagg, D. J., Hill, T. L. & Cammarano, A. The use of normal forms for analysing nonlinear mechanical vibrations. *Philos. Trans. R. Soc. A* **373**, 20140404 (2015).
61. Cirillo, G. I., Mauroy, A., Renson, L., Kerschen, G. & Sepulchre, R. A spectral characterization of nonlinear normal modes. *J. Sound Vib.* **377**, 284–301 (2016).
62. Cabré, X., Fontich, E. & de la Llave, R. The parameterization method for invariant manifolds I: Manifolds associated to non-resonant subspaces. *Indiana Univ. Math. J.* **52**, 283–328 (2003).
63. Kogelbauer, F. & Haller, G. Rigorous model reduction for a damped-forced nonlinear beam model: An infinite-dimensional analysis. *J. Nonlinear Sci.* **28**, 1109–1150 (2018).
64. Haro, A. & de la Llave, R. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: rigorous results. *J. Differential Eqs.* **228**, 530–579 (2006).
65. Haro, A., Canadell, M., Figueras, J.-L., Luque, A., & Mondelo, J.M. *The Parameterization Method for Invariant Manifolds: from Rigorous Results to Effective Computations*. (Springer, New York, 2016).
66. Szalai, R. Invariant spectral foliations with applications to model order reduction and synthesis. *Nonlinear Dyn.* **101**, 2645–2669 (2020).
67. Bishop, C.M. *Pattern Recognition and Machine Learning*. (Information Science and Statistics. Springer-Verlag New York, 2006).
68. Ponsioen, S., Pedergrana, T. & Haller, G. Analytic prediction of isolated forced response curves from spectral submanifolds. *Nonlinear Dyn.* **98**, 2755–2773 (2019).
69. Ponsioen, S., Jain, S. & Haller, G. Model reduction to spectral submanifolds and forced-response calculation in high-dimensional mechanical systems. *J. Sound Vib.* **488**, 115640 (2020).
70. Dankowicz, H. & Schilder, F. *Recipes for Continuation*. (Society for Industrial and Applied Mathematics, 2013).

## Acknowledgements

B.B. and K.A. acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the framework of the research unit FOR 2688 'Instabilities, Bifurcations and Migration in Pulsatile Flows' under Grant No. AV 156/1-1. K.A. acknowledges funding for an 'Independent Project for Postdocs' from the Central Research Development Fund of the University of Bremen.

## Author contributions

M.C. and G.H. designed the research. M.C. carried out the research. M.C. and J.A. developed the software and analyzed the examples. B.B. and K.A. performed the liquid sloshing experiments and participated in their analysis. M.C. and G.H. wrote the paper. G.H. lead the research team.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-28518-y>.

**Correspondence** and requests for materials should be addressed to George Haller.

**Peer review information** *Nature Communications* thanks Bernd Noack, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022