Experimental Implementation of Energy Closure Analysis for Reduced Order Modeling

Suparno Bhattacharyya[1]
Department of Engineering Science and
Mechanics,
Pennsylvania State University,
University Park, PA 16802
e-mail: sxb1086@psu.edu

Joseph. P. Cusumano[1]
Department of Engineering Science and
Mechanics,
Pennsylvania State University,
University Park, PA 16802
e-mail: jpc3@psu.edu

Reduced order models (ROMs) provide an efficient, kinematically condensed representation of computationally expensive high-dimensional dynamical systems; however, their accuracy depends crucially on the accurate estimation of their dimension. We here demonstrate how the energy closure criterion, developed in our prior work, can be experimentally implemented to accurately estimate the dimension of ROMs obtained using the proper orthogonal decomposition (POD). We examine the effect of using discrete data with and without measurement noise, as will typically be gathered in an experiment or numerical simulation, on estimating the degree of energy closure on a candidate reduced subspace. To this end, we used a periodically kicked Euler–Bernoulli beam with Rayleigh damping as the model system and studied ROMs obtained by applying POD to discrete displacement field data obtained from simulated numerical experiments. An improved method for quantifying the degree of energy closure is presented: the convergence of energy input to or dissipated from the system is obtained as a function of the subspace dimension, and the dimension capturing a predefined percentage of either energy is selected as the ROM dimension. This method was found to be more robust to data discretization error and measurement noise. The data-processing necessary for the experimental application of energy closure analysis is discussed in detail. We show how ROMs formulated from the simulated data using our approach accurately capture the dynamics of the beam for different sets of parameter values. [DOI: 10.1115/1.4054295]

Keywords: dynamical systems, modal analysis, structural vibration, reduced order modeling (ROM), proper orthogonal decomposition (POD), Physics-informed model reduction, energy closure analysis, energy balance, vibro-impact systems

Introduction

Reduced order models (ROMs) identify and capture the dominant features of high-dimensional dynamical systems and thereby provide a parsimonious representation of their dynamics. Derivation of such models makes "many-query" computational analyses—such as for uncertainty propagation, optimization, and control—feasible and more efficient. However, accurately estimating the dimension of such models remains a challenge and is one of the prime factors that determine their accuracy. For ROMs formulated using proper orthogonal decomposition (POD), which are obtained by projecting governing equations onto low-dimensional subspaces derived from system-generated data, the dimension is typically determined by selecting a subspace that captures a sufficiently large percentage of the data variance [1]. In Ref. [2], we studied a simple structural system subjected to impulsive loading and demonstrated that the standard variance-based dimension estimation led to inaccurate ROMs. As an alternative to the variance-based approach, we proposed a physics-based approach, in which the dimension is chosen so that the energy input and dissipated from the associated reduced subspace is approximately in balance. However, in the previous study, we used analytical solutions as the basis of our calculations. In this article, we show how this *energy closure analysis* can be implemented using discrete, noisy data of the type gathered in physical experiments or numerical studies. To do this, we again use the periodically kicked Euler–Bernoulli beam of Ref. [2] as our model system and simulate experiments in which spatially and temporally discrete measurements are made.

Proper orthogonal decomposition is one of the most widely used data-driven dimension reduction methods. It is used to identify the predominant low-dimensional embedded patterns that describe the dynamical activity of a high-dimensional system [1,3–8]. These patterns are extracted in the form of a hierarchical set of basis functions, also known as proper orthogonal modes (POMs), a subset of which provides the least-squares optimal, low-dimensional representation of the parent data. The dimension of the resulting subspace is chosen to capture a predefined percentage of the data variance. Unfortunately, however, while data projection error is guaranteed to decrease monotonically with the fraction of total variance captured, the same relationship is not guaranteed to hold for the modeling error of a ROM formulated on this subspace. As a result, the dimension of a subspace sufficient for accurate representation of high-dimensional *data* can still fail to adequately capture the *dynamics* that generated the data.

There have been research efforts to better incorporate physics into reduced order models obtained with POD, particularly in the fluid mechanics community. For example, in fluid systems, a fairly common approach to address the issue of modal truncation is to a add a closure model to an existing ROM, which accounts for the effect of discarded modes [9–19]; reviews of different ROM-closure models implemented in fluid systems can be found in Refs. [20,21]. A fundamental idea used in these efforts is that of eddy viscosity, in which the effect of small-scale eddies in a fluid flow is modeled by adding a term in the fluid flow equation, which obviates the need to numerically resolve the behaviors of small-scale eddies and leads to reasonably accurate yet fast simulations. In the context of control problems, balanced POD, or BPOD [22,23], is used instead of regular POD, where the modes are constructed from the system-generated data such that in the direction of the modes there is an optimal balance between controllability and observability of the reduced system.

For structural vibration problems, in Refs. [24–27], an approach was proposed for selecting the dimension of ROMs using the

---

[1] Corresponding authors.

Contributed by the Technical Committee on Vibration and Sound of ASME for publication in the Journal of Vibration and Acoustics. Manuscript received June 23, 2021; final manuscript received April 3, 2022; published online May 4, 2022. Assoc. Editor: Shahrzad Towfighian.

concept of “dynamical consistency.” This approach takes into account the state-space dynamics of the system generating the data, but it is fundamentally tied to the study of chaotic steady states, for which delay reconstruction can be applied. However, very limited literature exist on ROM closure in the context of structural vibrations. Guo and Przekop in Ref. [28] proposed an energy-based mode selection approach for reduced order modeling of vibrating plate structures, which involved estimating POMs and then constructed a subspace for model reduction using linear normal modes of the system that resembled the highest-energy POMs. This work used both smooth and proper orthogonal decomposition [29,30] to estimate the modal energies. However, in addition to assuming some resemblance between POMs and the explicitly available normal modes, which may be difficult to achieve in general, the approach did not address the physical requirement that energy be balanced on the subspace spanned by the selected modes.

In Ref. [2], we suggested an alternative, energy-based approach for estimating the dimension of ROMs obtained with POD. With a focus on structural vibration problems, we studied the model reduction of a periodically kicked Euler–Bernoulli beam via POD. We showed that energy closure analysis provides physical insight to the process of dimension selection. Specifically, we showed that to accurately capture the steady-state dynamics of the kicked beam with a ROM, it is crucial to select its dimension so that the associated reduced subspace (onto which the governing partial differential equations (PDEs) are projected to formulate the ROMs) accurately captures the total energy input and dissipated from the system and, on that subspace, both are approximately in balance. We found that ROMs formulated using this approach were able to accurately capture the dynamics of the true system more reliably and consistently than those formulated using a fixed variance, because the latter approach can easily ignore the contribution of modes with smaller variances, even when they are important to the system’s dynamics.

In the context of computational fluid dynamics, Balajewicz et al. [31,32] used the closely related idea of power balance on statistically stationary turbulent flows; however, in that study, power balance was used as a constraint to formulate a set of orthonormal basis functions that, like POD, optimally respresents the system-generated data. In contrast, our energy closure approach provides a way to accurately represent the energetics of the system in the model reduction process without making any changes to the already widely familiar algorithm of POD.

For dynamical systems that are nonlinear and/or have a complicated geometry, analytical solutions, such as those used for our previous study [2], are typically not available. Equations of motion of such systems are often solved numerically using finite difference, finite element, or finite volume methods where the solutions are discretized in space and time. Likewise, in physical experiments, the quantities of interest, such as displacements, are acquired via discrete measurements at a given sampling rate from specific locations in the problem’s domain and are furthermore contaminated by measurement noise. In this work, we address the challenges associated with energy balance calculations performed with such discrete, noisy data. We use the same kicked Euler–Bernoulli beam model system and perform simulated experiments to generate discrete kinematic data from it. Experimental measurement noise is also simulated and added to the data.

A hypothetical experimental setup is shown in Fig. 1, in which a beam is pinned at both ends and is kicked periodically over a small region for a short-time interval. The data flow in our analysis starts with the collection of displacement time series from specific locations on the beam using an array of sensors. For a given steady state, we first apply POD to acquired displacement data to calculate the POMs. By using the displacement data, velocity data are estimated, as needed to determine the energy balance error on any candidate reduced subspace spanned by the POMs. The dimension of the subspace for which the energy balance error goes below a specified tolerance is selected as the ROM dimension. The data processing needed to carry out these calculations must overcome inaccuracies arising from discretization, on the one hand, and additive measurement noise, on the other hand, as we discuss in what follows. In addition, we present an improved metric for estimating the energy balance error: compared to that used in our previous work, the new metric is more robust to discretization error and noise and is also easier to interpret.

Model Reduction Procedure

For the sake of completeness, we here briefly summarize the model reduction procedure for the Euler–Bernoulli beam, performed using data generated from its simulation. Further details

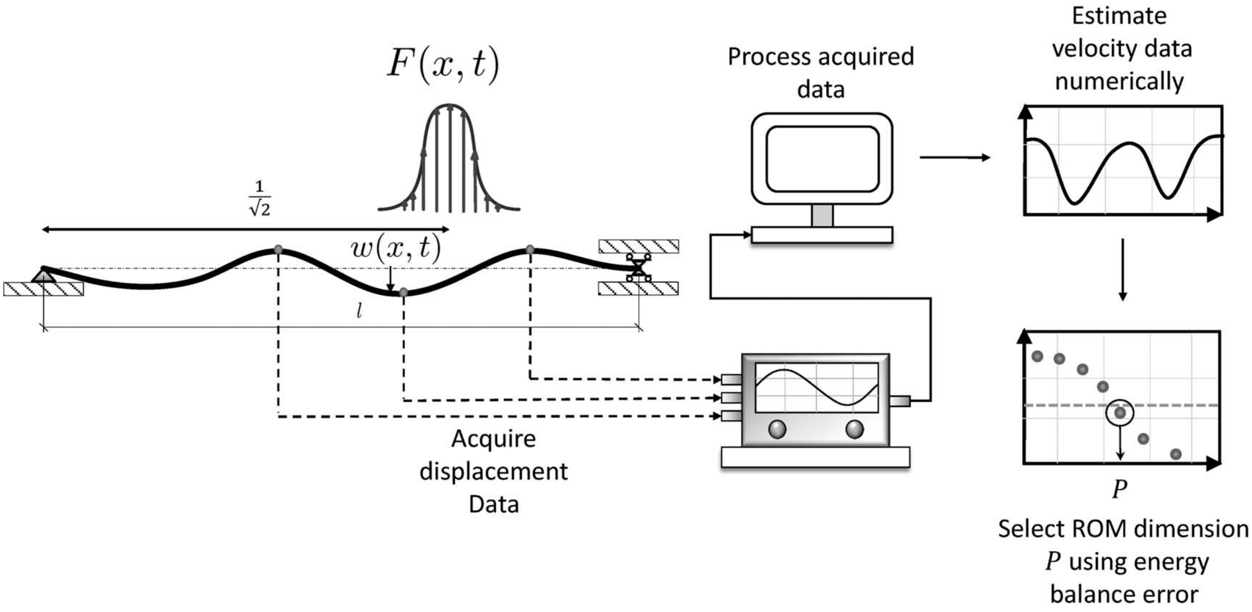

Fig. 1 **Schematic of the experimental setup used for the simulations in this article, showing the data collection and analysis used to estimate the dimension of a ROM of a kicked Euler–Bernoulli beam**

including a detailed description of the model and the derivation of its closed-form solutions is available in Ref. [2].

Data Generation. The Euler–Bernoulli beam considered here has uniform cross section with Rayleigh damping. It is pinned at both ends and is subjected to a periodic "impulsive" distributed load (see Fig. 1). We denote the transverse displacement of the beam, as measured from its neutral axis, as $w(x, t)$ and write the dimensionless equation of motion as follows:

$$w''''(x, t) + c_v \dot{w}(x, t) + c_m \dot{w}''''(x, t) + \ddot{w}(x, t) = F(x, t) \tag{1a}$$

with boundary conditions

$$\begin{gathered} w(0, t) = 0, \quad w''(0, t) + c_m \dot{w}''(0, t) = 0, \quad w(1, t) = 0, \\ w''(1, t) + c_m \dot{w}''(1, t) = 0 \end{gathered} \tag{1b}$$

where $0 \le x \le 1$ denotes the position along the length of the beam; t denotes time; primes and overdots denote spatial and temporal derivatives, respectively; c_v and c_m denote the nondimensional viscous and material damping coefficients [33], respectively.

The forcing term F is defined as follows:

$$F(x, t) = f(x; \epsilon, s) g(t; \tau, T) \tag{2}$$

with spatial factor

$$f(x; \epsilon, s) = \begin{cases} \dfrac{2}{\epsilon} \cos^2\!\left(\dfrac{\pi}{\epsilon}(x - s)\right) & \text{for } s - \dfrac{\epsilon}{2} \le x \le s + \dfrac{\epsilon}{2} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

in which the parameter s denotes the location on the beam, where f is maximum (here taken to be $s = 1/\sqrt{2}$), and the periodic temporal factor $g(t; \tau, T)$ represents the Fourier series of a periodic pulse train, defined as follows:

$$G(t; \tau, T) = \begin{cases} \dfrac{2}{\tau} \cos^2\!\left(\dfrac{\pi}{\tau}\left(t - \dfrac{T}{2}\right)\right) & \text{for } \left(k - \dfrac{1}{2}\right)T - \dfrac{\tau}{2} \le t \le \left(k - \dfrac{1}{2}\right)T + \dfrac{\tau}{2} \\ 0 & \text{for } (k-1)T \le t < \left(k - \dfrac{1}{2}\right)T - \dfrac{\tau}{2} \text{ and } \left(k - \dfrac{1}{2}\right)T + \dfrac{\tau}{2} < t \le kT \end{cases} \tag{4}$$

where each pulse G takes place over $(k-1)T \le t \le kT$, with $k = 1, 2, 3, \ldots, \infty$, and T is the forcing period. The force density $F(x, t)$ of Eq. (2) generates a unit impulse over one forcing period, for all values of ϵ and τ, so that

$$\int_0^1 \int_0^T F(x, t)\, \mathrm{d}x\, \mathrm{d}t = 1 \tag{5}$$

With smaller values of ϵ and τ, F becomes spatio-temporally more localized and its peak value increases until, in the limit, as ϵ and τ approach zero, and F takes the form of a delta function located at $x = 1/\sqrt{2}$. Thus, by tuning the parameters ϵ and τ, it is possible to control the "impulsivity" of F, and consequently, the number of excited modes in, and the dimensionality of, the system's response.

The "exact" steady-state periodic solution (with period same as the forcing period T) of Eq. (1a) is expressed as the linear combination of the first hundred mass-normalized normal modes of the beam [34,35]:

$$w(x, t) = \sum_{i=1}^{100} a_i(t) \phi_i(x) \tag{6}$$

where $\phi_k(x) = \sqrt{2}\, \sin(k\pi x)$. The time-dependent modal coordinates $a_i(t)$ are solved analytically from the set of decoupled ordinary differential equations (using the method of undetermined coefficients) obtained by substituting Eq. (6) in Eq. (1a), followed by using the Galerkin projection. The normal modes here are used as an admissible basis for the Galerkin projection, and the boundary conditions are incorporated in the resulting integral equation.

We used the analytical solutions $a_i(t)$ in Eq. (6) to generate spatio-temporally discretized displacement data for our simulated experiments. Considering N_Ω and N_t as the number of discrete measurements in the spatial and temporal domains, respectively, we represent the data so collected by the matrix $W \in \mathbb{R}^{N_\Omega \times N_t}$:

$$W = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_{N_t} \end{bmatrix} \tag{7}$$

with columns

$$\mathbf{w}_k := \begin{bmatrix} w(x_1, t_k) & w(x_2, t_k) & \cdots & w\left(x_{N_\Omega}, t_k\right) \end{bmatrix}^T \tag{8}$$

In each simulated experiment, 90 Fourier modes were used to model the periodic temporal factor g of the forcing function (Eq. (2)). The spatial domain of the beam was discretized using $N_\Omega = 45$ grid points after initial experimentation. Furthermore, the parameters of F were set to small values, $\tau = 0.001$ and $\epsilon = 0.002$, to generate an impulsive loading condition. The remaining parameters were varied from case to case.

Proper Orthogonal Decomposition. For continuous systems, POD is used to determine a finite set of hierarchical orthonormal basis functions $\psi(x) \in L^2(\Omega)$ $(\Omega = [0, 1])$, which are the proper orthogonal modes (POMs), that represent the displacement w optimally in the L^2 sense. That is, the projection of w onto a subspace spanned by a finite number of these POMs yields minimum projection error in comparison to any other subspace of the same dimension spanned by any other orthonormal basis. It can be shown [1] that this projection error minimization is equivalent to the following eigenvalue problem, from which the ψ's are calculated:

$$\int_\Omega r(x, y) \psi(y)\, dy = \lambda \psi(x) \tag{9}$$

where the integral kernel $r(x, y)$ represents the two-point spatial autocorrelation function

$$r(x, y) = \left\langle w(x, t) w(y, t) \right\rangle \tag{10}$$

in which the angle brackets $\langle \cdot \rangle$ indicate the time average over the interval $0 \le t \le T$.

The eigenvalue λ represents the variance of w along the corresponding POM $\psi(x)$, so the magnitude of the eigenvalues indicates the relative dominance of various "coherent structures" [1,36], that is, the POMs, in the dynamics. The POMs are sorted based on the magnitudes of their associated eigenvalues, and the first P of these modes are chosen to form a P-dimensional subspace.

When the displacement field w is discretized in space and time, as is the case here, and mean centered, the discrete version of the autocorrelation function gives the covariance matrix $R \in \mathbb{R}^{N_\Omega \times N_\Omega}$:

$$R = \frac{1}{N_t} W W^T \tag{11}$$

which results from employing trapezoidal integration to approximate the time average in Eq. (10) [4]. This converts Eq. (9) to the matrix eigenvalue problem

$$R\boldsymbol{\psi} = \lambda\boldsymbol{\psi} \tag{12}$$

where $\boldsymbol{\psi} \in \mathbb{R}^{N_\Omega \times 1}$ is the finite-dimensional eigenvector of R.

Reduced Order Model. Once the POMs are determined, some criterion must be used (as discussed in the following section) to select P, the number of POMs, $\{\psi_i\}_{i=1}^{P}$, for constructing a subspace onto which the governing equation Eq. (1a) is projected to obtain a P-dimensional ROM. For continuous systems, the P-dimensional approximation to w

$$w(x, t) \approx \widehat{w}(x, t) = \sum_{j=1}^{P} b_j(t)\psi_j(x) \tag{13}$$

is substituted in Eq. (1a), and using the boundary conditions Eq. (1b), a system of linear second-order ordinary differential equations (ODEs) is obtained in terms of proper orthogonal coordinates $b_j(t)$:

$$M\ddot{\mathbf{b}} + C\dot{\mathbf{b}} + K\mathbf{b} = \widehat{\mathbf{f}} \tag{14}$$

where $\mathbf{b} \in \mathbb{R}^{P\times 1}$, and the elements of the mass matrix $M \in \mathbb{R}^{P\times P}$, stiffness matrix $K \in \mathbb{R}^{P\times P}$ and the reduced forcing vectors $\widehat{\vec{f}} \in \mathbb{R}^{P\times 1}$ are given for continuous POMs by

$$M_{ij} = \int_0^1 \psi_i(x)\psi_j(x)\,\mathrm{d}x \tag{15a}$$

$$K_{ij} = \int_0^1 \psi_i''(x)\psi_j''(x)\,\mathrm{d}x \tag{15b}$$

$$\widehat{f}_i = g\int_0^1 f\psi_i(x)\,\mathrm{d}x \tag{15c}$$

with f and g as in Eq. (2). The damping matrix in Eq. (14) is defined as $C = c_v M + c_m K$.

For discrete displacement data contained in W, the POMs are no longer continuous functions of x, but N_Ω-dimensional vectors. To estimate the mass and stiffness matrices, one option is to evaluate the integrals in Eqs. (15a) and (15b) numerically. However, for a coarse spatial discretization (which may occur, for example, because of a limited number of sensors that can be placed on the structure), direct numerical integration may yield high error. Hence, to achieve a higher accuracy, we first interpolated the POMs using the mass-normalized normal modes $\phi_j(x)$ of the system. In general, should the normal modes of a system not be available, any other admissible basis functions that satisfy the essential boundary conditions could also be used.

We define the matrix $\Psi \in \mathbb{R}^{N_\Omega \times P}$ containing the first P vector POMs, $\boldsymbol{\psi}_i \in \mathbb{R}^{N_\Omega \times 1}$, as its columns

$$\Psi = \begin{bmatrix} \boldsymbol{\psi}_1 & \boldsymbol{\psi}_2 & \cdots & \boldsymbol{\psi}_P \end{bmatrix} \tag{16}$$

and relate the jth element of the ith column of Ψ, Ψ_{ji}, to a continuous function $\Psi_i(x)$ via modal interpolation:

$$\Psi_{ji} \equiv \Psi_i(x_j) = \sum_{k=1}^{N_r} H_{ki}\phi_k(x_j) \tag{17}$$

where $N_r = 40 \leq N_\Omega$. The coefficient matrix $H \in \mathbb{R}^{N_r \times P}$ with element H_{ki} is determined via linear regression:

$$H = \Phi^{\dagger}\Psi \tag{18}$$

where $\Phi^{\dagger}$ represents the Moore–Penrose inverse [37] of $\Phi \in \mathbb{R}^{N_\Omega \times N_r}$, which has elements $\phi_{jk} = \phi_k(x_j)$ denoting the value of the kth normal mode at x_j. Instead of regression, we could have also used interpolation with $N_r = N_\Omega$ normal modes, but in anticipation of the later analysis of noisy data in Sec. 4, we opted to use regression since the interpolation of noisy data generally leads to overfitting. The functional representation of $\psi_i(x)$ defined by Eq. (17) was then used to symbolically evaluate M, K, and $\widehat{\vec{f}}$ in Eq. (15).

The steady-state solution to the ROM thus obtained can be solved using standard libraries for solving ordinary differential equations (e.g., `ode45` in MATLAB). Since we are interested only in steady-state data, we avoided transients by first diagonalizing Eq. (14) using its generalized eigenvectors and then calculated the steady-state solution of the decoupled equations using symbolic algebra via method of undetermined coefficients [2]. The displacement and velocity fields of the ROMs were then estimated using the matrix expressions:

$$\begin{gathered} \widehat{W} = \Psi B \\ \dot{\widehat{W}} = \Psi \dot{B} \end{gathered} \tag{19}$$

where $B = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{N_t}]$ contains the values of $\mathbf{b}$ Eq. (14) at different sample times as its columns. Subsequently, we calculated the mean square error between displacement and velocity data predicted by the ROM and that of the true system and used these errors to characterize the accuracy of the ROM.

Energy Closure Analysis. Arguably, the most fundamental factor that determines the accuracy of a ROM is the choice of its dimension P. Traditionally, a variance-based approach is implemented in which P is selected so that the associated subspace captures a predefined percentage, p, of the total variance in the system-generated data. Recalling that the eigenvalues λ_i give the variance of data predicted in the direction of the ith POM, this criterion is defined mathematically as follows:

$$P = \underset{P}{\operatorname{argmin}}\left\{\frac{\sum_{i=1}^{P}\lambda_i}{\sum_{i=1}^{N_\Omega}\lambda_i} \geq p\right\} \tag{20}$$

However, even though p is often chosen to be large (99.9% or greater), the resulting subspace only ensures a low *data projection* error [1]: it does not guarantee that a ROM formulated on it will have a low *modeling* error. Being purely data driven, this approach often ignores the contribution of POMs with small variances that are, nevertheless, essential for capturing the true dynamics of the system, leading to poor ROMs. In such cases, p is adjusted by trial and error so as to include all the dynamically important modes.

However, this purely statistical approach does not provide any physical understanding that can be used to guide the dimension selection process. In our previous work [2], we showed that it fails to estimate an accurate ROM dimension for the kicked Euler–Bernoulli beam when subjected to impulsive loading conditions, leading to large modeling errors. In contrast, we found that accurate ROMs could be consistently obtained by choosing a value of P ensuring that the resulting subspace is closed with respect to energy flowing in and out of it [2]. This approach has the added benefit of providing a physics-based explanation as to why a given dimension is needed.

The energy dissipated from our full-order model system over one period T of its steady-state response is

$$W_d = c_v\int_0^T\int_0^1 \dot{w}^2(x, t)\,\mathrm{d}x\,\mathrm{d}t + c_m\int_0^T\int_0^1 \dot{w}''^2(x, t)\,\mathrm{d}x\,\mathrm{d}t \tag{21a}$$

whereas the energy input is

$$W_f = \int_0^T\int_0^1 f(x; \epsilon)g(t; \tau, T)\dot{w}(x, t)\,\mathrm{d}x\,\mathrm{d}t \tag{21b}$$

At steady state, W_d and W_f must be in balance:

$$W_d = W_f \tag{22}$$

To check whether W_d and W_f are in balance on a subspace of dimension P, we estimate them using the P-dimension projection of the original velocity data and its spatial derivatives in Eq. (21). We write these estimates as $\widetilde{W}_d$ and $\widetilde{W}_f$, respectively. For an insufficiently large value of P, $\widetilde{W}_d$ and $\widetilde{W}_f$ will not be in balance. In Ref. [2], we defined an energy balance residual e_W as follows:

$$e_W = \left| \widetilde{W}_d - \widetilde{W}_f \right| \tag{23}$$

which provides a measure of the energy imbalance on a given subspace.

In principle, following Eq. (22), we can select the value of P such that on the corresponding subspace, $e_W \approx 0$ and both $\widetilde{W}_d$ and $\widetilde{W}_f$ are close to their true values, which ensures approximate energy balance and indicates that all dynamically relevant dimensions (POMs) have been included. However, when using finite precision noisy data, the estimation of W_d and W_f, and the application of the energy residual Eq. (23), present challenges that must be overcome, as discussed in the following sections.

Estimating Reduced Order Model Dimension Using Discrete Experimental Data

The energy closure analysis outlined in the previous section relies on the determination of the energy dissipated and input to a given subspace. In this section, we discuss how to estimate e_W from discrete displacement data. The two main problems associated with this estimate are (1) the numerical error introduced by the discretization and (2) for physical experiments, measurement noise. We address the first in this section, leaving the effect of measurement noise for the next section.

In each simulated experiment, the sampling frequency for collecting the displacement data was chosen to be larger than the Nyquist frequency corresponding to the highest input frequency, 90/*T*, corresponding to the 90th temporal Fourier mode of the forcing function.

Estimating the Velocity Field. Good estimates of velocity field data are required for energy balance calculations Eq. (21). However, it may be difficult to measure velocity directly in an experimental setup, particularly in the vibrations of solids and structures. Although there exist methods of doing so (such as laser vibrometry), such apparatus is frequently not available, so we assumed that velocity data were not independently collected, thus providing a worst-case test of our approach. (We note that, in contrast with systems undergoing structural vibrations, in experimental studies of fluid systems, velocity is typically the primary measurement.)

Multiple approaches can be adopted to estimate the velocity data, for example, by using the finite difference method [38], fitting a spline through the displacement data and taking its derivative, or using spectral differentiation [4]. We implemented all three and found that spectral differentiation yielded the most accurate results. The accuracy associated with numerical differentiation schemes, such as with the forward Euler or the central differences, is limited by the size of the time-step and hence on the sampling frequency. As an example, for simulations with time period $T = 1/2\pi$, sampling frequency 3000/*T*, and damping coefficients $c_v = 1.000$ and $c_m = 0.001$, the central difference scheme gave a relative estimation error on the order of 10^{-5}. Similarly, first fitting B-splines to the data and then differentiating gave an estimation error that was dependent on both the sampling rate and the order of the B-spline. The lowest velocity estimation error we could achieve was with an eighth-order B-spline, which gave an error on the order of 10^{-11}. In contrast, for the same sampling rate, velocity data calculated using spectral differentiation showed an estimation error on the order of 10^{-13}, which is at least two orders of magnitude lower than other approaches.

To calculate the spectral derivative of the displacement data, the fast Fourier transform (FFT) was applied to time series $w(x_i, t_j)$ obtained at each location x_i on the beam. Then, to incorporate the effect of differentiation, the kth element of the FFT vector was multiplied by $2\pi i\, k/N_t$, where N_t is the total number of points on the time series. The inverse FFT of the resulting vector then gave the velocity time series corresponding to each x_i.

The acquired displacement field and the estimated velocity time series at the midpoint of the beam, corresponding to the parameter values mentioned in the previous paragraph, are shown in Figs. 2(*a*) and 2(*b*).

The estimated discrete velocity field data, denoted as $v(x_i, t_j)$, was assembled into a data matrix $V \in \mathbb{R}^{N_\Omega \times N_t}$ and its relative error e_v, was calculated by comparing it with the true values $\dot{w}(x_i, t_j)$ obtained from the simulation of the original system:

$$e_v = \frac{\langle \| \dot{w}(x,t) - v(x,t) \| \rangle}{\langle \| \dot{w}(x,t) \| \rangle} \tag{24}$$

where $\|\cdot\|$ denotes standard the L^2 norm. The integrations in Eq. (24) were again performed numerically using the trapezoidal scheme, as in Sec. 2.

Estimation of the Energy Balance Error. We estimated the energy balance residual, e_W, on a P-dimensional subspace by evaluating $\widetilde{W}_d$ and $\widetilde{W}_f$ Eq. (21) using the projected velocity data matrix $\widetilde{V}$:

$$\widetilde{V} = \Psi \Psi^T V \tag{25}$$

where the matrix Ψ is as in Eq. (16). To reduce discretization errors, we interpolated each column of $\widetilde{V}$, which represents the spatial component of the velocity data at every time sample, using n linear normal modes, $\{\phi_i(x)\}_{i=1}^n$, as was done to interpolate the POMs

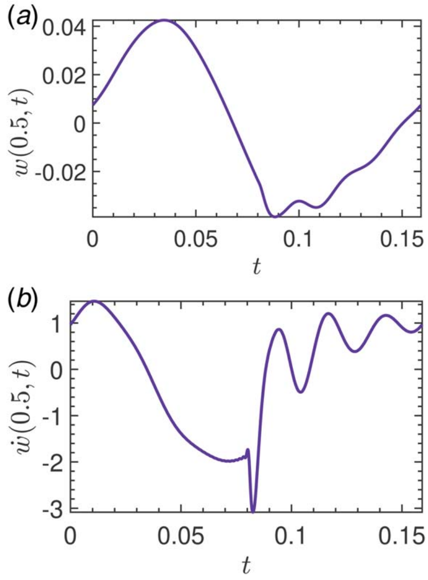

Fig. 2 Steady-state displacement and velocity time series over one period corresponding to the midpoint ($x = 0.5$) of the beam, which is forced ($\epsilon = 0.002$, $\tau = 0.001$) at its second natural frequency ($T = 1/2\pi \approx 0.16$): (*a*) the displacement time series, considered here as the primary measurement, is acquired at a sampling rate of 3000/*T* and (*b*) the velocity time series, which is estimated from the displacement time series via spectral differentiation with a relative error of $\approx 10^{-13}$ Eq. (24). The coefficients of viscous and material damping are $c_v = 1.000$ and $c_m = 0.001$, respectively.

Eq. (17):

$$\widetilde{V}_{ij} \equiv \widetilde{v}(x_i, t_j) = \sum_{k=1}^{n} A_{kj}\phi_k(x_i) \equiv \sum_{k=1}^{n} a_k(t_j)\phi_k(x_i) \tag{26}$$

We used $n = 40 \leq N_\Omega$, and the coefficient matrix elements $A_{ij} = a_i(t_j)$ were determined via linear regression as was done for Eqs. (16)–(18).

Using the fact that $\int_0^1 \phi_i(x)\phi_j(x)\,\mathrm{d}x = \delta_{ij}$ and $\int_0^1 \phi_i''(x)\phi_j''(x)\,\mathrm{d}x = \omega_i^2\delta_{ij}$, it is then possible to rewrite the expressions of $\widetilde{W}_d$ and $\widetilde{W}_f$ Eq. (21) as follows:

$$\widetilde{W}_d = c_v \sum_{i=1}^{n} \int_0^T a_i^2(t)\,\mathrm{d}t + c_m \sum_{i=1}^{n} \omega_i^2 \int_0^T a_i^2(t)\,\mathrm{d}t \tag{27a}$$

$$\widetilde{W}_f = \sum_{i=1}^{n} q_i\gamma_i \tag{27b}$$

where

$$q_i = \int_0^T g(t)a_i(t)\,\mathrm{d}t \tag{28a}$$

$$\gamma_i = \int_0^1 f(x)\phi_i(x)\,\mathrm{d}x \tag{28b}$$

We used Boole's algorithm [39], which is a sixth-order integration scheme, to accurately evaluate Eq. (28a) using its nodal values. The higher order method was necessary to overcome large integration errors induced by the impulsive nature of $g(t)$. The integral in Eq. (28b) was evaluated analytically, as both $f(x)$ and $\phi_i(x)$ are known.

By using Eq. (27), we evaluated and plotted e_W Eq. (23) against different values of subspace dimension P. In Fig. 3(a), we show plots of the projected energies $\widetilde{W}_d$ and $\widetilde{W}_f$, together with e_W, for a typical simulated experiment using the same system parameters as shown in Fig. 2. For comparison, we also plot e_W^a, which is the energy balance error calculated using the analytical solutions, as was done in Ref. [2]. Both e_W and e_W^a are plotted using a log scale on the left y-axis, whereas $\widetilde{W}_d$ and $\widetilde{W}_f$ are plotted using a linear scale on the right y-axis. Note that on the x-axis the maximum subspace dimension is $P_{\max} = 16$. This upper bound is calculated from the numerically estimated rank of the associated covariance matrix R Eq. (11): eigenvectors corresponding to $P_{\max} >> \mathrm{rank}(R)$ span the null space of R and, hence, do not contain information about the dynamics of the system.

We observe that both $\widetilde{W}_d$ and $\widetilde{W}_f$ increase monotonically up to about $P = 11$ and saturate afterward. Furthermore, we see that both e_W and e_W^a drop significantly at $P = 11$. Beyond $P = 11$, we observe that e_W^a continues to decay with a general downward trend toward zero, indicating an approximate energy balance is achieved; however, because of data discretization error, especially at smaller scales, numerical errors are introduced in e_W for larger values of P, and as a result, the decaying trend observed in e_W^a is lost. In addition, both e_W and e_W^a are negatively impacted by substantial floating-point error incurred while taking the difference between $\widetilde{W}_d$ and $\widetilde{W}_f$: because they are approximately equal for larger values of P, their difference becomes a poorly conditioned computation and, therefore, is highly sensitive to round off errors. Once we add measurement noise, as we do in the next section, we therefore expect even more error in the estimation of e_W.

To resolve this issue, we examined the convergence of $\widetilde{W}_f$ or $\widetilde{W}_d$ individually with increasing P. Equation (22) shows that once $\widetilde{W}_f$ or $\widetilde{W}_d$ is captured with sufficient accuracy, the physics *requires* that they approximately balance each other. Hence, there is no reason to use the poorly conditioned residual e_W. Instead, one can select the value of P for which either quantity converges to within a given tolerance of its true value, taken to be its saturated (maximum) value. We thus define the convergence errors for $\widetilde{W}_d$ and $\widetilde{W}_f$ as follows:

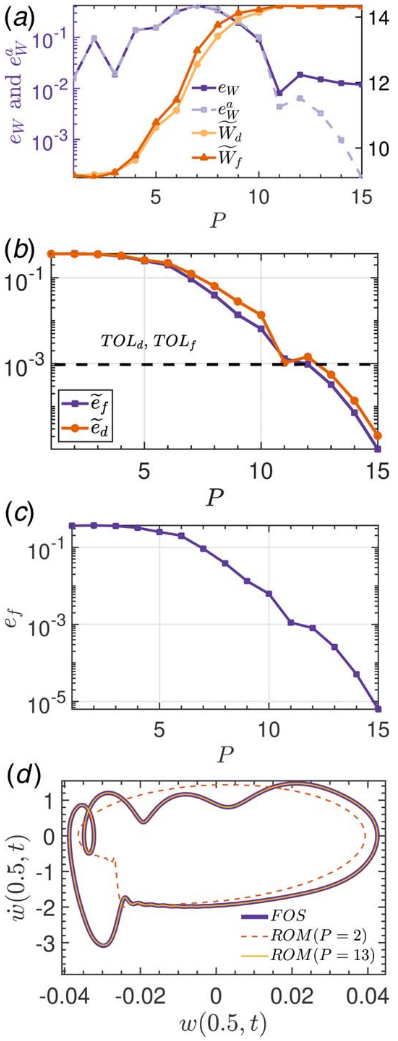

Fig. 3 Energy balance error and ROM error as a function of subspace dimension P, calculated using discrete data sets obtained from simulated numerical experiments. The subfigures were computed using the same system parameters as shown in Fig. 2. (*a*) The approximate energy dissipated from ($\widetilde{W}_d$), and input to ($\widetilde{W}_f$), a given subspace, calculated using the projection of the estimated velocity data, are plotted against the right *y*-axis. Their difference e_W (see Eq. (23)) is shown on the left *y*-axis using a log scale; on the same axis, for comparison, we show e_W^a, which is the same energy balance error calculated analytically [2]. (*b*) The projected energy convergence errors $\widetilde{e}_d$ and $\widetilde{e}_f$ (see Eq. (29)) versus P, as an alternative approach to measuring energy imbalance on the subspace: both curves decrease below the tolerance (dashed line) at $P = 13$. (*c*) The displacement ($\widehat{e}_d$) and velocity ($\widehat{e}_v$) ROM errors (see Eq. (30)) versus P: large filled squares correspond to the two-dimensional ROM, and large filled circles to the 13-dimensional ROM. Observe the "knee" in the plots at $P = 13$. (*d*) Phase plane plots of the full-order system compared to that of two ROMs with dimensions estimated using energy closure ($P = 13$) and 99.9% variance-based ($P = 2$) criteria.

$$\widetilde{e}_d = \left| 1 - \frac{\widetilde{W}_d}{\widetilde{W}_d^{P_{\max}}} \right| \tag{29a}$$

$$\widetilde{e}_f = \left| 1 - \frac{\widetilde{W}_f}{\widetilde{W}_f^{P_{\max}}} \right| \tag{29b}$$

respectively, where $\widetilde{W}_f^{P_{\max}}$ and $\widetilde{W}_d^{P_{\max}}$ are the values corresponding to $P = P_{\max}$. For the same experiment considered in Fig. 3(*a*), we plot $\widetilde{e}_f$ and $\widetilde{e}_d$ against P in Fig. 3(*b*). We observe that both curves decay in a nearly monotonic fashion as P is increased, indicating that $\widetilde{W}_d$ and $\widetilde{W}_f$ approximate the true value of W_d and W_f with increasing accuracy and, therefore, that the subspaces come into energy balance. Thus, either of the two curves, $\widetilde{e}_d$ or $\widetilde{e}_f$, can be used to choose P according to whether either quantity is smaller than a prescribed tolerance, which we denote as TOL_d and TOL_f, respectively.

We define the relative ROM-modeling errors in the displacement ($\widehat{w}$ Eq. (13)) and velocity ($\widehat{\dot{w}}$) fields generated by the ROMs as follows:

$$\widehat{e}_d = \frac{\langle \| w(x,t) - \widehat{w}(x,t) \| \rangle}{\langle \| w(x,t) \| \rangle} \tag{30a}$$

$$\widehat{e}_v = \frac{\langle \| \dot{w}(x,t) - \widehat{\dot{w}}(x,t) \| \rangle}{\langle \| \dot{w}(x,t) \| \rangle} \tag{30b}$$

We typically select some small values for TOL_d and TOL_f to ensure approximate energy balance. However, selecting too small a tolerance may lead to unnecessary inflation of the ROM dimension. It is therefore advisable to start with a reasonably small value of TOL_d and TOL_f (say, on the order of 10^{-2}) and examine the corresponding ROM errors. Should ROM errors be too high, the value of the tolerance may be systematically lowered until the desired level of model accuracy is achieved. Once this has happened, one can conclude that all relevant modes contributing to the energetics of the system have been included.

By using this process for the system parameters used in Fig. 3(*a*), we found that setting $\text{TOL}_d = \text{TOL}_f = 10^{-3}$ led to small modeling errors $\widehat{e}_d$ and $\widehat{e}_v$, on the order of 10^{-5} and 10^{-3}, respectively. By examining the intersection of the $\widehat{e}_d$ and $\widehat{e}_v$ curves with the horizontal dashed tolerance line in Fig. 3(*b*), we found that $P = 13$ was an appropriate subspace dimension for the ROM. In Fig. 3(*e*), we plot the ROM errors versus P, with the values of $\widehat{e}_d$ and $\widehat{e}_v$ at $P = 13$ indicated by filled large circles. An interesting, and potentially practically useful, feature of the ROM error can be observed: both curves display a "knee" at $P = 13$, with a steeper rate of decrease in the modeling error for $P < 13$ than for $P > 13$. This supports the determination that $P = 13$ is a minimal model of the system's behavior, in that adding additional degrees-of-freedom (DOFs) does not substantially increase the model's accuracy.

In contrast, by using the 99.9% variance-based criterion, we obtained $P = 2$: the resulting two-dimensional ROM was found to capture only ≈63% of the true energy input to the system, with modeling errors $\widehat{e}_d$ and $\widehat{e}_f$ as high as 3% and 11%, respectively, about three orders of magnitude higher than that for the 13-dimensional ROM, as is evident from Fig. 3(*e*). The modeling errors corresponding to the 2D ROM are shown in Fig. 3(*e*) by large square symbols.

The results discussed from Fig. 3(*e*) are also reflected in the phase plane plots of Fig. 3(*d*): the two-dimensional variance-based ROM solution differs significantly from the true solution, whereas the energy closure solution is virtually identical to it.

Effect of Measurement Noise

In physical experiments, in addition to having to account for the effect of discrete data, measurement noise will always be present. For this study, we modeled this source of measurement error as independent identically distributed (IID) Gaussian noise with zero mean and standard deviation σ. This noise was added to each element of the displacement data contained in W Eq. (7). Two noise amplitudes, specified as $\sigma = 0.1\%$ and 1% of the root-mean-square (RMS) beam displacement, were employed to study how noise-related errors propagate in calculations of $\widetilde{e}_d$ and $\widetilde{e}_f$ and, hence, in our ability to assess energy closure. Other than the added noise, for the results presented here, we kept the system parameters, external forcing, and boundary conditions unchanged from those used in the previous section.

In Fig. 4(*a*), we show an ensemble of single-period records of steady-state diplacement data, for the chosen parameter values, corresponding to the midpoint of the beam. The data in the figure are contaminated with 1% measurement noise. The single-period records, when stacked on top of each other, display a small amount of statistical variability (indicated by the orange band in the figure), corresponding to the normal distribution of noise around each true data point. Considering the displacement data as the primary measurement, velocity data were again estimated using spectral differentiation (Estimating the Velocity Field section). The result is shown in Fig. 4(*b*), where see that variability in the velocity estimates (again indicated by the orange band) is significantly larger than it is for the displacement time series (≈115% of the RMS velocity amplitude). Given that velocity field data are required to estimate the energetics of the system (via Eq. (21)), it is important to denoise the displacement data and prevent the propagation of its associated error in subsequent calculations as much as possible.

To this end, we applied synchronous averaging, a time-domain method that is useful for periodic, statistically stationary data. In physical experiments, data will typically be acquired for a large

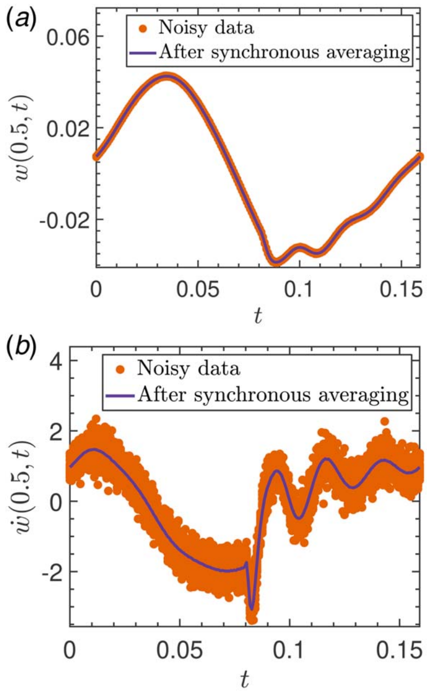

Fig. 4 Ensemble of single-period records of steady-state (*a*) displacement and (*b*) velocity time series data, stacked on top of each other, corresponding to the midpoint (*x* = 0.5) of the beam, contaminated with measurement noise (IID). The displacement data in (*a*) is acquired at a sampling rate of 200/*T* with noise amplitude $\sigma = 1\%$ of the RMS amplitude of the beam displacement. Each record of the velocity time series in (*b*) is estimated from each record of the displacement time series via spectral differentiation. In both (*a*) and (*b*), ensemble of data records is observed to form a noise band, shown using orange dots (lighter shade in black and white print), which corresponds to the normal distribution of the noise around each data point. Furthermore, in (*a*) and (*b*), using the purple curves, we show the ensemble average (synchronous) of the displacement and velocity time series data, which are used to calculate the energy balance error. The number of averages performed $N_{av} = 10^6$. Rest of the parameter values are same as in the previous figure.

number of periods, N_{av}, and an accompanying trigger pulse train synchronized with the data signal is used to mark individual single-period records. Averaging over all periods gradually eliminates additive measurement noise, which is not synchronized with the trigger, but emphasizes the underlying periodic signal. The signal-to-noise ratio increases as $\sqrt{N_{av}}$, indicating that the number of synchronous averages must be increased by two orders of magnitude to reduce the noise amplitude in the data by one.

The result of carrying out this process is illustrated by the purple curves in Figs. 4(*a*) and 4(*b*). The relative measurement errors for the averaged displacement and velocity field data are quite low, approximately 0.001% and 0.06%, respectively.

Temporal Upsampling. In the noise-free case, while acquiring data, one can use a sampling frequency well above the Nyquist frequency to achieve higher accuracy in the integrals needed for the energy balance calculations. However, if the data include measurement noise, a larger sampling rate leads to more noise in the data, which introduces larger error in subsequent calculations. To avoid this problem, the displacement data were initially acquired at a lower sampling frequency, just above the Nyquist frequency, in this case $2(90/T)$. Then, after noise reduction and spectral velocity estimation, the displacement and velocity data were upsampled temporally.

There are two advantages to upsampling. First, the consequent reduction in the effective time-step reduces the integration error during the numerical estimation of R Eq. (11), where trapezoidal integration is used to evaluate the time-averaging integral Eq. (11). This, in turn, increases the accuracy of the eigenvectors of R (the POMs). Second, in a similar fashion, the upsampled velocity field lowers the error in calculating W_d and W_f Eq. (21), which also involves numerical integration over the discretized spatial and temporal domains.

The first step toward upsampling the displacement or velocity data is to calculate the Fourier transform of the time series corresponding to each location of the beam. The number of points in the Fourier

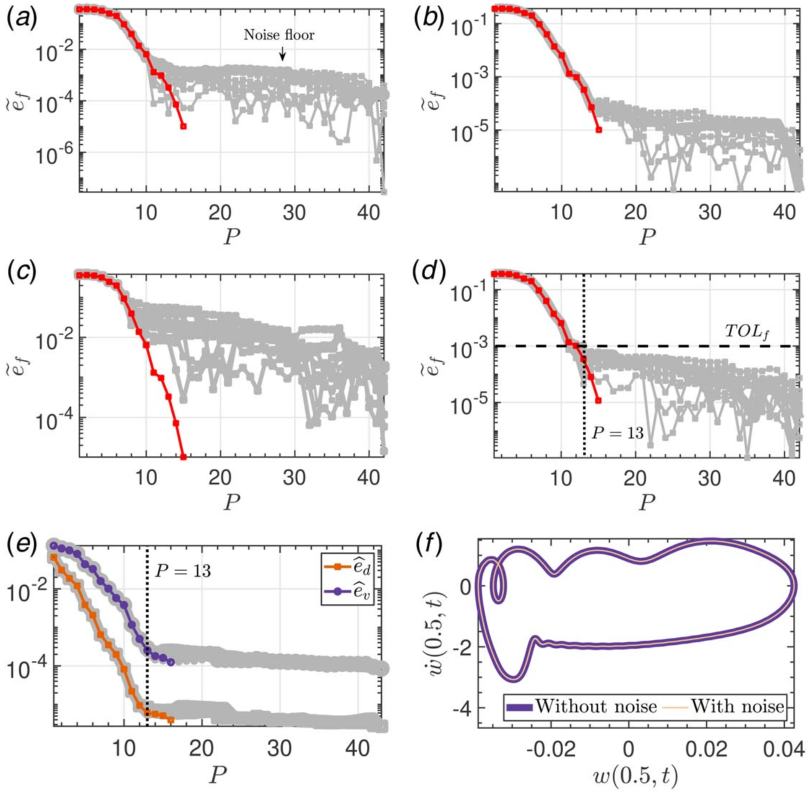

Fig. 5 The effect of the number of synchronous averages, performed for denoising displacement data, on the calculation of energy convergence error ($\tilde{e}_f$) and ROM errors ($\hat{e}_d$,$\hat{e}_v$). Except for the measurement noise amplitudes, all subfigures were computed with the same parameters as in the previous figures. In (*a*)–(*d*), using gray lines (lighter shade in black and white print), we show different $\tilde{e}_f$ plots generated from the data obtained from different simulated experiments (15, in each subfigure). The red curve (darker shade in black and white print) in each corresponds to the noise-free data and is used as a reference. The noise amplitude in (*a*) and (*b*) is *0.1*% of the RMS amplitudes of the beam displacement and the number of synchronous averaging (N_{av}) performed on the displacement data acquired from each experiment is 30 and 30, 000, respectively. In (*c*) and (*d*), the noise amplitude is *1*%, and N_{av} is 30 and 100, 000, respectively. In each case, the reduced subspace is considered to be close when $\tilde{e}_f < 10^{-3}$, and the smallest value *P* for which this occurs is taken as the ROM dimension, which in this case is $P = 13$. We plot the ROM errors for the case shown in (*a*) and (*b*), in (*e*), as a function of *P*. Like the other subfigures, the gray lines correspond to ROM errors from different experiments. (*f*), On the phase plane plot, we compare the solution from a 13-dimensional ROM, with that of the full-order system. The solution obtained from the ROM is significantly accurate, and the two phase plots are virtually indistinguishable. The ROM error plot and the phase plane plot for the *1*% noise were identical and are not shown here.

transform vector is then increased by zero-padding, and the inverse Fourier transform is taken to obtain the upsampled time series. As discussed in Ref. [40], this method works well with periodic signals. We upsampled the synchronously averaged displacement and velocity data by a factor of 15. These upsampled data sets were then used to calculate the projected energy errors $\widetilde{e}_f$ and $\widetilde{e}_d$ Eq. (29).

Selecting Reduced Order Model Dimension Using Energy Convergence. In Sec. 3, we discussed how either of the energy convergence errors $\widetilde{e}_f$ or $\widetilde{e}_d$ can, in principle, be used for selecting the ROM dimension P. However, at least for this system, with noisy data, $\widetilde{e}_d$ was observed to be quite sensitive to noise. We suspect that this is because $\widetilde{W}_d$ has multiple space and time derivatives that can result in substantial error. This required a relatively large number of synchronous averages N_{av} to attain a desired level of the energy convergence error $\widetilde{e}_d$. We observed that, in comparison, $\widetilde{e}_f$ required N_{av} to be one order of magnitude lower, which, in turn, leads to quicker data acquisition and lower data volume. We therefore used $\widetilde{e}_f$ for estimating the ROM dimension. While this was the best choice in this case, the relative sensitivity of $\widetilde{e}_d$ and $\widetilde{e}_f$ to measurement noise, and whether one is a better choice than the other, will likely depend on the exact forms of $\widetilde{W}_d$ and $\widetilde{W}_f$ Eq. (21) in any given problem.

Figure 5 shows how the number of averages performed for denoising affects the calculation of $\widetilde{e}_f$. Figures 5(*a*) and 5(*b*) correspond to a measurement noise amplitude $\sigma = 0.1\%$ of the RMS beam displacement, whereas Figs. 5(*c*) and 5(*d*) correspond to $\sigma = 1\%$. The remaining parameter values for all plots in Fig. 5 are the same as those used in Fig. 2.

In Fig. 5(*a*), $N_{av} = 30$, and in Fig. 5(*b*), $N_{av} = 30{,}000$: in both figures, the red line (darker shade in black and white print) corresponds to the noise-free data, which is included to validate our calculations with measurement noise. The gray lines (lighter shade in black and white print) in both figures show plots of $\widetilde{e}_f$ corresponding to data obtained from different simulated experimental runs. We see that the gray lines extend well beyond $P = 16$, which was the rank of the covariance matrix, $P_{\max}$, in the noise-free case of Fig. 3(*d*). This occurs because measurement noise dominates the lower eigenvalues of the covariance matrix R and changes its rank from 16 to N_Ω, which in this case is 43 (not counting the two pinned end points).

We also observe that the gray lines in Fig. 5(*a*) initially remain close to each other and tightly follow the red noise-free reference line; however, as the value of P goes beyond ≈10, the lines hit a noise floor. It is well known that the eigenvalue problem is not robust with respect to noise [41]: the variability in value of $\widetilde{e}_f$ for $10 < P < 17$ arises from the inaccuracy in the estimated eigenvalues, which are small in magnitude and are easily corrupted. Even worse, for $P > 16$, the true rank of the covariance matrix, all eigenvalues are spurious as they are supposed to be all equal to zero. Moreover, the velocity estimation process, as needed to calculate $\widetilde{e}_f$, amplifies the noise and further increases error in its estimate.

To properly identify a value of P, it is clear that the noise floor must remain below the required tolerance (TOL_f, in this case). In Fig. 5(*a*), we see that this is not the case: the noise floor is above our required tolerance of $\mathrm{TOL}_f = 10^{-3}$ (recall the results of Fig. 3(*b*)), so it is impossible to identify an appropriate value of P. We thus increased N_{av} from 30 to 30,000 to remove more noise in the data, as shown in Fig. 5(*b*), where it is now possible to identify $P = 13$ as the appropriate subspace dimension. Not surprisingly, finding P becomes more difficult when the noise amplitude is increased to 1% of the data's RMS amplitude. As shown in Fig. 5(*c*), the noise floor is substantially higher than that in Fig. 5(*a*), even though $N_{av} = 30$ for both, making it again impossible to estimate P to the necessary tolerance. In this case, it was necessary to set $N_{av} = 10^6$ to reduce the noise floor below $\mathrm{TOL}_f = 10^{-3}$, which yielded $P = 13$ as mentioned earlier, as shown in Fig. 5(*d*).

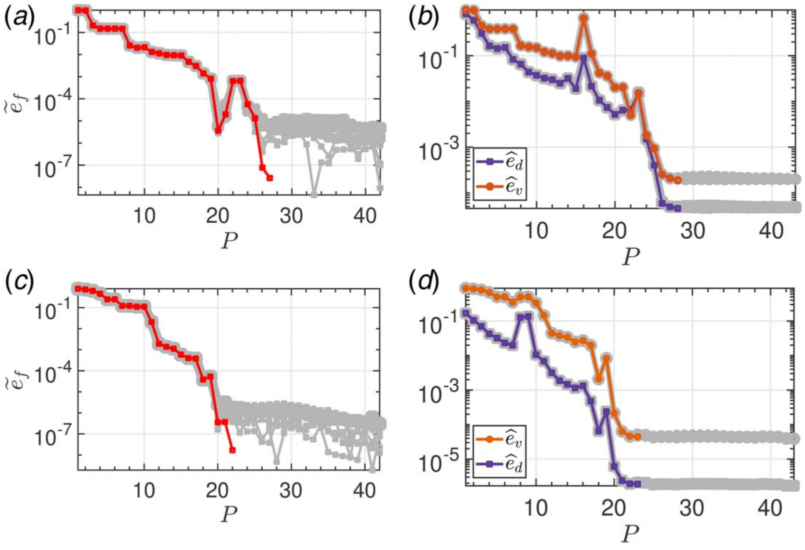

Fig. 6 Energy convergence error and ROM errors against different values of *P* for two cases, in which material damping was considered to be zero ($c_v = 1.000$, $c_m = 0$). The forcing frequency in (*a*) and (*b*) is subresonant ($T = 0.1$) and in (c) and (*d*), equals the second natural frequency ($T = 1/2\pi$) of the beam, with $\tau = 0.001$ and $\epsilon = 0.002$ in both cases. The noise amplitudes in both cases are *0.1*% of the respective RMS amplitudes of the beam displacements. In (*a*) (and c), like in the previous figure, the gray lines (lighter shade in black and white print) correspond to $\widetilde{e}_f$ curves obtained from different experiments, with $N_{av} = 10^4$ ($N_{av} = 10^6$), and the red line (darker shade in black and white print) correspond to the data without any noise. In (*b*) and (*d*), we show the modeling errors for each experiment, again using gray lines. The purple and orange curves correspond to noise-free data. In both cases, we found changing TOL_f to 10^{-4} from 10^{-3}, the value considered earlier, gave an accurate estimate of the ROM dimension: $P = 24$ and $P = 18$, respectively.

Having determined a good estimate for P, we constructed the ROM following the procedure outlined in Reduced Order Model section. In Fig. 5(*e*), we show the modeling errors associated with ROMs of different dimensions P. As shown in Fig. 3(*e*), we find that the ROM errors have a knee at the selected ROM dimension of $P = 13$, which is accentuated by the noise floor that extends across the figure for $P > 13$. Figure 5(*f*) displays the phase plane plot of the solution to the 13-dimensional ROM, obtained for the data with 0.1% noise amplitude, together with that of the full order system. As expected, they almost lie on top of each other, which demonstrates the accuracy of the ROM, for which $\widehat{e}_d$ and $\widehat{e}_f$ are on the order of 10^{-3} and 10^{-5}, respectively. The phase plane plot for the 1% noise case was found to be virtually indistinguishable, as well, so is not shown separately.

We tested the robustness of our procedure by generating simulated data for a number of other sets of parameter values. However, in contrast to the parameters used to generate Figs. 2, 3, and 5, the material damping coefficient was set to zero. Even though real structures always possess substantial material damping, our aim was to perform a "stress test" on our method for a system possessing only viscous damping, which is much weaker than material damping and can, therefore, be expected to result in higher-dimensional dynamics, all other things being equal. For these simulations, the noise amplitudes were set to 0.1% of the RMS amplitudes of their respective displacement fields. The results are shown in Fig. 6.

In Figs. 6(*a*) and 6(*b*), we considered a nonresonant dimensionless forcing frequency of 10. As in Fig. 5, the gray $\widetilde{e}_f$ curves were obtained from the averaged noisy data from different simulated experiments, whereas the red curve was obtained from the noise-free data and is included as a reference. As expected, in the absence of material damping, the number of DOFs excited in the beam were larger than found in Figs. 3 and 5 and, hence, more POMs needed to be included in the ROM. We found that in this case it was necessary to set $TOL_f = 10^{-4}$, in contrast to 10^{-3} as used in the previous cases. To achieve this tolerance level, it was necessary to set $N_{av} = 10^4$. This yielded a ROM dimension of $P = 24$. The modeling errors for different values of P are shown in Fig. 6(*b*): we observe that for $P = 24$, both the displacement and velocity errors for the ROM are on the order of 10^{-3}.

For the results of Figs. 6(*c*) and 6(*d*), we considered a forcing frequency equal to beam's second natural frequency and kept the damping the same as for Figs. 6(*a*) and 6(*b*). In this case, we also used $\mathrm{TOL}_f = 10^{-4}$; however, the number of synchronous averages needed in this case was $N_{av} = 10^6$. This resulted in a ROM dimension of $P = 18$, with associated displacement and velocity errors, as shown in Fig. 6(*d*), on the order of 10^{-4} and 10^{-3}, respectively.

Conclusion

This study builds upon the prior work in which we developed the energy closure criterion for estimating the dimension of ROMs obtained using the POD. In this article, we have shown that energy closure analysis can be applied to discrete, noisy data of the type gathered in simulations or experiments. In our previous work, the ROM dimension was selected by ensuring low balance error between the energy input and dissipated on a candidate subspace, whereas we here selected the dimension by setting a tolerance on the energy input to *or* dissipated from the original system, as captured by a given subspace. The steady-state dynamics guarantees that an accurate estimate of either quantity will automatically lead to a balance between the two. This new method for quantifying the degree of energy closure was found to be less sensitive to data discretization error and measurement noise while also being easier to interpret.

For this study, we used an Euler–Bernoulli beam with both viscous and material damping, which was subjected to a periodic impulsive load. We performed different simulations of the beam for different parameter values to generate steady-state discrete displacement field data (both with and without measurement noise) and applied POD on it to formulate reduced order models. We estimated the projected energy input ($\widetilde{W}_f$) on subspaces of varying dimension and set an energy input tolerance, which was then used to select an appropriate subspace to formulate a ROM. When the ROM errors, for the chosen tolerance, were not sufficiently low, we further lowered the tolerance, as needed, to achieve the desired level of model accuracy. While, in principle, we could have also estimated energy dissipated ($\widetilde{W}_d$) to accomplish the same goal, we found that, at least for this system, accurately estimating $\widetilde{W}_f$ required less noise removal and, therefore, less data. We suspect that the relative convenience of using either quantity will be system dependent. The important thing, however, is that either will work provided sufficient noise removal can be performed. Furthermore, in many cases, we suspect that careful experimentalists will want to check ROM dimension estimates made using the convergence of both quantities, for cross validation purposes.

We presented the results of applying our approach to data from three different simulations. In the first, the Euler–Bernoulli beam was subjected to both material damping and viscous damping and was excited periodically at its second natural frequency. We first implemented our method on the noise-free data set and obtained a smoothly decaying energy input curve as a function of the subspace dimension. From this curve, by using an energy input tolerance of $TOL_f = 10^{-3}$, we obtained a subspace dimension of 13, which gave a ROM with relative displacement and velocity errors on the order of 10^{-5} and 10^{-3}, respectively.

For the same parameters, but with the data subjected to measurement noise, we observed two distinct regimes of the resulting energy input convergence curve: a smooth downward trend, followed by a noise floor. The level of the noise floor became apparent after a few trials and was typically higher than the energy tolerance required for an accurate ROM. Thus, efficient noise removal was necessary to lower the noise floor below the desired tolerance so that the ROM dimension could be properly identified. To this end, we used synchronous averaging to denoise the data. We considered two different noise amplitudes in the displacement data, 0.1% and 1% of the RMS amplitude of the displacement, respectively, and found that 3×10^4 and 10^6 synchronous averages, respectively, were necessary to reveal the ROM dimension of 13 from the energy input convergence plot.

We performed two more studies where the beam was excited with resonant and nonresonant forcing frequencies, but in both cases subjected to only viscous damping, which is much weaker than material damping. For these viscous damping cases, the energy input tolerance needed to be set to $\mathrm{TOL}_f = 10^{-4}$, an order of magnitude lower than needed for the system with material damping, to achieve the desired level of model accuracy. Energy closure analysis provides a physical explanation for why this significantly lower tolerance was required: with viscous damping alone, the response of the system not only has more excited degrees-of-freedom, but it has energetically (and therefore dynamically) important modes with much smaller nonzero amplitudes than those observed with material damping. Such modes, even though small, must be included to obtain an accurate representation of the energy input (or dissipated) from the system and, therefore, to obtain the correct ROM dimension.

This study addressed the key challenges of data discretization and measurement noise to the application of energy closure analysis. While the model problem considered here was an impulsively driven Euler–Bernoulli beam, our results provide insights into the data-processing requirements for successfully implementing the energy closure criterion to structural vibration problems *in general*. In addition, the fact that the dynamics of this specific system could be obtained analytically (albeit, in this case, with the help of symbolic algebra) has no direct connection with the data-processing requirements or, for that matter, the idea of energy closure itself, which is agnostic to the exact nature of system, provided the system is at the steady state. Our approach, however, is model based in that it assumes the end-user has appropriate expressions for the relevant work quantities used to estimate energy closure.

There is scope for exploring other challenges. In actual experiments, the effect of sensor accuracy, resolution, and placement can be examined. Furthermore, the model system used for this study was linear, and we focused on using data only from its steady-state periodic responses. Provided the responses of interest are still periodic, the application of our approach to nonlinear systems should, in principle, be fairly straightforward. How to effectively carry out energy closure analysis for chaotic systems, systems with process noise (whether linear or nonlinear), or other systems exhibiting nonperiodic dynamical behaviors, will have to be studied in future efforts. However, our work demonstrates that energy closure analysis is a practical and systematic approach that can be applied in experimental contexts and may prove to be quite useful in formulating ROMs to accurately simulate the dynamics of complicated structures such as drill strings, gear boxes, structures with clearances or other piecewise components, or to simulate fluid–structure interaction problems for aircraft wings and wind turbines.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The authors attest that all data for this study are included in the paper.

References

[1] Holmes, P., Lumley, J. L., and Berkooz, G., 1996, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, UK.

[2] Bhattacharyya, S., and Cusumano, J. P., 2020, "An Energy Closure Criterion for Model Reduction of a Kicked Euler–Bernoulli Beam," ASME J. Vib. Acoust., **143**(4), p. 041001.

[3] Feeny, B. F., and Kappagantu, R., 1998, "On the Physical Interpretation of Proper Orthogonal Modes in Vibrations," J. Sound. Vib., **211**(4), pp. 607–616.

[4] Brunton, S. L., and Kutz, J. N., 2019, *Data-Driven Science and Engineering*, Cambridge University Press, Cambridge, UK.

[5] Liang, Y., Lin, W., Lee, H., Lim, S., Lee, K., and Sun, H., 2002, "Proper Orthogonal Decomposition and Its Applications—Part II: Model Reduction for MEMS Dynamical Analysis," J. Sound. Vib., **256**(3), pp. 515–532.

[6] Liang, Y., Lin, W., Lee, H., Lim, S., Lee, K., and Sun, H., 2002, "Proper Orthogonal Decomposition and Its Applications—Part II: Model Reduction for MEMS Dynamical Analysis," J. Sound. Vib., **256**(3), pp. 515–532.

[7] Cusumano, J. P., Sharkady, M. T., and Kimble, B. W., 1994, "Experimental Measurements of Dimensionality and Spatial Coherence in the Dynamics of a Flexible-Beam Impact Oscillator," Philos. Trans.: Phys. Sci. Eng., **347**(1683), pp. 421–438.

[8] Cusumano, J. P., and Bai, B. Y., 1993, "Period-Infinity Periodic Motions, Chaos, and Spatial Coherence in a 10 deg of Freedom Impact Oscillator," Chaos, Solitons Fractals, **3**(5), pp. 515–535.

[9] Sirisup, S., and Karniadakis, G. E., 2004, "A Spectral Viscosity Method for Correcting the Long-Term Behavior of Pod Models," J. Comput. Phys., **194**(1), pp. 92–116.

[10] Cazemier, W., Verstappen, R., and Veldman, A., 1998, "Proper Orthogonal Decomposition and Low-Dimensional Models for Driven Cavity Flows," Phys. Fluids., **10**(7), pp. 1685–1699.

[11] Couplet, M., Sagaut, P., and Basdevant, C., 2003, "Intermodal Energy Transfers in a Proper Orthogonal Decomposition–Galerkin Representation of a Turbulent Separated Flow," J. Fluid. Mech., **491**, pp. 275–284.

[12] Bergmann, M., Bruneau, C.-H., and Iollo, A., 2009, *Computational Fluid Dynamics 2008*, Springer, Berlin/Heidelberg, pp. 779–784.

[13] Bergmann, M., Bruneau, C.-H., and Iollo, A., 2009, "Enablers for Robust Pod Models," J. Comput. Phys., **228**(2), pp. 516–538.

[14] Borggaard, J., Iliescu, T., and Wang, Z., 2011, "Artificial Viscosity Proper Orthogonal Decomposition," Math. Comput. Model., **53**(1–2), pp. 269–279.

[15] Aubry, N., Holmes, P., Lumley, J. L., and Stone, E., 1988, "The Dynamics of Coherent Structures in the Wall Region of a Turbulent Boundary Layer," J. Fluid. Mech., **192**, pp. 115–173.

[16] Everson, R., and Sirovich, L., 1995, "Karhunen–Loeve Procedure for Gappy Data," JOSA A, **12**(8), pp. 1657–1664.

[17] Ravindran, S. S., 2000, "A Reduced-Order Approach for Optimal Control of Fluids Using Proper Orthogonal Decomposition," Inter. J. Numer. Methods Fluids, **34**(5), pp. 425–448.

[18] Amsallem, D., and Farhat, C., 2008, "Interpolation Method for Adapting Reduced-Order Models and Application to Aeroelasticity," AIAA. J., **46**(7), pp. 1803–1813.

[19] Peherstorfer, B., and Willcox, K., 2016, "Data-Driven Operator Inference for Nonintrusive Projection-Based Model Reduction," Comput. Methods. Appl. Mech. Eng., **306**, pp. 196–215.

[20] San, O., and Iliescu, T., 2014, "Proper Orthogonal Decomposition Closure Models for Fluid Flows: Burgers Equation," Inter. J. Numer. Anal. Model. Seri. B, **5**, pp. 285–305.

[21] Ahmed, S. E., Pawar, S., San, O., Rasheed, A., Iliescu, T., and Noack, B. R., 2021, "On Closures for Reduced Order Models—A Spectrum of First-Principle to Machine-Learned Avenues," Phys. Fluids., **33**(9), p. 091301.

[22] Rowley, C. W., 2005, "Model Reduction for Fluids, Using Balanced Proper Orthogonal Decomposition," Inter. J. Bifurcat. Chaos, **15**(3), pp. 997–1013.

[23] Willcox, K., and Peraire, J., 2002, "Balanced Model Reduction Via the Proper Orthogonal Decomposition," AIAA. J., **40**(11), pp. 2323–2330.

[24] Segala, D. B., and Chelidze, D., 2014, "Robust and Dynamically Consistent Model Order Reduction for Nonlinear Dynamic Systems," J. Dyn. Syst. Meas. Control., **137**(2), p. 021011.

[25] Chelidze, D., 2014, "Identifying Robust Subspaces for Dynamically Consistent Reduced-Order Models," Nonlinear Dynamics, Volume 2. Conference Proceedings of the Society for Experimental Mechanics Series, G. Kerschen, ed., Springer, Cham, pp. 123–130.

[26] Ilbeigi, S., and Chelidze, D., 2018, "A New Approach to Model Reduction of Nonlinear Control Systems Using Smooth Orthogonal Decomposition," J. Robust. Nonlinear. Control., **28**(15), pp. 4367–4381.

[27] Ilbeigi, S., and Chelidze, D., 2017, "Persistent Model Order Reduction for Complex Dynamical Systems Using Smooth Orthogonal Decomposition," Mech. Syst. Signal. Process., **96**, pp. 125–138.

[28] Guo, X., and Przekop, A., 2010, "Energy-Based Modal Basis Selection Procedure for Reduced-Order Nonlinear Simulation," 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Orlando, FL, Apr. 12–15, p. 2796.

[29] Chatterjee, A., Cusumano, J. P., and Chelidze, D., 2002, "Optimal Tracking of Parameter Drift in a Chaotic System: Experiment and Theory," J. Sound. Vib., **250**(5), pp. 877–901.

[30] Chelidze, D., and Zhou, W., 2006, "Smooth Orthogonal Decomposition-Based Vibration Mode Identification," J. Sound. Vib., **292**(3–5), pp. 461–473.

[31] Balajewicz, M. J., Dowell, E. H., and Noack, B. R., 2013, "Low-Dimensional Modelling of High-Reynolds-Number Shear Flows Incorporating Constraints From the Navier–Stokes Equation," J. Fluid. Mech., **729**, pp. 285–308.

[32] Balajewicz, M., Tezaur, I., and Dowell, E., 2016, "Minimal Subspace Rotation on the Stiefel Manifold for Stabilization and Enhancement of Projection-Based Reduced Order Models for the Compressible Navier–Stokes Equations," J. Comput. Phys., **321**, pp. 224–241.

[33] Banks, H. T., and Inman, D. J., 1991, "On Damping Mechanisms in Beams," J. Appl. Mech., **58**(3), pp. 716–723.

[34] Blevins, R. D., 2016, *Formulas for Dynamics, Acoustics and Vibration*, John Wiley and Sons Inc., Chichester, West Sussex, Hoboken, NY.

[35] Meirovitch, L., 2001, *Fundamentals of Vibrations*, McGraw-Hill, Boston, MA.

[36] Sirovich, L., 1987, "Turbulence and the Dynamics of Coherent Structures Part I: Coherent Structures," Q. Appl. Math., **45**(3), pp. 561–571.

[37] Penrose, R., 1955, "A Generalized Inverse for Matrices," Math. Proc. Cambridge Philos. Soc., **51**(3), pp. 406–413.

[38] Lomax, H., Pulliam, T. H., and Zingg, D. W., 2001, *Fundamentals of Computational Fluid Dynamics*, Springer-Verlag, Berlin/Heidelberg.

[39] Abramowitz, M., and Stegun, I. A., 1970, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, Vol. 55, US Government Printing Office, Washington, DC.

[40] Lyons, R. G., 2010, *Understanding Digital Signal Processing*, 3rd ed., Pearson, Upper Saddle River, NJ.

[41] Golub, G. H., and Van Loan, C. F., 1983, *Matrix Computations*, Vol. 3, Johns Hopkins University Press, Baltimore, MA.