Physica D 239 (2010) 2039–2049

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Dimensional implications of dynamical data on manifolds to empirical KL analysis

Erik M. Bollt [a,*], Chen Yao [a], Ira B. Schwartz [b]

[a] Department of Mathematics & Computer Science, Clarkson University, Potsdam, NY 13699-5815, United States

[b] Naval Research Laboratory, Plasma Physics Division, Nonlinear Dynamics System Section, Code 6792, Washington, DC 20375, United States

ARTICLE INFO

Article history:
Received 31 March 2008
Received in revised form
13 July 2009
Accepted 16 July 2010
Available online 21 August 2010
Communicated by I. Mezic

Keywords:
Dimension reduction
POD
Empirical KL analysis
Slow manifold
Singular perturbation

ABSTRACT

We explore the approximation of attracting manifolds of complex systems using dimension reducing methods. Complex systems having high-dimensional dynamics typically are initially analyzed by exploring techniques to reduce the dimension. Linear techniques, such as Galerkin projection methods, and nonlinear techniques, such as center manifold reduction are just some of the examples used to approximate the manifolds on which the attractors lie. In general, if the manifold is not highly curved, then both linear and nonlinear methods approximate the surface well. However, if the manifold curvature changes significantly with respect to parametric variations, then linear techniques may fail to give an accurate model of the manifold. This may not be a surprise in itself, but it is a fact so often overlooked or misunderstood when utilizing the popular KL method, that we offer this explicit study of the effects and consequences. Here we show that certain dimensions defined by linear methods are highly sensitive when modeled in situations where the attracting manifolds have large parametric curvature. Specifically, we show how manifold curvature mediates the dimension when using a linear basis set as a model. We punctuate our results with the definition of what we call, a "curvature induced parameter," d_{CI}. Both finite- and infinite-dimensional models are used to illustrate the theory.

1. Introduction

When considering a dynamical system with complex dynamics, one of the central problems in its analysis is first attempting to reduce the dimension of the attractor. For a given model with sufficient dissipation, there exists constructive methods for dimension reduction, such as a center manifold analysis and singular perturbation theory. For problems consisting of data generated from experimental or physical experiments, the techniques are fewer but still exist.

One very popular method adapted from the probability and statistics communities is that of principal component analysis (POD), which also goes by the name of Karhunen–Loeve (KL) analysis, among others. (See the very nice text [1] and the references therein.) KL methods have been applied to construct optimal basis functions which minimize the error in an L_2 norm, and also minimize the entropy [2]. The technique has been valuable in approximating the dynamics and data from many fields such as turbulence [3], sea surface temperatures and weather prediction [4], the visual system [5], facial detection and classification [6], and even analyzing voting patterns of the supreme court [7]. Since KL forms a complete orthonormal basis from the model or data, a finite-dimensional projection of the dynamical system or data set can be done with a truncated set of modes using a Galerkin type of expansion [8]. For classifying complexity, the spectrum is a direct measure of the variance of each mode, and can be used to compute the entropy of the system [2].

However, given the potential power of the KL technique for dimension reduction, a fundamental problem with the use of KL modes applied to dynamical systems [9–12] is that KL analysis, often called POD analysis, is fundamentally a linear analysis. Given a data set of high-dimensional randomly distributed data points, principle component analysis gives the principle axis of the time-averaged covariance matrix. That is, it treats data as an ellipsoidal cloud, and yields the major and minor axes. Details will be reviewed in Section 3. The aim of this paper is to remind explicitly how this linear point of view may not be appropriate for all of the many ways in which POD is applied to data collected from the evolution of dynamical data toward an underlying global attractor.

Since KL analysis is so widely used to reduce the dimension of high-dimensional and complicated models of evolution laws and dynamical systems, it is important to understand exactly what functions such an analysis does well, and what are its shortcomings. This paper is meant to understand better what KL

* Corresponding author. Tel.: +1 315 268 2307.
E-mail address: bolltem@clarkson.edu (E.M. Bollt).

0167-2789/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2010.07.004

analysis can do usefully with regard to dimension reduction, and how its nonperformance sometimes leads to misleading results. The problem is that the linear analysis is in some sense ill-equipped to describe the nonlinear manifold embedding a global attractor, but it can nonetheless be useful for approximating the evolution of the dynamical system in the short run, by a low-dimensional model. Specifically, we will show how the KL analysis misleads the choice of dimension due to simple scaling of some dynamical variables, in the case of a specific class of systems with a well understood stable invariant manifold. We will show how such systems can lead to errors of embedding dimension with topological errors, as well as numerical estimation errors; a well used modeling technique should be insensitive to such change of variables. We will punctuate our results by introduction of a definition which we call, a "curvature induced parameter", d_{CI}.

We will display our points regarding curvature of the slow manifold, and the corresponding KL dimension in terms of several singularly perturbed systems of increasing complexity, including a PDE system of a rod coupled to a pendulum. We wish to note that these KL results are complementary but contrast with any embedding results which may be derived by Takens' embedding theorem [13]. Takens' embedding theorem also has relevant and important implications regarding representation of a dynamical system with as few coordinates as possible, and this is a goal of any Galerkin method. The two goals sought may be: (1) to produce minimal coordinates sufficient for good simulation/predictions of the higher-dimensional system, or, (2) to faithfully (and even more specifically—diffeomorphically) represent the dynamical system in an alternative coordinate system for understanding the topological nature and perhaps bifurcation structure of the original system. Generally, here, we will be interested in the former. Takens' embedding theorem relies on time-delay sampling of a measurable function of a single scalar time series from the dynamical system, or combinations of several measurements in various modern variations of the theorem. In its original form, the theorem states that a delay vector of dimension $2d + 1$ is a sufficient embedding of the dynamics on the underlying manifold, the $2d+1$ coming from the Whitney embedding theorem from which Takens' theorem is derived. In the case of a PDE, for example, the scalar measurement could be the time varying Fourier coefficient from some (dominant) mode in some general basis. If the dimension d of this slow manifold is large, then the embedding may be large. However, the goal of the embedding analysis is of exact representation of the dynamics in terms of an alternative coordinate system, the time-delay coordinates, where the measurement is generally considered to be the same in terms of diffeomorphism. This analysis is in contrast to the KL analysis, which is a modal analysis whose goal is representation in terms of quality of prediction generally, and the modes are defined in terms of time-average optimal representation over a sampled time period. For a large enough truncation, we should expect that the two representations will become diffeomorphically equivalent, although this does not normally enter into the discussion of KL analysis, and it will not be explicitly described here.

2. Fast–slow systems as a model for stable invariant manifolds

In this section, we will briefly review the part of standard singular perturbation theory [14,15] necessary for our discussion, and then introduce our special restricted form and model problem. A general system with two distinct time scales is the following standard [14,15] fast–slow, or singularly perturbed system,

$$\begin{aligned} \dot{x} &= F(x, y), \\ \epsilon \dot{y} &= G(x, y) \end{aligned} \tag{1}$$

where $x \in \Re^m, y \in \Re^n, F : \Re^m \times \Re^n \to \Re^m$, and $G : \Re^m \times \Re^n \to \Re^n$. It is easy to see that for $0 < \epsilon \ll 1$, the $y(t)$-equation runs fast, relative to the slow dynamics of the first equation for evolution of $x(t)$. Such systems are called singularly perturbed, since if $\epsilon = 0$ we get a differential–algebraic equation

$$\begin{aligned} \dot{x} &= F(x, y), \\ G(x, y) &= 0. \end{aligned} \tag{2}$$

The second ODE becomes an algebraic constraint.

Under sufficient smoothness assumptions on the functions F and G so that the implicit function theorem can be applied in form of the Tikhonov theorem, [16], there is a function, or ϵ slow manifold,

$$y = h_\epsilon(x), \tag{3}$$

such that,

$$G(x, h_\epsilon(x)) = 0, \tag{4}$$

for a local neighborhood about $\epsilon = 0$. The singular perturbation theory concerns itself with continuation and persistence of stability of this manifold $h_\epsilon(x)$ within $O(\epsilon)$ of $h_\epsilon(x)|_{\epsilon=0}$, for $0 < \epsilon \ll 1$ and possibly even for larger ϵ.

To motivate our problem, we will concern ourselves with a special case of fast–slow systems with one way coupling in the special form,

$$\begin{aligned} \dot{x} &= f(x), \\ \epsilon \dot{y} &= y - \alpha g(x). \end{aligned} \tag{5}$$

For reasons of studying manifolds with relevant curvature, we shall assume that $g(x)$ is $o(|x|)$. Given an equation of this form, it is immediate that we can write the $\epsilon = 0$ slow manifold in the closed form,

$$h(x)|_{\epsilon=0} = \alpha g(x). \tag{6}$$

Eq. (6) gives us freedom to use this system to deliberately design a slow manifold with curvature properties which we use for comparisons between the nonlinear nature of curvature to the linear properties selected by POD. Note that our inclusion of the α-parameter is an explicit control over curvature of the slow manifold.

As an explicit example, consider a Duffing oscillator evolving in the x-variables, contracting transversally onto a slow manifold specified as a paraboloid in the y-variables, graphed over the slow variables,

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \sin(x_3) - a x_2 - x_1^3 + x_1, \\ \dot{x}_3 &= 1, \\ \epsilon \dot{y} &= y - \alpha(x_1^2 + x_2^2). \end{aligned} \tag{7}$$

If we choose, $a = 0.02$, $b = 3$, $\alpha = 1$, and $\epsilon = 0.001$, we get the chaotic data set shown projected onto a paraboloid, as in Fig. 1.

As an example application of KL analysis to expose its strengths and shortcomings, we take the data from Eq. (7),

$$\mathbf{z}(t_i) = \langle x_1(t_i), x_2(t_i), y(t_i) \rangle, \tag{8}$$

which is a $3 \times n$ matrix, shown in Fig. 1, as a parameterized curve in $\Re^3$. Also shown on the plane $y = 0$, in red is the Duffing oscillator data of the **x**-component.

Examination of the singular value spectrum, and large spectral splitting thereof, of the time-averaged covariance matrix is the usual basis for deciding a KL projection dimension [9–12]. More precisely, the KL dimension may be defined as the minimum of KL modes which approximates the dynamic variance to within a prescribed threshold, usually 95%. We show in Fig. 2 how the 3 eigenvalues of this simple example change with respect to α. We

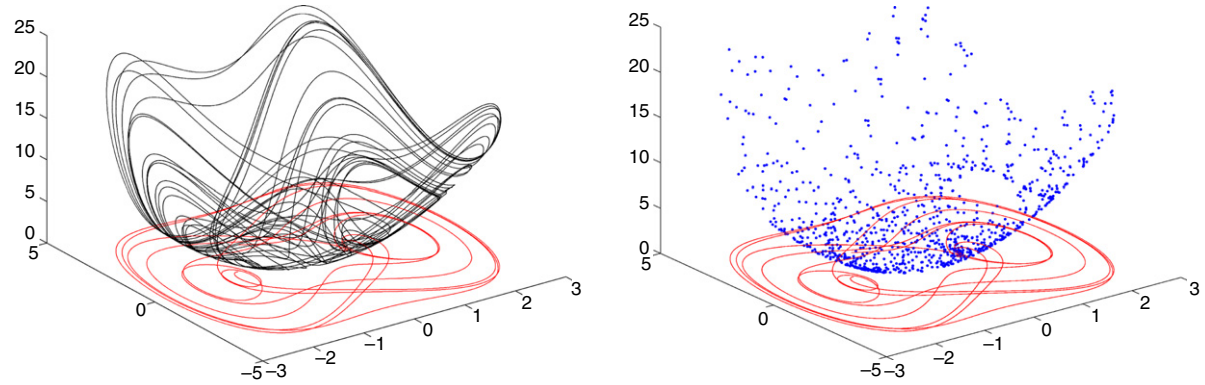

Fig. 1. A fast–slow Duffing oscillator on a paraboloid attracting submanifold, according to the singularly perturbed Eqs. (7). On the left a typical trajectory and its projection onto $x - y$, which is the familiar Duffing oscillator is shown. Right is a uniform sampling of the flow, which yields the dots on the paraboloid, which would be a typical data set to be processed by a KL method for learning the dimension reduction.

will review the calculation in the next section, but for now, note that the key point is the possible presence of a spectral gap, which we define to be,

$$n : \frac{\lambda_{d+1} - \lambda_d}{\lambda_d} > p \tag{9}$$

for some large criterion p. In practice, what is often used instead is a criterion that d is the first value such that the first n-modes capture $100q\%$ of the variance, stated,

$$d : \frac{\sum_{i=1}^{d} \lambda_i}{\sum_{i=1}^{N} \lambda_i} > q, \quad \text{but} \quad \frac{\sum_{i=1}^{d} \lambda_i}{\sum_{i=1}^{N} \lambda_i} < q, \quad d \leq N. \tag{10}$$

As shown in Fig. 2, we see that there are three regions in which we would interpret that $d = 1$, 2, or 3. In other words, all possible values could be validly concluded, depending on how α is chosen. It is easy to see that α can be controlled by scaling the variable y as follows. Let,

$$Y = sy, \tag{11}$$

then by substitution, it follows that Eqs. (5) become, as exemplified by Eqs. (7)

$$\begin{aligned} \dot{x} &= f(x), \\ \epsilon \dot{Y} &= Y - \alpha s g(x), \end{aligned} \tag{12}$$

written in terms of the new spacial dimension Y.

Emphasizing a major point of this work, we consider it to be an undesirable property, for many applications, for the value of the dimension of reduction to depend on the particular choice of units on the y-variable, say in cm if it were length, versus Y say in m. Therefore, given the wide-spread acceptance and use of the KL method in dynamical systems, we hope the we can offer a better understanding of this issue. It is our goal in the rest of this paper to better understand the effect of such dimension reductions, when they are appropriate, and when they are not. We will give analytic bounds, and also several applications to indicate the generality of the situation. We will argue that Eq. (5) represents a typical form for such behavior.

3. Review of KL analysis as a model reduction technique

Karhunen–Loeve (KL) modes [9,10], also known as empirical mode reduction and also principal component analysis (PCA), as well as proper orthogonal decomposition (POD), was first applied to spatiotemporal analysis by Lorenz [17] for weather prediction. Later Lumley [18] brought the technique to the study of fluid

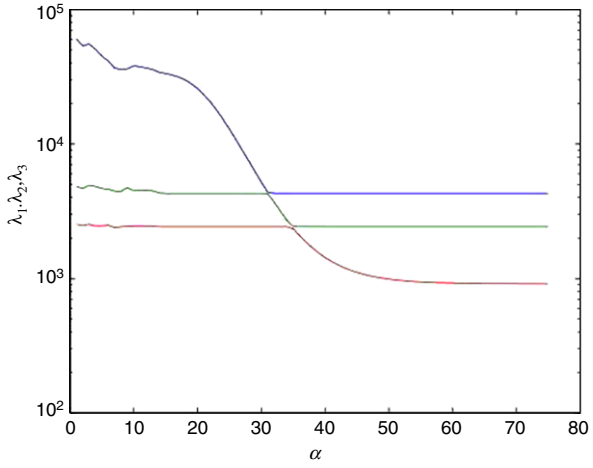

Fig. 2. Singular spectrum of time-averaged covariance matrix from the Duffing oscillator on paraboloid data from Eq. (7). α (horizontal) versus $\lambda_1 > \lambda_2 > \lambda_3$, singular eigenvalues. As α is varied, corresponding to a change of scale of the y-variable, as described by Eqs. (11)–(12), the embedding manifold's curvature is varied: the embedding paraboloid evolves from short and flat to tall and skinny, and thus according to the theory in Section 4, eigenvalues vary through three-dimensional regimes. In Region 1, when $\alpha < 20$, $\lambda_1 \gg \lambda_2, \lambda_3$, and the KL analysis concludes that the system is $n = 1$ dimensional. In Region 2, when $30 < \alpha < 40$, $\lambda_1 \sim \lambda_2, \lambda_3$ and we conclude a reduced model of dimension $m+n = 3$. In Region 3, when $\alpha > 50$, $\lambda_2, \lambda_3 > \lambda_1$, and we conclude a reduced model of dimension $m = 2$.

turbulence, as described in book [11]. The idea is that empirical modes form the basis which minimizes the L_2 error at any finite truncation. That is, we wish to maximize variance and minimize covariance at each finite truncation, which is a well known property of PCA [19].

The procedure requires a spatiotemporal pattern, such as a PDE solution, $u(x, t)$, sampled on a spatial grid in x, and in time t: $\{u^n(x)\} = \{u(x, t_n)\}_{n=1,M}$, from which the spatial mean has been subtracted. Then the KL modes are the eigenfunctions $\Psi_n(x)$ of the time-averaged covariance matrix,

$$K(x, x') = \langle u(x, t_n) u(x', t_n) \rangle, \tag{13}$$

which may be arrived at by a singular value decomposition [19]. Then u may be expanded in the resulting orthogonal basis,

$$u(x, t) = \sum_n a_n(t) \psi_n(x), \tag{14}$$

and this is the optimal basis in the sense of *time-averaged* projection:

$$\max_{\psi \in L^2(D)} \frac{\langle |(u, \psi)| \rangle}{\|\psi\|}, \tag{15}$$

[11], where $\langle . \rangle$ denotes time average. These functions are orthogonal in time, meaning in terms of time averaging,

$$\langle a_n(t) a_m(t) \rangle = \lambda_n \delta_{nm}, \tag{16}$$

in terms of eigenvalues of,

$$K : \lambda_n = \frac{(\psi_n, K\psi_n)}{\|\psi_n\|}. \tag{17}$$

Thus, the time varying Fourier coefficients $a_n(t)$ are decorrelated *in time average*. A computationally important approach [12] to solve this eigenvalue problem involves successive computation to maximize mean square variance. Formal substitution of a finite expansion of empirical modes $u(x,t) = \sum_n a_n(t)\psi_n(x)$ into the PDE, and then projection onto each basis element $\psi_m(x)$ produces an ODE which is expected to be a maximal variance model of the PDE. We give a continuum structure model of this behavior in Section 7.

In the next section, we discuss how the statistical geometry of the data samples justifies the dimension reductions which fall possibly into three distinct regimes depending upon the curvature of the slow manifold. This is an often overlooked truth of KL analysis which we highlight in this paper.

4. Statistical geometry justifying dimension reduction

The data set, $[u(x_i, t_j)]_{i=1,\ldots,N, j=1,\ldots,M}$ represents (treated as if random) M sample points in an N-dimensional space. In this interpretation, we have a data cloud. The time-averaged covariance matrix, Eq. (13), $K(x,x') = \langle u(x,t_n)u(x',t_n)\rangle$ has eigenvalues which can be interpreted as follows. If the data were distributed as an ellipsoid, with long major axis, and small minor axis, then the eigenvalues of K represent relative lengths of the eigenvectors of orthogonal (decorrelated) directions. This is standard within the POD theory, and it is straightforward to see that the spectral decomposition of the matrix K into a linear combination of rank-one operators $\Psi_n \otimes \Psi_n$ follows the spectral decomposition theorem in the case K is of finite rank [19], and Mercer's theorem [11,20] in the case of infinite rank, since it is straightforward to show that such covariance matrices are positive semidefinite and symmetric.

We will now compare explicitly these statements motivated by the POD theory to the reality of what we observed in the simple dynamical systems with the stable nonlinear invariant manifold, of Section 2.

In general, a zero-mean vector random variable $\mathbf{Z}$ has a covariance,

$$cov(\mathbf{Z}) = E[\mathbf{Z}\mathbf{Z}'], \tag{18}$$

and we require a diagonalizing orthogonal similarity transformation P, such that,

$$\mathbf{Y} = P'\mathbf{Z}, \tag{19}$$

and Y has a diagonal covariance matrix,

$$\begin{aligned} cov(\mathbf{Y}) &= E[\mathbf{Y}\mathbf{Y}'] = E[P'\mathbf{Z}\mathbf{Z}'P] \\ &= P'E[\mathbf{Z}\mathbf{Z}']P = P'cov[\mathbf{Z}]P \\ &= diag[\rho_1, \ldots, \rho_N]. \end{aligned} \tag{20}$$

Consider the following model example:

Example 1 (*Exact POD of a Bounding Box*)**.** Let,

$$\mathbf{Z} = U(B), \tag{21}$$

a uniform random variable over B, where B is a two-dimensional rectangle of sides $H \times L$. Thus, we may proceed to perform the POD in closed form for this simple example. In general, let $[\mathbf{z}]_i$ be the ith component of $\mathbf{z}$. Then the demeaned covariance matrix is,

$$C_{i,j} = \int_{\Re^2} ([\mathbf{z}]_i - \overline{[\mathbf{z}]_i})([\mathbf{z}]_j - \overline{[\mathbf{z}]_j})\chi_B(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{22}$$

where $\chi_B(\mathbf{z}) = 1$ if $\mathbf{z} \in B$, and 0 otherwise, is an indicator function representing the uniform random variable. In the case that $1 \le i, j \le 2$,

$$C_{i,j} = \int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{-\frac{L}{2}}^{\frac{L}{2}} ([\mathbf{z}]_i - \overline{[\mathbf{z}]_i})([\mathbf{z}]_j - \overline{[\mathbf{z}]_j})\mathrm{d}[\mathbf{z}]_i \mathrm{d}[\mathbf{z}]_j, \tag{23}$$

where $\overline{[\mathbf{z}]_i} = \int_{\Re^2} \mathbf{z}_i \chi_B(\mathbf{z})\mathrm{d}\mathbf{z}$ is the ith mean, from which we compute the eigenvalues,

$$\rho_{1,2} = \left\{ \frac{H^2}{12}, \frac{L^2}{12} \right\}. \tag{24}$$

Hence, the ratio of eigenvalues is simply,

$$r = \frac{H^2}{L^2}. \tag{25}$$

Likewise, it is straightforward and similar to show that the eigenvalues of the covariance matrix of a uniform random variable over an $L \times H \times W$ three-dimensional box are,

$$\rho_{1,2,3} = \left\{ \frac{H^2}{12}, \frac{L^2}{12}, \frac{W^2}{12} \right\}. \tag{26}$$

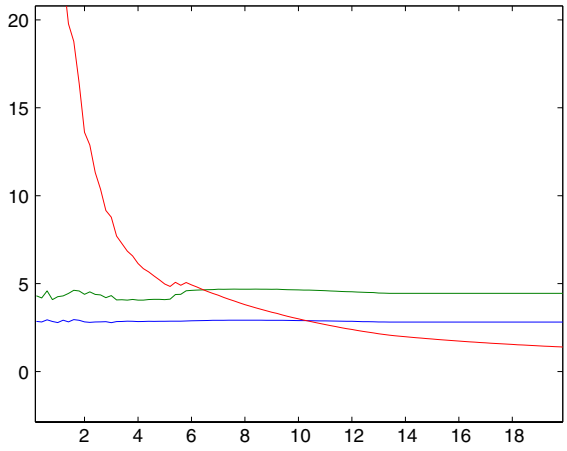

Fig. 3. Eigenvalues of the uniform bounding box closely match those of the Duffing oscillator on paraboloid data, according to Eq. (27).

Example 2 (*Comparison Between POD of Bounding Box and Singularly Perturbed Duffing System*)**.** The KL dimensions of uniform densities in boxes which trap the data from the family of singularly perturbed Duffing oscillators from system Eq. (7) shown in Fig. 1, are approximately,

$$\begin{aligned} W = X_1 &\equiv \sup_{Duffing} x_1 \approx 2.84, \\ L = X_2 &\equiv \sup_{Duffing} x_2 \approx 4.48, \\ H = Y_1 &\equiv \sup_{Duffing} y_1 = \alpha(X_1^2 + X_2^2) \approx \alpha 28.12, \end{aligned} \tag{27}$$

estimating the extreme X_1, X_2, and Y_1 values through simulation.

We can see in Fig. 3 that the analytically computed eigenvalues of a uniform distribution in a tight bounding box closely match those of time-averaged covariance matrix of data generated by

the singularly perturbed Duffing systems Eq. (7) with paraboloid slow manifolds. Thus, the curvature of the slow manifold dictates the dimensions of the bounding box, and the dimensions of the bounding box *approximates* the KL dimension.

Example 3 (*KL Dimension of a Delta Function Uniformly Distributed on a Paraboloid*)**.** For a better approximation of the time-averaged covariance of Duffing data on the paraboloid, we compute the covariance of data uniformly distributed on the same paraboloid. Note the difference between this computation and that of the singularly perturbed Duffing system. While we will use a delta function in the z-direction to restrict to the paraboloid, we use a uniform measure for the x and y directions. The true system does not use a uniform measure in the x and y directions, but instead there is a true, and not exactly computable, invariant measure of the Duffing system. So, we offer the uniform measure for its computability, and the fact that we believe that it gets to the heart of our point at hand.

We let,

$$x_2 = h(x_1) = 4H\frac{x_2^2}{L^2} - \frac{H}{2}, \tag{28}$$

giving a parabola whose corners are at the corners of an $H \times L$ rectangle, and whose minimum is at the bottom of $(0, \frac{H}{2})$. Therefore, the mathematical means of the uniform distribution on the parabola are computed,

$$\begin{aligned}
A &= \int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{\frac{-L}{2}}^{\frac{L}{2}} \delta(x_2 - h(x_1))\mathrm{d}x_1\mathrm{d}x_2 \\
&= \frac{L(-\sqrt{H-L} + \sqrt{H+L})}{\sqrt{2}\sqrt{H}}, \\
M_{x_1} &= \int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{\frac{-L}{2}}^{\frac{L}{2}} x_1\delta(x_2 - h(x_1))\mathrm{d}x_1\mathrm{d}x_2 = 0, \\
M_{x_2} &= \int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{\frac{-L}{2}}^{\frac{L}{2}} x_2\delta(x_2 - h(x_1))\mathrm{d}x_1\mathrm{d}x_2 \\
&= \frac{L(2H(\sqrt{H-L} - \sqrt{H+L}) + L(\sqrt{H-L} + \sqrt{H+L}))}{6\sqrt{2}\sqrt{H}},
\end{aligned} \tag{29}$$

in terms of the Dirac-delta function. Then, similarly to Eq. (30), but now using the Dirac density,

$$C_{i,j} = \frac{1}{A}\int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{\frac{-L}{2}}^{\frac{L}{2}} (x_i - M_i)(x_j - M_j)\delta(x_2 - h(x_1))\mathrm{d}x_1\mathrm{d}x_2, \tag{30}$$

from which it follows that

$$\begin{aligned}
C_{xx} &= \frac{L^2(H(-\sqrt{H-L} + \sqrt{H+L}) + L(\sqrt{H-L} + \sqrt{H+L}))}{24H(-\sqrt{H-L} + \sqrt{H+L})} \\
C_{yy} &= [-80H^{\frac{7}{2}}L + 60H^{\frac{3}{2}}L^3 \\
&\quad - 8\sqrt{2}H^3(3 + 5L^2)(\sqrt{H-L} - \sqrt{H+L}) + \cdots \\
&\quad + \sqrt{2}HL^2(-9 + 20L^2)(\sqrt{H-L} - \sqrt{H+L}) - \cdots \\
&\quad - 4\sqrt{2}H^2L(3 + 5L^2)(\sqrt{H-L} + \sqrt{H+L}) + \cdots \\
&\quad + 5(-4L^3\sqrt{H^3 - HL^2} + 16L\sqrt{H^7 - H^5L^2} \\
&\quad + \sqrt{2}L^5(\sqrt{H-L} + \sqrt{H+L}))]/\ldots \\
&\quad /[180\sqrt{2}H(-\sqrt{H-L} + \sqrt{H+L})] \\
C_{xy} &= C_{yx} = 0.
\end{aligned} \tag{31}$$

We see that the eigenvalues are the diagonal elements, $\lambda_{1,2} = \{C_{xx}, C_{yy}\}$. Fig. 4 shows that the eigenvalues of this uniform in x, delta function model closely match in character those of the data on paraboloid from the singularly perturbed Duffing system, as

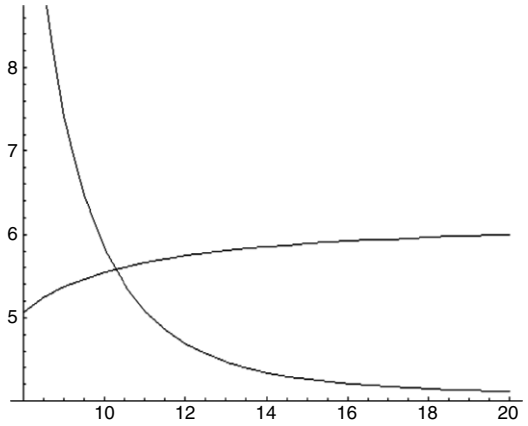

Fig. 4. Eigenvalues of the covariance matrix from a uniform distribution on a parabolic-delta function according to Eq. (31) and its precursors.

shown in Fig. 2. For specificity of the picture, we choose $L = 7$, and the horizontal axis is H. That the above is a two-dimensional calculation is not an important failure in comparison to the Duffing system, since the paraboloid-delta function version trapped in a $L \times H \times W$ box could also be easily computed, albeit with a more extensive and tedious algebraic solution. The major difference is the fact that we compute integrals against Lebesgue uniform density. However, the Duffing singularly perturbed system would call for integration against the Duffing $x - y$ invariant measure, to which we do not have analytic access, as this is generally not possible for realistic chaotic dynamical systems. If we were to resort to numerical approximation of the invariant measure, then that would be more or less equivalent to the eigenvalue covariance computation from data as we already performed leading to Fig. 2.

The point here is summarized by the following observations:

1. The spectrum of singular values, corresponding to the square root of eigenvalues of the time-averaged covariance matrix of the dynamical data, Eq. (13), is approximated by the lengths of the sides of a tight bounding box.
2. The dimension of an embedding manifold of the attractor may be quite different from that of a tight bounding box.
3. If singular vectors are used to decide what should be the embedding dimension, based on the usual KL method, then a change of variables, such as the dilation in Eq. (11), can easily change that concluded dimension dramatically.

The dimension of a reduced model should not be so easily dependent upon an implicitly chosen dilation (choice of units), as it is for the widely popular KL analysis. But since it is as we have shown, we suggest that at least this implication should be better and widely understood.

Our canonical form Eq. (5) is sufficiently general to any system Eq. (1) in variables $z = \langle x, y\rangle^t$ such that there is coordinate transformation,

$$Z = H(z), \tag{32}$$

where H is a diffeomorphism, $H : \Re^{m+n} \to \Re^{m+n}$, $Z = \langle X, Y\rangle^t$, and,

$$H \circ G = Z - \alpha g(X). \tag{33}$$

In other words, the example form is sufficient if there is a coordinate transformation (such as a rotation) where the invariant slow manifold is a Lipschitz graph over X. In such a case, the KL analysis will automatically tend to find a proper coordinate axis aligned with this axis when the linear part of g is zero. If this graph $g(X)$ has a bounded second derivative,

$$\sup_{x\in\omega(x)} |D^2 g(x)| = M, \tag{34}$$

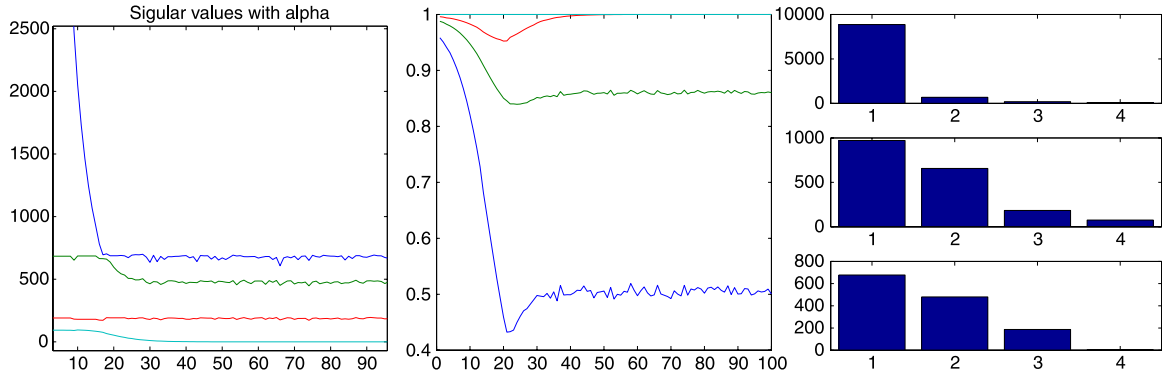

Fig. 5. (Left) The four eigenvalues of the time-averaged covariance matrix, with respect to varying α in Eqs. (37), much as was seen for the Duffing oscillator in Fig. 2. (Middle) Total variance as a function of the top i eigenvalues. In order are the curves, c_1 (blue), c_2 (green), c_3 (red), and $c_4 \equiv 1$, from Eq. (39). Thus by considering the percentage of variance captured due to truncation, we would conclude that for $\alpha < 15$, $d = 1$, for $15 < \alpha < 25$, $d = 4$, and for $\alpha > 25$, $d = 3$. (Right) A bar plot showing the 4 eigenvalues for each of these 3 regions, according to Eq. (35), of fixed α shows the spectra which leads to the conclusions of $d = 1$, 4, and 3, respectively, in order from top to bottom. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

then,

1. smaller α results in KL dimension n,
2. intermediate α results in KL dimension $m + n$,
3. larger α results in KL dimension m. (35)

This motivates us to summarize this relationship with the following definition of a closure parameter.

Definition. Given a KL dimension d_{KL} according to Eqs. (9)–(10), and d_M which is the standard manifold dimension in terms of charts, atlases, and homeomorphisms to Euclidean space [21], of an embedding manifold, then let the **curvature induced parameter**, d_{CI}, be defined by the equation,

$$d_{\mathrm{KL}} \equiv d_M + d_{CI}. \tag{36}$$

Note that following (35), d_{CI} can assume any sign, and therefore while it makes a convenient closure parameter, it may be awkward to interpret as a dimension.

5. An example of modeling by KL, subject to highly curved slow manifolds

As an example of how embedding problems lead to modeling problems, we will choose the following explicit quadratic example, from which to carry forward the full modeling parameter estimation program we specified in [22] to reconstruct equations of motion which approximately reproduce the data. The question we address is how well can we model by parameter estimation the dynamical system which produced the data, using dimension reduction methods in the three major α regimes discussed in the previous section.

Consider a four-dimensional system of ODEs, consisting of three Lorenz equations and a parabolic slow manifold,

$$\begin{aligned}
\dot{x}_1 &= \sigma(x_2 - x_1),\\
\dot{x}_2 &= r x_1 - x_2 - x_1 x_3,\\
\dot{x}_3 &= x_1 x_2 - b x_3,\\
\epsilon \dot{y} &= \alpha(x_1^2 + x_2^2 + x_3^2) - y,
\end{aligned} \tag{37}$$

with the usual, $\sigma = 10$, $b = 8/3$, $r = 28$, and we choose $\epsilon = 0.05$ for the simulations shown. See Fig. 5, showing results of the KL method, based on the usual method to truncate at $100q\%$ of total variance (for the sake of argument $100q\% = 95\%$ is chosen here), as already mentioned in Eq. (10). We choose the *smallest* d so that,

$$1 \geq c_d > q > 0, \tag{38}$$

where,

$$c_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i}. \tag{39}$$

For spectral analysis, we arrange a $4 \times N$ data matrix X,

$$\mathbf{Z}^{(i)} \equiv \mathbf{Z}(:, i) = \langle x_1(t_i), x_2(t_i), x_3(t_i), y(t_i)\rangle^T, \quad i = 0, 1, \ldots, N-2. \tag{40}$$

We highlight three different choices of the curvature of the slow manifold controlling parameter α which leads to the three different parameter regimes in Eq. (35), of what should be the dimension of the reduced model.

Recently, some of us [22] have studied the numerical analysis of nonlinear parameter estimation to fit differential equations from data $\mathbf{Z}$, which are meant to reproduce (predict) $\mathbf{Z}$. If we have reason to suspect that the model which reproduces $\mathbf{Z}$ is quadratic, then we could write the general quadratic ODE of appropriate dimension d,

$$\dot{\mathbf{Z}} = A_1 \mathbf{Z} + A_2 \mathbf{Q} + A_0 = [A_0 | A_1 | A_2] \begin{bmatrix} \mathbf{1} \\ \mathbf{Z} \\ \mathbf{Q} \end{bmatrix} \tag{41}$$

where $\mathbf{Q}$ is the $N(N-1)/2 \times 1$ matrix of all quadratic terms of data $\mathbf{Z}$, which for system Eq. (37) may be arranged,

$$\mathbf{Q} = \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ \cdots \\ y x_2 \\ y x_3 \\ y^2 \end{bmatrix}, \tag{42}$$

and $\mathbf{1}$ is a $1 \times M$ matrix of ones, the same size as $\mathbf{Z}(:, i)$, acting as a place holder for the affine shift part of the general quadratic equation. The goal in this paper is to discuss the consequences of different choices of d. In our recent paper, [22], we discuss convergence and stability issues of parameter estimation of the coefficients matrix,

$$A = [A_0 | A_1 | \ldots | A_d], \tag{43}$$

for general qth ordered polynomial models, by a least squares solution for the unknown parameters A in the undetermined

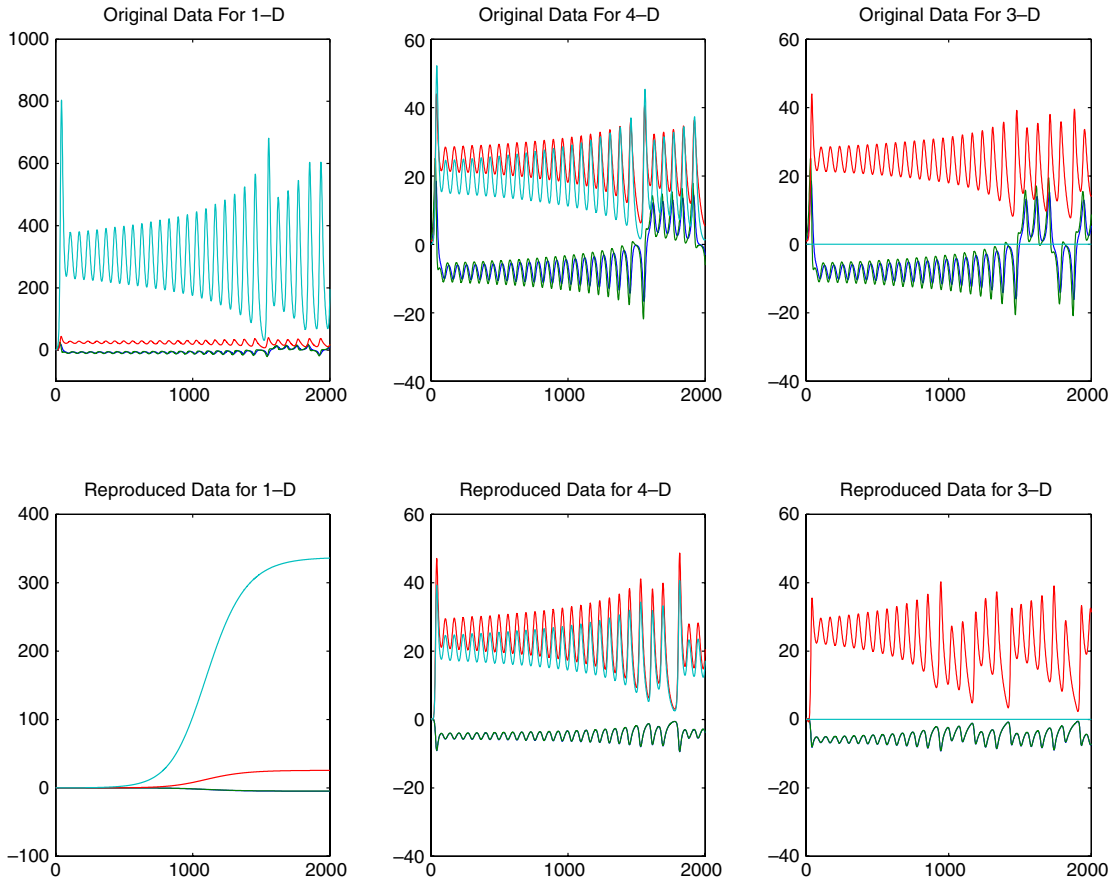

Fig. 6. The spectral analysis shown in Fig. 5 can mislead when it comes to modeling the data from Eqs. (37) by dimension reduction using the usual KL analysis dimension reduction methods. Data from parameter fitting techniques leads to a reduced model of dimension $d = 1$ (left) which is seen to poorly reproduce the data, $d = 4$ (middle) which well reproduces the data here (but overfitting can lead to numerical problems for higher-dimensional problems), while $d = 3$ (right) fits well. The values of α correspond to each of the three regions discussed in Fig. 5.

differential equation (41), of a generally overdetermined set of equations,

$$[\mathbf{Z}^{(1)} - \mathbf{Z}^{(0)}, \mathbf{Z}^{(2)} - \mathbf{Z}^{(1)}, \ldots, \mathbf{Z}^{(N)} - \mathbf{Z}^{(N-1)}] = h \cdot A \begin{bmatrix} \mathbf{1}^{(0)} & \mathbf{1}^{(1)} & & \mathbf{1}^{(N-1)} \\ \mathbf{Z} & \mathbf{Z} & \ldots & \mathbf{Z} \\ \mathbf{Q} & \mathbf{Q} & & \mathbf{Q} \end{bmatrix}, \tag{44}$$

here written for a general quadratic model. General cubic and higher qth ordered models are straightforward to pose, for which we refer [22].

Now we refer to Fig. 6 for conclusions of nonlinear parameter estimation to reproduce the data X, for the dimension d chosen to be the three different values $d = 1$, 4 and 3, respectively, suggested by α curvature controller as in Fig. 5. There are at least two different ways in which one might interpret differences between a dimension reduced system, and the original full-dimensional system.

5.1. Prediction and residual error

What is the error between data produced by modeled equations in the reduced dimension space as compared to the full model, in terms of the embedding norm? This is a main issue in finite-element analysis, and Galerkin's method, where error must be analytically controllable for (short) finite time.

We see in the first column of Fig. 6, that $d = 1$ results in poor reproduction by a poor model; this should not be a surprise with a priori knowledge of the original equations (37). However, without a priori knowledge of dimension for guidance, the KL analysis in Fig. 5 suggests that one dimension will be sufficient since for $\alpha < 15$, most of the variance is captured. Thus we see that indiscriminate use of KL analysis can lead to modeling a disaster, which as we point out here is due to an overly curved slow manifold. It should not be a surprise that for the next two columns of Fig. 6, that parameter estimation in both $d = 4$ and $d = 3$ dimensions reproduces the data well. But it is not always advisable to use the $d = m + n$ dimensions of the full system, since Eqs. (44) can lead to instability of the numerical least squares step as the shear size of the system grows exponentially with $m + n$, the dimension of the original system, and o, the polynomial order of the model. These issues of order, convergence, and stability of the model both for data residual, and well fitted parameters, are discussed in [22]. We could easily make an example system to accentuate this problem by choosing many y-variables corresponding to a higher-dimensional slow manifold, while maintaining a simple three-dimensional slow dynamics. For example, 3 slow x-variables, and 97 fast y-variables would result in such a high-dimensional least squares system to solve. It is easy to see the merits of reducing the order of the model as much as possible.

6. Nested reduced systems

So far, our examples have focused on the simplest case Eq. (5) in which one control parameter leads to legitimate KL reduced order models, sometimes giving significant errors. We now show a scenario of nested singularly perturbed systems which can lead to comparably legitimate multiple errors.

Consider a multiply nested version of singularly perturbed systems generalizing Eqs. (5), such as the two level nested system,

$$\begin{aligned}\dot{x} &= f(x),\\ \epsilon_1\dot{y} &= y - \alpha_1 g_1(x)\\ \epsilon_2\dot{z} &= z - \alpha_2 g_2(x, y)\end{aligned} \tag{45}$$

which can be formulated to have now 5 possible KL model reductions, based on the values of (α_1, α_2). Any level of complexity versions of this nesting form are possible, by appropriate design of nesting, leading to a complex degree of possible dimensions.

We have discussed at the end of the previous section that a high-dimensional ambient dimension and low-dimensional reduced system is possible. In this section, we showed that multiple-dimensional "confusions" are possible. In the next section, we will discuss how all of this can be possible in a very high-dimensional setting, spatiotemporal data from a partial differential equation.

Any spatiotemporal process which generates $u(x, t)$ and which is discretely sampled data in time,

$$t_i = i\Delta t, \quad \Delta t = t_{i+1} - t_i,\ i = 0, 1, \ldots, M-2, \tag{46}$$

and in space,

$$x_j = j\Delta x, \quad \Delta x = x_{j+1} - x_j,\ j = 0, 1, \ldots, N-2, \tag{47}$$

gives an $M \times N$ data matrix,

$$U_{j,i} = u(x_j, t_i), \tag{48}$$

which is meant to be modeled by the matrix of the data in Eq. (40). It is straightforward to index the spatial variable appropriately by raster scanning or multiscale methods in the case of more than one spatial dimension.

7. A singularly perturbed model in a continuum—oscillator mechanical models as a fast–slow system

In this section we describe an infinite-dimensional model consisting of multiple time scales which allows for a geometric dynamical splitting based on a singular perturbation parameter. In the following section, we will show how this system also displays the same variation of KL dimension embeddings, as listed in Eq. (35), as a natural parameter is varied.

The multiscale problems we consider here model linear continua coupled to nonlinear oscillators. Specifically, the problem class modeled is that of linear PDEs which are coupled to one or more nonlinear oscillators represented by ODEs, and are observed to exhibit nonlinear vibrations in experiments [23] as well as more complicated behavior in continuum systems with noise [24]. We restrict ourselves to models of linear elastica in one spatial dimension, which include cantilevered beams and extensible rods. Letting $W(\xi, t)$ denote a measure of displacement as a function of space (ξ) and time (t), and let $\varkappa_\mu(\xi, t)$ be a forcing function, then the general equations of motion may be represented as:

$$\begin{aligned}&L_\mu W(\xi, t) = \varkappa_\mu(\xi, t)\\ &\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + [1 + G(W_{,tt})]\sin\theta + \eta\frac{\mathrm{d}\theta}{\mathrm{d}t} = 0\end{aligned} \tag{49}$$

plus the appropriate boundary conditions. Here, L_μ is a linear differential operator. In Eq. (49), θ denotes the angular position of an attached pendulum at a free end of the elastica. Since there is an external driving body force on the structure, the function $G(W_{,tt})$ will also contain a time varying source, which will in general depend on another oscillator, such as a mechanical shaker or periodic electric potential.

7.1. Full PDE–ODE system

In formulating the dynamics of such a mutually coupled system, we follow [25,26] in formulating in detail a system based on a viscoelastic rod. We consider a specific mechanical system consisting of a vertically positioned viscoelastic linear rod of density ρ_r, with cross-section A_r and length L_r, with a simple pendulum of mass M_p and arm length L_p coupled at the bottom of the rod and where the rod is forced from the top harmonically with frequency Ω and magnitude α [26]. The rod obeys the Kelvin–Voigt stress–strain relation [27] and E_r and C_r denote the modulus of elasticity and the viscosity coefficient. C_p is the coefficient of viscosity (per unit length) of the pendulum and g is the gravitational constant of acceleration. The pendulum is restricted to a plane, and rotational motion is possible. The system is modeled by the following equations,

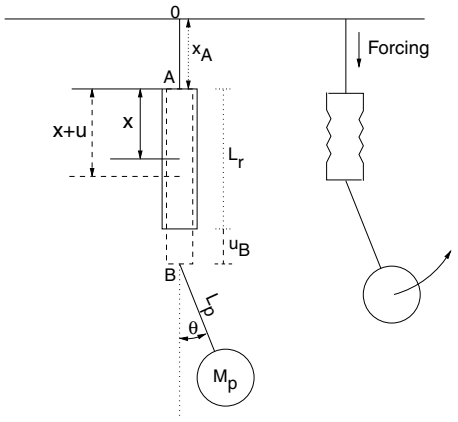

Fig. 7. Rod–pendulum configuration.

$$\begin{aligned}&M_pL_p\ddot{\theta} + M_p[g - \ddot{x}_A - \ddot{u}_B]\sin(\theta) + C_pL_p\dot{\theta} = 0\\ &A_r\rho_r\ddot{u}(x, t) - A_rE_ru''(x, t) - A_rC_r\dot{u}''(x, t) - A_r\rho_r(g - \ddot{x}_A) = 0,\end{aligned} \tag{50}$$

where $\dot{} \equiv \frac{\partial}{\partial t}$, and $' \equiv \frac{\partial}{\partial x}$, with boundary conditions

$$u(x = 0, t) = 0, \quad A_rE_r\left.\frac{\partial u}{\partial x}\right|_{x=L_r} = A_rE_r\frac{\partial u_B}{\partial x} = T_p\cos(\theta),$$

and where

$$T_p = M_pL_p\dot{\theta}^2 + M_p(g - \ddot{x}_A - \ddot{u}_B)\cos(\theta)$$

denotes the tension acting along the rigid arm of the pendulum. The variable $u(x, t)$ denotes the displacement field of the uncoupled rod with respect to the undeformed configuration at equilibrium, relative to the point A, while u_B denotes the relative position of the coupling end B of the rod with respect to point A. See Fig. 7 for a schematic of the rod and pendulum system. Note that the coupling in θ appears in the boundary conditions.

We further suppose that the drive at A, given by the function $x_A(t)$ in Eq. (50), is such that it comes from another oscillator. To keep the coupling bi-directional and general, we suppose that the oscillator is weakly coupled to the pendulum through its frequency. Specifically, we model the drive oscillator by

$$\begin{aligned}\dot{\Phi}_1 &= \Phi_1 + \Omega(1 + \Sigma P(\dot{u}(x, t)))\Phi_2 - \Phi_1(\Phi_1^2 + \Phi_2^2)\\ &\equiv F_1(\Phi_1, \Phi_2, \Sigma, \Omega)\\ \dot{\Phi}_2 &= -\Omega(1 + \Sigma P(\dot{u}(x, t)))\Phi_1 + \Phi_2 - \Phi_2(\Phi_1^2 + \Phi_2^2)\\ &\equiv F_2(\Phi_1, \Phi_2, \Sigma, \Omega),\end{aligned} \tag{51}$$

where P is a projection onto a Fourier mode (see below), and $|\Sigma| \ll 1$ is the coupling term that modulates the frequency. Note that when $\Sigma = 0$, the solution of Eq. (51) consists of sines and cosines of frequency ω given the appropriate initial conditions. In terms of the solutions to Eq. (51), note that $x_A(t) = \Phi_2(t, \Sigma)$.

The scaled PDE–ODE system is derived in the Appendix in Eq. (55), and may be represented as a fast–slow system with a

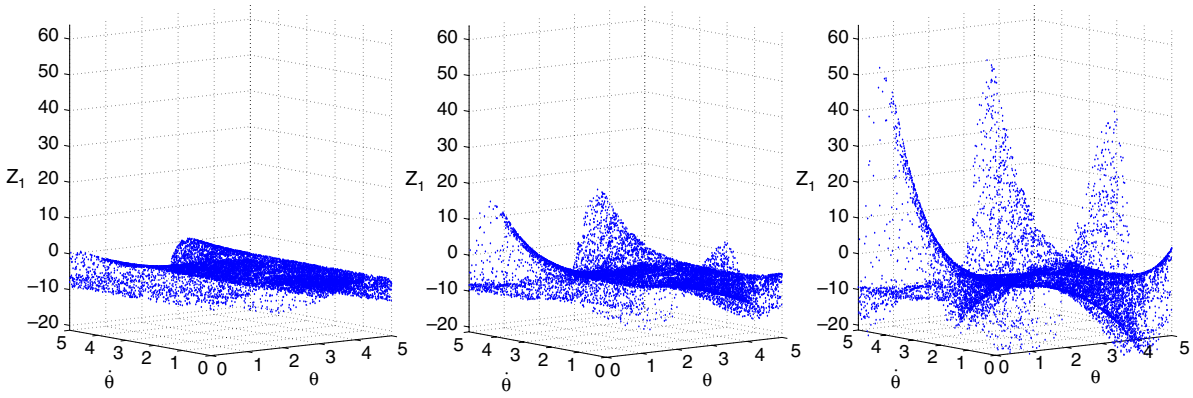

Fig. 8. The slow manifold of the viscoelastic rod system, Eq. (53), represented by Z_1 as a function of θ and $\mathrm{d}\theta/\mathrm{d}t$ with parameter $\kappa = 31, 60, 130$, left to right, respectively, in the three different parameter regimes Eq. (35), accordingly as seen in the dimension parameter study plots, Fig. 9. Here, we show these data sets in the same vertical scale, which would not be normally used for all three, but which lays bare the varying manifold curvature, leading to varying KL dimension, with the manifold curvature varying parameter κ in the rod system, Eq. (54). Note that the concluded KL in the three regimes of Eq. (35) switch roughly when $\kappa \sim 40$ and 65.

parameter, μ, defined in Eq. (52), which is the ratio of the slow pendulum frequency to the fast rod frequency. If μ is sufficiently small, it is then a singular perturbation parameter. Using the scaled equations and expansions in the Appendix, we let $\Psi = [\Psi_1, \Psi_2, \Psi_3, \Psi_4]^T$ and $Z = [Z_1, Z_2, \ldots, Z_{2m-1}, Z_{2m}]$, for $m = 1, 2, \ldots, N$, denote the phase space for the pendulum and drive, and the rod modes. Performing a similar analysis in [26], the slow manifold approximation can be computed by expanding the solution to the manifold equations to get:

$$Z = H^{\mu}(\Psi, N), \tag{52}$$

where $H^{\mu}(\Psi, N) = \kappa \sum_{j=0}^{\infty} \mu^j H^j(\Psi, N)$. Here κ is acting to amplify the nonlinear geometry of the surface. That is, the local curvature terms plus other terms of higher order will be controlled by the parameter κ. Since the leading order terms in Ψ will be quadratic in general for H^0, we expect the curvature on the manifold to have the largest increase as a function of κ.

For the examples we consider, we examine the slaved relationship of the dynamics of the tip of the rod to that of the dynamics of pendulum. The parameter $\mu = 0.025$, is fixed throughout the example. We omit the manifold expansion details, since similar equations have already been presented in [26].

7.2. Continuum KL analysis of manifolds for different curvatures

Here again, this time in a continuum model, the result is clearly just as it was for the constructed lower-dimensional singularly perturbed models of the previous sections. We see clearly in Figs. 8 and 9 the same scenario where the dimension concluded depends inherently on the curvature parameter κ which can be determined by something as arbitrary as a choice of measurement units.

8. Conclusion

While the KL method is a highly popular method for analysis of laboratory data, and empirical data, for producing reduced order models from high-dimensional systems. We have demonstrated a particular scenario, where a singularly perturbed system is expected to have a lower-dimensional representation of the flow data on a submanifold. When the KL method is applied to such systems, it may be expected that we might properly recover an appropriate dynamically relevant dimension either for modeling functionality or for performing prediction. These are typical goals when a model reduction program is undertaken.

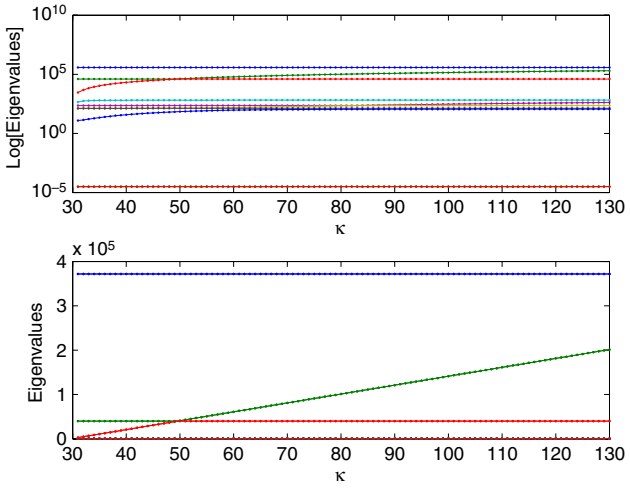

Fig. 9. (a) Log of eigenvalues of the time-averaged covariance as a function of curvature parameter, κ. (b). Top eigenvalues of panel (a) showing the crossover effect leading to a change in KL dimension as a function of κ.

However, while the KL method is so widely used, the degree to which certain simple data scalings, such as a change of units of one or some of the variables (such as a linear transformation like changing from inch to meters), can dramatically effect the curvature of the slow manifold may be overlooked. The implication to the KL analysis which we highlight here is that the concluded KL dimension can be varied through three distinct regimes: it is easy to accidentally choose m, n or $m + n$. This can cause dramatic differences in results whether they be for good models, or good predictions. We hope that this work will serve for a better understanding of scale and unit issues in model reduction techniques. We wish to point out here that an "energy-based" inner product was suggested in [28], where $\langle (x_1, y_1), (x_2, y_2) \rangle = x_1 x_2 + s^2 y_1 y_2$, with which it can be shown that scaling adjustments such as in Eq. (11) can be defined in terms of adjustments of the inner product weights, and so there is a connection to resolve scaling issues discussed here directly in the inner product; a change of units is then inversely proportional to the parameter in the inner product to form a scaling independent KL subspace.

Acknowledgements

EMB and CY are supported by the NSF under DMS-0404778. IBS is supported by the Office of Naval Research.

Appendix

Eqs. (50) and (51) are nondimensionalized by the following variable re-scalings

$$\xi = \frac{x}{L_r}, \quad \tau = \omega_p t,$$
$$X_A = \frac{x_A}{L_p}, \qquad U = \frac{u}{L_p}, \qquad U_B = \frac{u_B}{L_p},$$

and parameter re-scalings

$$\mu = \frac{\omega_p}{\omega_1}, \qquad \mu_m = \frac{\omega_1}{\omega_m} = \frac{1}{2m-1}, \qquad \beta = \frac{M_p}{A_r \rho_r L_r}$$
$$\zeta_p = \frac{1}{2\omega_p}\frac{C_p}{M_p}, \qquad \zeta_r = \frac{1}{2\omega_1}\frac{\pi^2 C_r}{4L_r^2 \rho_r},$$

where

$$\omega_p = \sqrt{\frac{g}{L_p}}, \qquad \omega_m = \frac{\pi(2m-1)}{L_r}\sqrt{\frac{E_r}{\rho_r}}, \quad m = 1, 2, \ldots, \infty,$$

are the natural frequency of the uncoupled pendulum and the spectrum of natural frequencies of the uncoupled flexible rod, respectively, while ζ_p and ζ_r denote their damping factors.

The stable and unstable static equilibrium configurations of the coupled rod and pendulum system are given by $(\theta_c, \hat{U})$ and $(\theta_{S_\pm}, \hat{U})$, where

$$\theta_c = 0, \qquad \theta_{S_\pm} = \pm\pi$$
$$\hat{U} = \frac{\mu^2\pi^2}{2}[2(1+\beta)\xi - \xi^2].$$

The normalized equations are thus

$$\begin{aligned}
&\ddot{\theta} + [1 - \ddot{V}_B(\tau) - \ddot{X}_A(\tau)]\sin(\theta) + 2\zeta_p\dot{\theta} = 0,\\
&\mu^2\pi^2\ddot{V}(\xi,\tau) - V''(\xi,\tau) - 8\zeta_r\mu\dot{V}''(\xi,\tau) = -\mu^2\pi^2\ddot{X}_A(\tau) \qquad (53)\\
&V(\xi=0,\tau) = 0, \qquad V'(\xi=1,\tau) = -\mu^2\beta\pi^2[1 - T\cos(\theta)],
\end{aligned}$$

where

$$V(\xi,\tau) = U(\xi,\tau) - \hat{U}(\xi), \quad 0 \le \xi \le 1, \quad -\infty < \tau < +\infty,$$

and note that we redefine $\dot{} \equiv \frac{\partial}{\partial\tau}$ and $' = \frac{\partial}{\partial\xi}$ for the remainder of the paper.

A.1. Projection onto a finite model

In carrying out our analysis, we will consider a reduction of the ODE/PDE system in Eq. (53). This reduction is obtained by performing a modal expansion of the rod equation, where the displacement V is expanded as $V(\xi,\tau) = \sum_{m=1}^{\infty}\eta_m(\tau)\phi_m(\xi)$. This results in an infinite system of coupled oscillators,

$$\ddot{\theta} = -\left[1 + \sum_{j=1}^{\infty}(-1)^{j+1}\ddot{\eta}_j - \ddot{X}_A(\tau)\right]\sin(\theta) - 2\zeta_p\dot{\theta}$$

$$\begin{aligned}
L_m(\theta)\ddot{\eta}_j \;=\; &-\frac{\eta_m}{4\eta_m^2} + 2\zeta_r\frac{\dot{\eta}_m}{\mu\mu_m^2}\\
&-(-1)^{m+1}2\beta[\dot{\theta}^2\cos(\theta) - \sin^2(\theta)]\\
&-\left[\frac{4\mu_m}{\pi} + (-1)^{m+1}2\beta\cos^2(\theta)\right]\ddot{X}_A(\tau), \qquad (54)
\end{aligned}$$

equivalent to Eq. (50), where $L_m(\theta)$ is the infinite linear operator

$$L_m(\theta) \equiv \sum_{j=1}^{\infty}[\delta_{mj} + (-1)^{m+j}2\beta\cos^2(\theta)].$$

See [25] for the details of this transformation.

Finally, consider the finite set of ordinary differential equations obtained from Eq. (54) by truncating to the first N rod modes and applying the additional re-scalings $\{\Psi_1, \Psi_2\} = \{\theta, \dot{\theta}\}$ and $\{\mu^2\mu_m^2 Z_{2m-1}, \mu\mu_m^2 Z_{2m}\} = \{\eta_m, \dot{\eta}_m\}$, obtaining

$$\begin{aligned}
&\dot{\Psi}_1 = \Psi_2\\
&\dot{\Psi}_2 = -\left[1 - \sum_{j=1}^{N}(-1)^{j+1}f_N(\Psi, Z) - \alpha\Psi_4\right]\sin(\Psi_1) + 2\zeta_p\Psi_2\\
&\dot{\Psi}_3 = F_1(\Psi_3, \Psi_4, \Sigma, \Omega) \qquad (55)\\
&\dot{\Psi}_4 = F_2(\Psi_3, \Psi_4, \Sigma, \Omega)\\
&\mu\dot{Z}_{2m-1} = Z_{2m}\\
&\mu\mu_m^2\dot{Z}_{2m} = f_N(\Psi, Z), \quad m = 1, 2, \ldots, N,
\end{aligned}$$

where

$$\begin{aligned}
f_N(\Psi, Z) \;= L_{m,N}^{-1}(\Psi_1)\Bigg[&-\frac{1}{4}Z_{2m-1} + 2\zeta_r Z_{2m}\\
&-(-1)^{m+1}2\beta[\Psi_2^2\cos(\Psi_1) - \sin^2(\Psi_1)]\\
&-\left[\frac{4\mu_m}{\pi} + (-1)^{m+1}2\beta\cos^2(\Psi_1)\right]\alpha\Psi_4\Bigg]
\end{aligned}$$

and $L_{m,N}^{-1}(\theta)$ is the inverse of the $N \times N$ truncation of operator $L_m(\theta)$. F_1 and F_2 are given by the right-hand sides of Eq. (51). Note that Eq. (55) is an autonomous system, and the cyclic variables, Ψ_3 and Ψ_4 are introduced to account for the periodic forcing, which has period Ω when the coupling parameter $\Sigma = 0$. For this example, we will assume $N = 32$ modes.

The primary parameter governing the coupling between the rod and pendulum is the ratio of the natural frequency of the pendulum to the frequency of the first rod mode, $\mu \equiv \omega_p/\omega_1$. In the limit $\omega_1 \to \infty$, the rod is perfectly rigid, $\mu \to 0$, and the system reduces to a forced and damped pendulum. For $0 < \mu \ll 1$ sufficiently small, global singular perturbation theory predicts that system motion is constrained to a slow manifold, and the (fast) linear rod modes are slaved to the slow pendulum motion [14]. For nonzero α (the amplitude of the periodic forcing) and $\Sigma = 0$, the slow manifold is a non-stationary (periodically oscillating) two-dimensional surface.

Note that Eq. (55) is now in a form which reveals the slow and fast components for a small parameter μ. One may now show how the mode amplitudes of the rod, Z_i are slaved to the slow manifold by the use of the center manifold theorem. The details of the construction of the manifold are carried out in [26].

References

[1] M. Kirby, Geometric Data Analysis, John Wiley and Sons, New York, 2001.
[2] S. Watanabe, Trans. 4th Prague Conf. Information Theory, vol. 635, 1965.
[3] L. Sirovich, Turbulence and the dynamics of coherent structures. 1. Coherent structures, Quart. Appl. Math. 45 (1987) 561–571.
[4] R. Everson, P. Cornillon, L. Sirovich, A. Webber, An empirical eigenfunction analysis of sea surface temperatures in the western north atlantic, J. Phys. Oceanogr. 27 (1997) 468–479.
[5] E.V. Obrien, D. Orbach, R. Everson, D. Samber, M. Rossetto, L. Sirovich, B. Knight, E. Kaplan, Principal component analysis of intrinsic optical signals in mammalian visual-cortex reveals functional architecture, Invest. Ophthalmol. Vis. Sci. 35 (1994) 1663–1663.
[6] M. Kirby, L. Sirovich, Application of the Karhunen–Loeve procedure for the characterization of human faces, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 103–108.
[7] L. Sirovich, A pattern analysis of the second rehnquist us supreme court, Proc. Natl. Acad. Sci. USA 100 (2003) 7432–7437.
[8] I. Triandaf, I.B. Schwartz, Phys. Rev. E 56 (1997) 204.
[9] K. Karhunen, Ann. Acad. Sci. Fenn. 34 (Ser. A1) (1946) 37.
[10] M.M. Loeve, Probability Theory, Van Nostrand, Princeton, NJ, 1955.
[11] P. Holmes, J.L. Lumley, G. Berkooz, Turbulence, Coherent Structures, Dynamical Systems, and Symetry, Cambridge Press, New York, 1996.
[12] L. Sirovich, Q. Appl. Math. XLV 561 (1987).
[13] F. Takens, Detecting strange attractors in turbulence, in: A. Dold, B. Eckmann (Eds.), Dynamical Systems and Turbulence, in: Lecture Notes in Mathematics, vol. 898, Springer, Berlin, 1981, pp. 366–381.

[14] N. Fenichel, Geometric singular perturbation theory for ordinary differential equations, J. Differential Equations 31 (1) (1979) 5398.
[15] J. Carr, Applications of Center Manifold Theory, Springer-Verlag, New York, 1981.
[16] A.N. Tikhonov, A.B. Vasileva, A.G. Sveshnikov, Differential Equations, Springer, 1985.
[17] E.N. Lorenz, Emprical Orthogonal Functions and Statistical Weather Prediction, MIT, Cambridge, 1956.
[18] J.L. Lumley, Stochastic Tools in Turbulence, Academic, New York, 1970.
[19] G.H. Golub, C.F. VanLoan, Matrix Computations, 2nd edition, The Johns Hopkins University Press, Baltimore, 1989.
[20] R. Courant, D. Hilbert, Methods of Mathematical Physics, vol. 1, Interscience, 1953.
[21] James R. Munkres, Analysis on Manifolds, Westview Press, 1997.
[22] Yao, Erik Bollt, Modeling and nonlinear parameter estimation with kronecker product representation for coupled oscillators and spatiotemporal systems, Physica D 227 (1) (2007) 78–99.
[23] I.T. Georgiou, I. Schwartz, E. Emaci, A. Vakakis, Trans. ASME J. Appl. Mech. 66 (1999) 448.
[24] I.B. Schwartz, D.S. Morgan, L. Billings, Y.C. Lai, Multi-scale continuum mechanics: from global bifurcations to noise induced high-dimensional chaos, Chaos 14 (2004) 373–386.
[25] D.M. Morgan, E. Bollt, I.B. Schwartz, Phys. Rev. E 68 (2003) 056210.
[26] I.T. Georgiou, I.B. Schwartz, SIAM J. Appl. Math. 59 (1999) 1178.
[27] L. Marven, Introduction to the Mechanics of a Continuous Medium, Prentice-Hall, Englewood Cliffs, 1969.
[28] C.W. Rowley, T. Colonius, R.M. Murray, Model reduction for compressible flows using POD and Galerkin projection.