Nonlinear model reduction for dynamical systems using sparse sensor locations from learned libraries

Syuzanna Sargsyan, Steven L. Brunton, and J. Nathan Kutz

Phys. Rev. E **92**, 033304 — Published 10 September 2015

DOI: 10.1103/PhysRevE.92.033304

Nonlinear Model Reduction for Dynamical Systems using Sparse Sensor Locations from Learned Libraries

Syuzanna Sargsyan*, Steven L. Brunton† and J. Nathan Kutz*
*Department of Applied Mathematics, University of Washington, Seattle, WA 98195-3925
†Department of Mechanical Engineering, University of Washington, Seattle, WA 98195
(Dated: August 14, 2015)

We demonstrate the synthesis of sparse sampling and dimensionality-reduction to characterize and model nonlinear dynamical systems over a range of bifurcation parameters. First, we construct modal libraries using the classical proper orthogonal decomposition in order to expose the dominant low-rank coherent structures. Here, libraries of the nonlinear terms are also constructed in order to take advantage of the discrete empirical interpolation method and projection that allows for the approximation of nonlinear terms from a sparse number of grid points. The selected grid points are shown to be effective sensing/measurement locations for characterizing the underlying dynamics, stability, and bifurcations of nonlinear dynamical systems. The use of empirical interpolation points and sparse representation facilitates a family of local reduced-order models for each physical regime, rather than a higher-order global model, which has the benefit of physical interpretability of energy transfer between coherent structures. The method advocated also allows for orders-of-magnitude improvement in computational speed and memory requirements. To illustrate the method, the discrete interpolation points and nonlinear modal libraries are used for sparse representation in order to classify and reconstruct the dynamic bifurcation regimes in the complex Ginzburg-Landau equation. It is also shown that point measurements of the nonlinearity are more effective than linear measurements when sensor noise is present.

I. INTRODUCTION

The theoretical study of nonlinear dynamical systems pervades the physical, biological and engineering sciences. Today, these studies are increasingly driven by computational simulations that are of growing complexity and dimension due to increasing computational power and resolution in numerical discretization schemes. Yet most dynamics of interest are known ultimately to be low-dimensional in nature [1], thus contrasting, and in antithesis to, the high-dimensional nature of scientific computing. Reduced-order models (ROMs) are of growing importance in scientific applications and computing as they help reduce the computational complexity and time needed to solve large-scale, engineering systems [2, 3]. Specifically, ROMs provide a principled approach to approximating high-dimensional spatio-temporal systems. However, the complexity of evaluating low-rank approximations remains challenging due to higher-order nonlinear terms [4, 5]. The empirical interpolation method (EIM), and the simplified discrete empirical interpolation method (DEIM) for the proper orthogonal decomposition (POD) [6, 7], overcomes this difficulty by providing a computationally efficient method for discretely (sparsely) sampling and evaluating the nonlinearity. These methods ensure that the computational complexity of ROMs scale favorably with the rank of the approximation, even for complex nonlinearities.

An alternative computational strategy for handling the nonlinearity is based upon machine learning techniques of dimensionality reduction whereby libraries of *learned* POD modes can be constructed and inner products precomputed for a number of distinct dynamical regimes of the nonlinear dynamical system [8–11]. This strategy also evokes the power of compressive sensing for efficiently identifying the active POD subspace necessary for a low-dimensional Galerkin-POD truncation [6, 7]. In this manuscript, we combine the power of the DEIM with the library building strategy. Specifically, we show that building libraries that encode the nonlinearities allows one to (i) take advantage of the DEIM to evaluate the nonlinearities, (ii) more robustly classify the dynamical regime the system is in, and (iii) identify the discrete sensor locations to construct a nonlinear model reduction. Our method allows for orders-of-magnitude improvement in computational speed and memory requirements, thus making it highly attractive for computational physics simulations.

We demonstrate the full integration of the methods on a canonical model of mathematical physics and nonlinear dynamical systems, the cubic-quintic Ginzburg-Landau (CQGLE) equation. However, the methodology can be more broadly applied to nonlinear dynamical systems. In biophysical applications, for instance, advances in neuroscience over the past decade have revealed two critical, and seemingly ubiquitous, phenomena: (i) that meaningful input/output of signals in high-dimensional networks of neurons are encoded in low-dimensional patterns of dynamic activity [12–17], and (ii) that sparsity plays a key role in encoding and decoding strategies, especially in neural computation [18, 19]. This typifies the state-of-the-art in neuroscience, which manifests the most sophisticated functionality in information processing and computation known in any system.

A. Dimensionality Reduction

Although a variety of dimensionality-reduction techniques exist, the ROM methodology considered here is based upon the proper orthogonal decomposition [6, 7]. The POD method is ubiquitous in the dimensionality reduction of physical systems. It is alternatively referred to as principal components analysis (PCA) [20], the Karhunen–Loève (KL) decomposition, empirical orthogonal functions (EOF) [21], or the Hotelling transform [22, 23]. Snapshots (measurements) of many nonlinear dynamical system often exhibit low-dimensional phenomena [1], so that the majority of variance/energy is contained in a few modes computed from a singular value decomposition (SVD). For such a case, the POD basis is typically truncated at a pre-determined cut-off value, such as when the modal basis contain 99% of the variance, so that only the first r-modes (r-rank truncation) are kept. There are numerous additional criteria for the truncation cut-off, and recent results derive a hard-threshold value for truncation that is optimal for systems with well-characterized noise [24]. The SVD acts as a filter, and often the truncated modes correspond to random fluctuations and disturbances. If the data considered are generated by a dynamical system (nonlinear system of ordinary differential equations of order n), it is then possible to substitute the truncated POD expansion into the governing equation and obtain Galerkin projected dynamics on the rank-r basis modes [7, 11]. Recall that we are assuming that the nonlinear dynamical systems under consideration exhibit low-dimensional attractors, thus the Galerkin truncation with only a few modes should provide an accurate prediction of the evolution of the system [8, 9]. Note that it has also been shown recently that it is possible to obtain a *sketched*-SVD by randomly projecting the data initially and then computing the SVD [25–27].

B. Sparse Sampling

Efficiently managing the computation of the nonlinearity (inner products) in dimensionality reduction schemes is of paramount importance. This was recognized early on in the reduced order modeling community, and a variety of techniques where proposed to accomplish the task. Among the first methods used was the technique of Everson and Sirovich developed for gappy data [28]. In their sparse sampling scheme, random measurements were used to perform reconstruction tasks of inner products. Willcox [29] and Karniadakis [30] built on these ideas by advocating principled approaches for selecting sampling locations for Gappy POD.

The EIM was also developed for the purpose of efficiently managing the computation of the nonlinearity. And as with Gappy POD, principled techniques for sparse measurements where also advocated early on in its history [31]. A variant of this techniques, the DEIM method, was specifically tailored to POD with Galerkin projection. Indeed, the DEIM approximates the nonlinearity by using a small, discrete sampling of spatial points that are determined in an algorithmic way. This ensures that the computational cost of evaluating the nonlinearity remains proportional to the rank of the reduced POD basis. As an example, consider the case of an r-mode POD-Galerkin truncation. A simple cubic nonlinearity requires that the POD-Galerkin approximation be cubed, resulting in r^3 operations to evaluate the nonlinear term. The DEIM approximates the cubic nonlinearity by using $O(r)$ discrete sample points of the nonlinearity, thus preserving a low-dimensional ($O(r)$) computation, as desired. The DEIM approach combines projection with interpolation. Specifically, the DEIM uses selected interpolation indices to specify an interpolation-based projection for a nearly optimal ℓ_2 subspace approximating the nonlinearity. The EIM/DEIM are not the only methods developed to reduce the complexity of evaluating nonlinear terms, see for instance the missing point estimation (MPE) [32], "best points" method [31], or gappy POD [28–30, 33] methods already mentioned. However, they have been successful in a large number of diverse applications and models [5]. In any case, the MPE, gappy POD, and EIM/DEIM use a small selected set of spatial grid points to avoid evaluation of the expensive inner products required to evaluate nonlinear terms (See background section on POD modeling).

The discrete sampling points given by the DEIM to evaluate the nonlinearity get a new interpretation in the current work. Specifically, we show them to be the effective locations for placing sensors in the nonlinear dynamical system in order to (i) determine the dynamic regime of the system, (ii) reconstruct the current state of the system, and (iii) produce a POD-Galerkin prediction (nonlinear model reduction) of the future state of the system. Such tasks are accomplished by using ideas of sparse representation [34] and compressive sensing [35–42]. In particular, the theory of compressive sensing shows that a small number of measurements are sufficient to perform a reconstruction provided there exists a sparse representation (or basis) of the data. Sparsity techniques have also been shown to be highly effective for numerical solution schemes [43, 44]. In our case, the sparse basis is generated from a library learning procedure of a nonlinear dynamical system that exhibits low-rank dynamics. More than that, however, we also build libraries of the *nonlinearities*, thus pre-computing the low-dimensional structures observed in the different dynamical states of the nonlinear dynamical system. This allows for more robust dynamical classification as well as allowing easy evaluation of the nonlinear terms through the DEIM. The combination of library building, compressive sensing and the DEIM is demonstrated to be a highly effective and intuitively appealing methodology for scientific computing applications. It further highlights the need in modern scientific computing of nonlinear dynamical systems to integrate a variety of data-driven modeling strategies, many

of which are being developed under the aegis of dimensionality reduction, in order to most efficiently simulate large-scale systems.

C. Physical Interpretation

The ideas presented here are more than just numerical efficiencies. Indeed, the methodology identifies the underlying modal structures that drive the dynamics of the nonlinear dynamical system, thus helping to understand the fundamental interactions and physics of the system. Throughout the development of 20th-century physics and engineering sciences, the understanding of many canonical problems has been driven by recasting the problem into its *natural* basis (mode) set. The majority of classical problems from mathematical physics are linear Sturm-Liouville problems whose ideal modal representations are generated from eigenfunction decompositions, i.e. special functions. In quantum mechanics, for instance, Gauss-Hermite (denoted by $H_n(x)$) polynomials are the natural basis elements for understanding the harmonic oscillator. Likewise, spherical harmonics (denoted by $Y_l^m(\theta, \varphi)$) are critical in the computation of atomic orbital electron configurations as well as in representation of gravitational fields, the magnetic fields of planetary bodies and stars, and the characterization of the cosmic microwave background radiation.

For modern dynamical systems, nonlinearity plays a dominant role and shapes the underlying modes, thus necessitating a new approach, such as that presented here, for extracting these critical spatio-temporal structures. Remarkably, although nonlinearity creates new modal structures, it does not destroy the underlying low-dimensional nature of the dynamics. Distinct physical regimes may be obtained by varying bifurcation parameters, and these regimes will typically have different local bases and physical interactions. Instead of developing a global interpolated model, which may obscure these distinct physical mechanisms, we advocate a hierarchy of models along with sparse sampling and library learning to classify and characterize the system parameters from a few online measurements. Methods that take advantage of such underlying structure are critical for developing theoretical understanding and garnering insight into the fundamental interactions of the physical, engineering and biological systems under consideration.

The paper is outlined as follows. In Sec. II, an overview of the mathematical framework of the POD method and the the DEIM is given. This is followed up in Sec. III with an introduction of the cubic-quintic Ginzburg-Landau equation, to which the methods proposed here will be applied. The library building procedure that encodes the various dynamical regimes of our model equation are discussed in Sec. IV. Once the libraries are constructed, the DEIM points, or sensor locations, are computed in Sec. V and their ability to classify dynamical regimes is evaluated in Sec. VI. The reconstruction of the dynamics and future state projection is discussed in Sec. VII. A summary of our findings and an outlook on the method are given in the concluding Sec. VIII.

II. BACKGROUND FOR MODEL REDUCTION

Our innovations are built upon two key methods which are used for model reduction and approximating nonlinear dynamical systems. The first approach is the well-known POD-Galerkin method, which is used to reduce the dimension of systems in a principled way. However, computing the form of the nonlinearity in the reduced-order system is an expensive offline computation, as inner products of the full high-dimensional system must still be computed. Online evaluation of the nonlinear terms in the reduced-order model may remain expensive, as these typically involve dense matrix or tensor operations of the same order as the degree of nonlinearity. The second approach highlighted is the DEIM algorithm [5] which reduces the complexity of evaluating the nonlinear terms. In particular, it gives a principled way to sparsely sample the nonlinearity in order to approximate the nonlinear terms in a low-dimensional way.

A. POD

Consider a high-dimensional system of nonlinear differential equations that can arise, for example, from the finite-difference discretization of a partial differential equation:

$$\frac{d\mathbf{u}(t)}{dt} = L\mathbf{u}(t) + N(\mathbf{u}(t)), \tag{1}$$

where $\mathbf{u}(t) = [u_1(t)\ u_2(t)\ \cdots\ u_n(t)]^T \in \mathbb{R}^n$ and $n \gg 1$. Typically under discretization of a single spatial variable, $u_j(t) = u(x_j, t)$ is the value of the field of interest at the spatial location x_j. The linear part of the dynamics is given by $L \in \mathbb{R}^{n\times n}$ and the nonlinear terms are in the vector $N(\mathbf{u}(t)) = [N_1(\mathbf{u}(t))\quad N_2(\mathbf{u}(t))\quad \cdots\quad N_n(\mathbf{u}(t))]^T \in \mathbb{R}^n$. The nonlinear function is evaluated component-wise at the n spatial grid points used for discretization.

For achieving high-accuracy solutions, n is typically required to be a very large number, thus making the computation of the solution expensive and/or intractable. The POD-Galerkin method is a principled dimensionality-reduction scheme that approximates the function $\mathbf{u}(t)$ with rank-r-optimal basis functions where $r \ll n$. These optimal basis functions are computed from a singular value decomposition of a series of temporal snapshots of the nonlinear dynamical system. Specifically, suppose snapshots of the state, $\mathbf{u}(t_j)$ with $j = 1, 2, \cdots, p$, are collected. The snapshot matrix $\mathbf{X} = [\mathbf{u}(t_1)\ \mathbf{u}(t_2)\ \cdots\ \mathbf{u}(t_p)] \in \mathbb{R}^{n\times p}$ is constructed and the SVD of $\mathbf{X}$ is computed: $\mathbf{X} = \mathbf{\Phi\Sigma W}^*$. The r-dimensional basis for optimally approximating $\mathbf{u}(t)$ is given by the

TABLE I: DEIM algorithm for finding approximation basis for the nonlinearity and its interpolation indices.

DEIM algorithm	
Basis	
• collect data, construct snapshot matrix	$\mathbf{X} = [\mathbf{u}(t_1)\ \mathbf{u}(t_2)\ \cdots\ \mathbf{u}(t_p)]$
• construct nonlinear snapshot matrix	$\mathbf{N} = [N(\mathbf{u}(t_1))\ N(\mathbf{u}(t_2))\ \cdots\ N(\mathbf{u}(t_p))]$
• singular value decomposition of $\mathbf{N}$	$\mathbf{N} = \mathbf{\Xi}\mathbf{\Sigma}_N\mathbf{W}_N^*$
• construct approximating basis (first m columns)	$\mathbf{\Xi}_m = [\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_m]$
Interpolation Indices (Iteration Loop)	
• choose the first index (initialization)	$[\rho, \gamma_1] = \max\|\boldsymbol{\xi}_1\|$
• approximate $\boldsymbol{\xi}_j$ by $\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_{j-1}$ at indices $\gamma_1, ..., \gamma_{j-1}$	Solve for $\mathbf{c}$: $\mathbf{P}^T\boldsymbol{\xi}_j = \mathbf{P}^T\mathbf{\Xi}_{j-1}\mathbf{c}$ with $\mathbf{P} = [\mathbf{e}_{\gamma_1}\ \cdots\ \mathbf{e}_{\gamma_{j-1}}]$
• select γ_j and loop ($j = 2, 3, ..., m$)	$[\rho, \gamma_j] = \max\|\boldsymbol{\xi}_j - \mathbf{\Xi}_{j-1}\mathbf{c}\|$

first r columns of matrix $\mathbf{\Phi}$, denoted by $\mathbf{\Phi}_r$. Thus the POD-Galerkin approximation is given by

$$\mathbf{u}(t) \approx \mathbf{\Phi}_r\mathbf{a}(t) \tag{2}$$

where $\mathbf{a}(t) \in \mathbb{R}^r$ is the time-dependent coefficient vector and $r \ll n$. Plugging this modal expansion into the governing equation (1) and applying orthogonality (multiplying by $\mathbf{\Phi}_r^T$) gives the dimensionally reduced evolution

$$\frac{d\mathbf{a}(t)}{dt} = \mathbf{\Phi}_r^T L\mathbf{\Phi}_r\mathbf{a}(t) + \mathbf{\Phi}_r^T N(\mathbf{\Phi}_r\mathbf{a}(t)). \tag{3}$$

By solving this system of much smaller dimension, the solution of a high-dimensional nonlinear dynamical system can be approximated.

This standard POD procedure [7] has been a ubiquitous algorithm in the reduced-order modeling community. However, it also helps illustrate the need for innovations such as the DEIM, Gappy POD and/or MPE. Consider the nonlinear component of the low-dimensional evolution (3): $\mathbf{\Phi}_r^T N(\mathbf{\Phi}_r\mathbf{a}(t))$. For a simple nonlinearity such as $N(u(x,t)) = u(x,t)^3$, consider its impact on a spatially-discretized, two-mode POD expansion: $u(x,t) = a_1(t)\phi_1(x) + a_2(t)\phi_2(x)$. The algorithm for computing the nonlinearity would require the evaluation of

$$u(x,t)^3 = a_1^3\phi_1^3 + 3a_1^2a_2\phi_1^2\phi_2 + 3a_1a_2^2\phi_1\phi_2^2 + a_2^3\phi_2^3\,. \tag{4}$$

The dynamics of $a_1(t)$ and $a_2(t)$ would then be computed by projecting onto the low-dimensional basis set by taking the inner product of this nonlinear term with respect to both ϕ_1 and ϕ_2. Thus the number of computations not only doubles, but the inner products must be computed with the n-dimensional vectors. Methods such as the DEIM overcome this high-dimensional computation and instead produce an $O(r)$ dimensional evaluation of the nonlinear terms.

B. DEIM

As outlined in the previous section, the shortcomings of the POD-Galerkin method are generally due to the evaluation of the nonlinear term $N(\mathbf{\Phi}_r\mathbf{a}(t))$. To avoid this difficulty, the DEIM approximates $\mathbf{N} = N(\mathbf{\Phi}_r\mathbf{a}(t))$ through projection and interpolation instead of evaluating it directly. A considerable reduction in complexity is achieved by the DEIM because evaluating the approximate nonlinear term does not require a prolongation of the reduced state variables back to the original high dimensional state approximation required to evaluate the nonlinearity in the POD approximation. The DEIM therefore improves the efficiency of the POD approximation and achieves a complexity reduction of the nonlinear term with a complexity proportional to the number of reduced variables. The DEIM constructs these specially selected interpolation indices that specify an interpolation-based projection to provide a nearly ℓ_2 optimal subspace approximation to the nonlinear term without the expense of orthogonal projection [5].

In the DEIM, a low-rank representation of the nonlinearity, taken as snapshots in time, is computed from the singular value decomposition

$$\mathbf{N} = [N(\mathbf{u}_1)\ N(\mathbf{u}_2)\ \cdots\ N(\mathbf{u}_p)] = \mathbf{\Xi}\mathbf{\Sigma}_N\mathbf{W}_N^* \tag{5}$$

where the matrix $\mathbf{\Xi}$ contains the optimal (in an ℓ_2 sense) basis set for spanning the nonlinearity. Specifically, we consider the rank-m basis set $\mathbf{\Xi}_m = [\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_m]$ that approximates the nonlinear function ($m \ll n$ and $m \sim r$). The approximation to the nonlinearity $\mathbf{N}$ is given by:

$$\mathbf{N} \approx \mathbf{\Xi}_m\mathbf{c}(t) \tag{6}$$

where $\mathbf{c}(t)$ is similar to $\mathbf{a}(t)$ in (2). Since this is a highly overdetermined system, a suitable vector $\mathbf{c}(t)$ can be found by selecting only m rows of the system. The DEIM algorithm was specifically developed to identify which m rows to evaluate.

The DEIM algorithm begins by considering the vectors $\mathbf{e}_{\gamma_j} \in \mathbf{R}^n$ which are the γ_j-th column of the n dimensional identity matrix. We can then construct the projection matrix $\mathbf{P} = [\mathbf{e}_{\gamma_1}\ \mathbf{e}_{\gamma_2}\ \cdots\ \mathbf{e}_{\gamma_m}]$ which is chosen so that $\mathbf{P}^T\mathbf{\Xi}_m$ is nonsingular. Then $\mathbf{c}(t)$ is uniquely defined from

$$\mathbf{P}^T\mathbf{N} = \mathbf{P}^T\mathbf{\Xi}_m\mathbf{c}(t), \tag{7}$$

and thus,

$$\mathbf{N} \approx \mathbf{\Xi}_m(\mathbf{P}^T\mathbf{\Xi}_m)^{-1}\mathbf{P}^T\mathbf{N}. \tag{8}$$

The tremendous advantage of this result for nonlinear model reduction is that the term $\mathbf{P}^T\mathbf{N}$ requires evaluation of nonlinearity only at m indices, where $m \ll n$. The DEIM further proposes a principled method for choosing the basis vectors $\boldsymbol{\xi}_j$ and indices γ_j. The DEIM algorithm, which is based upon a greedy-like search, is detailed in [5] and further demonstrated in Table I.

C. Application to ROMs

The POD and DEIM provide a number of advantages for nonlinear model reduction of nonlinear dynamical systems. POD provides a principled way to construct an r-dimensional subspace $\boldsymbol{\Phi}_r$ characterizing the dynamics. The DEIM augments the POD method by providing a method to evaluate the problematic nonlinear terms using an m-dimensional subspace $\boldsymbol{\Xi}_m$ that represents the nonlinearity. Thus a small number of points, specifically m, can be sampled to approximate the nonlinear terms in the ROM.

The method proposed here capitalizes on these methods by building low-dimensional libraries associated with the full nonlinear system dynamics as well as the specific nonlinearities. Moreover, the sparse measurement locations computed by the DEIM are found to be highly effective for sensor placement. Such sensors, as will be shown in what follows, can be used with sparse representation and compressive sensing to (i) identify dynamical regimes, (ii) reconstruct the full state of the system, and (iii) provide an efficient nonlinear model reduction and POD-Galerkin prediction for the future state. Moreover, we show that point measurements of the nonlinearity of the dynamical system can be much more robust to noise for accomplishing the above tasks.

The concept of library building of low-rank "features" from data is well established in the computer science community. In the reduced-order modeling community, it has recently become an issue of intense investigation. Indeed, a variety of recent works, for instance from Amsallem, Charbel and co-workers [45, 46] and Peherstorfer and Willcox [47–49], have produced libraries of ROM models that can be selected and/or interpolated through measurement and classification (typically clustered with k-means type algorithms). Alternatively, cluster-based reduced order models use a k-means clustering to build a Markov transition model between dynamical states [50]. These recent innovations are similar to the ideas advocated here. However, the focus of this work is on determining nearly optimal sparse sensor locations that work across all the libraries. Further, we build two sets of libraries: one for the full dynamics and a second for the nonlinearity so as to make it computationally efficient with the DEIM strategy. Before these more formal techniques based upon machine learning were developed, it was already realized that parameter domains could be decomposed into subdomains and a local ROM/POD computed in each subdomain. Patera *et al.* [51] used a partitioning based on a binary tree whereas Amsallem *et al.* [52] used a Voronoi Tessellation of the domain. Such methods were closely related to the work of Du and Gunzburger [53] where the data snapshots were partitioned into subsets and multiple reduced bases computed. The multiple bases were then recombined into a single basis, so it doesn't lead to a library per se. For a review of these domain partitioning strategies, please see Ref. [54].

III. MODEL PROBLEM

One of the canonical nonlinear PDEs in mathematical physics and pattern forming systems is the Ginzburg-Landau (GL) equation and its many-variants [1]. It has been used to model a variety of physical systems from condensed matter to biological waves. Here we consider a variant of the GL equation arising in mode-locked laser theory that has cubic and quintic nonlinear terms and a fourth-order derivative [55]:

$$iU_t + \left(\frac{1}{2} - i\tau\right) U_{xx} - i\kappa U_{xxxx} + (1 - i\mu)|U|^2 U + (\nu - i\varepsilon)|U|^4 U - i\gamma U = 0, \tag{9}$$

where $U(x,t)$ is a complex valued function of space and time. Under discretization of the spatial variable, $U(x,t)$ becomes a vector $\mathbf{u}$ with n components, i.e. $\mathbf{u}_j(t) = U(x_j,t)$ with $j = 1, 2, \cdots n$.

An efficient and exponentially accurate numerical solution to (9) can be found using standard spectral methods [11]. Specifically, the equation is solved by Fourier transforming in the spatial dimension and then time-stepping with an adaptive 4th-order Runge-Kutta method. The extent of the spatial domain is $x \in [-20, 20]$ with $n = 1024$ discretized points. Note that in what follows, the indices for evaluation of the nonlinear term correspond to the collocation points away from the center spatial point of the computational domain so that x_0 is the 513th point in the domain. Here, we allow the parameters $\beta = (\tau, \kappa, \mu, \nu, \epsilon, \gamma)$ to vary in order to discover various dynamical regimes that exhibit low-rank structure and stable attractors. Table II shows six different parameter regimes that have unique low-dimensional attractors (see [10]). The evolution of the system for parameter regimes β_1, β_3 and β_5 is illustrated in Fig. 1. Such stereotypical low-dimensional behaviors, which are commonly observed in pattern forming systems [1], will serve as the basis for our library building methodology, especially in regards to using a small number of measurements to identify the β_j regime, reconstruct the solution, and project a future state. Although our results are demonstrated on this specific PDE, the methodology is quite general.

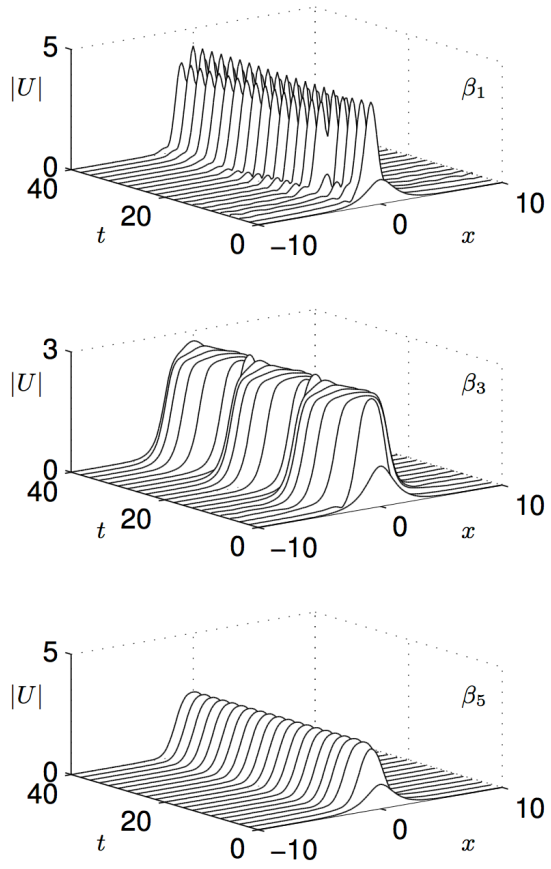

FIG. 1: Evolution dynamics of (9) for the parameter regimes β_1, β_3 and β_5 over the time interval $t \in [0, 40]$. The initial transients are quickly attenuated away, leaving the stable attractor for the given β_j regime. Sampling of the dynamics for library building occurs once the transients have decayed.

	τ	κ	μ	ν	ϵ	γ	description
β_1	-0.3	-0.05	1.45	0	-0.1	-0.5	3-hump, localized
β_2	-0.3	-0.05	1.4	0	-0.1	-0.5	localized, side lobes
β_3	0.08	0	0.66	-0.1	-0.1	-0.1	breather
β_4	0.125	0	1	-0.6	-0.1	-0.1	exploding soliton
β_5	0.08	-0.05	0.6	-0.1	-0.1	-0.1	fat soliton
β_6	0.08	-0.05	0.5	-0.1	-0.1	-0.1	dissipative soliton

TABLE II: Values of the parameters from equation (9) that lead to six distinct dynamical regimes. To exemplify our algorithm, the first, third and fifth regimes will be discussed in this paper.

IV. LIBRARIES

As can be seen from Fig. 1 and Table II, generic initial conditions evolve towards a variety of low-dimensional attractors. This suggests that each dynamic regime, with a given β_j, can be approximated by a small number of modes via a POD reduction. These modes will constitute our *library modes* in what follows. For each of the six regimes β_j in Table II, we build a library of POD modes. The number of POD modes r is selected to capture 99% of the total variance (energy). For the β_1, β_2, β_5 and β_6 regimes, only a single mode is required so that $r = 1$. For the β_3 regime $r = 6$, whereas for the β_4 regime, $r = 14$ in order to capture the fluctuations observed. Figure 3(a) illustrates the library POD modes in differing colors for all of the β_j regimes except β_4. The exclusion of the β_4 modes in this visualization is simply due to the large number ($r = 14$) necessary in comparison to the other dynamical regimes. As illustrated in Fig. 2, library building is the first step in a training module aimed at *learning* the low-rank dynamical behavior of a nonlinear dynamical system.

In practice, a dynamical system such as (9) may change over time due to evolution or modulation of the parameters β_j. Thus the dynamics may evolve from one attractor to another with some prescribed transition time (typically on the order of $O(1)$ time for (9)). One of the primary goals of this and previous [8, 56] work is to find *optimal and sparse sensor locations* whereby limited measurements of the system are taken in order to classify the dynamical regime. Interestingly, the previous efforts [8] used expert-in-the-loop knowledge to help select the optimal measurement positions. For the simple model considered here, such expert knowledge can be acquired from familiarity with the POD library modes and considering locations of maximal variance. However, for a more general system, this is a difficult task that could greatly benefit from a more principled mathematical approach. The DEIM algorithm will provide this approach. Moreover, as required by the DEIM, we also build low-rank libraries for the cubic and quintic terms associated with the dynamical regimes β_j. In doing so, we not only find effective sensor locations, but we also circumvent the computational difficulties of the POD in evaluating the nonlinear terms.

To library build, consider the following linear and nonlinear functions associated with the governing equations (9) for a given parameter regime β_j:

$$N_L(U) = U \tag{10a}$$
$$N_3(U) = |U|^2U \tag{10b}$$
$$N_5(U) = |U|^4U \tag{10c}$$
$$N_{NL}(U) = (i+\mu)|U|^2U + (i\nu+\epsilon)|U|^4U\,, \tag{10d}$$

where the second and third terms are the standard cubic and quintic nonlinearities of (9) and the last term enforces their prescribed relative weighting.

Associated with each nonlinearity (10) are a set of measurements and snapshot matrices. For a snapshot matrix sampled at p temporal locations $[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_p] \in \mathbb{R}^{n\times p}$, we can construct the nonlinear $\mathbb{R}^{n\times p}$snapshot matrices

$$\mathbf{N}_L = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_p] \tag{11a}$$
$$\mathbf{N}_3 = [N_3(\mathbf{u}_1)\ N_3(\mathbf{u}_2)\ \cdots\ N_3(\mathbf{u}_p)] \tag{11b}$$
$$\mathbf{N}_5 = [N_5(\mathbf{u}_1)\ N_5(\mathbf{u}_2)\ \cdots\ N_5(\mathbf{u}_p)] \tag{11c}$$
$$\mathbf{N}_{NL} = [N_{NL}(\mathbf{u}_1)\ N_{NL}(\mathbf{u}_2)\ \cdots\ N_{NL}(\mathbf{u}_p)]. \tag{11d}$$

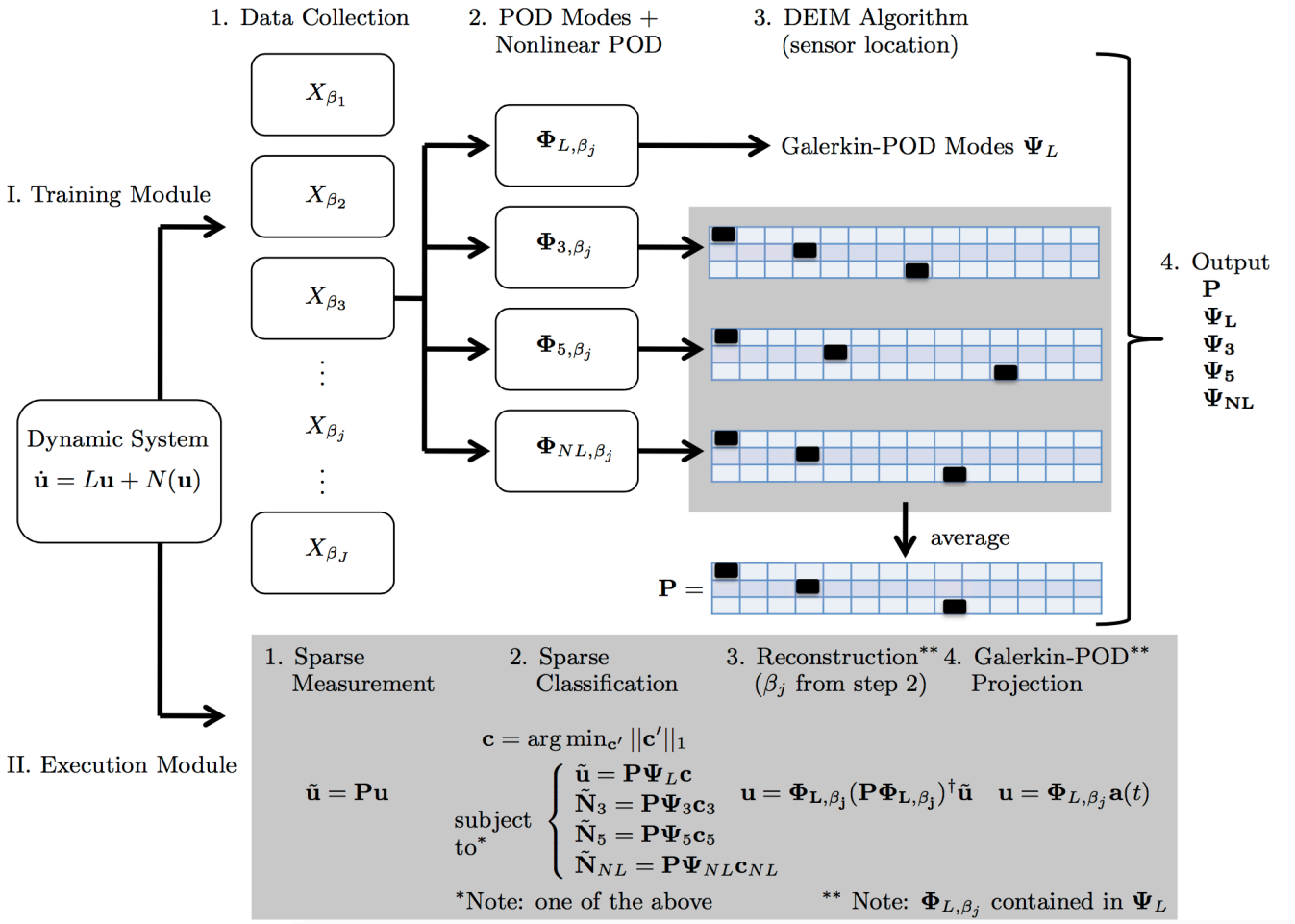

FIG. 2: Training and execution modules for the library learning and sensor location optimization with the DEIM. The training module samples the various dynamical regimes $(\beta_1, \beta_2, \cdots, \beta_J)$ through snapshots. For each dynamical regime, low-rank libraries are constructed for the nonlinearities of the nonlinear dynamical system $(\mathbf{\Phi}_{L,\beta_j}, \mathbf{\Phi}_{3,\beta_j}, \mathbf{\Phi}_{5,\beta_j}, \mathbf{\Phi}_{NL,\beta_j})$. The DEIM algorithm is then used to select sparse sampling locations and construct the projection matrix $\mathbf{P}$. The execution module uses the sampling locations to classify the dynamical regime β_j of the nonlinear dynamical system, reconstruct its full state $(\mathbf{u} = \mathbf{\Phi}_{\mathbf{L},\beta_\mathbf{j}}(\mathbf{P}\mathbf{\Phi}_{\mathbf{L},\beta_\mathbf{j}})^\dagger\tilde{\mathbf{u}})$, and provide a low-rank Galerkin-POD approximation for its future $(\mathbf{u} = \mathbf{\Phi}_{L,\beta_j}\mathbf{a}(t))$. Note that $(\mathbf{P}\mathbf{\Phi}_{\mathbf{L},\beta_\mathbf{j}})^\dagger$ denotes the Moore-Penrose pseudo-inverse of $(\mathbf{P}\mathbf{\Phi}_{\mathbf{L},\beta_\mathbf{j}})$.

The singular value decomposition of these matrices will give a basis for approximation of each of the nonlinearities for a given β_j as well as the standard snapshot matrix of POD. Specifically, the SVD gives the library of modes: $\mathbf{\Phi}_{L,\beta_j}$, $\mathbf{\Phi}_{3,\beta_j}$, $\mathbf{\Phi}_{5,\beta_j}$ and $\mathbf{\Phi}_{NL,\beta_j}$ (See Fig. 2).

The POD modes can be arranged in a collection of library elements, $\mathbf{\Psi}_L$, $\mathbf{\Psi}_3$, $\mathbf{\Psi}_5$ or $\mathbf{\Psi}_{NL}$, by concatenating the POD modes from each of the different β_j regimes. Thus the construction of multiple libraries would take the form

$$\mathbf{\Psi}_L = [\mathbf{\Phi}_{L,\beta_1}\ \mathbf{\Phi}_{L,\beta_2}\ \cdots\ \mathbf{\Phi}_{L,\beta_6}] \tag{12a}$$
$$\mathbf{\Psi}_3 = [\mathbf{\Phi}_{3,\beta_1}\ \mathbf{\Phi}_{3,\beta_2}\ \cdots\ \mathbf{\Phi}_{3,\beta_6}] \tag{12b}$$
$$\mathbf{\Psi}_5 = [\mathbf{\Phi}_{5,\beta_1}\ \mathbf{\Phi}_{5,\beta_2}\ \cdots\ \mathbf{\Phi}_{5,\beta_6}] \tag{12c}$$
$$\mathbf{\Psi}_{NL} = [\mathbf{\Phi}_{NL,\beta_1}\ \mathbf{\Phi}_{NL,\beta_2}\ \cdots\ \mathbf{\Phi}_{NL,\beta_6}]\,. \tag{12d}$$

The number of basis elements (rank) for the cubic and quintic terms in a given POD library coincides with the rank r required for each β_j, i.e. $r = m$. Note that the library $\mathbf{\Psi}_L$ is the library containing the POD modes used for POD-Galerkin projections of the future state. It is also the only library constructed in previous work [8, 9]. Figure 3(b,c) shows the cubic and quintic library modes for (9). They can be compared to the standard POD modes illustrated in Fig. 3(a). Although the modes look quite similar, we will show that the classification can be improved using the libraries of nonlinearities. Further, evaluation of the nonlinearities through the DEIM now remains a low-order computation.

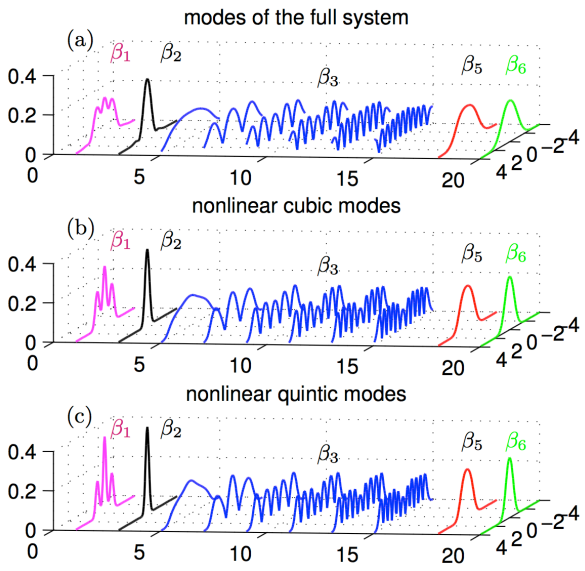

FIG. 3: Library modes for (a) the full system, (b) the cubic nonlinearity, and (c) the quintic nonlinearity. The modes are color coded by their dynamical regime from β_1 to β_6 as given in Table II. The rank-r for each library is chosen by selecting the modes that comprise 99% of the total variance for a given dynamical regime.

V. DEIM FOR SENSOR LOCATIONS

The idea of using a limited (sparse) number of sensors to characterize the dynamics has previously been considered in [8–10]. However, no algorithm was specified to determine the best locations for the sensors, although optimal sensor placement has been investigated in the context of categorical decisions [56]. Indeed, the previous work relied on expert-in-the-loop selection of the sensors in order to classify the dynamics. Interestingly, the DEIM algorithm gives a principled way to discretely and sparsely sample the nonlinearity in order to evaluate the various inner products for a POD reduction. These same DEIM spatial sampling locations make good sensor locations for classification and reconstruction. Since the interpolation indices from the DEIM algorithm [5] correspond to the entries with largest magnitude of the residual error between the chosen basis and its approximation at each step (see last line of the table I), it becomes interesting to see what the classification/reconstruction will be if we pick these locations for sensors. As demonstrated in Fig. 2, determining the sensor locations is part of a training module.

We apply the DEIM algorithm outlined in Table I on the nonlinear POD (SVD) library modes ($\mathbf{\Psi}_3$, $\mathbf{\Psi}_5$ or $\mathbf{\Psi}_{NL}$) computed from (11) and (12). The application of the algorithm yields the DEIM interpolation locations which we will call our *sensor locations*. Note that the indices indicate the distance away from the center of the computational grid. Thus $x_0 = 0$, $x_{\pm 1} = dx$, $x_{\pm 2} = 2dx$, etc. Or more generally, the index n corresponds to $x_n = n\,dx$. Thus the indices depend on the specific discretization of the domain. Sensor locations are computed for each of the nonlinearities: $\mathbf{\Phi}_{3,\beta_j}$, $\mathbf{\Phi}_{5,\beta_j}$ and $\mathbf{\Phi}_{NL,\beta_j}$ for $j = 1, 2, 3$. Each dynamical regime β_j and nonlinear library gives a unique set of sensor locations. Our goal is to evaluate the placement of 3 sensors. Table III and its accompanying figure gives a vector of the indices for the locations $\mathbf{x}_{\beta_j}$ of the 3 sensors found for three regimes β_1, β_3 and β_5 using the libraries $\mathbf{\Phi}_{3,\beta_j}$, $\mathbf{\Phi}_{5,\beta_j}$ and $\mathbf{\Phi}_{NL,\beta_j}$. Also represented are the 3 sensor locations when all three β_j regimes are combined into a single library, i.e. the best sensor locations for the combined dynamic library is identified. This regime is represented in Table III by $\mathbf{x}_{\beta_{\text{all}}}$.

	Cubic $\lvert U\rvert^2U$				Quintic $\lvert U\rvert^4U$				Nonlinear $N(U)$			
Sensor	$\mathbf{x}_{\beta_1}$	$\mathbf{x}_{\beta_3}$	$\mathbf{x}_{\beta_5}$	$\mathbf{x}_{\beta_{all}}$	$\mathbf{x}_{\beta_1}$	$\mathbf{x}_{\beta_3}$	$\mathbf{x}_{\beta_5}$	$\mathbf{x}_{\beta_{all}}$	$\mathbf{x}_{\beta_1}$	$\mathbf{x}_{\beta_3}$	$\mathbf{x}_{\beta_5}$	$\mathbf{x}_{\beta_{all}}$
one	0	0	0	0	0	0	0	0	0	9	0	0
two	5	15	12	6	4	13	10	6	6	21	6	6
three	13	26	17	22	13	23	15	20	13	32	15	13

TABLE III: Summary of sensor location vectors (indices for evaluation) from the DEIM algorithm. The table summarizes the findings from Fig. 4, giving precise grid cells to be used in evaluating the nonlinear inner products in the Galerkin-POD approximation.

Application of the DEIM algorithm results in the measurement matrix $\mathbf{P}$ of (8). For 3 sensors, generically it takes the form

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & \cdots & & & & & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & & \cdots & 0 \\ 0 & \cdots & & & \cdots & 0 & 1 & 0 \cdots & 0 \end{bmatrix} \tag{13}$$

where the specific columns containing the nonzero entries are given by the indices found from the DEIM and shown in Table III. More precisely, this matrix is *exactly* the output of the DEIM algorithm. In our scenario, the construction of the P matrix is made for each nonlinearity as well as for each dynamical regime β_j. This gives the sensor locations for the sparse sensing scheme presented in the next section. Figure 4 illustrates the locations of the sensors and the value of library modes at the prescribed locations for both the cubic and quintic nonlinearities. Figure 5 shows a histogram of the various sensor location evaluations. In particular, the three dominant locations are at the indices $n = 0, 6$ and 13. Thus they will be used in what follows.

VI. CLASSIFICATION

Our goal is to make use of recent innovations in sparse sampling and compressive sensing [35–42] for characterizing the nonlinear dynamical system [8–10]. Specifically, we wish to use a limited number of sensors for classifying the dynamical regime of the system. With this classification, a reconstruction of the full state space can

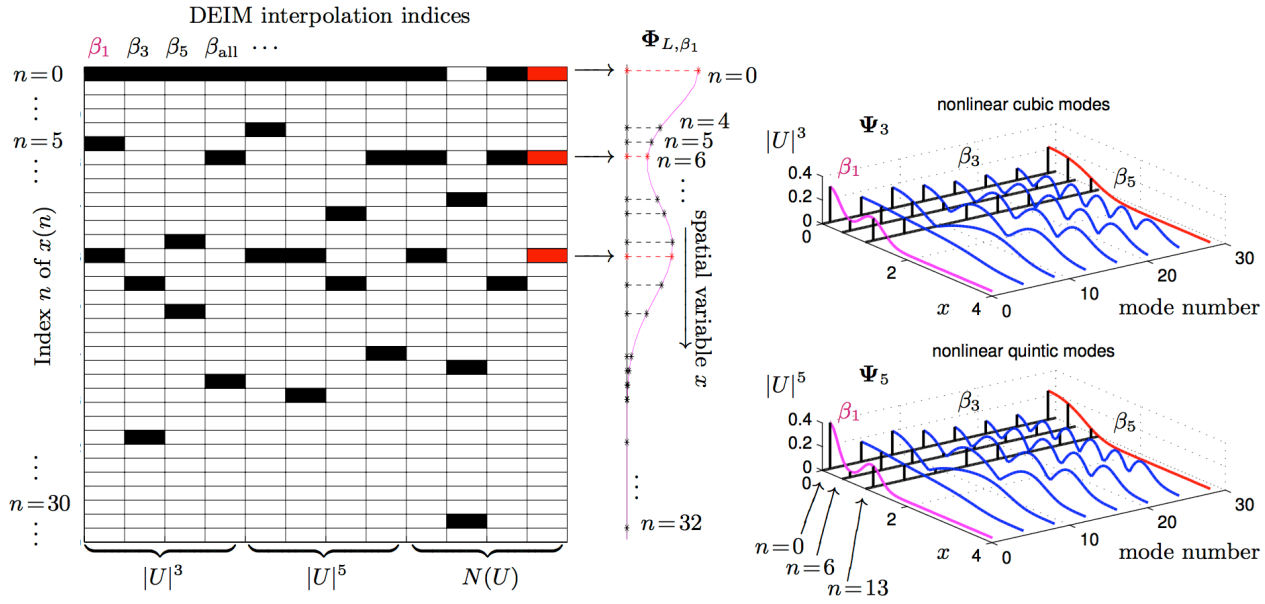

FIG. 4: Location of indices determined by the DEIM for the libraries of nonlinearities $|U|^3$, $|U|^5$ and $N(U)$. The spatial domain $x \in [-20, 20]$ is discretized on a periodic domain with $n = 1024$ points. The center point of the domain corresponds to $x(0) = 0$. The index values are the number of grid points ndx away from the center grid point, e.g. $x(5) = 5dx$. The left grid shows the location of the DEIM indices (black boxes) determined by the algorithm in Table I for the regimes β_1, β_3 and β_5 as well as the combination of all three regimes together β_{all}. The middle panel shows the library mode Φ_{L,β_1} (laid out vertically) as a function of the spatial variable $x(n)$. Indicated on this transverse mode are the measurement locations for the different DEIM nonlinearities and β_j regimes. The right two panels show the β_1, β_3 and β_5 modes with the black lines indicating the measurement locations for $n = 0, 6$ and 13. This allows one to visualize where the measurement occur on the mode structures.

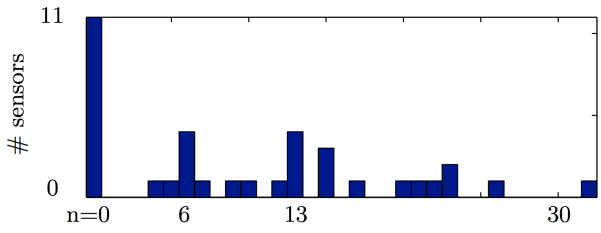

FIG. 5: Histogram of sensor placement locations based upon Fig. 4. The top three sensor location indices are located at $n = 0, 6$ and 13. Thus we use these locations in our simulations for classification and reconstruction.

be accomplished and a POD-Galerkin prediction can be computed for its future. In general, if we have a sparse measurement $\tilde{\mathbf{u}} \in \mathbf{R}^q$, where q is the number of measurements, then

$$\tilde{\mathbf{u}} = \mathbf{P}\mathbf{u}\,, \tag{14}$$

where $\mathbf{u}$ is the full state vector and $\mathbf{P}$ is the sampling matrix determined by the DEIM given by (13). In the previous section, we constructed the matrix $\mathbf{P}$ for $q = 3$.

The full state vector $\mathbf{u}$ can be approximated with the POD library modes ($\mathbf{u} = \mathbf{\Psi}_L\mathbf{c}$), therefore

$$\tilde{\mathbf{u}} = \mathbf{P}\mathbf{\Psi}_L\mathbf{c}, \tag{15}$$

where $\mathbf{\Psi}_L$ is the low-rank matrix whose columns are POD basis vectors concatenated across all β regimes and $\mathbf{c}$ is the coefficient vector giving the projection of $\mathbf{u}$ onto these POD modes. If $\mathbf{P}\mathbf{\Psi}_L$ obeys the restricted isometry property [57] and $\mathbf{u}$ is sufficiently sparse in $\mathbf{\Psi}_L$, then it is possible to solve the highly-underdetermined system (15) with the sparsest vector $\mathbf{c}$. Mathematically, this is equivalent to the optimization problem

$$\mathbf{c} = \arg\min_{\mathbf{c}'} ||\mathbf{c}'||_0, \quad \text{subject to} \quad \tilde{\mathbf{u}} = \mathbf{P}\mathbf{\Psi}_L\mathbf{c}.$$

Minimizing the l_0 norm is computationally an np-hard problem. However, It has been proven that under certain conditions, a sparse solution of equation (15) can be found by minimizing the l_1 norm instead [36, 38] so that

$$\mathbf{c} = \arg\min_{\mathbf{c}'} ||\mathbf{c}'||_1, \quad \text{subject to} \quad \tilde{\mathbf{u}} = \mathbf{P}\mathbf{\Psi}_L\mathbf{c}. \tag{16}$$

The last equation can be solved through standard convex optimization methods such as the CVX package for Matlab.

To classify the dynamical regime from limited measurements $\tilde{\mathbf{u}}$ (specifically 3 spatial measurements), we use the

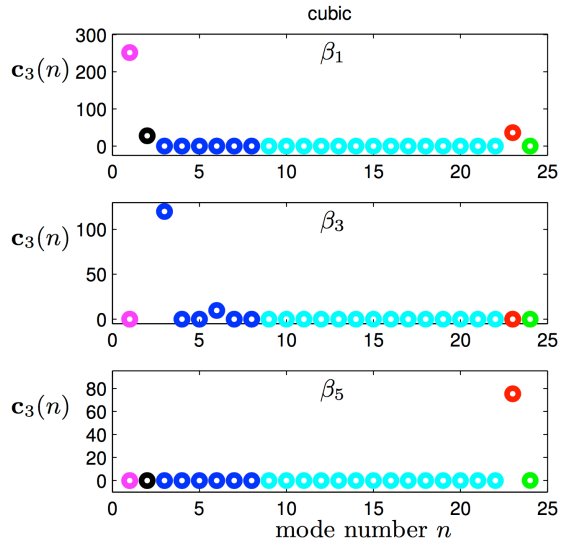

FIG. 6: The values of the 24×1 projection vector $\mathbf{c}$ from solving using a cubic measurement $\tilde{\mathbf{u}}_3 = |\tilde{\mathbf{u}}|^2\tilde{\mathbf{u}}$ and the cubic library $\mathbf{\Psi}_3$ in (17a). The three panels show the dominant vector component to be in the β_1, β_3 and β_5 regime respectively, thus showing that it correctly identifies each dynamical regime from 3 measurement locations. The values of the colored circles correspond to the expression strength of the different library elements of Fig. 3.

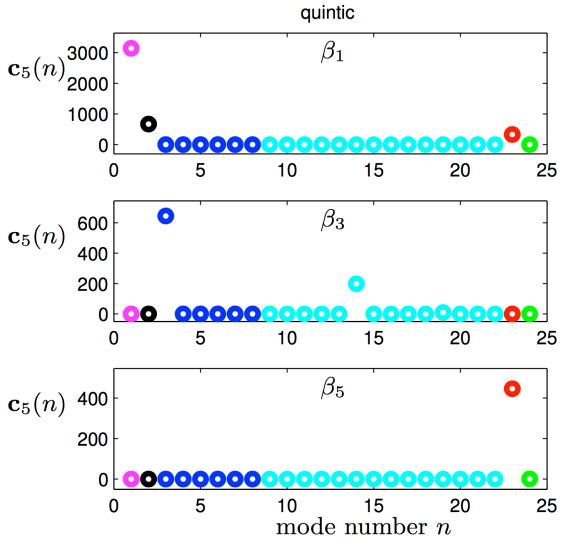

FIG. 7: The values of the 24×1 projection vector $\mathbf{c}$ from solving using a quintic measurement $\tilde{\mathbf{u}}_5 = |\tilde{\mathbf{u}}|^4\tilde{\mathbf{u}}$ and the quintic library $\mathbf{\Psi}_5$ in (17b). The three panels show the dominant vector component to be in the β_1, β_3 and β_5 regime respectively, thus showing again that point measurements of the nonlinearity correctly identify each dynamical regime from 3 measurement locations. The values of the colored circles correspond to the expression strength of the different library elements of Fig. 3.

sensor locations matrix $\mathbf{P}$ found from the DEIM on the libraries of nonlinearities. Here, the sensor locations used for $\mathbf{P}$ are from all the library elements combined and the nonlinearity $N(U)$ (See the last column in Table III remarked with red boxes), i.e. $n = 0, 6$ and 13. Suppose we have a linear measurement $\tilde{\mathbf{u}}$, then we can construct the vectors $\tilde{\mathbf{u}}_3 = |\tilde{\mathbf{u}}|^2\tilde{\mathbf{u}}$ and $\tilde{\mathbf{u}}_5 = |\tilde{\mathbf{u}}|^4\tilde{\mathbf{u}}$ and classify them using the libraries of nonlinearities. Specifically, the nonlinear classification is accomplished with:

$$\mathbf{c}_3 = \arg\min_{\mathbf{c}_3'} ||\mathbf{c}_3'||_1, \quad \text{subject to} \quad \tilde{\mathbf{u}}_3 = \mathbf{P}\mathbf{\Psi}_3\mathbf{c}_3 \tag{17a}$$

$$\mathbf{c}_5 = \arg\min_{\mathbf{c}_5'} ||\mathbf{c}_5'||_1, \quad \text{subject to} \quad \tilde{\mathbf{u}}_5 = \mathbf{P}\mathbf{\Psi}_5\mathbf{c}_5\,. \tag{17b}$$

Figures 6 and 7 show the coefficient vectors $\mathbf{c}_3$ and $\mathbf{c}_5$ respectively for measurements performed in the β_1, β_3 and β_5 regimes. The vectors $\mathbf{c}_3$ and $\mathbf{c}_5$ clearly act as accurate indicator functions for the dynamical regime, better than even simple linear measurements siads,Bright:2013. Indeed, the DEIM algorithm for sensor location does as well as expert-in-the-loop selections [8–10], but requires no extensive and pre-existing knowledge about the dynamical libraries. We can also make a categorical decision, with similar results, about the dynamical regime the dynamics belongs to by computing error of projection onto a given library and considering which has the smallest error. This is the same as sparse representation used for image classification [34].

The above analysis assumes that there is no noise in the measurements or the system itself. However, most sensors are subject to noise fluctuations which can impact the ability of a scheme such as this to correctly identify β_j. As a consequence, we also perform the classification task with noisy data. First, assume that we collect linear measurements which have additive noise. Denote this data by

$$\bar{\mathbf{u}} = \tilde{\mathbf{u}} + \mathcal{N}(0, \sigma^2) \tag{18}$$

where $\mathcal{N}(0, \sigma^2)$ is a Gaussian distributed noise term with variance σ^2 .

In order to evaluate the classification, we need to once again compute the nonlinear terms and run the optimization algorithm for computing the library coefficients and the associated dynamical regime. The statistical result for 400 trials when $\sigma = 0.2$ is shown in Table IV. One can see that the noise introduces misclassification errors to the original 100%-accurate classification scheme. However, multiple measurements still give an accurate classification overall with the exception of using the quintic library in the β_3 regime.

Interestingly, if point measurements of the nonlinearity are considered, then the results can improve drastically. For instance, in optics, measurements (full field or point measurements) are made of the intensity of the field rather than the field itself. This represents a simple form of a nonlinear measurement. Thus consider the

β_1 regime	β_1	β_2	β_3	β_4	β_5	β_6
$\|\bar{u}\|^2\bar{u}$	**98.75**	0	1.25	0	0	0
$\|\bar{u}\|^4\bar{u}$	**91**	6.5	2.5	0	0	0
$\bar{u}_3$	**100**	0	0	0	0	0
$\bar{u}_5$	**100**	0	0	0	0	0
β_3 regime	β_1	β_2	β_3	β_4	β_5	β_6
$\|\bar{u}\|^2\bar{u}$	2.5	0	**61.75**	18	17.5	0.25
$\|\bar{u}\|^4\bar{u}$	5.5	0	**38**	34.5	21.75	0.25
$\bar{u}_3$	0	0	**100**	0	0	0
$\bar{u}_5$	0	0	**100**	0	0	0
β_5 regime	β_1	β_2	β_3	β_4	β_5	β_6
$\|\bar{u}\|^2\bar{u}$	5.25	0.75	7.5	5	**62**	19.5
$\|\bar{u}\|^4\bar{u}$	6.75	2	6.25	2.5	**61.25**	21.25
$\bar{u}_3$	0	0	0	0	**100**	0
$\bar{u}_5$	0	0	0	0	**100**	0

TABLE IV: Classification accuracy with noisy measurements ($\sigma = 0.2$) using 400 realizations in the β_1, β_3 and β_5 regimes. The accuracy of classification for the correct regime is denoted by the bold numbers, whereas the other percentages denote to what extent and where misclassifications occur. The accuracy of the classification schemes are evaluated using linear measurements ($\bar{u}$ in (18)) with the cubic and quintic libraries illustrated in Figs. 6 and 7. Also shown are classification results using point measurements of the nonlinearity ($\bar{\mathbf{u}}_3$ and $\bar{\mathbf{u}}_5$ in 19). Point measurements of the nonlinearity, if possible, offer significant accuracy improvement and robustness to noise.

nonlinear point (spatial) measurements subject to noise:

$$\bar{\mathbf{u}}_3 = |\tilde{\mathbf{u}}|^2\tilde{\mathbf{u}} + \mathcal{N}(0,\sigma^2) \tag{19a}$$
$$\bar{\mathbf{u}}_5 = |\tilde{\mathbf{u}}|^4\tilde{\mathbf{u}} + \mathcal{N}(0,\sigma^2)\,. \tag{19b}$$

The classification results for this case are also shown in Table IV. Note the clear improvement (100% accuracy) in using point measurements of the nonlinearity for classification tasks. Thus if the noise is driven by the sensor itself, then point measurements of the nonlinearity may be quite advantageous.

VII. RECONSTRUCTION AND THE GALERKIN-POD APPROXIMATION

The classification step of the last section identifies the dynamical regime of the nonlinear dynamical system by using sparsity promoting ℓ_1 optimization on the learned libraries. Once the correct β_j regime is determined, reconstruction of the solution and a future state prediction can be achieved through the POD-Galerkin approximation. Specifically, once the dynamical regime β_j has been identified, then a subset of modes $\mathbf{\Psi}_L \to \mathbf{\Phi}_{L,\beta_j}$ form the correct modal basis for a POD-Galerkin approximation.

To be more precise, recall that only a limited number of measurements are made as in (14). But now $\mathbf{u} = \mathbf{\Phi}_{L,\beta_j}\mathbf{c}$ where the vector $\mathbf{c}$ is now the projection onto the smaller set of library modes associated with a single β_j. Thus instead of (15), we now we have

$$\tilde{\mathbf{u}} = \mathbf{P}\mathbf{\Phi}_{L,\beta_j}\mathbf{c}\,. \tag{20}$$

Unlike the classification step, we can now determine $\mathbf{c}$ by simply solving the above equation using a standard Moore-Penrose pseudo-inverse operator † [58] so that $\mathbf{c} = (\mathbf{P}\mathbf{\Phi}_{L,\beta_j})^\dagger\tilde{\mathbf{u}}$, i.e. it solves for $\mathbf{c}$ by minimizing the ℓ_2 norm. With $\mathbf{c}$ determined, the reconstruction of the solution thus follows:

$$\mathbf{u} = \mathbf{\Phi}_{L,\beta_j}(\mathbf{P}\mathbf{\Phi}_{L,\beta_j})^\dagger\tilde{\mathbf{u}} \tag{21}$$

This is the reconstruction of the system given the sparse measurement vector $\tilde{\mathbf{u}}$ and a classification β_j. The POD-Galerkin approximation for the future state can then be accomplished by using (3) and with the DEIM algorithm for evaluating the nonlinearities (8). The initial condition for the POD-Galerkin is given from (21). Thus as advocated in previous work [8, 9], accurate classification is accomplished with ℓ_1 optimization (decoding) while the more standard ℓ_2 norm is used for reconstruction and POD-Galerkin projection (encoding). Figure 2 illustrates the execution state outlined here for classification, reconstruction and projection.

The computational efficiency of the proposed method can be evaluated. Of course, there is no computational savings in the training stage of the algorithm. Indeed, there is some computational overhead associated with building the libraries (an SVD evaluation of $O(N^3)$) and running the DEIM algorithm (of $O(N)$). However, once the training stage is done, then the compressive sensing, which is a small ℓ_1 optimization procedure can be used to identify the correct POD modes and project into the future with a Galerkin-POD approximation. In the case of the CQGLE considered here, this provides a computational savings of three orders of magnitude, both in computational time and memory requirements. This highlights the potentially transformative use of reduced-order models in computational physics and for simulations of multi-scale nonlinear dynamical systems.

VIII. CONCLUSIONS AND OUTLOOK

In conclusion, we advocate a general theoretical framework for nonlinear dynamical systems whereby low-rank libraries representing the optimal modal basis are constructed, or learned, from snapshot sampling of the dynamics. In order to make model reduction methods such as POD computationally efficient, especially in evaluating the nonlinear terms of the governing equations, libraries of nonlinearities are also constructed during the learning stage. This allows for the application of the discrete empirical interpolation method which identifies a limited number of spatial sampling locations that can

allow for reconstruction of the nonlinear terms in a low-dimensional manner. Such sparse sampling of the nonlinearity is directly related to compressive sensing strategies whereby a small number of sensors can be used to characterize the dynamics of the nonlinear system. Indeed, the POD method, when combined with the DEIM and compressive sensing, can (i) correctly identifying the dynamical parameter regime, (ii) reconstruct the full state dynamics and (iii) produce a low-rank prediction of the future state of the nonlinear dynamical system. All of these tasks are accomplished in a low-dimensional way, unlike standard POD-Galerkin models whose nonlinearities can prove to be computationally inefficient.

To be more precise about our learning algorithm for the nonlinear dynamical system, we construct the library modes representing the dynamics by the ℓ_2-optimal proper orthogonal decomposition. Several libraries are constructed: one for linear snapshot measurements, one for each nonlinear term, and one which combines all the nonlinear terms together with their prescribed weightings. The DEIM algorithm then allows us to identify sparse measurement locations capable of both classifying the dynamics regime of the nonlinear dynamical system and efficiently evaluating the inner products of the nonlinear terms for a POD-Galerkin projection of the system. Indeed, the dynamical state is identified from limited noisy measurements using the sparsity promoting ℓ_1 norm and the compressive sensing architecture. The strategy for building modal libraries by concatenating truncated POD libraries across a range of relevant bifurcation parameters may be viewed as a simple dimensionality-reduction implementation of machine learning. The resulting modal libraries are a natural sparse basis for the application of compressive sensing. After the expensive one-time library-building procedure, accurate identification, projection, and reconstruction may be performed entirely in a low-dimensional framework.

The method is effective for nonlinear dynamical systems where POD approximations are relevant. Thus it can be applied only to systems where low-dimensional attractors are the key dynamical features of interest. Specifically, it fails if such attractors are not present, if wave propagation is the dominant dynamical feature (although modifications exist to account for this), and/or intrinsically high-dimensional systems such as turbulent flows. More broadly, the methodology can also be extended to complex systems that exhibit low-dimensional attractors in their repertoire of dynamical behaviors. As an example, one need only consider encoding schemes in neuro-sensory systems whereby the collective behavior of networked neurons produce emergent low-dimensional patterns of activity for given stimulus [17, 59].

With three DEIM determined sensor locations, it is possible to accurately classify bifurcation regimes, reconstruct the low-dimensional content, and simulate the Galerkin projected dynamics of the complex Ginzburg Landau equation. In addition, we investigate the performance of sparse representation with the addition of sensor noise. For moderate noise levels, the method accurately classifies the correct dynamic regime. Point measurements of the nonlinearity dramatically improve the classification procedure. Interestingly, the DIEMs algorithm not only provides efficient sensor positioning, it also helps perform POD-Galerking truncations in a fully low-rank manner, thus avoiding the computational expense of evaluating nonlinear terms using the POD methodology. Overall, the combination of ℓ_2 low-rank representations and ℓ_1 sparse sampling enables efficient characterization and manipulation of low-rank dynamical systems.

For modern nonlinear dynamical systems, it is known that nonlinearity plays a dominant role and shapes the underlying spatio-temporal dynamics and modal structures, thus necessitating a new approach, such as that presented here, for extracting these critical structures. As has been demonstrated, although nonlinearity drives new modal structures, it does not destroy the underlying low-dimensional nature of the dynamics. Methods that take advantage of such underlying structure are critical for developing theoretical understanding and garnering insight into the fundamental interactions of a vast array of physical, engineering and biological systems.

Acknowledgements

We are grateful for discussions with David Amsallem, Ido Bright, Bingni W. Brunton, Kevin Carlberg, Xing Fu, Josh Proctor and Jonathan Tu. J. N. Kutz acknowledges support from the U.S. Air Force Office of Scientific Research (FA9550-09-0174).

[1] M. Cross and P. Hohenberg. Pattern formation out of equilibrium. *Reviews of Modern Physics*, 65:851–1112, 1993.

[2] A. Quarteroni and G. Rozza Eds. *Reduced Order Methods for Modeling and Computational Reduction*, (Springer, 2014)

[3] P. Benner, S. Gugercin and K. Willcox, "A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems," SIAM Review , to appear, 2015.

[4] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, "An 'empirical interpolation' method: Application to efficient reduced-basis discretization of partial differential equations," C. R. Math. Acad. Sci. Paris, 339 (2004), pp. 667?672.

[5] S. Chaturantabut, D. Sorensen, "Nonlinear Model Re-

duction via Discrete Empirical Interpolation," SIAM J. SCI. COMPUT. **32**, 2737-2764 (2010).

[6] J. L. Lumley. *Stochastic Tools in Turbulence*. Academic Press, 1970.

[7] P. J. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge Monographs in Mechanics. Cambridge University Press, Cambridge, England, 2nd edition, 2012.

[8] S. L. Brunton, J. H. Tu, I. Bright, J. N. Kutz, "Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems," SIAM J. App. Dyn. Sys., **13**(4): 1716–1732, 2014.

[9] I. Bright, G. Lin, and J. N. Kutz. Compressive sensing and machine learning strategies for characterizing the flow around a cylinder with limited pressure measurements. *Physics of Fluids*, 25:127102–1–127102–15, 2013.

[10] J.L. Proctor, S.L. Brunton, B.W. Brunton and J.N. Kutz "Exploiting sparsity and equation-free architectures in complex systems," European Journal of Physics, **223**: 2665–2684, 2014.

[11] J. N. Kutz. *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*. Oxford University Press, 2013.

[12] L.M. Jones, A. Fontanini, B.F. Sadacca, P. Miller, and D.B. Katz, "Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles," Proc. Natl. Acad. Sci. **104**:18772?18777 (2007).

[13] M. Rabinovich, A. Volkovskii, P. Lecanda, R. Huerta, H.D.I. Abarbanel, and G. Laurent, "Dynamical encoding by networks of competing neuron groups: Winnerless competition," Phys. Rev. Lettl. **87**:068102 (2001).

[14] M. Rabinovich, R. Huerta, P. Varona, and V.S. Afraimovich. "Transient cognitive dynamics, metastability, and decision making," PLoS Comput. Biol. **4**:e1000072 (2008).

[15] G. Laurent, M. Stopfer, R.W. Friedrich, M.I. Rabinovich, A. Volkovskii, and H.D.I. Abarbanel, "Odor encoding as an active, dynamical process: Experiments, computation and theory," Annu. Rev. Neurosci. **24**:263?297 (2001).

[16] J.I. Gold and M.N. Shadlen, "Representation of a perceptual decision in developing oculomotor commands," Nature **404**:390?394 (1999).

[17] E. Shlizerman, J. Riffell, and J. N. Kutz, "Data-driven inference of network connectivity for modeling the dynamics of neural codes in the insect antennal lobe," Front. Neuro. **18** (2014) article 70, 1-15.

[18] S. Ganguli and H. Sompolinsky, "Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis," Annual Review of Neuroscience **35**:485?508 (2012).

[19] B. Olshausen and D. Field, "Sparse coding of sensory inputs," Current Opinion in Neurobiology, **14**:481? 487 (2004).

[20] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(7–12):559–572, 1901.

[21] E. N. Lorenz. Empirical orthogonal functions and statistical weather prediction. Technical report, Massachusetts Institute of Technology, December 1956.

[22] H. Hotelling. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24:417–441, September 1933.

[23] H. Hotelling. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24:498–520, October 1933.

[24] M. Gavish and D. L. Donoho. The optimal hard threshold for singular values is $4/\sqrt{3}$. *ArXiv e-prints*, 2014.

[25] J. E. Fowler. Compressive-projection principal component analysis. *IEEE Transactions on Image Processing*, 18(10):2230–2242, 2009.

[26] A. C. Gilbert, J. Y. Park, and M. B. Wakin. Sketched SVD: Recovering spectral features from compressive measurements. *ArXiv e-prints*, 2012.

[27] H. Qi and S. M. Hughes. Invariance of principal components under low-dimensional random projection of the data. IEEE International Conference on Image Processing, October 2012.

[28] R. Everson and L. Sirovich, "Karhunen-Loéve procedure for gappy data," J. Opt. Soc. Am. A **12**, 1657-1664 (1995).

[29] K. Willcox, "Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition," Computers and Fluids **35**: 208-226 (2006).

[30] B. Yildirim, C. Chryssostomidis and G.E. Karniadakis, "Efficient sensor placement for ocean measurements using low-dimensional concepts," Ocean Modeling, **273**(3-4), 160-173, (2009).

[31] N.C. Nguyen, A. T. Patera, J. Peraire, "A "best points" interpolation method for efficient approximation of parametrized functions." Int. J. Num. Methods Eng. **73**, 521–543 (2008).

[32] P. Astrid, "Fast reduced order modeling technique for large scale LTV systems," in Proc. 2004 Am. Control Conf. **1**, 762-767 (2004).

[33] K. Carlberg, C. Farhat, J. Cortial, and D. Amsallem. The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows. *Journal of Computational Physics*, 242:623–647, 2013.

[34] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 31(2):210–227, 2009.

[35] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[36] D. L. Donoho. "For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution." Communications on pure and applied mathematics, **59**(6):797-829, 2006.

[37] E. J. Candès. Compressive sensing. *Proceedings of the International Congress of Mathematics*, 2006.

[38] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[39] E. J. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications in Pure and Applied Mathematics*, 8(1207–1223), 59.

[40] E. J. Candès and T. Tao. Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006.

[41] R. G. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, 24(4):118–120, 2007.

[42] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde. Model-based compressive sensing. *IEEE Transactions on*

Information Theory, 56(4):1982–2001, 2010.

[43] H. Schaeffer, R. Caflisch, C. D. Hauck, and S. Osher. Sparse dynamics for partial differential equations. *Proceedings of the National Academy of Sciences USA*, 110(17):6634–6639, 2013.

[44] A. Mackey, H. Schaeffer, and S. Osher. On the compressive spectral method. *UCLA CAM Report* 14–33, 2014.

[45] D. Amsallem, R. Tezaur and C. Farhat "Real-time solution of computational problems using databases of parametric linear reduced-order models with arbitrary underlying meshes," arXiv:1506.07153 (2015).

[46] Y. Choi, D. Amsallem and C. Farhat "Gradient-based Constrained Optimization Using a Database of Linear Reduced-Order Models," arXiv:1506.07849 (2015).

[47] B. Peherstorfer and K. Willcox, "Online Adaptive Model Reduction for Nonlinear Systems via Low-Rank Updates," SIAM Journal on Scientific Computing , to appear, 2015.

[48] B. Peherstorfer and K. Willcox, "Dynamic data-driven reduced-order models," Computer Methods in Applied Mechanics and Engineering," **291**, 21-41 (2015).

[49] B. Peherstorfer and K. Willcox, "Detecting and Adapting to Parameter Changes for Reduced Models of Dynamic Data-driven Application Systems," Procedia Computer Science **51**, 2553-2562, (2015).

[50] E. Kaiser, B. R. Noack, L. Cordier, A. Spohn, M. Segond, M. Abel, G. Daviller, J. Osth, S. Krajnovic and R. K. Niven, "Cluster-based reduced-order modelling of a mixing layer," J. Fluid Mech. **754**, 365–414, (2014).

[51] J. L. Eftang, A. T. Patera, and E. M. Rnquist, "An hp certified reduced-basis method for parameterized elliptic PDEs," Siam SISC, 2010.

[52] D. Amsallem, J. Cortial, and C. Farhat. "On demand CFD-based aeroelastic predictions using a database of reduced-order bases and models," AIAA Conference, 2009.

[53] Q. Du and M. Gunzburger, "Model reduction by proper orthogonal decomposition coupled with centroidal Voronoi tessellation," 2002.

[54] D. Amsallem, M. J. Zahr and K. Washabaugh, "Fast local reduced basis updates for the efficient reduction of nonlinear systems with hyper-reduction," Advances in Computational Mathematics, February 2015, DOI 10.1007/s10444-015-9409-0

[55] J. N. Kutz, "Mode-locked soliton lasers," SIAM Rev. **48**:629-678, 2006.

[56] B. W. Brunton, S. L. Brunton, J. L. Proctor, and J. N. Kutz. Optimal sensor placement and enhanced sparsity for classification. *ArXiv e-prints*, 2014.

[57] E. J. Candes and T. Tao. Decoding by linear programming. Information Theory, IEEE Transactions on, 51(12):4203-4215 (2005).

[58] N. Trefethen and D. Bau III, *Numerical Linear Algebra* (SIAM, Philadelphia, 1997).

[59] J. Kunert, E. Shlizerman and J. N. Kutz, "Low-dimensional functionality of complex network dynamics: Neuro-sensory integration in the Caenorhabditis elegans connectome," Phys. Rev. E **89**, 052805 (2014).