Compressive Sensing and Low-Rank Libraries for Classification of Bifurcation Regimes in Nonlinear Dynamical Systems[*]

Steven L. Brunton[†], Jonathan H. Tu[‡], Ido Bright[†], and J. Nathan Kutz[†]

Abstract. We show that for complex nonlinear systems, model reduction and compressive sensing strategies can be combined to great advantage for classifying, projecting, and reconstructing the relevant low-dimensional dynamics. L_2-based dimensionality reduction methods such as the proper orthogonal decomposition are used to construct separate modal libraries and Galerkin models based on data from a number of bifurcation regimes. These libraries are then concatenated into an overcomplete library, and L_1-sparse representation in this library from a few noisy measurements results in correct identification of the bifurcation regime. This technique provides an objective and general framework for classifying the bifurcation parameters and, therefore, the underlying dynamics and stability. After classifying the bifurcation regime, it is possible to employ a low-dimensional Galerkin model, only on modes relevant to that bifurcation value. These methods are demonstrated on the complex Ginzburg–Landau equation using sparse, noisy measurements. In particular, three noisy measurements are used to accurately classify and reconstruct the dynamics associated with six distinct bifurcation regimes; in contrast, classification based on least-squares fitting (L_2) fails consistently.

1. Introduction. Nonlinear dynamical systems are ubiquitous in characterizing the behavior of physical, biological, and engineering systems. With few exceptions, nonlinearity impairs our ability to construct analytically tractable solutions, and we instead rely on experiments and high-performance computation to study a given system. Numerical discretization can often yield a system of equations with millions or billions of degrees of freedom. Thus, both simulations and experiments can generate enormous datasets that strain computational resources and confound one's understanding of the underlying dynamics. Fortunately, many high-dimensional systems exhibit dynamics that evolve on a slow-manifold and/or a low-dimensional attractor (e.g., pattern forming systems [16]). We propose a data-driven modeling strategy that represents low-dimensional dynamics using *dimensionality reduction*

methods such as proper orthogonal decomposition (POD) [26] and classifies/reconstructs the observed low-dimensional manifolds with *compressive (sparse) sensing* [17, 12, 5, 52]. Thus dynamic structures are represented efficiently with the L_2 norm and identified from sparse measurements with the L_1 norm.

The application of machine learning and compressive sensing to dynamical systems is synergistic in that underlying low-rank structures facilitate sparse measurements [31]. This combination has the potential to transform a number of challenging fields. Such a strategy may enhance nonlinear estimation and control, where real-time analysis is critical. Moreover, adaptive time-stepping algorithms can take advantage of the low-dimensional embedding for greatly reduced computational costs [30, 42]. Additionally, the interplay of sparsity and complex systems has been investigated with the goal of overcoming the curse of dimensionality associated with neuronal activity and neurosensory systems [22]. Compressive sensing may also play a role in similar statistical learning, library-based, and/or information theory methods [15, 7] used in fluid dynamics [8, 2], climate science [24, 7], and oceanography [1]. Indeed, compressive sensing is already playing a critical role in model building and assessment in the physical sciences [35, 45, 53, 48]. These challenging open problems would benefit from a paradigm shift in modeling and analysis, whereby low-dimensional coherence is leveraged for use with sparse sampling techniques.

1.1. Challenges of POD-Galerkin models across parameter regimes. Galerkin-POD is a well-known [26] dimensionality reduction method for complex systems. In the context of fluid dynamics, Galerkin projection of the Navier–Stokes equations onto a truncated POD mode basis is an effective method of model-order reduction, resulting in a system of ordinary differential equations (ODEs). These equations may be orders of magnitude more computationally efficient than the full simulation. However, Galerkin projection onto POD modes obtained across a range of parameter values, the so-called global POD [49, 44], often results in unstable and/or inaccurate models. There have been a number of modifications to POD-Galerkin models that seek to address this issue, but it remains a major challenge of low-order modeling in fluids.

A modified method that uses interpolated angles of multiple POD subspaces has been demonstrated to capture F-16 parameterized dynamics [32]. Including additional modes, such as the shift mode [36], to capture transients between qualitatively different flow regimes has resulted in additional methods such as double POD [46], the Gauss–Newton with approximated tensors (GNAT) method [14], and trust-region POD [19, 6]. Alternative methods for stabilizing POD by adding additional modes and closure terms have been investigated [3, 37]. In each case, the objective is to construct a dimensionally reduced set of dynamics that accurately represents the underlying complex system and that does not suffer from instabilities. There are additional methods based on POD manifolds that are useful for multiple parameter regimes [50, 51]. A recent method that combines transition matrix models with dynamic regimes that cluster, the so-called cluster reduced-order modeling (CROM), is also interesting for these multiregime problems [29].

1.2. Current approach. To avoid a single POD-Galerkin model defined across dynamical regions, we instead develop a classification scheme to determine which dynamic region our system is in, and then use a Galerkin model defined only on modes in that region. The

procedure advocated here involves two main steps. First, a modal library is constructed that is representative of a number of distinct dynamical regimes; i.e., the low-dimensional attractors are approximated by their optimal bases. Second, compressive sensing techniques are applied using this learned library. The goals are threefold: (1) classify the dynamic regime, (2) project the measurements onto the correct modal amplitudes, and (3) reconstruct the low-dimensional dynamics through the efficient Galerkin projection [26]. Here we concatenate POD bases to construct the library, although generalizing the library building strategy into a broader machine learning context [18] is interesting and may yield even more efficient strategies. There are many ways to build a library, especially considering the three goals above. In this case, we keep distinct POD bases for each dynamic regime, since this is better for the Galerkin projection step. The classification scheme, using L_1 minimization in an overcomplete library, is closely related to sparse representation from image classification [54].

The paper is outlined as follows. In section 2 a brief review is provided of the compressive sensing architecture and its relationship to L_1 convex optimization. Also reviewed are the basic ideas behind the POD for L_2 dimensionality reduction. These methods are combined in section 3 to form the key contributions of this work. Namely, the L_2 norm provides the sparse basis modes used by the L_1 norm for sparse representation. Section 4 demonstrates the use of these techniques on one of the classical models of mathematical physics: the Ginzburg–Landau equation. An outlook of the advantages and general applicability of the method to complex systems is given in the concluding section 5.

2. Background. In the following subsections, we introduce two well-established techniques that will be combined in this paper. The first method is compressive sensing, whereby a signal that is sparse in some basis may be recovered using proportionally few measurements by solving for the L_1-minimizing solution to an underdetermined system. The second method is the POD, which allows a dataset to be reduced optimally in an L_2 sense.

Both theories have been applied to a range of problems. In this paper, we advocate combining these methods since the L_2 basis obtained from POD is a particularly good choice of a sparse basis for compressive sensing. The underlying reason for this is that the data is obtained from the low-dimensional attractors of the governing complex system.

2.1. L_1-based sparse sensing. Consider a high-dimensional measurement vector $\mathbf{U} \in \mathbb{R}^n$, which is sparse in some space, spanned by the columns of a matrix $\mathbf{\Psi}$:

$$\mathbf{U}(x,t) = \mathbf{\Psi a}. \tag{2.1}$$

Here, sparsity means that $\mathbf{U}$ may be represented in the transform basis $\mathbf{\Psi}$ by a vector of coefficients $\mathbf{a}$ that contains mostly zeros. More specifically, *K-sparsity* means that there are K nonzero elements. In this sense, sparsity implies that the signal is *compressible*.

Consider a sparse measurement $\hat{\mathbf{U}} \in \mathbb{R}^m$, with $m \ll n$:

$$\hat{\mathbf{U}} = \mathbf{\Phi U}, \tag{2.2}$$

where $\mathbf{\Phi}$ is a measurement matrix that maps the full-state measurement $\mathbf{U}$ to the compressed measurement vector $\hat{\mathbf{U}}$. Details of this measurement matrix will be given shortly. Plugging (2.1) into (2.2) yields an underdetermined system:

$$\hat{\mathbf{U}} = \mathbf{\Phi \Psi a}. \tag{2.3}$$

We may then solve for the sparsest solution $\mathbf{a}$ to the underdetermined system of equations in (2.3). The L_0 norm measures how many nonzero elements a vector has (i.e., the cardinality of the vector); therefore, it measures sparsity. Solving for the solution $\mathbf{a}$ that has the smallest $|\mathbf{a}|_0$ norm is a combinatorially hard problem. However, this problem may be relaxed to a convex problem, whereby the $|\mathbf{a}|_1$ norm is minimized, which may be solved in polynomial time [12, 17]. The L_1 norm is the sum of the absolute values of every element of a vector, and under some conditions, minimizing this norm will yield results similar to those obtained when minimizing the L_0 norm. The specific minimization problem is

$$\arg\min |\hat{\mathbf{a}}|_1 \text{ such that } \boldsymbol{\Phi}\boldsymbol{\Psi}\hat{\mathbf{a}} = \hat{\mathbf{U}}.$$

There are other algorithms that result in sparse solution vectors, such as orthogonal matching pursuit [52].

This procedure, known as *compressive sensing*, is a recent development that has had widespread success across a range of problems. There are technical issues that must be addressed. For example, the number of measurements m in $\hat{\mathbf{U}}$ should be on the order of $K \log(n/K)$, where K is the degree of sparsity of $\mathbf{a}$ in $\boldsymbol{\Psi}$ [10, 11, 4]. In addition, the measurement matrix $\boldsymbol{\Phi}$ must be *incoherent* with respect to the sparse basis $\boldsymbol{\Psi}$, meaning that the columns of $\boldsymbol{\Phi}$ and the columns of $\boldsymbol{\Psi}$ are uncorrelated. Interestingly, significant work has gone into demonstrating that Bernouli and Gaussian random measurement matrices are almost certainly incoherent with respect to a given basis [13].

Typically a generic basis such as Fourier or wavelets is used in conjunction with sparse measurements consisting of random projections of the state. However, in many engineering applications, it is unclear how random projections may be obtained without first starting with a dense measurement of the state. In this work, we constrain the measurements to be point measurements of the state, so that $\boldsymbol{\Phi}$ consists of rows of a permutation matrix. Our primary motivation for such point measurements arises from physical considerations in such applications as ocean or atmospheric monitoring where point measurements are physically relevant. Moreover, sparse sensing is highly desirable as each measurement device is often prohibitively expensive, thus motivating many of our efforts in using sparse measurements to characterize the complex dynamics.

2.2. L_2-based dimensionality reduction. POD [34, 26] is a tool with ubiquitous use in dimensionality reduction of physical systems.[1] Data snapshots $\mathbf{U}(x,t_1), \mathbf{U}(x,t_2), \ldots, \mathbf{U}(x,t_q)$ are collected into columns of a matrix $\mathbf{A} \in \mathbb{R}^{n \times q}$. We then compute the singular value decomposition (SVD) of $\mathbf{A}$:

$$\mathbf{A} = \boldsymbol{\Psi}\boldsymbol{\Sigma}\mathbf{W}^*.$$

Columns of the matrix $\boldsymbol{\Psi}$ are POD modes,[2] and they are ordered according to the variance that they capture in the data $\mathbf{A}$; if the columns of $\mathbf{A}$ are velocity measurements, then the POD modes are ordered in terms of the energy that they capture. This variance/energy content

---

[1]POD is sometimes referred to as principal component analysis [38], the Karhunen–Loève decomposition, empirical orthogonal functions [33], or the Hotelling transform [27].

[2]Often, POD modes are given by the matrix $\boldsymbol{\Phi}$. However, we choose $\boldsymbol{\Psi}$ for the POD basis and $\boldsymbol{\Phi}$ for the sparse measurement matrix for consistency with the compressive sensing literature. This is not to be confused with notation from balanced POD, where $\boldsymbol{\Phi}$ are direct modes and $\boldsymbol{\Psi}$ are adjoint modes [40].

is quantified by the entries of the diagonal matrix $\mathbf{\Sigma}$, which are called singular values and appear in descending order.

When the size of each snapshot (n) is much larger than the number of snapshots (q) collected, i.e., $n \gg q$, as in high-dimensional fluid systems, there are at most q nonzero singular values, and it is beneficial to use the *method of snapshots* [47]. In this method, we solve the following eigenvalue problem:

$$\mathbf{A}^* \mathbf{A} \mathbf{W} = \mathbf{W} \mathbf{\Sigma}_q^2,$$

where $\mathbf{\Sigma}_q$ is the $q \times q$ upper-left block of $\mathbf{\Sigma}$. It is then possible to find the first q POD modes corresponding to nontrivial singular values by

$$\mathbf{\Psi}_q = \mathbf{A} \mathbf{W} \mathbf{\Sigma}_q^{-1}.$$

The snapshots often exhibit low-dimensional phenomena, so that the majority of the variance/energy is contained in a few modes, smaller than the number of snapshots collected. In this case, the POD basis is typically truncated at a predetermined cut-off value, such as when the columns contain 99% of the variance, so that only the first r modes 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 [23]. The SVD acts as a filter, and so the truncated modes often correspond to random fluctuations and disturbances. If the data in the matrix $\mathbf{A}$ is generated by a dynamical system (nonlinear system of ODEs of order n), it is then possible to substitute the truncated POD expansion for the state $\mathbf{U}$ into the governing equation and obtain Galerkin projected dynamics on the r basis modes [26]. Recall that we are assuming that the complex system under consideration exhibits low-dimensional attractors; thus the Galerkin truncation with only a few modes should provide an accurate prediction of the evolution of the system. 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 [20, 25, 39].

3. Methods—Combining L_1 and L_2. The major contribution of this work is the combination of library building techniques (depicted schematically in Figure 1) based on the L_2-optimal POD with the L_1-based compressive sensing architecture (depicted schematically in Figure 2) for classification and reconstruction.

Consider a complex system that evolves according to the partial differential equation (PDE)

$$(3.1) \qquad \mathbf{U}_t = \mathbf{N}(\mathbf{U}, \mathbf{U}_x, \mathbf{U}_{xx}, \dots, x, t, \beta),$$

where $\mathbf{U}(x,t)$ is a vector of physically relevant quantities and the subscripts t and x denote partial differentiation in time and space, respectively. Note that higher spatial dimensions may be considered without loss of generality. The function $\mathbf{N}(\cdot)$ can be a complicated, nonlinear function of the quantity $\mathbf{U}$, its derivatives, and both space and time. The parameter β is a bifurcation parameter with respect to which the solution of the governing PDE changes markedly. We assume a spatial discretization of (3.1), which yields a high-dimensional system of degree n.
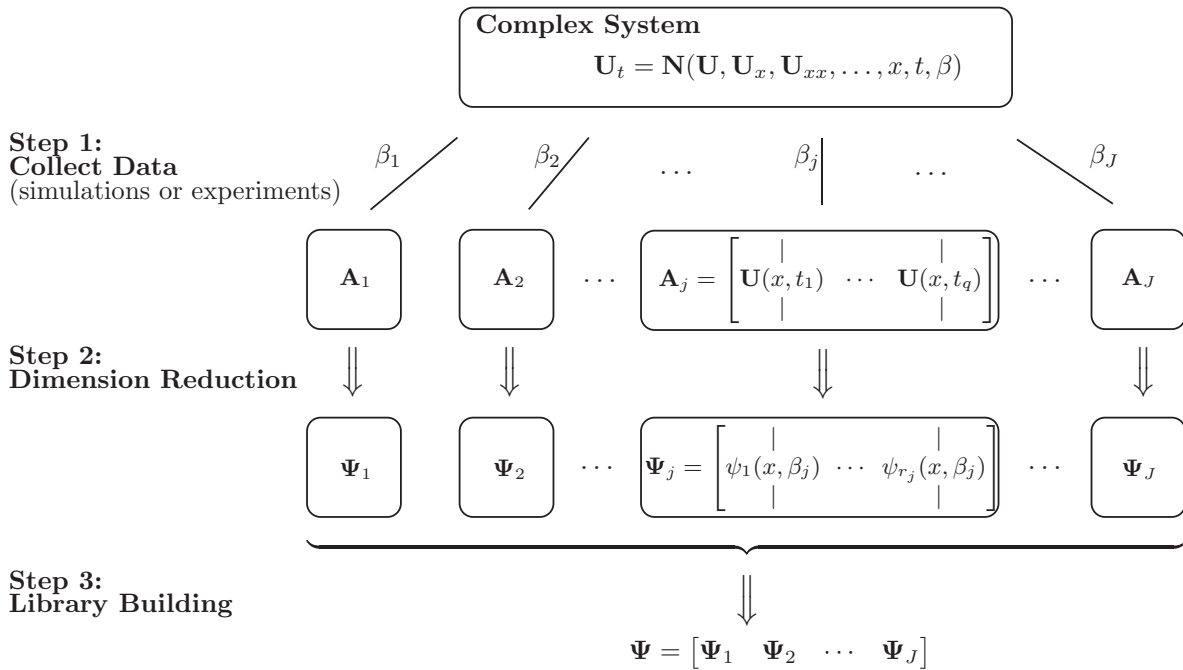
Complex System

$$\mathbf{U}_t = \mathbf{N}(\mathbf{U}, \mathbf{U}_x, \mathbf{U}_{xx}, \ldots, x, t, \beta)$$

Step 1:
Collect Data
(simulations or experiments)

β_1 $\quad$ β_2 $\quad$ $\ldots$ $\quad$ β_j $\quad$ $\ldots$ $\quad$ β_J

$$\mathbf{A}_1 \quad \mathbf{A}_2 \quad \cdots \quad \mathbf{A}_j = \begin{bmatrix} | & & | \\ \mathbf{U}(x, t_1) & \cdots & \mathbf{U}(x, t_q) \\ | & & | \end{bmatrix} \quad \cdots \quad \mathbf{A}_J$$

Step 2:
Dimension Reduction

$$\mathbf{\Psi}_1 \quad \mathbf{\Psi}_2 \quad \cdots \quad \mathbf{\Psi}_j = \begin{bmatrix} | & & | \\ \psi_1(x, \beta_j) & \cdots & \psi_{r_j}(x, \beta_j) \\ | & & | \end{bmatrix} \quad \cdots \quad \mathbf{\Psi}_J$$

Step 3:
Library Building

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{\Psi}_1 & \mathbf{\Psi}_2 & \cdots & \mathbf{\Psi}_J \end{bmatrix}$$

Figure 1. *Schematic of L_2 strategy for library building. Data $\mathbf{A}_j$ is collected for many values of the bifurcation parameter β_j, and the principal components of this data are computed and truncated in $\mathbf{\Psi}_j$. Although each basis $\mathbf{\Psi}_j$ is truncated to contain only the most energetically relevant structures, the concatenated library $\mathbf{\Psi}$ is overcomplete.*

We would like to use measurements of the system (3.1) to determine its state. However, full-state measurements are impractical for the high-dimensional system generated by discretization. Instead, m measurements are taken, where $m \ll n$; thus the measurements are *sparse*. In this paper, we consider *spatially localized* or *point* measurements, $\hat{\mathbf{U}}$, as discussed in section 2.1. In this case, the matrix $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ from (2.2) is composed of rows of the identity matrix corresponding to the measurement locations. These m-dimensional sparse observations are used to reconstruct the full n-dimensional state vector $\mathbf{U}$.

Our approach is to *learn* a library of low-rank dynamical approximations in which the dynamics are *sparse* and then apply compressive sensing to reconstruct the dynamics from $m \ll n$ measurements. First, we explore the full system (3.1) and collect dense measurements for various values $\beta_1, \beta_2, \ldots, \beta_J$ of interest, making sure to cover a number of unique dynamical regimes. For each case, snapshots of data from simulations or experiments are taken at a number of instances in time and organized into a data matrix describing the evolution of the full-state system:

$$\mathbf{A}_j = \begin{bmatrix} | & | & & | \\ \mathbf{U}(x, t_1) & \mathbf{U}(x, t_2) & \cdots & \mathbf{U}(x, t_q) \\ | & | & & | \end{bmatrix},$$

where q is the number of snapshots taken.

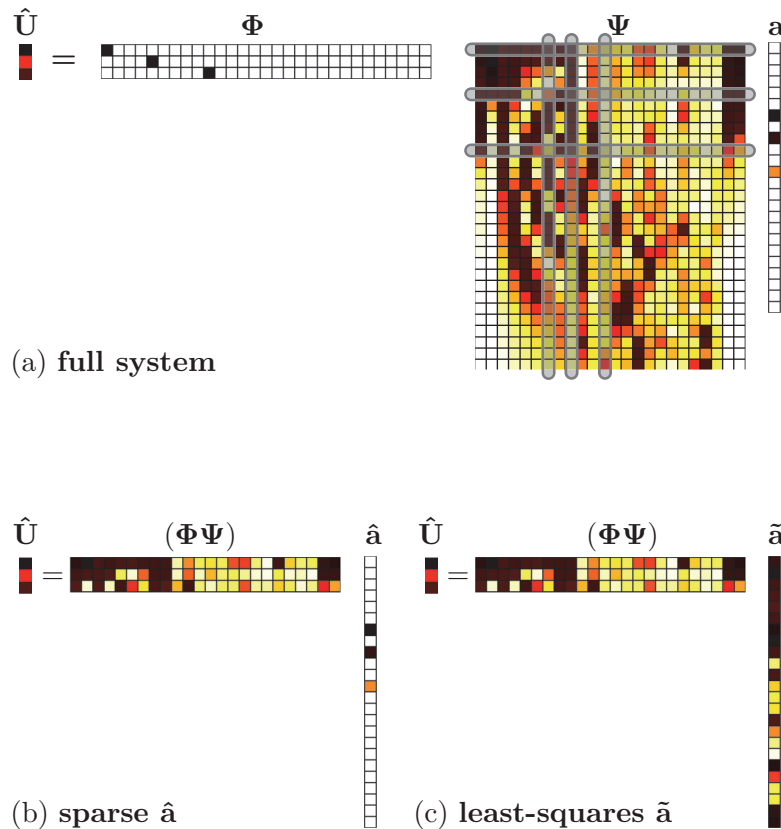Once the data matrix is constructed for a given β_j, its POD modes, or principal compo-

Figure 2. *Schematic of identification of sparse mode amplitudes* $\hat{\mathbf{a}}$ *by* L_1 *minimization.* (a) *Illustration of measurement matrix* $\boldsymbol{\Phi}$ *and sparse basis* $\boldsymbol{\Psi}$. *The underdetermined matrix* $(\boldsymbol{\Phi}\boldsymbol{\Psi})$ *admits a sparse solution* $\hat{\mathbf{a}}$ (b) *and a least-squares solution* $\tilde{\mathbf{a}} = (\boldsymbol{\Psi}^*\boldsymbol{\Phi}^*\boldsymbol{\Phi}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^*\boldsymbol{\Phi}^*\hat{\mathbf{U}}$ (c). *This type of diagram was introduced by Baraniuk in* [4]. *The data used in this figure is from the cubic-quintic Ginzburg–Landau equation (CQGLE) system in section* 4.

nents, $\boldsymbol{\Psi}_j$, are identified through an SVD: $\boldsymbol{\Psi}_j = \{\psi_i(x, \beta_j)\}_{i=1}^{r_j}$. The POD modes are orthogonal and ordered by energy content. The number of modes retained, r_j, is determined by a cut-off criterion; for instance, one might specify that modes comprising 99% of the energy be kept, or a hard threshold may be implemented [23].

With the modes identified for each β_j, an overcomplete library $\boldsymbol{\Psi}$ is constructed that contains all of the low-rank approximations for each dynamic regime:

$$(3.2) \qquad \boldsymbol{\Psi} = \begin{bmatrix} | & & | & & | & & | \\ \psi_1(x,\beta_1) & \cdots & \psi_{r_1}(x,\beta_1) & \cdots & \psi_1(x,\beta_J) & \cdots & \psi_{r_J}(x,\beta_J) \\ | & & | & & | & & | \end{bmatrix}.$$

The library $\boldsymbol{\Psi} \in \mathbb{R}^{n \times p}$ contains the representative low-rank modes for all of the dynamical behavior of the governing system that we explored in simulations or experiments. This is the *supervised learning* portion of the analysis, resulting in a small number ($p \ll n$) of library elements; note that $p = \sum_{j=1}^{J} r_j$. The p library modes are *not* orthogonal, but rather come in

groups of POD modes for each different β_j. The dynamics at any given time will belong to a specific β_j regime so that the instantaneous dynamics are *sparse* in the library basis, allowing for a sparse representation [54]. This *overcomplete library building* procedure is summarized in Figure 1.

With the library (3.2), we can expand the state $\mathbf{U}$ using the low-rank POD representation

$$(3.3) \qquad \mathbf{U}(x,t) = \sum_{j=1}^{J} \sum_{r=1}^{r_j} a_{jr}(t)\psi_r(x,\beta_j) = \mathbf{\Psi a}.$$

The solution is now represented in the p library elements constructed for the various values of β, and by construction, we expect $\mathbf{a}$ to be sparse in the basis $\mathbf{\Psi}$. This is because for any particular β_j, only a small subset of library elements is required to represent the solution.

Equation (3.3) is of the form in (2.1). To determine the vector $\mathbf{a}$ from a sparse data measurement $\hat{\mathbf{U}} = \mathbf{\Phi U}$, insert (3.3) into (2.2) and solve the underdetermined linear system $\hat{\mathbf{U}} = (\mathbf{\Phi\Psi})\mathbf{a}$ from (2.3) which has m equations (constraints) and p unknowns (modal coefficients), with $m \ll p$. We solve for a sparse $\hat{\mathbf{a}}$ using compressive sensing (L_1 minimization). This approach is natural because it promotes sparsity, an expected property of $\mathbf{a}$. Further, solving for $\mathbf{a}$ using L_1 minimization in the reduced-order library basis is significantly more efficient than solving for $\mathbf{U}$ in the full space since $p \ll n$. The sparsity-promoting compressed sensing procedure is illustrated in Figure 2.

The library construction (a one-time cost) and sparse sensing combine to give an efficient algorithm for approximating the low-rank dynamics of the full PDE (3.1) using a limited number of sensors and an empirically determined, overcomplete database. Specifically, the full state of the system $\mathbf{U}$ at any given time t is achieved by evaluating $\mathbf{a}$. There are a number of immediate advantages to this method for characterizing complex dynamical systems:

(i) Once the library is constructed from extensive simulations, future prediction of the system is efficient since the correct POD modes for any dynamical regime β_j have already been computed.

(ii) The algorithm works equally well with experimental data in an equation-free context, for instance, by using dynamic mode decomposition [41, 43] or equation-free modeling [30] in place of POD.

(iii) Given the low-rank space in which the algorithm works, it is ideal for use with control strategies, which are practical only for real-time application with low-dimensional systems.

Sparse sensing is significantly less expensive in the learned library $\mathbf{\Psi}$ since the high-dimensional state has been replaced with a truncated POD representation. Additionally, less information is required to categorize a signal than is required to fully reconstruct the signal, as in the compressive sensing paradigm. This combination of classification and reconstruction in a concatenated set of truncated POD bases using L_1 minimization is appealing on a number of levels. There is also a benefit to keeping the individual POD bases $\mathbf{\Psi}_j$ for reconstruction once the bifurcation regime β has been identified.

4. Results. To illustrate the aforementioned strategy, consider the complex Ginzburg–Landau model [16], which is ubiquitous in mathematical physics. Here it is modified to

<div align="center">

Table 1

Parameter regimes β_j for the complex Ginzburg–Landau equation (4.1) *(see Figure* 3*). The low-rank approximations of these parameter regimes are used to construct the elements of the library* $\boldsymbol{\Psi}$*.*

</div>

	τ	κ	μ	ν	ε	γ	Description
β_1	-0.3	-0.05	1.45	0	-0.1	-0.5	three-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

include both quintic terms and a fourth-order diffusion term much like the Swift–Hohenberg equation:

$$(4.1) \qquad i\mathbf{U}_t + \left(\frac{1}{2} - i\tau\right)\mathbf{U}_{xx} - i\kappa\mathbf{U}_{xxxx} + (1 - i\mu)|\mathbf{U}|^2\mathbf{U} + (\nu - i\varepsilon)|\mathbf{U}|^4\mathbf{U} - i\gamma\mathbf{U} = 0,$$

where $\mathbf{U}(x,t)$ is a complex function of space and time. Interesting solutions to this governing equation abound, characterized by the parameter values $\beta = (\tau, \kappa, \mu, \nu, \varepsilon, \gamma)$. In particular, we consider six regimes that illustrate different dynamical behaviors, described in Table 1. The CQGLE is useful for describing different qualitative phenomena associated with ultrashort pulse lasers. Both the fourth-order diffusion and the quintic term serve to regularize the dynamics, providing a broader range of interesting parameter regimes that remain stable. Note that there are many interesting regimes in the complex Ginzburg–Landau equation [28].

Figure 3 illustrates the corresponding low-rank behavior produced in the simulations. Figure 4 shows the singular values when POD is performed on the whole library, concatenated from each regime. The rank required for 99% is $r = 15$, which is similar to the β_4 regime. However, this library is not useful for Galerkin projection, since it is not stable for any of the regimes. This is a typical problem of applying the standard POD-Galerkin method across a range of parameters. However, there are advanced methods based on POD manifolds that are flexible for multiple parameter regimes [50, 51].

It is important to note that the six regimes described above are representative for characterizing the behavior of a mode-locked laser. However, these regimes are not exhaustive, and there are many additional parameter values that exhibit interesting phenomena. It is important to note that even if the library is incomplete to begin with, it is possible to augment it on-the-fly when entirely new dynamic regimes are sampled [9, 21].

As is common in many complex dynamical systems, especially those of a dissipative nature, low-dimensional attractors are embedded in the high-dimensional space. The simulations from each of these dynamic regimes exhibit low-dimensional structures which are spontaneously formed from generic, localized initial data. The low-dimensional structures allow for the low-rank POD approximations used in the library construction of Figure 5, as described in (3.2) and Figure 1.

To highlight the role of compressive sensing in identification and reconstruction for dynamical systems, we allow the bifurcation parameter $\beta = \beta(t)$ to vary in time so that the dynamics switch between attractors as β changes. Consider an example where $\beta = \beta_1$ for
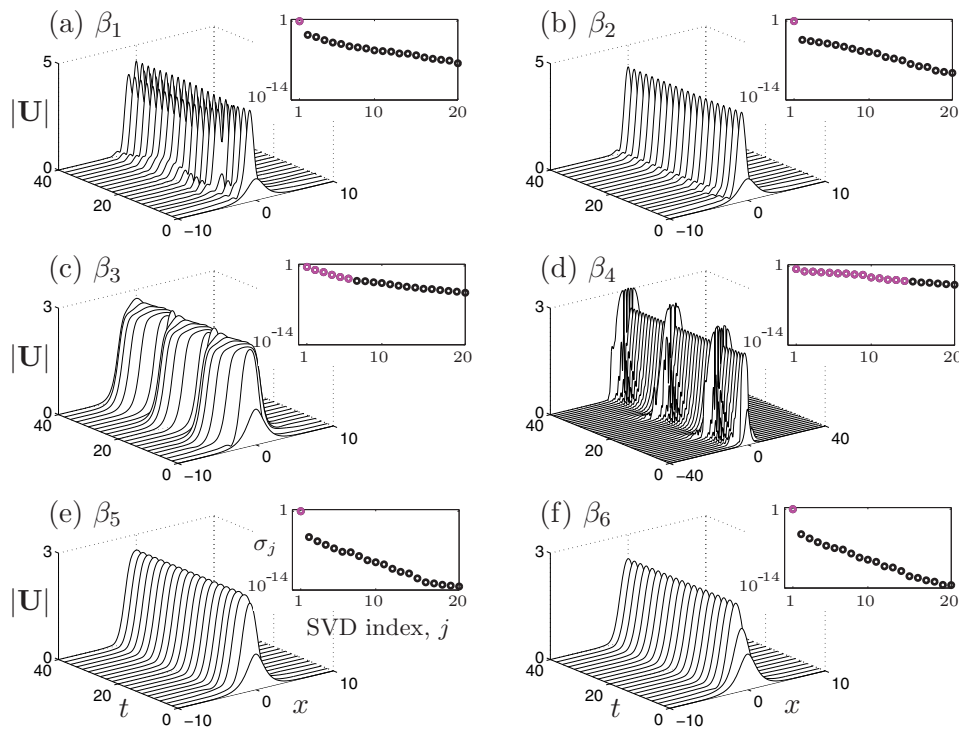
Figure 3. *Evolution dynamics of* (4.1) *for the six parameter regimes given in Table* 1*: (a) β_1, (b) β_2, (c) β_3, (d) β_4, (e) β_5, and (f) β_6. All parameter regimes exhibit stable, low-dimensional attractors as evidenced by the singular values (inset). The SVD sampling occurs for every $\Delta t = 1$ in the interval $t \in [40, 80]$. Magenta circles represent the modes that comprise 99% of the data and are used for the library $\mathbf{\Psi}$.*
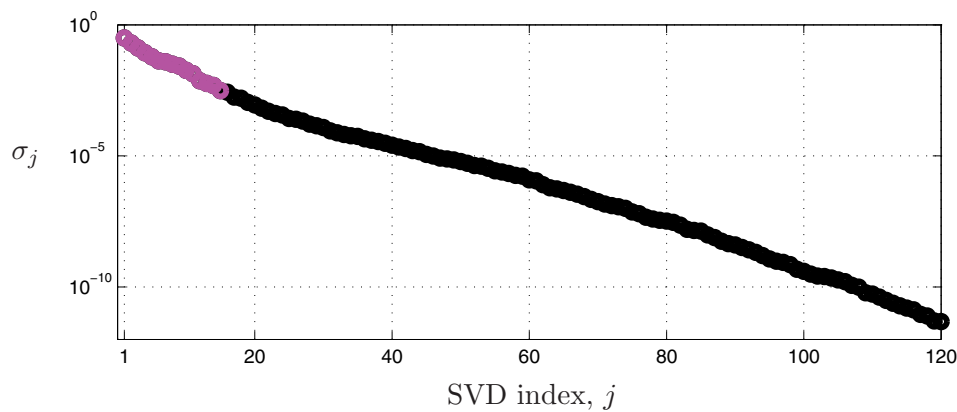

Figure 4. *Singular values for concatenated library with all regimes:* $\mathbf{A}_{cat} = [\mathbf{A_1} \ \mathbf{A_2} \ \mathbf{A_3} \ \mathbf{A_4} \ \mathbf{A_5} \ \mathbf{A_6}]$.

$t \in [0, 100)$, $\beta = \beta_3$ for $t \in [100, 200)$, and $\beta = \beta_5$ for $t \in [200, 300]$. The evolution dynamics for this case are illustrated in Figure 6(a).

We measure the state at either three (x_1–x_3) or five (x_1–x_5) locations $x_1 = 0$, $x_2 = 0.7$, $x_3 = 1.4$, $x_4 = 1.8$, $x_5 = 2.2$ (shown at the bottom of Figure 5) taking data only at the
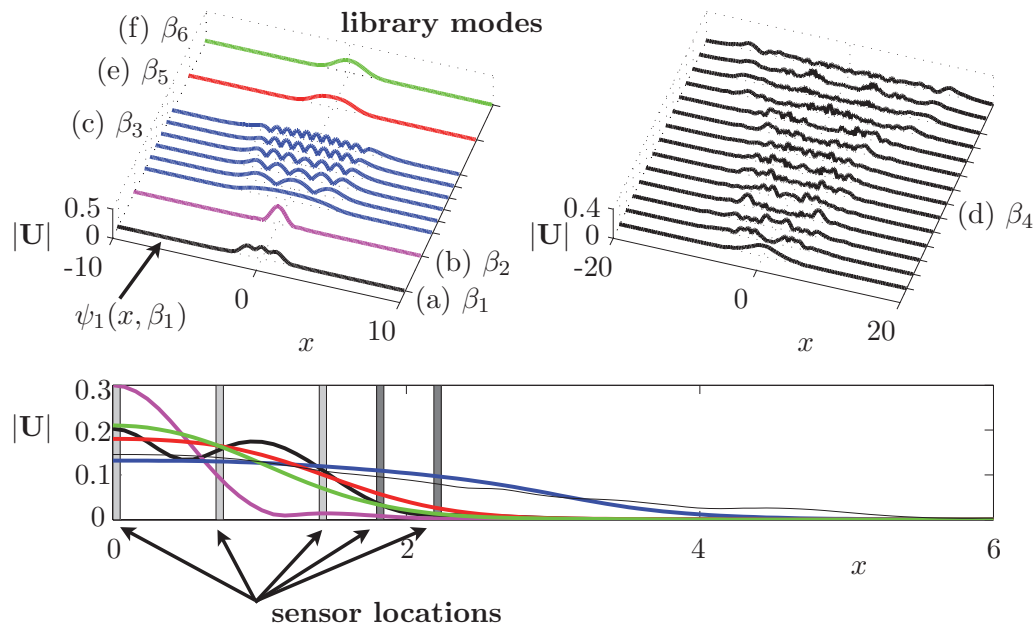
Figure 5. *Library* $\boldsymbol{\Psi}$ *of the dominant modes. The groupings, identified by their* β_j *value, are associated with the different dynamical regimes* (a)–(f) *in Figure* 3. *Note that the modes of the exploding dissipative solution* (d) *have been included separately in the right-hand panel as there are 14 modes required to capture 99% of the dynamics in this regime. A sample cross-section of the first mode of each library element* $\psi_1(x, \beta_j)$ *(j = 1, 2, 3, 4, 5, 6) is shown in the bottom panel, color-coded with the top panels. The bottom panel also shows the spatial location of the three sensors (light gray) and five sensors (including dark gray) used for sparse sampling.*

times $t_1 = 25$, $t_2 = 125$, and $t_3 = 225$; these times are chosen 25 units after the bifurcation value switches so that transients have decayed. At each instance, we take sparse measurements and perform classification, projection, and forward simulation (Galerkin reconstruction), while working exclusively in the low-dimensional POD library. The procedure is as follows:

   (i) *Classification.* From a sparse set of measurements (three or five), the modes corresponding to the specific β_j are identified and extracted.

  (ii) *Projection.* The sparse measurements are projected, through a standard pseudoinverse operation, onto the modes $\boldsymbol{\Psi}_j$ for the particular parameter β_j to determine initial values of a_n.

 (iii) *Reconstruction.* The extracted library modes are evolved according to the POD-Galerkin projection technique by using the spatial modes from the library $\boldsymbol{\Psi}_j$ in conjunction with their time dynamics $a_n(t)$ [26].

Figure 6(b) shows the resulting dynamic reconstruction, and Figure 6(c) shows the coefficients of the specific sparse vector $\hat{\mathbf{a}}$ identified at each time: t_1, t_2, and t_3. Indeed, the proposed algorithm using only three measurements reproduces the dynamics with remarkable success. Although the recognition and reconstruction begin 25 time units after the bifurcation parameter switches, this may be very fast in absolute time units, especially for optical systems.
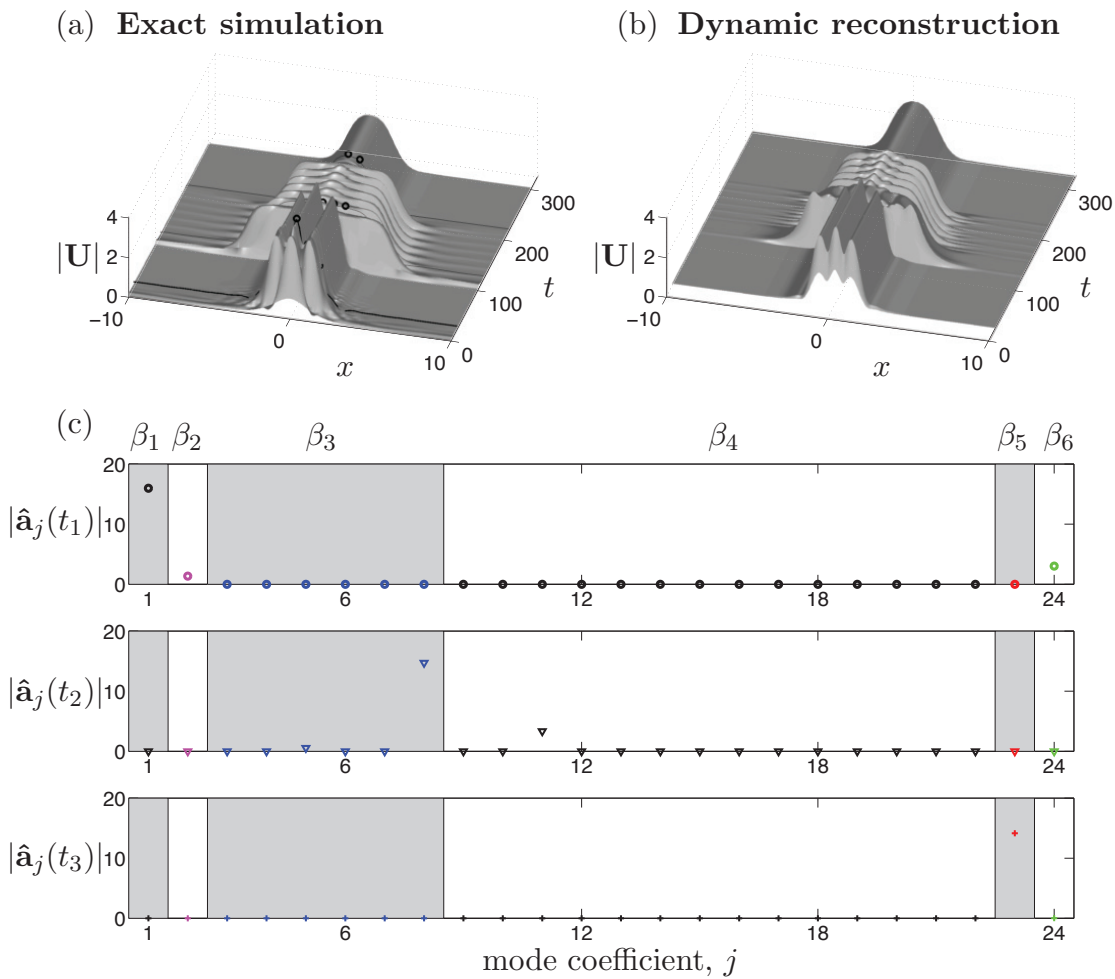
(a) **Exact simulation**   (b) **Dynamic reconstruction**

Figure 6. *Full evolution dynamics (a) and the low-rank POD dynamic reconstruction using compressive sensing and Galerkin projection (b). The black lines in (a) at $t = 25$, 125, and 225 represent the sampling times, while the three black circles represent the three sparse measurement locations. From the three samples, panel (b) is reconstructed by identifying the correct POD modes and using Galerkin projection. Panel (c) shows the modal coefficient vector $\hat{\mathbf{a}}$ evaluated via convex L_1 optimization for the three sampling times. Correct identification is achieved of the β_1 regime at $t_1 = 25$ (circles, $\circ$), of the β_3 regime at $t_2 = 125$ (triangles, $\triangle$), and of the β_5 regime at $t_3 = 225$ (pluses, $+$). The $|\hat{\mathbf{a}}_j|$ are color-coded according to the library elements depicted in Figure 5. For ease of viewing, the different β_j regimes are separated by shaded/nonshaded regions and are further identified at the top of panel (c).*

To improve the recognition rate, and therefore the bandwidth of closed-loop control built on these estimated states, it may be possible to augment the POD library with POD of the transient dynamics. This is the subject of current work by the authors. Note that the number of measurements $m = 3$, the number of library elements $p = 24$, and the original size of the system $n = 1024$ are ordered so that $m \ll p \ll n$. Consequently, the matrix $\mathbf{\Phi\Psi}$ in the underdetermined system (2.3) is a 3×24 matrix, yielding an efficient L_1 convex optimization

problem for the sparse identification.

For this example, we choose effective sampling locations based on the library modes of $\mathbf{\Psi}$. If poor choices are made, i.e., not aligning the sensors with maxima and minima observed in the POD library modes [55], then the dominant modes are often misidentified. This sensitivity to sensor location suggests that sensor placement should be carefully considered. Moreover, it is assumed that measurements of the system are perfect. However, noise is inherent in the detector and/or model.

The POD-Galerkin simulation is many orders of magnitude faster than the full simulation, as long as only a few modes are included. The dynamic regimes that include the fourth-order diffusion are numerically stiff and take hours to simulate at high resolution in MATLAB. Using an exponential time-stepping scheme would reduce this computation by an order of magnitude. However, POD-Galerkin simulations take on the order of seconds, resulting in many orders of magnitude speed-up. Reduced-order models based on POD-Galerkin have many well-documented advantages, the foremost being computational efficiency over the full simulation. The goal of this paper, however, is not to re-emphasize the benefits of POD-Galerkin but to demonstrate the advantage of combining L_1 and L_2 techniques. Since there are numerous subtle choices for the computation of the full simulation, we do not compare benchmarks but rather comment that the reduced-order model is much faster.

To quantify the impact of noise on the classification and reconstruction, (2.2) is modified to $\hat{\mathbf{U}} = \mathbf{\Phi}\mathbf{U} + \mathcal{N}(0, \sigma^2)$, where Gaussian distributed, white-noise error $\mathcal{N}$ with variance σ^2 is added to account for measurement error. Figure 7 shows statistical results of 400 trials using three or five sensors for noise strength $\sigma = 0.2$ or 0.5. With moderate noise ($\sigma = 0.2$), both the three- and five-sensor scenarios identify the correct regime quite accurately. For stronger noise ($\sigma = 0.5$), both three and five sensors lose a great deal of accuracy in the identification process. It is also observed that having more sensors actually hinders the evaluation of the β_1 parameter regime, although the β_3 and β_5 cases improve. Indeed, numerical simulations indicate that the three sensors placed at $x = 0$, 0.7, and 1.4 are robust and are not easily improved on by varying placement or quantity. Further study is needed to determine optimal sensor location.

These results suggest that multiple samplings in time can be used to reach a statistical conclusion about the correct parameter regime, thus avoiding misidentification. For example, we already wait 25 time units after β switches to take sparse measurements, so that transients decay. If, instead of sampling a single time unit at $t = 25$, we accumulate information over 5–10 time units, the effect of sensor noise is significantly attenuated.

We also investigate the least-squares estimate $\tilde{\mathbf{a}}$ for the mode amplitudes based on the three-sensor and five-sensor configurations. In every single case, for no noise, as well as for noise levels $\sigma = 0.2$ and $\sigma = 0.5$, the least-squares solution $\tilde{\mathbf{a}}$ results in the misidentification of the β_1 and β_5 regimes, instead identifying the incorrect β_3 regime. The collapse of L_2 minimization for identifying the bifurcation parameter regime highlights the success of the sparse sampling strategy, centered around L_1 minimization.

5. Discussion. In conclusion, we advocate a general theoretical framework for complex systems whereby low-rank structures are represented by the L_2-optimal POD and then identified from limited noisy measurements using the sparsity-promoting L_1 norm and the compressive sensing architecture. The strategy for building a modal library by concatenating
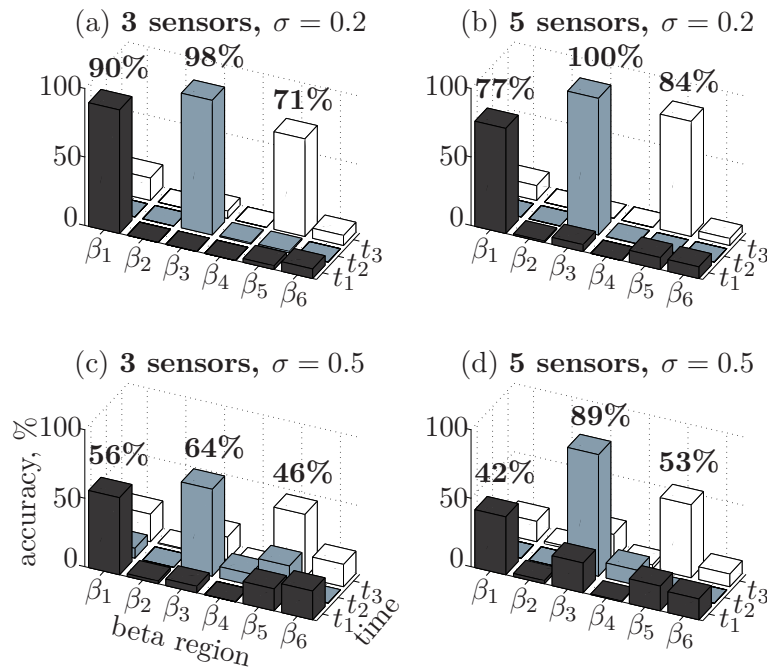
Figure 7. *Accuracy of region classification for the system described in Figure* 6. *The bifurcation parameter* β *switches from* β_1 *at* $t_1 = 25$ *to* β_3 *at* $t_2 = 125$ *to* β_5 *at* $t_3 = 225$. *The bar charts illustrate which bifurcation regime* β_1–β_6 *is classified from sparse measurements by the* L_1-*minimization procedure described above. Three or five sensors are considered under moderate (*$\sigma = 0.2$*) and strong (*$\sigma = 0.5$*) error measurements using* 400 *realizations. More sensors improve the region identification performance for regions* β_3 *and* β_5, *but decrease performance for* β_1.

truncated POD libraries across a range of relevant bifurcation parameters may be viewed as a simple machine learning implementation. The resulting modal library is 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.

These results are among the first, along with [8], to combine even simple machine learning concepts and compressive sensing to complex systems for both

  (i) correctly identifying the dynamical parameter regime, and

  (ii) reconstructing the associated low-rank dynamics.

Pairing a low-dimensional learned library, in which the dynamics have a sparse representation, with compressive sensing provides a powerful new architecture for studying dynamical systems that exhibit coherent behavior.

With three sensors, it is possible to accurately classify the bifurcation regime, reconstruct the low-dimensional content, and simulate the Galerkin projected dynamics of the complex Ginzburg–Landau equation. In addition, we investigate the performance of compressed sensing with the addition of sensor noise and the addition of more sensors. For moderate noise levels, the method accurately classifies the correct dynamic regime, although performance drops for larger noise values. The addition of more sensors does not significantly improve performance, although the sensor placement was not exhaustive. In contrast, classification based on least-squares fails to identify the β_1 and β_5 regions for all noise levels, in every trial.

There are a number of important directions that arise from this work. The library building procedure discussed in Figure 1 is quite general, and it will be interesting to investigate additional library building techniques and machine learning strategies. For example, is it possible to remove features that are common to all of the dynamic regimes to enhance contrast between categories in the L_1 classification step? It will also be interesting to investigate optimal sensor placement based on the principle of maximizing incoherence with respect to the overcomplete basis. To improve the classification time, it may also be possible to augment the POD libraries with transient dynamics. Finally, it may be possible to use coherence between each pair of local bases $(\mathbf{\Psi}_i, \mathbf{\Psi}_j)$ as a means of constructing an induced metric on the space of bifurcation parameters. This may facilitate the accurate categorization of dynamic regimes that have not been directly explored in the training step. Finally, the procedure above is promising for use with data assimilation techniques, which typically incorporate new measurements using least-squares fitting (L_2).

As these directions unfold, we believe that the combination of L_2 low-rank representations and L_1-sparse sampling will enable efficient characterization and manipulation of low-rank dynamical systems. The ultimate goal is to always work in a measurement space with dimension on the order of the underlying low-dimensional attractor. It will be interesting to leverage these efficient techniques for closed-loop feedback control.

REFERENCES

[1] A. AGARWAL AND P. F. J. LERMUSIAUX, *Statistical field estimation for complex coastal regions and archipelagos*, Ocean Modelling, 40 (2011), pp. 164–189.

[2] Z. BAI, T. WIMALAJEEWA, Z. BERGER, G. WANG, M. GLAUSER, AND P. K. VARSHNEY, *Physics based compressive sensing approach applied to airfoil data collection and analysis*, in Proceedings of the 51st AIAA Aerospace Sciences Meeting, 2013, AIAA Paper 2013-0772.

[3] M. BALAJEWICZ AND E. H. DOWELL, *Stabilization of projection-based reduced order models of the Navier-Stokes*, Nonlinear Dynam., 70 (2012), pp. 1619–1632.

[4] R. G. BARANIUK, *Compressive sensing*, IEEE Signal Processing Mag., 24 (2007), pp. 118–120.

[5] R. G. BARANIUK, V. CEVHER, M. F. DUARTE, AND C. HEGDE, *Model-based compressive sensing*, IEEE Trans. Inform. Theory, 56 (2010), pp. 1982–2001.

[6] M. BERGMANN AND L. CORDIER, *Optimal control of the cylinder wake in the laminar regime by trust-region methods and POD reduced-order models*, J. Comput. Phys., 227 (2008), pp. 7813–7840.

[7] M. BRANICKI AND A. J. MAJDA, *Quantifying Bayesian filter performance for turbulent dynamical systems through information theory*, Commun. Math. Sci., 12 (2014), pp. 901–978.

[8] I. BRIGHT, G. LIN, AND J. N. KUTZ, *Compressive sensing based machine learning strategy for characterizing the flow around a cylinder with limited pressure measurements*, Phys. Fluids, 25 (2013), 127102.

[9] S. L. BRUNTON, X. FU, AND J. N. KUTZ, *Self-tuning fiber lasers*, IEEE J. Sel. Topics Quantum Electron., 20 (2014), 1101408.

[10] E. J. CANDÈS, *Compressive sensing*, in Proceedings of the International Congress of Mathematicians, European Mathematical Society Publishing House, Zürich, 2006, pp. 1433–1452.

[11] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52 (2006), pp. 489–509.

[12] E. J. Candès, J. K. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59 (2006), pp. 1207–1223.

[13] E. J. Candès and T. Tao, *Near optimal signal recovery from random projections: Universal encoding strategies?*, IEEE Trans. Inform. Theory, 52 (2006), pp. 5406–5425.

[14] 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*, J. Comput. Phys., 242 (2013), pp. 623–647.

[15] R. R. Coifman and M. V. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Trans. Inform. Theory, 28 (1992), pp. 713–718.

[16] M. Cross and P. Hohenberg, *Pattern formation out of equilibrium*, Rev. Modern Phys., 65 (1993), pp. 851–1112.

[17] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, 52 (2006), pp. 1289–1306.

[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley-Interscience, New York, 2001.

[19] M. Fahl, *Trust-Region Methods for Flow Control Based on Reduced Order Modeling*, Ph.D. thesis, Trier University, Trier, Germany, 2000.

[20] J. E. Fowler, *Compressive-projection principal component analysis*, IEEE Trans. Image Process., 18 (2009), pp. 2230–2242.

[21] X. Fu, S. L. Brunton, and J. N. Kutz, *Classification of birefringence in mode-locked fiber lasers using machine learning and sparse representation*, Opt. Express, 22 (2014). pp. 8585–8597.

[22] S. Ganguli and H. Sompolinsky, *Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis*, Ann. Rev. Neurosci., 35 (2012), pp. 485–508.

[23] M. Gavish and D. L. Donoho, *The Optimal Hard Threshold for Singular Values Is $4/\sqrt{3}$*, preprint, arXiv:1305.5870v3 [stat.ME], 2014.

[24] D. Giannakis, A. J. Majda, and I. Horenko, *Information theory, model error, and predictive skill of stochastic models for complex nonlinear systems*, Phys. D, 241 (2012), pp. 1735–1752.

[25] A. C. Gilbert, J. Y. Park, and M. B. Wakin, *Sketched SVD: Recovering spectral features from compressive measurements*, preprint, arXiv:1211.0361v1 [cs.IT], 2012.

[26] P. J. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, 2nd ed., Cambridge Monogr. Mech., Cambridge University Press, Cambridge, UK, 2012.

[27] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, J. Educ. Psychol., 24 (1933), pp. 417–441, 498–520.

[28] S. M. Houghton, E. Knobloch, S. M. Tobias, and M. R. E. Proctor, *Transient spatio-temporal chaos in the complex Ginzburg-Landau equation on long domains*, Phys. Lett. A, 374 (2010), pp. 2030–2034.

[29] 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 (2014), pp. 365–414.

[30] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis*, Commun. Math. Sci., 1 (2003), pp. 715–762.

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

[32] T. Lieu and C. Farhat, Adaptation of POD-based aeroelastic ROMs for varying Mach number and angle of attack: Application to a complete F-16 configuration, in Proceedings of the U.S. Air Force T&E Days, 2005, AIAA Paper 2005-7666.

[33] E. N. Lorenz, *Empirical Orthogonal Functions and Statistical Weather Prediction*, Technical report, Massachusetts Institute of Technology, Cambridge, MA, 1956.

[34] J. L. Lumley, *Stochastic Tools in Turbulence*, Academic Press, New York, London, 1970.

[35] L. J. Nelson, G. L. Hart, F. Zhou, and V. Ozoliņš, *Compressive sensing as a paradigm for building physical models*, Phys. Rev. B, 87 (2013), 035125.

[36] B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele, *A hierarchy of low-dimensional models for the transient and post-transient cylinder wake*, J. Fluid Mech., 497 (2003), pp. 335–363.

[37] J. Östh, B. R. Noack, S. Krajnović, D. Barros, and J. Borée, *On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body*, J. Fluid Mech., 747 (2014), pp. 518–544.

[38] K. Pearson, *On lines and planes of closest fit to systems of points in space*, Philos. Mag., 2 (1901), pp. 559–572.

[39] H. Qi and S. M. Hughes, *Invariance of principal components under low-dimensional random projection of the data*, in Proceedings of the 19th IEEE International Conference on Image Processing, 2012, pp. 937–940.

[40] C. W. Rowley, *Model reduction for fluids, using balanced proper orthogonal decomposition*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 997–1013.

[41] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, *Spectral analysis of nonlinear flows*, J. Fluid Mech., 641 (2009), pp. 115–127.

[42] H. Schaeffer, R. Caflisch, C. D. Hauck, and S. Osher, *Sparse dynamics for partial differential equations*, Proc. Natl. Acad. Sci. USA, 110 (2013), pp. 6634–6639.

[43] P. J. Schmid, *Dynamic mode decomposition of numerical and experimental data*, J. Fluid Mech., 656 (2010), pp. 5–28.

[44] R. F. Schmidt and M. N. Glauser, *Improvements in low dimensional tools for flow-structure interaction problems: Using global POD*, in Proceedings of the 42nd AIAA Aerospace Sciences Meeting and Exhibit, 2004, AIAA 2004-0889.

[45] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, *Efficient measurement of quantum dynamics via compressive sensing*, Phys. Rev. Lett., 106 (2011), 100401.

[46] S. G. Siegel, J. Seidel, C. Fagley, D. M. Luchtenburg, K. Cohen, and T. McLaughlin, *Low-dimensional modelling of a transient cylinder wake using double proper orthogonal decomposition*, J. Fluid Mech., 610 (2008), pp. 1–42.

[47] L. Sirovich, *Turbulence and the dynamics of coherent structures. Parts* I–III, Quart. Appl. Math., 45 (1987), pp. 561–590.

[48] A. B. Tayler, D. J. Holland, A. J. Sederman, and L. F. Gladden, *Exploring the origins of turbulence in multiphase flow using compressed sensing MRI*, Phys. Rev. Lett., 108 (2012), 264505.

[49] J. A. Taylor and M. N. Glauser, *Towards practical flow sensing and control via POD and LSE based low-dimensional tools*, J. Fluids Eng., 126 (2004), pp. 337–345.

[50] F. Terragni and J. M. Vega, *On the use of POD-based ROMS to analyze bifurcations in some dissipative systems*, Phys. D, 241 (2012), pp. 1393–1405.

[51] F. Terragni and J. M. Vega, *Construction of bifurcation diagrams using POD on the fly*, SIAM J. Appl. Dyn. Syst., 13 (2014), pp. 339–365.

[52] J. A. Tropp and A. C. Gilbert, *Signal recovery from random measurements via orthogonal matching pursuit*, IEEE Trans. Inform. Theory, 53 (2007), pp. 4655–4666.

[53] W. X. Wang, R. Yang, Y. C. Lai, V. Kovanis, and C. Grebogi, *Predicting catastrophes in nonlinear dynamical systems by compressive sensing*, Phys. Rev. Lett., 106 (2011), 154101.

[54] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, *Robust face recognition via sparse representation*, IEEE Trans. Pattern Anal. Mach. Intell., 31 (2009), pp. 210–227.

[55] B. Yildirim, C. Chryssostomidis, and G. E. Karniadakis, *Efficient sensor placement for ocean measurements using low-dimensional concepts*, Ocean Modelling, 27 (2009), pp. 160–173.