Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems

Steven L. Brunton^{‡*}

Jonathan H. Tu[†] Ido Bright^{*} J. Nathan Kutz^{*}

December 14, 2013

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 over-complete 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 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.

Keywords: Dynamical systems, bifurcations, classification, compressive sensing, sparse representation, proper orthogonal decomposition.

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 data sets 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 [15]). We propose a data-driven modeling strategy that represents low-dimensional dynamics using *dimensionality reduction* methods such as the proper orthogonal decomposition (POD) [23] and classifies/reconstructs the observed low-dimensional manifolds with *compressive (sparse) sensing* (CS) [16, 11, 5, 46]. Thus dynamic structures are represented efficiently with the ℓ_2 norm and identified from sparse measurements with the ℓ_1 norm.

^{*}Department of Applied Mathematics, University of Washington, Seattle, WA. 98195-2420. [‡] (sbrunton@uw.edu). Questions, comments, or corrections to this document may be directed to that email address.

[†]Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544

The application of machine-learning and compressive sensing to dynamical systems is synergistic, in that underlying low-rank structures facilitate sparse measurements [27]. 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 [26, 38]. 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 neuro-sensory systems [20]. Compressive sensing may also play a role in similar statistical learning, library-based, and/or information theory methods [14, 7] used in fluid dynamics [8, 2], climate science [21, 7] and oceanography [1]. Indeed, compressive sensing is already playing a critical role in model building and assessment in the physical sciences [31, 41, 47, 44]. 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 [23] 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. However, Galerkin projection onto POD modes obtained across a range of parameter values, the so-called global POD [45, 40], 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 [28]. Including additional modes, such as the shift mode [32], to capture transients between qualitatively different flow regimes has resulted in additional methods such as double POD [42], and the GaussNewton with approximated tensors (GNAT) method [13] and trust-region POD [18, 6]. Alternative methods for stabilizing POD by adding additional modes and closure terms have been investigated [3, 33]. In each case, the objective is to construct a dimensionally reduced set of dynamics that accurately represent the underlying complex system and that does not suffer from instabilities.

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 Galerkin projection [23]. Here we concatenate POD bases to construct the library, although generalizing the library building strategy into a broader machine learning context [17] 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 a over-complete library, is closely related to sparse representation from image classification [48].

The paper is outlined as follows: In Sec. 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 proper orthogonal decomposition (POD) for ℓ_2 dimensionality reduction. These methods are combined in Sec. 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 (CS), 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 proper orthogonal decomposition (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 Ψ :

$$\mathbf{U}(x,t) = \mathbf{\Psi}\mathbf{a}.\tag{1}$$

Here, sparsity means that **U** may be represented in the transform basis Ψ by a vector of coefficients **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 $m \ll n$:

$$\hat{\mathbf{U}} = \mathbf{\Phi} \mathbf{U},\tag{2}$$

where Φ is a measurement matrix that maps the full state measurement **U** to the sparse measurement vector $\hat{\mathbf{U}}$. Details of this measurement matrix will be given shortly. Plugging (1) into (2) yields an underdetermined system:

$$\hat{\mathbf{U}} = \boldsymbol{\Phi} \boldsymbol{\Psi} \mathbf{a}. \tag{3}$$

We may then solve for the sparsest solution **a** to the underdetermined system of equations in (3). Sparsity is measured by the ℓ_0 norm, and solving for the solution **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 [11, 16]. The specific minimization problem is:

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

There are other algorithms that result in sparse solution vectors, such as orthogonal matching pursuit [46].

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 Ψ [9, 10, 4]. In addition, the measurement matrix Φ must be *incoherent* with respect to the sparse basis Ψ , meaning that the columns of Φ and the columns of Ψ 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 [12].

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 Φ 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 much of our efforts in using sparse measurements to characterize the complex dynamics.

2.2 ℓ_2 -based dimensionality reduction

The proper orthogonal decomposition (POD) [30, 23] is a tool with ubiquitous use in dimensionality reduction of physical systems¹. Data snapshots $\mathbf{U}(x, t_1), \mathbf{U}(x, t_2), \cdots \mathbf{U}(x, t_q)$ 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} :

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

Columns of the matrix Ψ are POD modes², and they are ordered according to the variance that they capture in the data **A**; if the columns of **A** are velocity measurements, then the POD modes are ordered in terms of the kinetic energy that they capture. This variance/energy content is quantified by the entries of the diagonal matrix Σ , 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, $n \gg q$, as in high-dimensional fluid systems, there are at most q non-zero singular values, and it is beneficial to use the *method of snapshots* [43]. In this method, we solve the following eigenvalue problem:

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

where Σ_q is the $q \times q$ upper-left block of Σ . It is then possible to find the first q POD modes corresponding to non-trivial singular values by:

$$\Psi_q = \mathbf{AW} \Sigma_q^{-1}.$$

The snapshots often exhibit low-dimensional phenomena, so that the majority of 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 pre-determined cut-off value, such as when the columns contain 99% of the variance, so that only the first r modes are kept. The SVD acts as a filter, and so often the truncated modes correspond to random fluctuations and disturbances. If the data in the matrix **A** is generated by a dynamical system (nonlinear system of ordinary differential equations

¹POD is sometimes referred to as principal components analysis [34], the Karhunen–Loève decomposition, empirical orthogonal functions [29], or the Hotelling transform [24, 25].

²Often, POD modes are given by the matrix $\mathbf{\Phi}$. However, we choose Ψ for the POD basis and $\mathbf{\Phi}$ for the sparse measurement matrix for consistency with compressive sensing literature. This is not to be confused with notation from balanced POD, where $\mathbf{\Phi}$ are direct modes and Ψ are adjoint modes [36].



Figure 1: Schematic of ℓ_2 strategy for library building. Data \mathbf{A}_j are collected for many values of the bifurcation parameter β_j , and the principal components of this data are computed and truncated in Ψ_j . Although each basis Ψ_j is truncated to contain only the most energetically relevant structures, the concatenated library Ψ is overcomplete.

of order n), It is then possible to substitute the truncated POD expansion for the state U into the governing equation and obtain Galerkin projected dynamics on the r basis modes [23]. Recall that we are assuming that the complex systems 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 [19, 22, 35].

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 proper orthogonal decomposition (POD) with the ℓ_1 -based compressive sensing (CS) architecture (depicted schematically in Figure 2) for classification and reconstruction.

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

$$\mathbf{U}_t = \mathbf{N}(\mathbf{U}, \mathbf{U}_x, \mathbf{U}_{xx}, \cdots, x, t, \beta), \tag{4}$$

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



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 CQGLE system in Sec. 4.

with respect to which the solution of the governing PDE changes markedly. We assume a spatial discretization of (4), which yields a high-dimensional system of degree n.

We would like to use measurements of the system (4) to determine its state. However, fullstate 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 work, we consider *spatially localized* or *point* measurements, $\hat{\mathbf{U}}$, as discussed in Sec. 2.1. In this case, the matrix $\boldsymbol{\Phi} \in \mathbb{R}^{m \times n}$ from (2) is comprised 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 (4) and collect dense measurements for various values

 $\beta_1, \beta_2, \dots, \beta_J$ 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 components, Ψ_j are identified through a singular value decomposition (SVD): $\Psi_j = \{\psi_i(x, \beta_j)\}_{i=1}^{r_j}$. 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.

With the modes identified for each β_j , an overcomplete library Ψ is constructed that contains all of the low-rank approximations for each dynamic regime:

$$\Psi = \begin{bmatrix} | & | & | & | \\ \psi_1(x,\beta_1) \cdots \psi_{r_1}(x,\beta_1) \cdots \psi_1(x,\beta_J) \cdots \psi_{r_J}(x,\beta_J) \\ | & | & | & | \end{bmatrix}.$$
(5)

The library $\Psi \in \mathbb{R}^{n \times p}$ 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 orthogonal POD modes for each given β_j . The dynamics at any given time will belong to a specific β_j regime so that that the instantaneous dynamics are *sparse* in the library basis, allowing for a sparse representation [48]. This *overcomplete library building* procedure is summarized in Figure 1.

With the library (5), we can expand the state U using the low-rank POD representation

$$\mathbf{U}(x,t) = \sum_{j=1}^{J} \sum_{r=1}^{r_j} a_{jr}(t)\psi_r(x,\beta_j) = \mathbf{\Psi}\mathbf{a}.$$
(6)

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

Equation (6) is of the form in (1). To determine the vector **a** from a sparse data measurement $\hat{\mathbf{U}} = \mathbf{\Phi}\mathbf{U}$, insert (6) into (2) and solve the under-determined linear system $\hat{\mathbf{U}} = (\mathbf{\Phi}\mathbf{\Psi})\mathbf{a}$ from (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 (4) 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,

Table 1: Parameter regimes β_j for the complex Ginzburg-Landau equation (7) (See Figure 3). The low-rank approximations of these parameter regimes are used to construct the elements of the library Ψ .

	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

- (ii) The algorithm works equally well with experimental data in an equation-free context, for instance by using dynamic mode decomposition [37, 39] or equation-free modeling [26] 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 only practical for real-time application with low-dimensional systems.

Sparse sensing is significantly less expensive in the learned library Ψ 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 benefit to keeping the individual POD bases Ψ_j for reconstruction once the bifurcation regime β has been identified.

4 Results

To illustrate the aforementioned strategy, consider the complex Ginzburg-Landau model [15], which is ubiquitous in mathematical physics. Here it is modified to include both quintic terms and a fourth-order diffusion term much like the Swift-Hohenberg equation:

$$i\mathbf{U}_t + \left(\frac{1}{2} - i\tau\right)\mathbf{U}_{xx} - i\kappa\mathbf{U}_{xxxx} + (1 - i\mu)|\mathbf{U}|^2\mathbf{U} + (\nu - i\varepsilon)|\mathbf{U}|^4\mathbf{U} - i\gamma\mathbf{U} = 0,\tag{7}$$

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.

Figure 3 illustrates the corresponding low-rank behavior produced in the simulations.

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 4, as described in (5) 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 $t \in [0, 100)$, $\beta = \beta_3$ for



Figure 3: Evolution dynamics of (7) 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 Ψ .

 $t \in [100, 200)$ and $\beta = \beta_5$ for $t \in [200, 300]$. The evolution dynamics for this case are illustrated in Figure 5 (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 4) taking data only at the 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_i are identified and extracted,
- (ii) [**projection**] the sparse measurements are projected, through a standard pseudo-inverse operation, onto the modes Ψ_j for the particular parameter β_j to determine initial values of a_n ,
- (iii) [reconstruction] the extracted library modes are evolved according the the POD-Galerkin projection technique by using the spatial modes from the library Ψ_j in conjunction with their time dynamics $a_n(t)$ [23].

Figure 5 (b) shows the resulting dynamic reconstruction, and Figure 5 (c) shows the coefficients of



Figure 4: Library Ψ 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 including separately in the right 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 shaded) and five sensors (addition dark shaded) used for sparse sampling.

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. Note that the number of measurements m = 3, 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 $\Phi \Psi$ in the underdetermined system (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 Ψ . If poor choices are made, i.e. not aligning the sensors with maxima and minima observed in the POD library modes [49], 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.

To quantify the impact of noise on the classification and reconstruction, (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 6 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



Figure 5: Full evolution dynamics (a) and the low-rank POD dynamic reconstruction using compressive sensing and Galerkin projection (b). The black lines on the left panel 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, the right panel is reconstructed by identifying the correct POD modes and using Galerkin projection. The bottom panel 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 $t_3 = 225$ (plusses, +). The $|\hat{\mathbf{a}}_j|$ are color coded according to the library elements depicted in Figure 4. For ease of viewing, the different β_j regimes are separated by shaded/non-shaded regions and are further identified at the top of panel (c).

that 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 mis-identification. 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.



Figure 6: Accuracy of region classification for the system described in Figure 5. 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 .

We also investigate the least-squares estimate $\tilde{\mathbf{a}}$ for the mode amplitudes based on the 3-sensor and 5-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 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 proper orthogonal decomposition, 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 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.

To our knowledge, these results are the first 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 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 β_1 and β_5 regions for all noise levels, on 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. Finally, it may be possible to use coherence between each pair of local bases (Ψ_i, Ψ_j) as a means to construct 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.

Acknowledgment

We are grateful for discussions with Bingni W. Brunton, Joshua L. Proctor, and Xing Fu. J. N. Kutz acknowledges support from the National Science Foundation (DMS-1007621) and the U.S. Air Force Office of Scientific Research (FA9550-09-0174).

References

- A. Agarwal and P. F. J. Lermusiaux. Statistical field estimation for complex coastal regions and archipelagos. *Ocean Modelling*, 40:164–189, 2011.
- [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. AIAA Paper 2013-0772, 51st Aerospace Sciences Meeting, January 2013.
- [3] M. Balajewicz and E. H. Dowell. Stabilization of projection-based reduced order models of the Navier-Stokes. *Nonlinear Dynamics*, 70:1619–1632, 2012.
- [4] R. G. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, 24(4):118–120, 2007.

- [5] 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.
- [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. *Journal of Computational Physics*, 227(16):7813–7840, 2008.
- [7] M. Branicki and A. J. Majda. Quantifying Bayesian filter performance for turbulent dynamical systems through information theory. to appear, 2013.
- [8] 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.
- [9] E. J. Candès. Compressive sensing. Proceedings of the International Congress of Mathematics, 2006.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. IEEE Transactions on Information Theory, 28(2):713–718, 1992.
- [15] M. Cross and P. Hohenberg. Pattern formation out of equilibrium. Reviews of Modern Physics, 65:851–1112, 1993.
- [16] D. L. Donoho. Compressed sensing. IEEE Transactions on Information Theory, 52(4):1289– 1306, 2006.
- [17] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern Classification. Wiley-Interscience, 2000.
- [18] M. Fahl. Trust-region methods for flow control based on Reduced Order Modeling. PhD thesis, Trier University, 2000.
- [19] J. E. Fowler. Compressive-projection principal component analysis. IEEE Transactions on Image Processing, 18(10):2230-2242, 2009.
- [20] 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.
- [21] D. Giannakis, A. J. Majda, and I. Horenko. Information theory, model error, and predictive skill of stochastic models for complex nonlinear systems. *Physica D*, 241(20):1735–1752, 2012.

- [22] A. C. Gilbert, J. Y. Park, and M. B. Wakin. Sketched SVD: Recovering spectral features from compressive measurements. ArXiv e-prints, 2012.
- [23] 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.
- [24] H. Hotelling. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol., 24:417–441, September 1933.
- [25] H. Hotelling. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol., 24:498–520, October 1933.
- [26] 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. *Communications in Mathematical Science*, 1(4):715–762, 2003.
- [27] J. N. Kutz. Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data. Oxford University Press, 2013.
- [28] 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. AIAA Paper 2005-7666, 2005.
- [29] E. N. Lorenz. Empirical orthogonal functions and statistical weather prediction. Technical report, Massachusetts Institute of Technology, December 1956.
- [30] J. L. Lumley. Stochastic Tools in Turbulence. Academic Press, 1970.
- [31] L. J. Nelson, G. L. Hart, F. Zhou, and V. Ozoliņš. Compressive sensing as a paradigm for building physical models. *Physical Review B*, 87:035125–1–035125–12, 2013.
- [32] B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele. A hierarchy of lowdimensional models for the transient and post-transient cylinder wake. *Journal of Fluid Mechanics*, 497:335–363, 2003.
- [33] J. Osth, B. R. Noack, and S. Krajnovic. On the need for a nonlinear subscale turbulence term in pod models as exemplified for a high reynolds number flow over an ahmed body. *submitted for publication*, 2013.
- [34] K. Pearson. On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 2(7-12):559-572, 1901.
- [35] 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.
- [36] C. W. Rowley. Model reduction for fluids using balanced proper orthogonal decomposition. International Journal of Bifurcation and Chaos, 15(3):997–1013, 2005.
- [37] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson. Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, 641:115–127, 2009.

- [38] 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.
- [39] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. Journal of Fluid Mechanics, 656:5–28, August 2010.
- [40] R. Schmidt and M. Glauser. Improvements in low dimensional tools for flow-structure interaction problems: using global POD. AIAA Paper 2004-0889, 2004.
- [41] 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. *Physical Review Letters*, 106:100401–1–100401–4, 2011.
- [42] S. G. Siegel, J. Seidel, C. Fagley, D. M. Luchtenburg, K. Cohen, and T. McLaughlin. Lowdimensional modelling of a transient cylinder wake using double proper orthogonal decomposition. *Journal of Fluid Mechanics*, 610:1–42, 2008.
- [43] L. Sirovich. Turbulence and the dynamics of coherent structures, parts I-III. Q. Appl. Math., XLV(3):561–590, 1987.
- [44] 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. *Physical Review Letters*, 108(26):264505-1-264505-5, 2012.
- [45] J. Taylor and M. Glauser. Towards practical flow sensing and control via POD and LSE based low-dimensiona tools. 2002 ASME Fluids Engineering Division Summer Meeting, 2002.
- [46] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, 2007.
- [47] W. X. Wang, R. Yang, Y. C. Lai, V. Kovanis, and C. Grebogi. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Physical Review Letters*, 106:154101–1– 154101–4, 2011.
- [48] 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.
- [49] B. Yildirim, C. Chryssostomidis, and G. E. Karniadakis. Efficient sensor placement for ocean measurements using low-dimensional concepts. *Ocean Modelling*, 27:160–173, 2009.