

COMPRESSED SENSING AND DYNAMIC MODE DECOMPOSITION

STEVEN L. BRUNTON

Department of Mechanical Engineering
University of Washington
Seattle, WA 98195, USA

JOSHUA L. PROCTOR

Institute for Disease Modeling
Intellectual Ventures Laboratory
Bellevue, WA 98004, USA

JONATHAN H. TU

Department of Electrical Engineering and Computer Science
University of California, Berkeley
Berkeley, CA 94720, USA

J. NATHAN KUTZ

Department of Applied Mathematics
University of Washington
Seattle, WA 98195, USA

(Communicated by the associate editor name)

ABSTRACT. This work develops compressed sensing strategies for computing the dynamic mode decomposition (DMD) from heavily subsampled or compressed data. The resulting DMD eigenvalues are equal to DMD eigenvalues from the full-state data. It is then possible to reconstruct full-state DMD eigenvectors using ℓ_1 -minimization or greedy algorithms. If full-state snapshots are available, it may be computationally beneficial to compress the data, compute DMD on the compressed data, and then reconstruct full-state modes by applying the compressed DMD transforms to full-state snapshots.

These results rely on a number of theoretical advances. First, we establish connections between DMD on full-state and compressed data. Next, we demonstrate the invariance of the DMD algorithm to left and right unitary transformations. When data and modes are sparse in some transform basis, we show a similar invariance of DMD to measurement matrices that satisfy the restricted isometry property from compressed sensing. We demonstrate the success of this architecture on two model systems. In the first example, we construct a spatial signal from a sparse vector of Fourier coefficients with a linear dynamical system driving the coefficients. In the second example, we consider the double gyre flow field, which is a model for chaotic mixing in the ocean.

2010 *Mathematics Subject Classification.* Primary: 65P99, 37M99, 37M10, 68P30, 94A99; Secondary: 76M99, 90C25, 94A08.

Key words and phrases. Compressive sampling, compressed sensing, dynamic mode decomposition, dynamical systems, unitary operators.

1. Introduction. Dynamic mode decomposition (DMD) is a powerful new technique introduced in the fluid dynamics community to isolate spatially coherent modes that oscillate at a fixed frequency [54, 49, 52]. DMD differs from other dimensionality reduction techniques such as the proper orthogonal decomposition (POD) [37, 9, 30, 43], where modes are selected to minimize the ℓ_2 -projection error of the data onto modes. In particular, DMD results in spatial modes that are constrained to each evolve at a single frequency and/or growth or decay rate in time. Thus, DMD not only provides modes, but also a linear model for how the modes evolve in time, in terms of these frequencies and growth/decay rates [60].

There are many immediate benefits to this formulation of coherent modes. First, it is possible to extract coherent structures that are known to oscillate at a fixed temporal frequency, which is common in fluid dynamics [64]. In addition, DMD modes with zero frequency and zero growth rate are *background* modes, providing efficient foreground/background separation in videos [25]. Finally, having a low-order model that describes the evolution of DMD modes may be useful for closed-loop feedback control.

The DMD is a data-driven and equation-free method that applies equally well to data from experiments or simulations. An underlying principle is that even if the data is high-dimensional, it may be described by a low-dimensional attractor subspace defined by a few coherent structures. When the data is generated by a nonlinear dynamical system, then the DMD modes are closely related to eigenvectors of the infinite-dimensional Koopman operator [35, 41, 49, 12, 40]. In Ref. [63], DMD has been shown to be equivalent to linear inverse modeling (LIM) [46, 47] from climate science, under certain conditions, and it also has deep connections to the eigensystem realization algorithm (ERA) [29, 33, 38].

DMD has been used to study various fluid experiments [53, 55], shock turbulent boundary layer interaction [26], the cylinder wake [4], and foreground/background separation in videos [25]. In the context of fluid dynamics, DMD typically relies on time-resolved, full-state measurements of a high-dimensional fluid vector field. For complex, turbulent flows, it may be prohibitive to collect data across all spatial and temporal scales required for this analysis.

The present work leverages tools from compressed sensing [20, 14, 16, 15, 7, 8] to facilitate the collection of considerably fewer measurements, resulting in the same dynamic mode decomposition, as illustrated in Figure 2. This reduction in the number of measurements may have a broad impact in situations where data acquisition is expensive and/or prohibitive. In particular, we envision these tools being used in particle image velocimetry (PIV) to reduce the data transfer requirements for each snapshot in time, increasing the maximum temporal sampling rate. Other applications include ocean and atmospheric monitoring, where individual sensors are expensive. Even if full-state measurements are available, the proposed method of compressed DMD will be computationally advantageous in many situations where there is low-rank structure in the high-dimensional data.

1.1. Previous work on sparsity in dynamics. There are a few examples of prior work utilizing sparsity for the dynamic mode decomposition. In [32], a sparsity-promoting variant of the dynamic mode decomposition was introduced whereby an ℓ_1 -penalty term on the number of DMD modes balanced the tradeoff between the number of modes and the quality of the DMD representation. Other algorithms have been developed to obtain only a fixed number of modes, but these have involved global minimization techniques that may not scale with large problems [18].

In [5, 65, 6, 27], compressed sensing has been used to design non-time resolved sampling strategies motivated by particle image velocimetry (PIV) of a fluid velocity field; in [65, 27], this sampling is specifically used to compute DMD utilizing sparsity of mode coefficients in time. These experimental methods are based on the fact that temporally sparse signals may be sampled considerably less often than suggested by the Shannon-Nyquist sampling frequency [44, 56].

In [57], compressed sensing is paired with the theory of linear dynamical systems to obtain higher temporal sampling resolution and accurate reconstruction of video MRI. Their work is based on prior studies relating compressed sensing, linear dynamical systems, and video MRI [50, 45]. Incoherent measurements are used to estimate an underlying snapshot matrix of hidden-Markov states, as well as an embedding from this low-dimensional attractor into the high-dimensional image pixel space. For the first part, they use system identification based on Hankel matrices and minimal realization theory [29]. Based on the heavy use of Hankel matrices in ERA, and the established connections between DMD and ERA, it will be interesting to see how the present work connects to [57] in the future.

1.2. Contribution of this work. This work deviates from the prior studies combining compressed sensing and DMD [32, 65], in that we utilize sparsity of the *spatial* coherent structures to reconstruct full-state DMD modes from few measurements. This method results in full-state DMD modes from spatially compressed or subsampled measurements using compressed sensing. The eigenvalues of the DMD on compressed data are shown to be equivalent to the full-state DMD eigenvalues under some conditions, so that we obtain the same low-dimensional DMD model to advance mode coefficients.

These results highlight the ability to perform DMD with significantly less data acquisition when the data and modes are sparse in some transform basis. If full-state snapshots are available, it is also possible to pre-compress the data, compute DMD on the compressed data, and then reconstruct full-state DMD modes as a linear combination of the original full-state data. Performing the DMD on compressed data requires significantly less computational overhead when compared with traditional DMD, because the most expensive step, computing a singular value decomposition (SVD), is performed on a much smaller data matrix. These methods are described in Section 3.3 as various paths we can take in Figure 2, depending on the available of full-state or compressed initial data.

Our results rely on a number of theoretical advances that may be useful more broadly. First, we establish connections between DMD on full-state and compressed data. We then show that DMD is invariant to left and right unitary transformations of the data. This implies that the DMD computed in the spatial domain, Fourier domain, or in a POD coordinate system will all be closely related, since these coordinate systems are related by unitary transformations. We then show that when data and modes are sparse in some basis, we obtain a similar invariance of the DMD when our measurement matrix and sparse basis satisfy the restricted isometry property.

The methods in this paper are illustrated on two examples that are relevant to fluid mechanics. However, we believe that there is broad applicability of compressed sensing techniques and dynamic mode decomposition in dynamical systems more generally.

2. Background. Dynamic mode decomposition is a method of modal extraction from full-state snapshot data that results in spatial-temporal coherent structures oscillating with a fixed frequency and damping rate. This theory has recently been generalized and extended to a larger class of datasets [63], and it is discussed in Sec. 2.1.

The present work is centered around the use of compressed sensing to compute the dynamic mode decomposition from very few spatial measurements. In compressed sensing, a high-dimensional signal may be reconstructed from few measurements as long as the signal is sparse in some transform basis. We discuss compressed sensing in Sec. 2.2.

2.1. Dynamic mode decomposition (DMD). The dynamic mode decomposition (DMD) is a new tool in dynamical systems that has been introduced in the fluid dynamics community [54, 49, 52]. The DMD provides the eigenvalues and eigenvectors of the best-fit linear system relating a snapshot matrix and a time-shifted version of the snapshot matrix at some later time.

Consider the following data snapshot matrices:

$$\mathbf{X} = \begin{bmatrix} | & | & | \\ \mathbf{x}_0 & \mathbf{x}_1 & \cdots & \mathbf{x}_{m-1} \\ | & | & & | \end{bmatrix}, \quad \mathbf{X}' = \begin{bmatrix} | & | & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \\ | & | & & | \end{bmatrix}.$$

Here, $\mathbf{x}_k \in \mathbb{R}^n$ is the k^{th} snapshot, and typically $n \gg m$. \mathbf{x} is often the state of a high-dimensional dynamical system, such as a fluid flow. We consider snapshots that are spaced evenly in time, and we also allow for these measurements to be discrete-time samples of a continuous-time signal, so that $\mathbf{x}_k = \mathbf{x}(k\Delta t)$.

The dynamic mode decomposition involves the decomposition of the best-fit linear operator \mathbf{A} relating the matrices above:

$$\mathbf{X}' \approx \mathbf{A}\mathbf{X}. \tag{1}$$

The system $\mathbf{X}' = \mathbf{A}\mathbf{X}$ is generally underdetermined for $n > m$, and \mathbf{A} is the least-squares solution, which is chosen to minimize the Frobenius norm of $\|\mathbf{X}' - \mathbf{A}\mathbf{X}\|_F$. In the atypical case when Eq. (1) is overdetermined (e.g., $m > n$), it is possible to solve for \mathbf{A} as a minimum-norm solution [63]. When ambiguous, we may refer to \mathbf{A} in Eq. (1) as $\mathbf{A}_{\mathbf{X}}$. It is important to note that Eq. (1) may not be exactly satisfied by the data, and exact equality is not required at any step in the results that follow, hence the least-squares solution.

In practice, the matrix \mathbf{A} may be extremely large, making it difficult or impossible to store in memory and analyze. However, DMD seeks the dominant eigenvectors of \mathbf{A} , corresponding to non-zero eigenvalues, of which there are at most m when $n \geq m$. Thus it is possible to analyze a low-rank projection of \mathbf{A} onto modes defined by the singular value decomposition (SVD) of \mathbf{X} . The eigendecomposition of this projected matrix is then used to reconstruct full-state DMD modes. The exact DMD algorithm proceeds as follows:

Algorithm 1. The method of *exact DMD* was recently defined [63], and it is repeated here for convenience. The exact DMD algorithm is given by the following:

1. Collect data \mathbf{X}, \mathbf{X}' and compute the SVD of \mathbf{X} :

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^*. \tag{2}$$

In general, \mathbf{U} and \mathbf{V} are square unitary matrices of sizes $n \times n$ and $m \times m$, respectively. The rank of the SVD is $r \leq m$, and so we only keep the first r

columns of \mathbf{U} and \mathbf{V} and the first $r \times r$ terms in Σ ; thus \mathbf{U} and \mathbf{V} become rectangular matrices. In practice, even if the matrix has full column rank m , one may choose to truncate the SVD at $r < m$, either at some pre-determined cutoff threshold or at an optimal threshold for noise reduction [22].

2. Compute the least-squares fit $\tilde{\mathbf{A}}$ that satisfies $\mathbf{X}' = \tilde{\mathbf{A}}\mathbf{X}$ and project onto POD/PCA modes \mathbf{U} :

$$\tilde{\mathbf{A}} = \mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{U}^* \mathbf{X}' \mathbf{V} \Sigma^{-1}. \quad (3)$$

If the matrix \mathbf{U} contains all m left-singular vectors of \mathbf{X} arranged as columns, then $\tilde{\mathbf{A}}$ is projected so that it acts on the range of \mathbf{X} . However, in practice, \mathbf{U} may be truncated, in which case $\tilde{\mathbf{A}}$ is projected onto the leading POD modes of \mathbf{X} .

3. Compute the eigen-decomposition of $\tilde{\mathbf{A}}$:

$$\tilde{\mathbf{A}} \mathbf{W} = \mathbf{W} \Lambda. \quad (4)$$

Λ are the DMD eigenvalues.

4. Compute the DMD modes Φ :

$$\Phi = \mathbf{X}' \mathbf{V} \Sigma^{-1} \mathbf{W}. \quad (5)$$

It is proven in [63] (Theorem 1) that the diagonal elements of Λ are eigenvalues of A with corresponding eigenvectors given by columns of Φ .

Remark 1. The first three steps in the algorithm above are identical to those in [52]. However, the last step differs in the computation of DMD modes. In [52], the modes are given by $\Phi = \mathbf{U}\mathbf{W}$. In [63], this formula $\Phi = \mathbf{U}\mathbf{W}$ is still used to compute modes corresponding to zero eigenvalues. The new definition of exact DMD has the benefit that modes Φ are in the column space of \mathbf{X}' , whereas traditional DMD modes are in the column space of \mathbf{X} (i.e., they are a linear combination of POD modes \mathbf{U}). In many fluid systems, the column space of \mathbf{X}' and \mathbf{X} are either close or identical, because the system is sampled on an attractor, so that either convention may be used to compute DMD modes. None of the major theoretical results in this paper depend on defining DMD modes using Eq. (5) instead of $\Phi = \mathbf{U}\mathbf{W}$. However, we use the definition in Eq. (5) to be consistent with [63].

The data \mathbf{X}, \mathbf{X}' may come from a nonlinear system

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k),$$

in which case the DMD modes are related to eigenvectors of the infinite-dimensional Koopman operator \mathcal{K} which acts as the pull-back operator on observable functions [35, 1, 49, 40]. In particular, \mathcal{K} acts on observable functions \mathbf{g} as:

$$\mathcal{K}\mathbf{g}(\mathbf{x}_k) = \mathbf{g}(\mathbf{f}(\mathbf{x}_k)) = \mathbf{g}(\mathbf{x}_{k+1}).$$

The connection between DMD and the Koopman operator justify the application of this method in a variety of contexts. We may interpret DMD as a model reduction technique if data is acquired from a high-dimensional model, or a method of system identification if the data comes from measurements of an uncharacterized system. In the latter case, the resulting DMD model is data-driven and may be used in conjunction with equation-free methods [34]. The hierarchy of structure in the data is illustrated in Figure 1. Recently, the assumption of evenly spaced snapshots was relaxed, so that the columns of \mathbf{X} may be sampled at any times, as long as the columns of \mathbf{X}' are sampled a fixed Δt later [63].

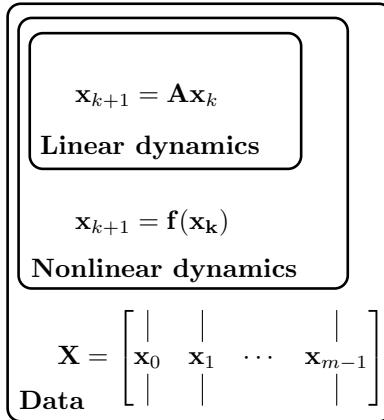


FIGURE 1. Schematic of various assumptions of dynamic structure underlying data \mathbf{X}, \mathbf{X}' .

2.2. Compressed Sensing. Consider a signal $\mathbf{x} \in \mathbb{R}^n$, which is sparse in some orthogonal basis given by the columns of Ψ , so that

$$\mathbf{x} = \Psi\mathbf{s}, \quad (6)$$

and \mathbf{s} is a vector containing mostly zeros. The signal \mathbf{s} is K -sparse if it has exactly K nonzero elements.

Most natural signals are sparse in some basis. For example, natural images and audio signals are sparse in Fourier or wavelet bases, resulting in a high-degree of *compressibility*. If we take the Fourier or Wavelet transform of an image, most of the coefficients will be small and can be neglected without resulting in much loss of image quality. Truncating in Fourier or Wavelet bases is the foundation of JPEG-2000 image compression and MP3 audio compression. Similarly, many high-dimensional nonlinear PDEs have sparse solutions, allowing for accelerated computational methods [51, 39].

The theory of compressed sensing [20, 14, 16, 15, 7, 8] suggests that instead of measuring the high-dimensional signal \mathbf{x} and then compressing, it is possible to measure a low-dimensional subsample or random compression of the data and then directly solve for the few non-zero coefficients in the transform basis. Consider the measurements $\mathbf{y} \in \mathbb{R}^p$, with $K < p \ll n$:

$$\mathbf{y} = \mathbf{C}\mathbf{x}.$$

The measurement matrix \mathbf{C} is often denoted by Φ in the compressed sensing literature. However, Φ is already used to represent DMD modes in Eq. (5).

If \mathbf{x} is sparse in Ψ , then we may solve the underdetermined system of equations

$$\mathbf{y} = \mathbf{C}\Psi\mathbf{s} \quad (7)$$

for \mathbf{s} and then reconstruct \mathbf{x} . Since there are infinitely many solutions to this system of equations, we seek the sparsest solution $\hat{\mathbf{s}}$,

$$\hat{\mathbf{s}} = \underset{\mathbf{s}'}{\operatorname{argmin}} \|\mathbf{s}'\|_0, \text{ such that } \mathbf{y} = \mathbf{C}\Psi\mathbf{s}'. \quad (8)$$

However, this amounts to a brute force combinatorial search, which is infeasible for even moderately large problems. Under certain conditions on the measurement

matrix \mathbf{C} , Eq. (8) may be relaxed to a convex ℓ_1 -minimization [15, 20]:

$$\hat{\mathbf{s}} = \underset{\mathbf{s}'}{\operatorname{argmin}} \|\mathbf{s}'\|_1, \text{ such that } \mathbf{y} = \mathbf{C}\Psi\mathbf{s}'. \quad (9)$$

Specifically, the measurement matrix \mathbf{C} must be *incoherent* with respect to the sparse basis Ψ , so that rows of \mathbf{C} are uncorrelated with columns of Ψ . The incoherence of the measurement matrix \mathbf{C} with respect to the sparse basis Ψ is quantified by the parameter $\mu(\mathbf{C}, \Psi)$:

$$\mu(\mathbf{C}, \Psi) = \sqrt{n} \max_{j,k} |\langle \mathbf{c}_k, \psi_j \rangle|,$$

where \mathbf{c}_k is the k -th row of the matrix \mathbf{C} . When measurements are incoherent, the matrix $\mathbf{C}\Psi$ satisfies a *restricted isometry property* (RIP) for sparse vectors \mathbf{s} ,

$$(1 - \delta_K)\|\mathbf{s}\|_2^2 \leq \|\mathbf{C}\Psi\mathbf{s}\|_2^2 \leq (1 + \delta_K)\|\mathbf{s}\|_2^2,$$

with restricted isometry constant δ_K [17]. The constant δ_K is defined as the smallest number that satisfies the above inequality for *all* K -sparse vectors \mathbf{s} . When δ_K is small, then $\mathbf{C}\Psi$ acts as a near isometry on K -sparse vectors \mathbf{s} . In practice, it is extremely difficult to compute δ_K directly; moreover, the measurement matrix \mathbf{C} may be chosen to be random, so that it is more desirable to derive statistical properties about the bounds on δ_K for a family of measurement matrices \mathbf{C} , rather than to compute δ_K for a specific \mathbf{C} . Generally, increasing the number of measurements will decrease the constant δ_K , improving the property of $\mathbf{C}\Psi$ to act isometry on sparse vectors. In addition to taking incoherent measurements, we must take on the order of $K \log(n/K)$ measurements to accurately determine the K nonzero elements of the n -length vector \mathbf{s} [13, 14, 7]. In this case, there are bounds on the constant δ_K that guarantee exact signal reconstruction for noiseless data. A more in-depth discussion of incoherence and the RIP can be found in [7, 17]. We will use the fact that $\mathbf{C}\Psi$ acts as a near isometry on K -sparse vectors to combine sparsity and dynamic mode decomposition for data with low-rank structure.

Typically a generic basis such as Fourier or wavelets is used to represent the sparse signal \mathbf{s} . The Fourier transform basis is particularly attractive for engineering purposes since single-pixel measurements are incoherent, exciting broadband frequency content. If a signal is K -sparse in the Fourier domain, we may then reconstruct the full state from $\mathcal{O}(K \log(n/K))$ single-pixel measurements at random spatial locations. This is especially beneficial when individual measurements are expensive, for example in ocean and atmospheric sampling, among other applications.

Another major result of compressed sensing is that Bernoulli and Gaussian random measurement matrices \mathbf{C} will satisfy the RIP with high probability for a generic basis Ψ [16]. There is also work describing incoherence with sparse matrices and generalizations to the RIP [23]. Recent work has shown the advantage of pairing compressed sensing with a data-driven POD/PCA basis, in which the data is optimally sparse [36, 10, 5, 6, 11]. The use of a POD/PCA basis results in a more computationally efficient signal reconstruction from fewer measurements.

In addition to the ℓ_1 minimization described in Eq. (9), there are a host of *greedy algorithms* [61, 62, 42, 24] that iteratively determine the sparse solution to the underdetermined system in Eq. (7). There has also been significant work on compressed SVD and PCA based on the Johnson-Lindenstrauss (JL) lemma [31, 21, 48, 24]. The JL lemma is closely related to the RIP, and it states when it is possible to embed a set of high-dimensional vectors in a low-dimensional space while preserving the spectral properties.

3. Compressive DMD. In this section, we combine ideas from compressed sensing to compute the dynamic mode decomposition from a few spatially incoherent measurements. In Sec. 3.1, we establish basic connections between the DMD on full-state and compressed data. These connections facilitate the two main applied results of this work:

- 1) It is possible to compute DMD on compressed data and reconstruct full-state DMD modes through compressed sensing. This will be referred to as the compressed sensing dynamic mode decomposition (CSDMD).
- 2) If full-state measurements are available, it is advantageous to compress the data, compute the compressed DMD, and then compute full-state DMD modes by linearly combining snapshots according to the compressed DMD transforms. This will be referred to as the compressed dynamic mode decomposition (CDMD).

In both cases, the full-state and compressed DMD eigenvalues are equal. These two approaches are described in Section 3.3 as various pathways to take in Figure 2, depending on what the initial data is.

Next, in Sec. 3.2, we demonstrate the invariance of the DMD algorithm to left or right unitary transformations of the data. We then discuss how the condition of unitarity may be relaxed to a transformation satisfying a restricted isometry property, as long as the data is sparse in a basis that is incoherent with respect to the measurements. This strengthens the connection to compressed sensing. Full-state DMD modes are then reconstructed from DMD modes on compressed data using compressed sensing; in particular, we use the method of compressive sampling matching pursuit (CoSaMP) [42].

3.1. DMD eigenvalues and eigenvectors on compressed data. It is possible to either collect data \mathbf{X}, \mathbf{X}' , or compressed data \mathbf{Y}, \mathbf{Y}' , where $\mathbf{Y} = \mathbf{C}\mathbf{X}$, $\mathbf{Y}' = \mathbf{C}\mathbf{X}'$ and $\mathbf{C} \in \mathbb{R}^{p \times n}$ is the measurement matrix.

Definition 3.1. We refer to \mathbf{X} and \mathbf{X}' as the *full-state snapshot matrices* and \mathbf{Y} and \mathbf{Y}' as the *compressed snapshot matrices*.

Similar to Eq. (1) above, we seek to find a best-fit operator $\mathbf{A}_\mathbf{Y}$ that relates \mathbf{Y} and \mathbf{Y}' as

$$\mathbf{Y}' \approx \mathbf{A}_\mathbf{Y} \mathbf{Y}. \quad (10)$$

Again, note that exact equality in the expression above is not required, nor is it likely to be satisfied by the data. We may also rely on the following three assumptions of sparsity of data and incoherence of measurements.

Assumption 1. The columns of \mathbf{X} and \mathbf{X}' are sparse in a transform basis Ψ so that $\mathbf{X} = \Psi \mathbf{S}$ and $\mathbf{X}' = \Psi \mathbf{S}'$, where the columns of \mathbf{S}, \mathbf{S}' are sparse (mostly zeros).

Assumption 2. The measurement matrix \mathbf{C} is *incoherent* with respect to Ψ and there are enough measurements (i.e., $p > K \log(n/K)$) so that $\mathbf{C}\Psi$ acts as a near isometry on K -sparse vectors. In this case, there are bounds on the restricted isometry constant δ_K that ensure, with high probability, that exact sparse signal reconstruction is possible for noiseless data.

Remark 2. Under Assumptions 1 and 2, it is possible to reconstruct \mathbf{x} from $\mathbf{y} = \mathbf{C}\mathbf{x}$, and therefore \mathbf{X}, \mathbf{X}' from \mathbf{Y}, \mathbf{Y}' using compressed sensing. However, this is laborious and inelegant.

Assumption 3. In addition to sparsity of the columns of \mathbf{X} and \mathbf{X}' , we may also require each of the columns to be in the same sparse subspace of the basis Ψ . The POD modes $\mathbf{U}_\mathbf{X}$ and the DMD modes $\Phi_\mathbf{X}$ are then guaranteed to be in this sparse subspace. This condition is reasonable when the dynamical system evolves on a low-dimensional subspace.

Assumption 4. The measurement matrix \mathbf{C} preserves the temporal information encoded in \mathbf{X} so that $\mathbf{V}_\mathbf{Y}\mathbf{V}_\mathbf{Y}^*\mathbf{V}_\mathbf{X} = \mathbf{V}_\mathbf{X}$. Recall that the matrix $\mathbf{V}_\mathbf{Y}$ may be truncated, so that it has size $r \times m$. If the system has rank $r = m$, then $\mathbf{V}_\mathbf{Y}\mathbf{V}_\mathbf{Y}^*$ will be the identity matrix and the assumption is trivially satisfied. However, if $r < m$, then the columns of $\mathbf{V}_\mathbf{X}$ are required to be in the column space of $\mathbf{V}_\mathbf{Y}$; of course, this will only be approximately satisfied if there is measurement noise.

Lemma 3.2. *If Assumption 4 holds, then the full-state and compressed DMD matrices $\mathbf{A}_\mathbf{X}$ from Eq. (1) and $\mathbf{A}_\mathbf{Y}$ from Eq. (10) are related as:*

$$\mathbf{C}\mathbf{A}_\mathbf{X}\mathbf{U}_\mathbf{X} = \mathbf{A}_\mathbf{Y}\mathbf{C}\mathbf{U}_\mathbf{X}. \quad (11)$$

Proof. We first use Assumption 4 to derive a useful identity:

$$\begin{aligned} \mathbf{Y} &= \mathbf{C}\mathbf{X} \\ \mathbf{U}_\mathbf{Y}\Sigma_\mathbf{Y}\mathbf{V}_\mathbf{Y}^* &= \mathbf{C}\mathbf{U}_\mathbf{X}\Sigma_\mathbf{X}\mathbf{V}_\mathbf{X}^* \\ \mathbf{V}_\mathbf{Y}^*\mathbf{V}_\mathbf{X}\Sigma_\mathbf{X}^{-1} &= \Sigma_\mathbf{Y}^{-1}\mathbf{U}_\mathbf{Y}^*\mathbf{C}\mathbf{U}_\mathbf{X} \\ \mathbf{V}_\mathbf{Y}\mathbf{V}_\mathbf{Y}^*\mathbf{V}_\mathbf{X}\Sigma_\mathbf{X}^{-1} &= \mathbf{V}_\mathbf{Y}\Sigma_\mathbf{Y}^{-1}\mathbf{U}_\mathbf{Y}^*\mathbf{C}\mathbf{U}_\mathbf{X} \\ \mathbf{V}_\mathbf{X}\Sigma_\mathbf{X}^{-1} &= \mathbf{V}_\mathbf{Y}\Sigma_\mathbf{Y}^{-1}\mathbf{U}_\mathbf{Y}^*\mathbf{C}\mathbf{U}_\mathbf{X}. \end{aligned} \quad (12)$$

Then

$$\begin{aligned} \mathbf{C}\mathbf{A}_\mathbf{X}\mathbf{U}_\mathbf{X} &= \mathbf{C}(\mathbf{X}'\mathbf{V}_\mathbf{X}\Sigma_\mathbf{X}^{-1}\mathbf{U}_\mathbf{X}^*)\mathbf{U}_\mathbf{X} \\ &= \mathbf{Y}'\mathbf{V}_\mathbf{X}\Sigma_\mathbf{X}^{-1} \\ &= \mathbf{Y}'(\mathbf{V}_\mathbf{Y}\Sigma_\mathbf{Y}^{-1}\mathbf{U}_\mathbf{Y}^*\mathbf{C}\mathbf{U}_\mathbf{X}) \\ &= \mathbf{A}_\mathbf{Y}\mathbf{C}\mathbf{U}_\mathbf{X}. \end{aligned}$$

□

The above lemma allows us to prove the following theorem, which establishes the central connection between DMD on full and compressed data.

Theorem 3.3. *Let ϕ_x be an eigenvector of $\mathbf{A}_\mathbf{X}$ that lies in the column space of \mathbf{X} . Then $\mathbf{C}\phi_x$ is an eigenvector of $\mathbf{A}_\mathbf{Y}$ with the same eigenvalue λ .*

Proof. Since ϕ_x lies in the column space of \mathbf{X} , we have $\mathbf{U}_\mathbf{X}\mathbf{U}_\mathbf{X}^*\phi_x = \phi_x$. Then starting with Lemma 3.2, we find that

$$\begin{aligned} \mathbf{C}\mathbf{A}_\mathbf{X}\mathbf{U}_\mathbf{X}\mathbf{U}_\mathbf{X}^* &= \mathbf{A}_\mathbf{Y}\mathbf{C}\mathbf{U}_\mathbf{X}\mathbf{U}_\mathbf{X}^* \\ \mathbf{C}\mathbf{A}_\mathbf{X}\mathbf{U}_\mathbf{X}\mathbf{U}_\mathbf{X}^*\phi_x &= \mathbf{A}_\mathbf{Y}\mathbf{C}\mathbf{U}_\mathbf{X}\mathbf{U}_\mathbf{X}^*\phi_x \\ \mathbf{C}\mathbf{A}_\mathbf{X}\phi_x &= \mathbf{A}_\mathbf{Y}\mathbf{C}\phi_x \\ \lambda\mathbf{C}\phi_x &= \mathbf{A}_\mathbf{Y}\mathbf{C}\phi_x, \end{aligned}$$

so $\mathbf{C}\phi_x$ is an eigenvector of $\mathbf{A}_\mathbf{Y}$ with eigenvalue λ . □

For real data, it is likely that DMD eigenvectors of \mathbf{A}_X will be approximately in column space of \mathbf{X} , since \mathbf{X} and \mathbf{X}' will generally have similar column spaces; however, if there is measurement error, the two matrices will not have identical column spaces. If \mathbf{C} is chosen poorly, so that ϕ_x is in the null-space of \mathbf{C} , then Theorem 3.3 applies trivially. Theorem 3.3 does not guarantee that every eigenvector ϕ_y of \mathbf{A}_Y is the projection of an eigenvector of \mathbf{A}_X through \mathbf{C} . The rank r of \mathbf{X} is guaranteed to be less than or equal to m , the number of columns, when $n \geq m$. In many cases, r is chosen by a threshold criteria to truncate noise [22], or to capture 99% of the variance in \mathbf{X} . If the noise-truncated rank r is equal to m , then it may be necessary to collect more snapshots in order to adequately sample the attractor. The assumption of low-rank structure is implicit in most dimensionality reduction strategies. For compressed sensing DMD, we also require that there are at least as many output measurements p as the rank r , so that $p \geq r$. As long as the columns of \mathbf{X} are not in the null-space of \mathbf{C} and the rank of $\mathbf{Y} = \mathbf{C}\mathbf{X}$ is also r , then the r nontrivial DMD eigenvalues Λ_Y of \mathbf{Y}, \mathbf{Y}' will be equal to Λ_X . Similarly, the projected (compressed) DMD eigenvectors are related to full-state DMD eigenvectors as described in Theorem 3.3.

Corollary 1. *Given Assumptions 1-4, we may reconstruct ϕ_x from the eigenvector $\phi_y = \mathbf{C}\phi_x = \mathbf{C}\Psi\phi_s$ from Theorem 3.3 by compressed sensing.*

Remark 3. Even starting with full-state measurements \mathbf{X}, \mathbf{X}' , it is beneficial to compress, compute the DMD, and reconstruct the modes according to:

$$\tilde{\Phi}_X = \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y. \quad (13)$$

We refer to this as *compressed DMD*, as opposed to the *compressive-sampling DMD* above. Under certain conditions, compressed DMD is equivalent to DMD.

Theorem 3.4. *If Assumption 4 holds and the column space of \mathbf{X}' is contained in that of \mathbf{X} , then the compressed DMD modes are DMD modes.*

Proof. By assumption, we have $\mathbf{U}_X \mathbf{U}_X^* \mathbf{X}' = \mathbf{X}'$. Together with Eq. (12), which follows from Assumption 4, this allows us to write

$$\begin{aligned} \mathbf{A}_X \tilde{\Phi}_X &= (\mathbf{X}' \mathbf{V}_X \Sigma_X^{-1} \mathbf{U}_X^*) (\mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y) \\ &= \mathbf{X}' (\mathbf{V}_Y \Sigma_Y^{-1} \mathbf{U}_Y^* \mathbf{C} \mathbf{U}_X) \mathbf{U}_X^* \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y \\ &= \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{U}_Y^* \mathbf{C} \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y \\ &= \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{U}_Y^* \mathbf{Y}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y \\ &= \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \tilde{\mathbf{A}}_Y \mathbf{W}_Y \\ &= \mathbf{X}' \mathbf{V}_Y \Sigma_Y^{-1} \mathbf{W}_Y \Lambda_Y \\ &= \tilde{\Phi}_X \Lambda_Y. \end{aligned}$$

Thus, the columns of $\tilde{\Phi}_X$ are eigenvectors of \mathbf{A}_X , making them DMD modes (as well as compressed DMD modes). \square

Remark 4. It is important to note that neither Lemma 3.2 nor Theorem 3.3 require exact equality in Eqs. (1) and (10). However, often $\mathbf{X}' \approx \mathbf{A}_X \mathbf{X}$, where $\mathbf{A}_X = \mathbf{X}' \mathbf{X}^\dagger \triangleq \mathbf{X}' \mathbf{V}_X \Sigma_X^{-1} \mathbf{U}_X^*$. In this case, we require that $(\mathbf{C}\mathbf{X})^\dagger \approx \mathbf{X}^\dagger \mathbf{C}^*$, which holds either when \mathbf{C} satisfies the Johnson-Lindenstrauss theorem, or $\mathbf{X} = \Psi \mathbf{S}$ and $\mathbf{C}\Psi$ satisfies the restricted isometry property, as in the next section.

3.2. Invariance of DMD to unitary transformations. In this section, we show that the dynamic mode decomposition is invariant to left and right unitary transformations of the data \mathbf{X} and \mathbf{X}' . We first assume that we have an $n \times n$ unitary measurement matrix \mathbf{C} with the number of measurements p equal to n . We will then relax the condition of \mathbf{C} being a square unitary measurement matrix, and instead require that our data is sparse in some basis and the $p \ll n$ measurements are incoherent with respect to that basis.

This section relies on the fact that the singular value decomposition of $\mathbf{Y} = \mathbf{C}\mathbf{X}$ is related to the singular value decomposition of $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^*$ if \mathbf{C} is unitary:

$$\mathbf{Y} = \mathbf{C}\mathbf{U}\Sigma\mathbf{V}^* \quad (14)$$

Now \mathbf{CU} are the left-singular vectors of \mathbf{Y} . Similarly, if $\mathbf{Y} = \mathbf{XP}$ for unitary \mathbf{P} , then $\mathbf{Y} = \mathbf{U}\Sigma\mathbf{V}^*\mathbf{P}$. This is discussed in more detail in the Appendix.

Theorem 3.5. *The DMD eigenvalues and eigenvectors are invariant to right-transformations \mathbf{P} of the columns of \mathbf{X} and \mathbf{X}' if \mathbf{P} is unitary.*

Proof. Let \mathbf{P} be a $m \times m$ unitary matrix that acts on the columns of \mathbf{X} and \mathbf{X}' as: $\mathbf{Y} = \mathbf{XP}$, and $\mathbf{Y}' = \mathbf{X}'\mathbf{P}$. The four steps of the exact DMD algorithm proceed as:

1. $\mathbf{Y} = \mathbf{U}\Sigma\mathbf{V}^*\mathbf{P}$
2. $\tilde{\mathbf{A}}_{\mathbf{Y}} = \mathbf{U}^*\mathbf{Y}'\mathbf{P}^*\mathbf{V}\Sigma^{-1} = \mathbf{U}^*\mathbf{X}'\mathbf{V}\Sigma^{-1} = \tilde{\mathbf{A}}_{\mathbf{X}}$
3. $\tilde{\mathbf{A}}_{\mathbf{Y}}\mathbf{W}_{\mathbf{Y}} = \mathbf{W}_{\mathbf{Y}}\Lambda_{\mathbf{Y}}$ is equivalent to $\tilde{\mathbf{A}}_{\mathbf{X}}\mathbf{W}_{\mathbf{X}} = \mathbf{W}_{\mathbf{X}}\Lambda_{\mathbf{X}}$, so that $\mathbf{W}_{\mathbf{Y}} = \mathbf{W}_{\mathbf{X}}$ and $\Lambda_{\mathbf{Y}} = \Lambda_{\mathbf{X}}$.
4. $\Phi_{\mathbf{Y}} = \mathbf{Y}'\mathbf{P}^*\mathbf{V}\Sigma^{-1}\mathbf{W}_{\mathbf{Y}} = \Phi_{\mathbf{X}}$.

Therefore, the DMD eigenvalues and eigenvectors are invariant to right-unitary transformations of the data. \square

Corollary 2. *The DMD eigenvalues and eigenvectors are invariant to permutations of the columns of \mathbf{X} and \mathbf{X}' , since a permutation matrix \mathbf{P} is unitary.*

Theorem 3.6. *The DMD eigenvalues are invariant to left-transformations \mathbf{C} of data \mathbf{X} if \mathbf{C} is unitary, and the resulting DMD modes are projected through \mathbf{C} .*

Proof. Again, let $\mathbf{Y} = \mathbf{C}\mathbf{X}$ and $\mathbf{Y}' = \mathbf{C}\mathbf{X}'$. Exact DMD proceeds as follows:

1. $\mathbf{Y} = \mathbf{C}\mathbf{U}\Sigma\mathbf{V}^*$
2. $\tilde{\mathbf{A}}_{\mathbf{Y}} = \mathbf{U}^*\mathbf{C}^*\mathbf{Y}'\mathbf{V}\Sigma^{-1} = \mathbf{U}^*\mathbf{X}'\mathbf{V}\Sigma^{-1} = \tilde{\mathbf{A}}_{\mathbf{X}}$
3. The eigendecomposition $(\Lambda_{\mathbf{Y}}, \mathbf{W}_{\mathbf{Y}})$ is equal to $(\Lambda_{\mathbf{X}}, \mathbf{W}_{\mathbf{X}})$ since $\tilde{\mathbf{A}}_{\mathbf{Y}} = \tilde{\mathbf{A}}_{\mathbf{X}}$.
4. $\Phi_{\mathbf{Y}} = \mathbf{Y}'\mathbf{V}\Sigma^{-1}\mathbf{W}_{\mathbf{Y}} = \mathbf{C}\Phi_{\mathbf{X}}$.

Therefore, the DMD modes $\Phi_{\mathbf{Y}}$ are the projection of $\Phi_{\mathbf{X}}$ through \mathbf{C} : $\Phi_{\mathbf{Y}} = \mathbf{C}\Phi_{\mathbf{X}}$. \square

Corollary 3. *The DMD computed in the spatial domain is related to the DMD computed in the Fourier domain or in a coordinate system defined by principle components. Both the discrete Fourier transform \mathcal{F} and the principal component coordinate transformation \mathbf{U} are unitary transformations, and so Theorem 3.6 applies with $\mathbf{C} = \mathcal{F}$ or $\mathbf{C} = \mathbf{U}$, respectively. The DMD eigenvalues will be unchanged, and the DMD eigenvectors will be projected into the new coordinates.*

Now we may consider data \mathbf{X}, \mathbf{X}' that are sparse in some basis Ψ , and **relax the condition that \mathbf{C} is unitary**. Instead, \mathbf{C} must be incoherent with respect to the sparse basis, so that the product $\mathbf{C}\Psi$ satisfies a restricted isometry property with small constant δ_K . This allows us to compute DMD on the compressed data

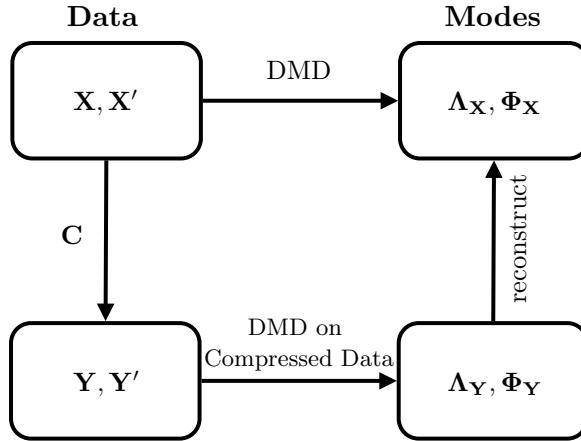


FIGURE 2. Schematic of compressive-sampling and compressed DMD as they relate to data \mathbf{X}, \mathbf{X}' and compressed data \mathbf{Y}, \mathbf{Y}' . \mathbf{C} is a projection down to measurements that are incoherent with respect to sparse basis.

$(\mathbf{Y}, \mathbf{Y}') = (\mathbf{C}\mathbf{X}, \mathbf{C}\mathbf{X}') = (\mathbf{C}\Psi\mathbf{S}, \mathbf{C}\Psi\mathbf{S}')$ and reconstruct full-state DMD modes by compressed sensing on $\Phi_{\mathbf{Y}} = \mathbf{C}\Phi_{\mathbf{X}} = \mathbf{C}\Psi\Phi_{\mathbf{S}}$. The invariances above persist if $\mathbf{C}\Psi$ acts as an isometry on sparse vectors and the columns of \mathbf{X}, \mathbf{X}' are sparse in Ψ . This can be verified by working through the method of snapshots in the Appendix on the compressed data.

3.3. Various approaches and algorithms. There are a number of algorithms that arise from various paths in Figure 2 depending on what data we have access to. The primary paths are: Path 1B (compressed DMD) and Path 2B (compressed sensing DMD). A schematic of the data flow is shown in Fig. 3.

Path 1: We start with full-state data \mathbf{X}, \mathbf{X}' .

Option A. Compute DMD to obtain $(\Lambda_{\mathbf{X}}, \Phi_{\mathbf{X}})$.

Option B. Compress data first:

- (i) Compress \mathbf{X}, \mathbf{X}' to \mathbf{Y}, \mathbf{Y}' .
- (ii) Compute DMD to obtain $(\Lambda_{\mathbf{Y}}, \Phi_{\mathbf{Y}})$ and $\mathbf{W}_{\mathbf{Y}}$.
- (iii) Reconstruct $\tilde{\Phi}_{\mathbf{X}} = \mathbf{X}'\mathbf{V}_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-1}\mathbf{W}_{\mathbf{Y}}$.
- (iv) [alternative to (iii)] Perform ℓ_1 -minimization on $\Phi_{\mathbf{Y}} = \mathbf{C}\Psi\Phi_{\mathbf{S}}$ to reconstruct $\Phi_{\mathbf{S}}$ and then construct $\Phi_{\mathbf{X}} = \Psi\Phi_{\mathbf{S}}$.

Path 2: We only have compressed data \mathbf{Y}, \mathbf{Y}' .

Option A. First reconstruct \mathbf{X}, \mathbf{X}' using compressed sensing.

- (i) Perform ℓ_1 -minimization on $\mathbf{Y} = \mathbf{C}\Psi\mathbf{S}$ to solve for \mathbf{S} , and hence \mathbf{X} (same for \mathbf{X}').
- (ii) Compute DMD on \mathbf{X} (or \mathbf{S}).

Option B. Compute DMD on compressed data and only reconstruct r modes using compressed sensing.

- (i) Compute DMD on \mathbf{Y}, \mathbf{Y}' to obtain $(\Lambda_{\mathbf{Y}}, \Phi_{\mathbf{Y}})$
- (ii) Perform ℓ_1 -minimization on $\Phi_{\mathbf{Y}}$ to solve $\Phi_{\mathbf{Y}} = \mathbf{C}\Psi\Phi_{\mathbf{S}}$ for $\Phi_{\mathbf{S}}$ and $\Phi_{\mathbf{X}}$.

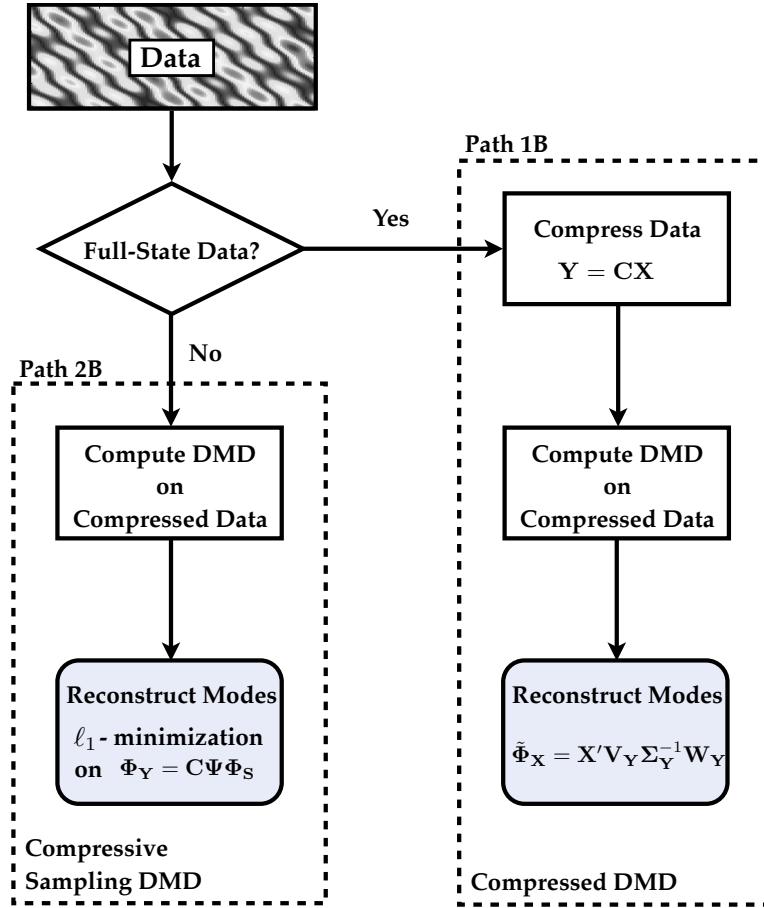


FIGURE 3. [New Figure] Flow-chart illustrating compressed DMD and compressed sensing DMD.

As a general rule, if full data \mathbf{X}, \mathbf{X}' is available, Path 1B (i)-(iii) is the most efficient option. This is because the most expensive step, consisting of an SVD on full data, is replaced by an SVD on compressed data; the computational cost of the SVD is reduced from $\mathcal{O}(nm^2)$ (full SVD) to either $\mathcal{O}(pm^2)$ (if $p > m$) or $\mathcal{O}(mp^2)$ (if $p < m$). We refer to this as *compressed DMD*. If only the compressed data \mathbf{Y}, \mathbf{Y}' is available, Path 2B is more computationally efficient than Path2A, since there will be fewer DMD modes than columns of \mathbf{Y} and \mathbf{Y}' , resulting in fewer ℓ_1 minimization problems. We refer to this as *compressed sensing DMD*. Path 2 is, in general, quite computationally expensive, especially for large problems, because of the ℓ_1 -minimization. However, when full-state data is unavailable, such as for examples in ocean or atmospheric sampling problems, the cost of compressed sensing is secondary, since the alternative is a prohibitively expensive NP-hard brute force search for DMD modes. Note that alternative (iv) in Path 1B is presented for completeness, although it is not recommended because of the unnecessary expense of ℓ_1 -minimization when full-state data is available.

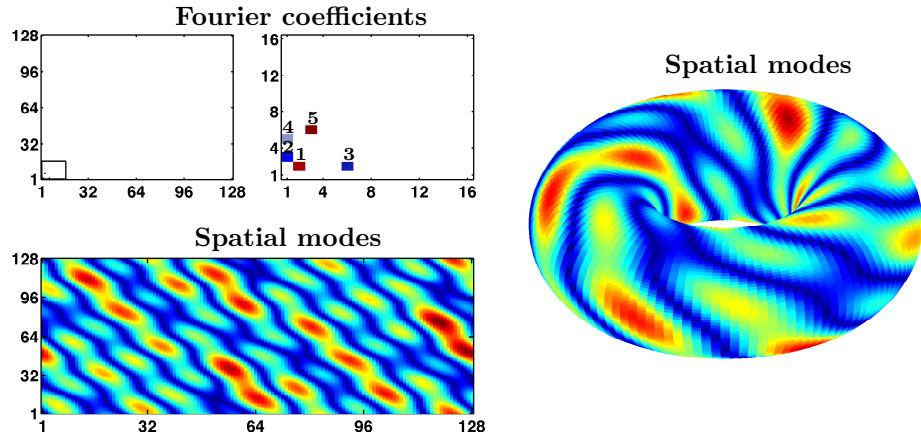


FIGURE 4. Illustration of the dynamical system in Example 1. $K = 5$ Fourier coefficients are driven with a random linear dynamical system, resulting in the spatial dynamics shown. Fifteen point sensors are placed randomly in space. For video, see <http://faculty.washington.edu/sbrunton/simulations/csdmd/Torus.mp4>.

4. Results. We illustrate the above methods on two example problems that are relevant for fluid dynamics. In the first example, we construct a spatially evolving system that has sparse dynamics in the Fourier domain. In the second example, we consider the time-varying double-gyre, which has been used as a model for ocean mixing. Both examples are coded in Matlab, using the built-in `svd` command, which is based on LAPACK routines [2]. In each example, the sparse basis Ψ is the two-dimensional inverse discrete Fourier transform (iDFT2) matrix.

4.1. Example 1: Sparse linear system in Fourier domain. This system is designed to test to compressive DMD algorithms in a well-controlled numerical experiment. We impose sparsity by creating a system with $K = 5$ non-zero 2D spatial Fourier modes; all other modes are exactly zero. It is also possible to allow the other Fourier modes to be contaminated with Gaussian noise and impose a very fast stable dynamic in each of these directions. We then define a stable linear, time-invariant dynamical system on the K modes. This is done by randomly choosing a temporal oscillation frequency and small but stable damping rate for each of the modes independently. In this way, we construct a system in the spatial domain that is a linear combination of coherent spatial Fourier modes that each oscillate at a different fixed frequency. Table 1 contains the specific values for this example. Figure 4 shows a snapshot of this system at $t = 2$. We see the five large Fourier mode coefficients generate distinct spatial coherent patterns. Figure 5 shows the five Fourier modes (real and imaginary parts) that contribute to the spatial structures in Figure 4.

This example is constructed to be an ideal test case for compressed sampling dynamic mode decomposition. The linear time-invariant system underlying these dynamics is chosen at random, so the data matrix \mathbf{X} will contain significant energy from many of the modes. Therefore, POD does not separate the spatial modes, as shown in Figure 6. Instead, POD extracts structures based on their variance in the data. Since POD is invariant to permutation of the columns, the underlying

TABLE 1. Parameters of system in for Example 1, visualized in Figure 4. Individual modes are obtained by constructing an oscillating and decaying Fourier mode (with eigenvalue $\lambda = d + i\omega$), $\hat{x}(I, J) = e^{\lambda t} = e^{dt} (\cos(\omega t) + i \sin(\omega t))$, and taking the real part of the inverse Fourier transform: $x(t) = \text{real}(\mathcal{F}^{-1}(\hat{x}))$.

Mode #	I (x -wavenumber)	J (y -wavenumber)	$d = \text{Real}(\lambda)$ (damping)	$\omega = \text{Imag}(\lambda)$ (frequency)	Initial Condition
1	2	2	-0.0880	4.7893	-0.5357
2	1	3	-0.0524	8.2214	0.9931
3	6	2	-0.0258	9.7120	-0.5434
4	1	5	-0.0637	6.8953	-0.4914
5	3	6	-0.0175	7.1518	-2.0610

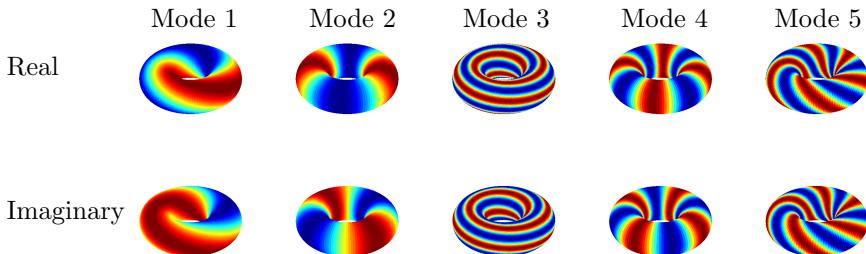


FIGURE 5. Spatial-temporal coherent modes corresponding to non-zero Fourier modes.

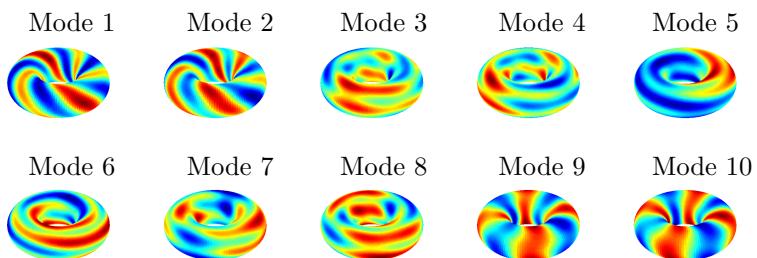


FIGURE 6. POD modes obtained from data. Energetic modes mix the underlying Fourier modes.

temporal structure is not reflected in the POD structures. Since each of our Fourier modes is oscillating at a fixed and distinct frequency, this is ideal for dynamic mode decomposition, which isolates the spatially coherent Fourier modes exactly, as shown in Figure 7.

In the case of no background noise, the compressed sensing DMD algorithm (Path 2B in Sec. 3.3) works extremely well, as seen in the mode reconstruction in Figures 8. Additionally, the method of compressed DMD (Path 1B in Sec. 3.3)

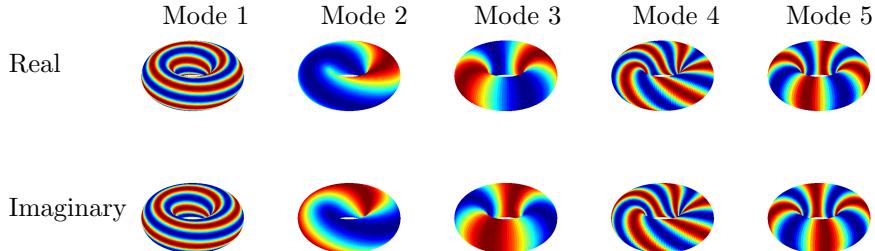


FIGURE 7. DMD modes correctly isolate spatially coherent modes.

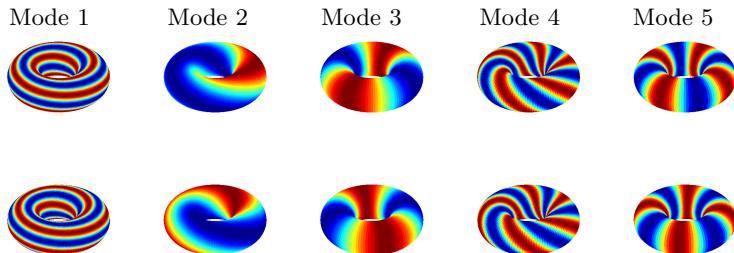
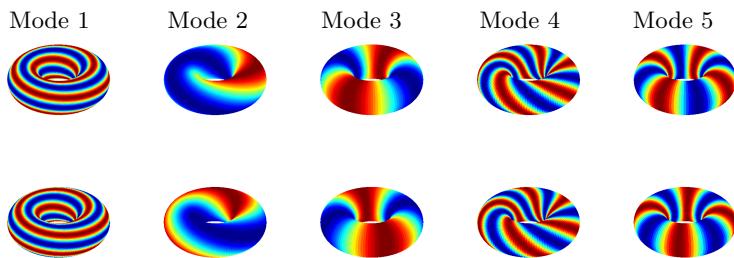


FIGURE 8. Compressive-sampling DMD modes, using matching pursuit. (Path 2B, Sec. 3.3)

FIGURE 9. DMD modes from compressed data, using $\tilde{\Phi}_X = X'V\Sigma^{-1}W_Y$. (Path 1B, Sec. 3.3)

starting with full-state snapshots, compressing, performing DMD, and then reconstructing using formula in Eq. (13) results in accurate reconstruction, shown in Figure 9. Both compressive-sampling DMD and compressed DMD match the true eigenvalues nearly exactly, as shown in Figure 10.

The \mathbf{X}, \mathbf{X}' data matrices are obtained on a 128×128 spatial grid from times 0 to 2 in increments of $\Delta t = 0.01$, so that $\mathbf{X} \in \mathbb{R}^{16384 \times 200}$. We use $p = 15$ measurements; randomly placed single-pixel measurements and measurements obtained by taking the dot product of a Gaussian or Bernoulli random vector with the state vector \mathbf{x} all yield accurate DMD eigenvalues and modes.

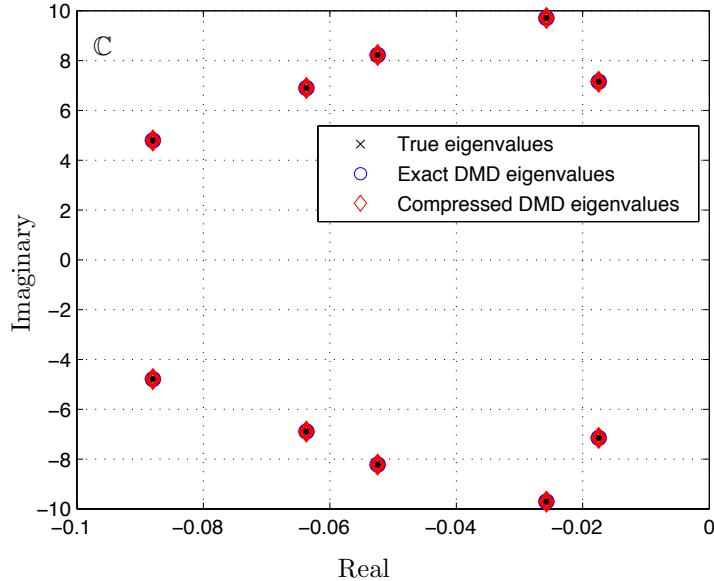


FIGURE 10. DMD modes captures dynamics faithfully.

TABLE 2. Accuracy of eigenvalues and modes for various DMD methods in Example 1.

		Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Eigenvalue error $ \lambda_k - \tilde{\lambda}_k $	Standard	1.635 e-13	1.025 e-12	1.091 e-12	6.238 e-13	1.676 e-13
	Compressed	1.358 e-13	5.266 e-12	7.597 e-13	9.729 e-13	4.115 e-13
	Compressive Sampling	1.358 e-13	5.266 e-12	7.597 e-13	9.729 e-13	4.115 e-13
Mode Error $\ \tilde{\phi}_k - \phi_k\ _2$	Standard	1.219 e-13	4.566 e-13	1.227 e-13	1.661 e-11	1.370 e-12
	Compressed	7.333 e-14	6.269 e-13	2.347 e-13	1.004 e-11	5.727 e-12
	Compressive Sampling	7.542 e-14	7.954 e-13	2.289 e-13	8.142 e-12	3.858 e-12

The accuracy of the compressed and compressed sensing DMD methods in reconstructing DMD modes and eigenvalues are summarized in Table 2. The accuracy of the standard DMD method is also included, since this example was constructed from a true low-rank linear model on FFT modes, providing an explicit solution. Since the DMD modes in this example are in complex conjugate pairs, and since these modes have a phase difference from the true FFT modes, we compare the magnitudes. For each of the five modes, the correct FFT modal wavenumber is identified, and all other wavenumbers are zero. The mode error in Table 2 arises

TABLE 3. Computation time for DMD methods in Example 1. Each method was run 1000 times to compute an average run-time.

		Standard DMD	Compressed DMD	Compressed Sensing DMD
SVD	Time [s]	1.254 e-1	2.459 e-4	2.459 e-4
	Speed-up	1	509.9	509.9
DMD	Time [s]	1.463 e-1	3.271 e-3	1.130 e-2
	Speed-up	1	44.73	12.95

from small differences in the magnitude of the single non-zero wavenumber. The computational time of each method has also been benchmarked and summarized in Table 3. For this example, the compressed sensing method is still faster than standard DMD because of the sparsity of the solution vector.

When we add small to moderate amounts of background noise (2-5% RMS) in the Fourier domain, a number of things change. The modes and frequencies are still very well characterized by both methods of compressed DMD and compressive-sampling DMD. However, the resulting DMD damping is often inflated; this effect is more pronounced when the full-state DMD eigenvalues are strongly damped. Recent results predict the change in DMD spectrum with small additive noise [3], and our results appear to be consistent. There are currently efforts to understand and correct for the effect of noise in the DMD computation [19, 28]. However, if the goal of a DMD model is closed-loop feedback control, then one may desire some controller robustness with respect to uncertain damping rates.

4.2. Example 2: Double gyre flow. The double gyre is a simple two-dimensional model [59] that is often used to study mixing between ocean basins. The double-gyre flow is given by the following stream-function

$$\begin{aligned} \psi(x, y, t) &= A \sin(\pi f(x, t)) \sin(\pi y) \\ f(x, t) &= \epsilon \sin(\omega t) x^2 + x - 2\epsilon \sin(\omega t) x, \end{aligned} \tag{15}$$

which results in the following time-periodic vector field

$$\begin{aligned} u &= -\frac{\partial \psi}{\partial y} = -\pi A \sin(\pi f(x)) \cos(\pi y) \\ v &= \frac{\partial \psi}{\partial x} = \pi A \cos(\pi f(x)) \sin(\pi y) \frac{df}{dx} \end{aligned} \tag{16}$$

on the closed and bounded domain $[0, 2] \times [0, 1]$. Typical parameter values are $A = 0.1, \omega = 2\pi/10, \epsilon = 0.25$.

Figure 11 illustrates the double gyre vector field, with color representing vorticity and arrows showing the vector field. The vorticity field is highly compressible, in that 99% of the Fourier coefficients may be zeroed with little effect on the reconstructed vorticity field (panel c).

The datasets \mathbf{X}, \mathbf{X}' are constructed by sampling the vorticity (curl of velocity field) on a 512×256 spatial grid at times 0 to 10 with $\Delta t = 0.05$. The POD modes

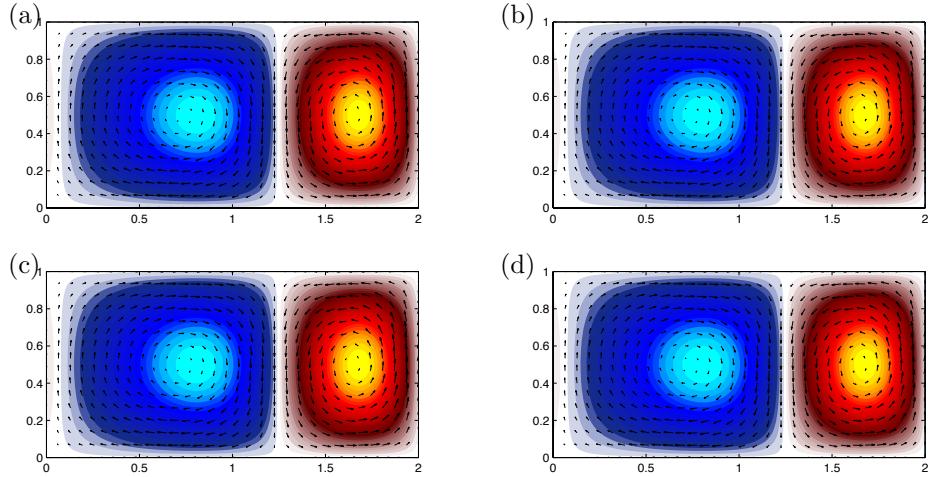


FIGURE 11. (a) Double gyre vector field overlaid onto vorticity for $t = 0.2$, $\epsilon = 0.25$, $A = 0.1$, $\omega = 2\pi/10$. (b) Logarithmically scaled Fourier coefficient magnitudes of the vorticity. (c) Compressed double gyre vorticity field using 1.0% of the Fourier coefficients. (d) Top 1.0% largest magnitude Fourier coefficients.

for the double gyre data set are shown in Figure 12 (a). The DMD modes are shown in Figure 12 (b). Taking 2500 single pixel measurements (i.e., constructing Y and Y' by sampling 2500 random rows of X and X'), which accounts for under 2% of the total pixels, the reconstructed modes are shown in Figure 12 (c) and (d). The modes exhibit good agreement. Moreover, the compressive DMD eigenvalues are close to the exact DMD eigenvalues, shown in Figure 13.

A quantitative comparison of the compressed DMD eigenvalues with the exact DMD eigenvalues is provided in Fig. 14. Only the first five modes are shown, since modes 6-9 are conjugate pairs with modes 1-4. The agreement between eigenvalues is surprisingly close, even for $p = 10$ compressed measurements. The compressed DMD eigenvalues continue to converge as the number of measurements p is increased. It is interesting to note that the lower frequency modes have significantly smaller errors in the eigenvalue approximations. Similar convergence plots of the L^2 error of the compressed DMD modes are shown in Fig. 15.

The compressed sensing DMD method yields the same eigenvalues as the compressed DMD algorithm, so Fig. 14 also applies to both methods. The convergence plots of the L^2 error of the compressed sensing DMD modes are shown in Fig. 16 for increasing numbers of measurements p . The convergence of modes is much slower than in compressed DMD, since full-state snapshots are not available for reconstruction. However, this is consistent with results from the compressive sensing literature. For the CoSaMP algorithm [42], we use 10 iterations and a desired sparsity of $K \sim p/4(1 + \log_{10} n - \log_{10} p)$, approximated from $p \sim 4K \log_{10}(n/K)$.

A benchmark comparison of the computational efficiency of the various algorithms is provided in Fig. 17. The compressed DMD algorithm is 30-40 times faster than the standard DMD algorithm for a small number of measurements ($p = 10$ to $p = 1250$), although the method is only 13 times faster for $p = 10000$ measurements.

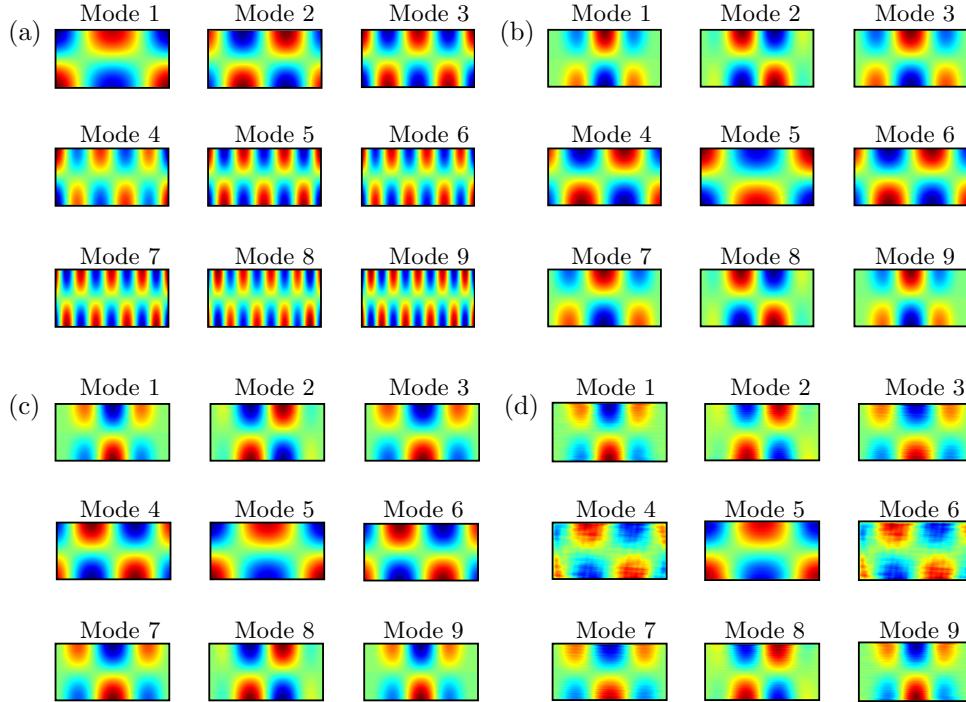


FIGURE 12. Modal decompositions for the double gyre: (a) POD modes, (b) DMD modes, (c) compressed DMD modes (Path 1B, Sec. 3.3), and (d) compressive-sampling DMD modes using matching pursuit (Path 2B, Sec. 3.3). Parts (c) and (d) use $p = 2500$.

Compressed sensing DMD is faster than DMD for a small number of measurements ($p < 300$), though it rapidly becomes inefficient, as is generally the case with compressed sensing. However, if full-state measurements are unavailable, the additional computational cost may be justified to approximate full-state DMD modes.

All results in Figs. 14–17 use p randomly placed single-pixel measurements. We have also investigated Gaussian random compression (i.e., `>>C=randn(p,n)`; in Matlab) and uniform random compression (i.e., `>>C=rand(p,n)`; in Matlab) of the data X , which yield similar results. Although the compressed measurements are chosen randomly in this work, there are recent studies that investigate optimal sensor placement for improved sparse approximation [11]. Compressed sensing works well with random measurements, because aliasing effects is highly unlikely.

DMD has been widely used for data generated by fluid dynamical systems, because of the ability to isolate spatially coherent modes that oscillate at a fixed frequency in time. For example, in [64], DMD modes are used to understand interactions between the shear layer and wake behind a finite-thickness flat plate. Recently, DMD has also been effectively used to find *background* structures by identifying DMD modes corresponding to $\lambda = 1$ (in discrete time) [25]. The DMD analysis on the double gyre flow identifies such a coherent background mode, corresponding to Mode 5. Without the DMD analysis, it would be unclear that this is coherent and unchanging in time.

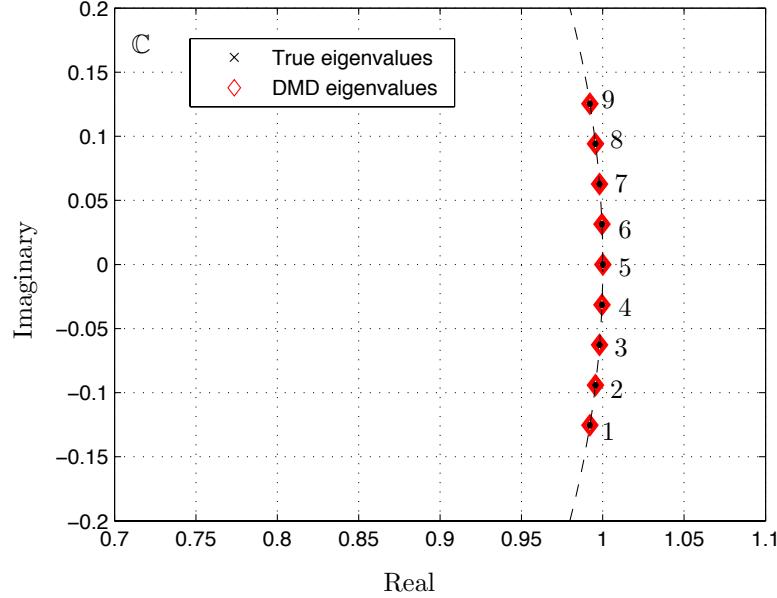


FIGURE 13. DMD modes captures dynamics faithfully. Modes 1-9 are labeled.

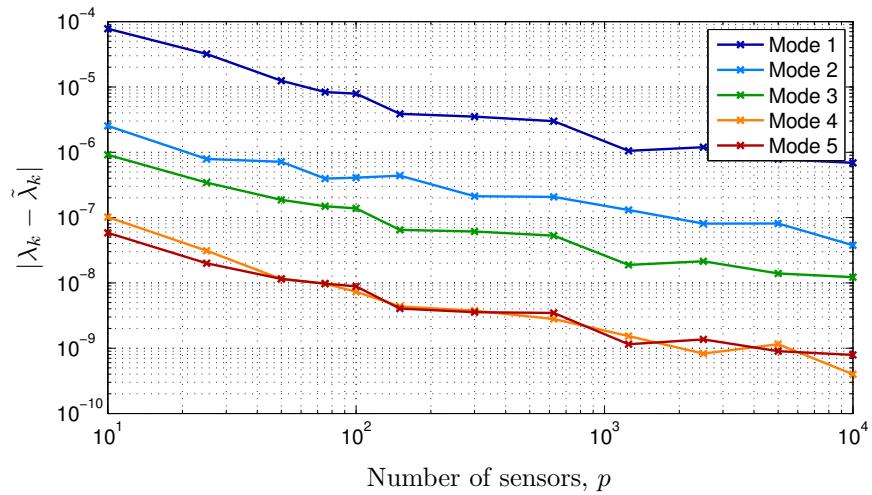


FIGURE 14. Convergence of compressed (and thus compressed sensing) DMD eigenvalues for increasing number of measurements p in Example 2.

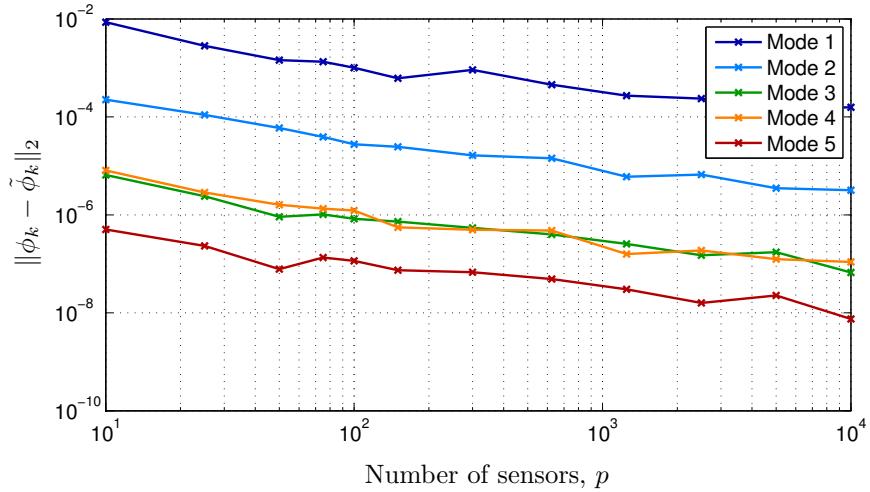


FIGURE 15. Convergence of compressed DMD modes for increasing number of measurements p in Example 2.

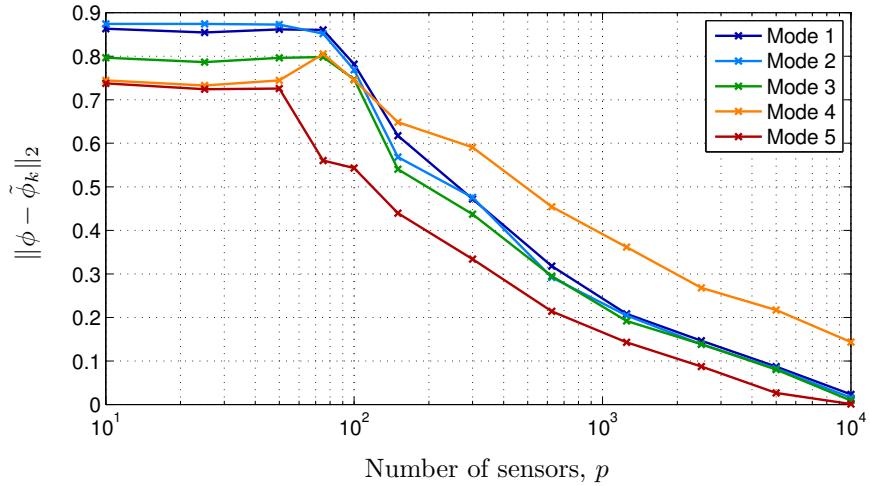


FIGURE 16. Convergence of compressed sensing DMD modes for increasing number of measurements p in Example 2.

5. Discussion. In this work we have developed a new method that leverages compressed sensing to compute the dynamic mode decomposition (DMD) for spatially subsampled or compressed data. There are two key advances resulting from this work. First, it is possible to reconstruct full-state DMD modes from heavily subsampled or compressed data using compressed sensing. This is called *compressive-sampling DMD*, and is illustrated in Sec. 3.3 and Figs. 2-3 as Path 2B. Second, if full-state snapshots are available, it is possible to first compress the data, perform DMD, and then reconstruct by taking a linear combination of the snapshot data,

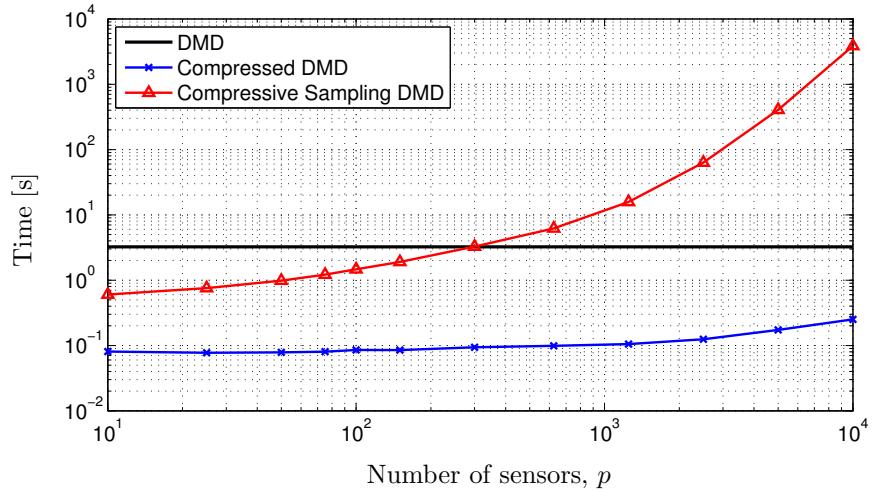


FIGURE 17. Algorithm run-time plotted against the number of measurements p in Example 2. Compressed DMD is between 13 and 40 times faster than standard DMD on this example.

determined by the DMD on compressed data. This is denoted by *compressed DMD*, and is explained in Path 1B in Sec. 3.3 and Figs 2-3. The theory relies on relationships between DMD on full-state and compressed data established in Sec. 3.1. We also show that DMD is invariant to left and right unitary transformations. We then relax this condition and use the restricted isometry property that is satisfied when incoherent measurements are applied to a signal that is sparse in some basis.

Both of these methods are demonstrated to be effective on two example problems with relevance to fluid dynamics and oceanographic/atmospheric sciences. In the first example, a low-order linear dynamical system is evolved on a few Fourier coefficients, establishing a high-dimensional, time-varying spatial flow with underlying low-rank patterns. In the second example, we consider the double gyre flow, which is a model for ocean mixing. The accurate and efficient reconstruction of the DMD eigenvalues and modes are demonstrated on both examples using few spatial measurements and compressed/compressive-sampling DMD.

There are a number of interesting directions that arise from this work. First, it will be a natural extension to apply these methods to high-dimensional systems in fluid dynamics and to oceanographic/atmospheric measurements. The reduced burden of spatial sampling may also allow for increased temporal sampling rates in particle image velocimetry (PIV), similar to recent advances in MRI [50, 45, 57]. In PIV, the data transfer from camera to RAM is a limiting factor, although one may envision compressing the full-state data before transferring to memory. It is also important to characterize and address the effect of noise on DMD [3, 19, 28], both full-state and compressed. The success of the compressed DMD method, even for very few measurements p , suggests that further theoretical developments are required, based on random matrix theory. Providing error bounds and theoretical limits for compressed DMD will be the subject of ongoing research. Finally, it may be possible to combine the spatial compressed sensing strategy advocated here with the temporal sampling strategy in [65] for greater efficiency gains.

Appendix: SVD and Unitary transformations. In this appendix we demonstrate that left or right-multiplication of a data matrix \mathbf{X} by a unitary transformation preserves all of the terms in the singular value decomposition except for the corresponding left or right unitary matrix \mathbf{U} or \mathbf{V} , respectively. These matrices are simply multiplied by the new unitary transformation.

5.1. Method of snapshots. Typically when computing the SVD of a large data matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$, with $n \gg m$, we use the method of snapshots [58]. In this method, we compute the $m \times m$ matrix $\mathbf{X}^* \mathbf{X}$ and solve the following eigendecomposition:

$$\begin{aligned}\mathbf{X}^* \mathbf{X} &= \mathbf{V} \boldsymbol{\Sigma}^2 \mathbf{V}^* \\ \implies \mathbf{X}^* \mathbf{X} \mathbf{V} &= \mathbf{V} \boldsymbol{\Sigma}^2\end{aligned}$$

With \mathbf{V} and $\boldsymbol{\Sigma}$ computed, it is possible to construct \mathbf{U} :

$$\mathbf{U} = \mathbf{X} \mathbf{V} \boldsymbol{\Sigma}^{-1}$$

Thus, we have the singular value decomposition: $\mathbf{X} = \mathbf{U}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}} \mathbf{V}_{\mathbf{X}}^*$. We have added the subscript \mathbf{X} to denote that the \mathbf{U} , $\boldsymbol{\Sigma}$, and \mathbf{V} are computed from \mathbf{X} .

5.2. Left unitary transformation. Now consider the data matrix $\mathbf{Y} = \mathbf{C} \mathbf{X}$, where \mathbf{C} is unitary. We find that

$$\mathbf{Y}^* \mathbf{Y} = \mathbf{X}^* \mathbf{C}^* \mathbf{C} \mathbf{X} = \mathbf{X}^* \mathbf{X},$$

so the compressed data has the same eigendecompositon and results in the same $\mathbf{V}_{\mathbf{X}}$ and $\boldsymbol{\Sigma}_{\mathbf{X}}$. Now, constructing $\mathbf{U}_{\mathbf{Y}}$, we have:

$$\mathbf{U}_{\mathbf{Y}} = \mathbf{Y} \mathbf{V}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} = \mathbf{C} \mathbf{X} \mathbf{V}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} = \mathbf{C} \mathbf{U}_{\mathbf{X}}.$$

Therefore, we have the following singular value decomposition: $\mathbf{Y} = \mathbf{C} \mathbf{U}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}} \mathbf{V}_{\mathbf{X}}^*$. In other words, $\mathbf{C} \mathbf{U}_{\mathbf{X}}$ are the new left-singular vectors.

5.3. Right unitary transformation. Similarly, consider $\mathbf{Y} = \mathbf{X} \mathbf{P}^*$, where \mathbf{P} is a unitary matrix. We find

$$\mathbf{Y}^* \mathbf{Y} = \mathbf{P} \mathbf{X}^* \mathbf{X} \mathbf{P}^* = \mathbf{P} \mathbf{V}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}}^2 \mathbf{V}_{\mathbf{X}}^* \mathbf{P}^*.$$

This results in the following eigendecomposition:

$$\mathbf{Y}^* \mathbf{Y} \mathbf{P} \mathbf{V}_{\mathbf{X}} = \mathbf{P} \mathbf{V}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}}^2.$$

Therefore, $\mathbf{V}_{\mathbf{Y}} = \mathbf{P} \mathbf{V}_{\mathbf{X}}$ and $\boldsymbol{\Sigma}_{\mathbf{Y}} = \boldsymbol{\Sigma}_{\mathbf{X}}$. It is then possible to construct $\mathbf{U}_{\mathbf{Y}}$:

$$\mathbf{U}_{\mathbf{Y}} = \mathbf{Y} \mathbf{P} \mathbf{V}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} = \mathbf{X} \mathbf{V}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} = \mathbf{U}_{\mathbf{X}}.$$

The singular value decomposition $\mathbf{Y} = \mathbf{U}_{\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}} \mathbf{V}_{\mathbf{X}}^* \mathbf{P}^*$ has $\mathbf{P} \mathbf{V}_{\mathbf{X}}$ as the new right-singular vectors.

Acknowledgements. S. L. Brunton acknowledges support from the University of Washington department of Mechanical Engineering and the eScience Institute. J. N. Kutz acknowledges support from the U.S. Air Force Office of Scientific Research (FA9550-09-0174). J. L. Proctor acknowledges support by Intellectual Ventures Laboratory. The authors thank Bingni Brunton for discussions on compressed sensing and for suggesting greedy algorithms.

REFERENCES

- [1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, 1988.
- [2] Edward Anderson, Zhaojun Bai, Christian Bischof, Susan Blackford, James Demmel, Jack Dongarra, Jeremy Du Croz, Anne Greenbaum, S Hammerling, Alan McKenney, et al. *LAPACK Users' guide*, volume 9. Siam, 1999.
- [3] S. Bagheri. Effects of small noise on the DMD/Koopman spectrum. *Bulletin Am. Phys. Soc.*, 58(18):H35.0002, p. 230, 2013.
- [4] S. Bagheri. Koopman-mode decomposition of the cylinder wake. *Journal of Fluid Mechanics*, 726:596–623, 2013.
- [5] 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.
- [6] Zhe Bai, Thakshila Wimalajeewa, Zachary Berger, Guannan Wang, Mark Glauser, and Pramod K Varshney. Low-dimensional approach for reconstruction of airfoil data via compressive sensing. *AIAA Journal*, pages 1–14, 2014.
- [7] R. G. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, 24(4):118–120, 2007.
- [8] 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.
- [9] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 23:539–575, 1993.
- [10] 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.
- [11] B. W. Brunton, S. L. Brunton, J. L. Proctor, and J. N. Kutz. Optimal sensor placement and enhanced sparsity for classification. *submitted for publication*, 2013.
- [12] Marko Budisic, Ryan Mohr, and Igor Mezić. Applied koopmanism a). *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(4):047510, 2012.
- [13] E. J. Candès. Compressive sensing. *Proceedings of the International Congress of Mathematics*, 2006.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, pages 21–30, 2008.
- [18] K. K. Chen, J. H. Tu, and C. W. Rowley. Variants of dynamic mode decomposition: Boundary condition, Koopman, and Fourier analyses. *Journal of Nonlinear Science*, 22(6):887–915, 2012.
- [19] Scott Dawson, Maziar Hemati, Matthew Williams, and Clarence Rowley. Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition. *Bulletin of the American Physical Society*, 59, 2014.
- [20] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [21] J. E. Fowler. Compressive-projection principal component analysis. *IEEE Transactions on Image Processing*, 18(10):2230–2242, 2009.
- [22] M. Gavish and D. L. Donoho. The optimal hard threshold for singular values is $4/\sqrt{3}$. *ArXiv e-prints*, 2014.
- [23] A. C. Gilbert and P. Indyk. Sparse recovery using sparse matrices. *Proceedings of the IEEE*, 98(6):937–947, 2010.
- [24] A. C. Gilbert, J. Y. Park, and M. B. Wakin. Sketched SVD: Recovering spectral features from compressive measurements. *ArXiv e-prints*, 2012.
- [25] J. Gosek and J. N. Kutz. Dynamic mode decomposition for real-time background/foreground separation in video. *submitted for publication*, 2013.

- [26] M. Grilli, P. J. Schmid, S. Hickel, and N. A. Adams. Analysis of unsteady behaviour in shockwave turbulent boundary layer interaction. *Journal of Fluid Mechanics*, 700:16–28, 2012.
- [27] F. Gueniat, L. Mathelin, and L. Pastur. A dynamic mode decomposition approach for large and arbitrarily sampled systems. *Physics of Fluids*, 27(2):025113, 2015.
- [28] Maziar S Hemati and Clarence W Rowley. De-biasing the dynamic mode decomposition for applied koopman spectral analysis. *arXiv preprint arXiv:1502.03854*, 2015.
- [29] B. L. Ho and R. E. Kalman. Effective construction of linear state-variable models from input/output data. In *Proceedings of the 3rd Annual Allerton Conference on Circuit and System Theory*, pages 449–459, 1965.
- [30] 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.
- [31] W. B Johnson and J. Lindenstrauss. Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics*, 26(189–206):1, 1984.
- [32] M. R. Jovanović, P. J. Schmid, and J. W. Nichols. Low-rank and sparse dynamic mode decomposition. *Center for Turbulence Research*, 2012.
- [33] J. N. Juang and R. S. Pappa. An eigensystem realization algorithm for modal parameter identification and model reduction. *Journal of Guidance, Control, and Dynamics*, 8(5):620–627, 1985.
- [34] 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.
- [35] B. O. Koopman. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931.
- [36] J. N. Kutz. *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*. Oxford University Press, 2013.
- [37] J. L. Lumley. *Stochastic Tools in Turbulence*. Academic Press, 1970.
- [38] Z. Ma, S. Ahuja, and C. W. Rowley. Reduced order models for control of fluids using the eigen-system realization algorithm. *Theoretical and Computational Fluid Dynamics*, 25(1):233–247, 2011.
- [39] Alan Mackey, Hayden Schaeffer, and Stanley Osher. On the compressive spectral method. *Multiscale Modeling & Simulation*, 12(4):1800–1827, 2014.
- [40] I. Mezić. Analysis of fluid flows via spectral properties of the Koopman operator. *Annual Review of Fluid Mechanics*, 45:357–378, 2013.
- [41] Igor Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1-3):309–325, 2005.
- [42] D. Needell and J. A. Tropp. CoSaMP: iterative signal recovery from incomplete and inaccurate samples. *Communications of the ACM*, 53(12):93–100, 2010.
- [43] 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. *Journal of Fluid Mechanics*, 497:335–363, 2003.
- [44] H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the A. I. E. E.*, pages 617–644, FEB 1928.
- [45] V. M. Patel and R. Chellappa. *Sparse representations and compressive sensing for imaging and vision*. Briefs in Electrical and Computer Engineering. Springer, 2013.
- [46] C. Penland. Random forcing and forecasting using Principal Oscillation Pattern analysis. *Mon. Weather Rev.*, 117(10):2165–2185, October 1989.
- [47] C. Penland and T. Magorian. Prediction of Niño 3 sea-surface temperatures using linear inverse modeling. *J. Climate*, 6(6):1067–1076, June 1993.
- [48] 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.
- [49] 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.
- [50] A. C. Sankaranarayanan, P. K. Turaga, R. G. Baraniuk, and R. Chellappa. Compressive acquisition of dynamic scenes. In *Computer Vision–ECCV*, pages 129–142, 2010.
- [51] 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.

- [52] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, August 2010.
- [53] P. J. Schmid. Application of the dynamic mode decomposition to experimental data. *Experiments in Fluids*, 50:1123–1130, 2011.
- [54] P. J. Schmid and J. Sesterhenn. Dynamic mode decomposition of numerical and experimental data. In *61st Annual Meeting of the APS Division of Fluid Dynamics*. American Physical Society, November 2008.
- [55] P. J. Schmid, D. Violato, and F. Scarano. Decomposition of time-resolved tomographic PIV. *Experiments in Fluids*, 52:1567–1579, 2012.
- [56] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.
- [57] J. V. Shi, W. Yin, A. C. Sankaranarayanan, and R. G. Baraniuk. Video compressive sensing for dynamic MRI. *submitted for publication*, 2013.
- [58] L. Sirovich. Turbulence and the dynamics of coherent structures, parts I-III. *Q. Appl. Math.*, XLV(3):561–590, 1987.
- [59] T. H. Solomon and J. P. Gollub. Chaotic particle transport in time-dependent Rayleigh-Bénard convection. *Physical Review A*, 38(12):6280–6286, 1988.
- [60] Gilles Tissot, Laurent Cordier, Nicolas Benard, and Bernd R Noack. Model reduction using dynamic mode decomposition. *Comptes Rendus Mécanique*, 2014.
- [61] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231–2242, 2004.
- [62] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk. Beyond Nyquist: Efficient sampling of sparse bandlimited signals. *IEEE Transactions on Information Theory*, 56(1):520–544, 2010.
- [63] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz. On dynamic mode decomposition: theory and applications. *Journal of Computational Dynamics*, 1(2):391–421, 2014.
- [64] Jonathan H Tu, Clarence W Rowley, Ehsan Aram, and Rajat Mittal. Koopman spectral analysis of separated flow over a finite-thickness flat plate with elliptical leading edge. *AIAA Paper 2011*, 2864, 2011.
- [65] Jonathan H Tu, Clarence W Rowley, J Nathan Kutz, and Jessica K Shang. Spectral analysis of fluid flows using sub-Nyquist rate PIV data. *Experiments in Fluids*, 55(9):1–13, 2014.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: sbrunton@uw.edu

E-mail address: joproctor@intven.com

E-mail address: jhtu@berkeley.edu

E-mail address: kutz@uw.edu