Low-rank and sparse dynamic mode decomposition

By M. R. Jovanović, P. J. Schmid AND J. W. Nichols

1. Motivation and objectives

Even though fluid flows possess an exceedingly high number of degrees of freedom, their dynamics often can be approximated reliably by models of low complexity. This observation has given rise to the notion of coherent structures – organized fluid elements that, together with dynamic processes, are responsible for the bulk of momentum and energy transfer in the flow. Recent decades have seen great advances in the extraction of coherent structures from experiments and numerical simulations. Proper orthogonal decomposition (POD) modes (Lumley 2007; Sirovich 1987), global eigenmodes, frequential modes (Sipp *et al.* 2010), and balanced modes (Moore 1981; Rowley 2005) have provided useful insight into the dynamics of fluid flows. Recently, dynamic mode decomposition (DMD) (Rowley *et al.* 2009; Schmid 2010) has joined the group of feature extraction techniques. Both POD and DMD are snapshot-based post-processing algorithms which may be applied equally well to data from simulations or experiments. By enforcing orthogonality, POD modes possess multi-frequential temporal content; on the other hand, DMD modes are characterized by a single temporal frequency. DMD modes may potentially be non-normal, but this non-normality may be essential to capturing certain types of dynamical effects. For an in-depth discussion of the connection between DMD and other data decomposition methods, please refer to Schmid (2010).

By projecting the full system onto the extracted modes, the governing equations may be replaced by a dynamical system with fewer degrees of freedom. This facilitates computationally tractable investigations of stability or receptivity as well as a model-based control design. In many situations, however, it is not trivial to identify a subset of modes that have the strongest impact on the flow dynamics. For example, spatial non-orthogonality of the DMD modes may introduce a poor quality of approximation of experimentally or numerically generated snapshots when only a subset of modes with the largest amplitude is retained. Recent attempts at extracting only a subset of desired frequencies and basis vectors (Chen *et al.* 2012) rely on a non-convex optimization problem whose solution in general requires an intractable combinatorial search.

In order to strike a balance between the quality of approximation (in the least-squares sense) and the number of modes that are used to approximate the given fields, this brief develops low-rank and sparsity-promoting versions of the standard DMD algorithm. To achieve this objective, we combine tools and ideas from linear algebra and convex optimization with the emerging area of compressive sensing (Candès & Tao 2006; Candès *et al.* 2006; Donoho 2006). We demonstrate that both low-rank and sparsity-promoting formulations yield convex optimization problems, for which efficient algorithms for finding a globally optimal solution can be developed, even for large-scale problems.

Our presentation is organized as follows. In Section 2, we formulate the problem and provide a brief overview of the dynamic mode decomposition and of the optimal selection of amplitudes of extracted DMD modes. In Section 3, we introduce a low-rank DMD

algorithm to identify an a priori specified number of modes that provide optimal approximation of experimental or numerical snapshots at a certain time interval. In Section 4, we use a sparsity-promoting framework to select a subset of DMD modes which strikes a balance between the approximation error (with respect to the full data sequence) and the number of extracted modes. In Section 5, we use the database resulting from the unstructured large-eddy simulation of a supersonic jet to illustrate the utility of the developed methods. We conclude our presentation in Section 6.

2. Problem formulation

2.1. *Dynamic mode decomposition*

The dynamic mode decomposition (DMD) is a data processing algorithm that extracts coherent structures with a single frequency from a (numerical or experimental) data-sequence. In what follows, we briefly outline the key steps of DMD.

We begin by collecting a sequence of snapshots from numerical simulations or physical experiments and form a data matrix whose columns represent the individual data samples. Even though we confine our attention to temporal evolution processes, note that the DMD-framework can accommodate a variety of "evolution coordinates" (e.g., spatial directions, or curved base-flow streamlines). Furthermore, we assume that the data are equispaced in time, with a time step Δt,

$$\{\psi_0, \psi_1, \ldots, \psi_N\},$$

where each ψ_i is, in general, a complex vector with M components (measurement points), i.e., $\psi_i \in \mathbb{C}^M$.

Next, we postulate that the snapshots have been generated by a discrete-time linear time-invariant system which maps a given snapshot to the subsequent snapshot

$$\psi_{t+1} \;=\; A\,\psi_t, \quad t \;=\; \{0, \ldots, N-1\}. \tag{2.1}$$

For fluid systems, the matrix A typically contains a large number of entries (which are in general complex numbers). DMD furnishes a procedure for determining a low-order representation of the matrix $A \in \mathbb{C}^{M\times M}$ that captures the dynamics inherent in the data sequence. The next computational step consists in forming two data matrices from the snapshot sequence according to

$$\begin{aligned} \Psi_0^{N-1} \;&=\; \left[\begin{array}{cccc} \psi_0 & \psi_1 & \cdots & \psi_{N-1} \end{array}\right] \;\in\; \mathbb{C}^{M\times N}, \\ \Psi_1^{N} \;&=\; \left[\begin{array}{cccc} \psi_1 & \psi_2 & \cdots & \psi_{N} \end{array}\right] \;\in\; \mathbb{C}^{M\times N} \end{aligned}$$

where the subscript and the superscript identify the first and the last elements of the sequence, respectively. In most fluid problems, the number of components (measurement points) in each snapshot ψ_i is typically much larger than the number of snapshots, $M \gg N$, thereby implying that Ψ_0^{N-1} and Ψ_1^N are tall rectangular matrices. The two data matrices Ψ_0^{N-1} and Ψ_1^N are linked via the matrix A, and using the linear relation (2.1) between the snapshots at two consecutive time steps, we can express Ψ_1^N as

$$\begin{aligned} \Psi_1^N \;&=\; \left[\begin{array}{cccc} \psi_1 & \psi_2 & \cdots & \psi_{N} \end{array}\right] \\ &=\; \left[\begin{array}{cccc} A\,\psi_0 & A\,\psi_1 & \cdots & A\,\psi_{N-1} \end{array}\right] \\ &=\; A\,\Psi_0^{N-1}. \end{aligned} \tag{2.2}$$

For sufficiently large data sequences, it is reasonable to postulate a linear dependence

of recent snapshots on previous ones (Schmid 2010). This assumption concurs with the supposition that the dynamical process that generated the snapshot sequence is indeed low dimensional and that it can be expressed by a projection onto a substantially smaller set of basis vectors. Consequently, the final snapshot ψ_N in our sequence can be expressed as a linear combination of $\{\psi_0, \psi_1, \ldots, \psi_{N-1}\}$, which yields

$$\psi_N \;=\; -a_0\,\psi_0 \;-\; a_1\,\psi_1 \;-\; \ldots \;-\; a_{N-1}\,\psi_{N-1} \;+\; r,$$

where r denotes the residual vector. This relation can be used to write Ψ_1^N as

$$\Psi_1^N \;=\; \Psi_0^{N-1}\,A_c \;+\; r\,e_N^T, \tag{2.3}$$

where e_N is the Nth unit-vector in $\mathbb{R}^N$, and A_c is a companion matrix determined by

$$A_c \;=\; \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & -a_{N-2} \\ & & & 1 & -a_{N-1} \end{bmatrix} \;\in\; \mathbb{C}^{N\times N}.$$

Using (2.2) and (2.3), we obtain the expression

$$\begin{aligned} \Psi_1^N \;&=\; A\,\Psi_0^{N-1} \\ &=\; \Psi_0^{N-1}\,A_c \;+\; r\,e_N^T \;\in\; \mathbb{C}^{M\times N}, \end{aligned} \tag{2.4}$$

which is reminiscent of an Arnoldi decomposition. The above expression implies that we can think of A_c as a representation of the matrix A on an N-dimensional subspace of $\mathbb{C}^M$ that is spanned by the columns of Ψ_0^{N-1}. In contrast to the classical Arnoldi method, this subspace is not spanned by a set of orthonormal basis vectors; this apparent disadvantage, however, allows the processing of data sequences without access to the matrix A.

The companion matrix A_c can be determined from the matrices of snapshots Ψ_0^{N-1} and Ψ_1^N by minimizing the Frobenius norm of the residual in (2.4)

$$\text{minimize} \quad \|\Psi_1^N \;-\; \Psi_0^{N-1}\,A_c\|_F^2. \tag{2.5}$$

We recall that the Frobenius norm of the matrix Q is determined by

$$\|Q\|_F^2 \;=\; \operatorname{trace}\left(Q^*\,Q\right) \;=\; \operatorname{trace}\left(Q\,Q^*\right).$$

For improved convergence and algorithmic robustness, we employ an economy-sized singular value decomposition of $\Psi_0^{N-1} \in \mathbb{C}^{M\times N}$

$$\Psi_0^{N-1} \;=\; U\,\Sigma\,V^*, \tag{2.6}$$

where Σ is an $N \times N$ diagonal matrix with non-zero singular values $\{\sigma_1, \ldots, \sigma_N\}$ on its main diagonal, and the matrices U and V are unitary, i.e.,

$$\begin{aligned} U \;&\in\; \mathbb{C}^{M\times N} \quad \text{with} \quad U^*\,U \;=\; I, \\ V \;&\in\; \mathbb{C}^{N\times N} \quad \text{with} \quad V^*\,V \;=\; V\,V^* \;=\; I. \end{aligned}$$

Upon substitution of (2.6) into (2.4) we arrive at

$$A\,U\,\Sigma\,V^* \;=\; U\,\Sigma\,V^*\,A_c \;+\; r\,e_N^T,$$

and multiplying the resulting expression with U^* (from the left) and with $V\,\Sigma^{-1}$ (from the right), we obtain a representation of A in the basis spanned by the POD modes of Ψ_0^{N-1}

$$U^*AU \;=\; \Sigma\,V^*A_c\,V\,\Sigma^{-1} \;+\; U^*r\,e_N^T\,V\,\Sigma^{-1}.$$

From this last expression we can determine the transformed version of the companion matrix A_c,

$$F \;=\; \Sigma\,V^*A_c\,V\,\Sigma^{-1} \;\in\; \mathbb{C}^{N\times N}.$$

In anticipation of modifications to treat the low-rank and sparse version of the above algorithm, we reformulate the DMD problem in terms of an optimization problem. A modest amount of straightforward algebraic manipulation brings the optimization problem (2.5) into the following form

$$\underset{F}{\text{minimize}} \quad \|\left(U^*\,\Psi_1^N\,V\,\Sigma^{-1} \;-\; F\right)\Sigma\,\|_F^2, \tag{2.7}$$

which implies that the optimal rank-N solution to (2.7) is determined by

$$F_{\text{dmd}} \;=\; U^*\,\Psi_1^N\,V\,\Sigma^{-1}.$$

This expression is identical to the expression given by Schmid (2010) and it concludes the robust implementation of the DMD algorithm starting from matrices of data snapshots Ψ_0^{N-1} and Ψ_1^N.

2.2. *Optimal amplitudes of DMD modes*

The matrix $F_{\text{dmd}} \in \mathbb{C}^{N\times N}$ is a low-dimensional representation of the inter-snapshot mapping $A \in \mathbb{C}^{M\times M}$ and it contains information about the spectrum and associated coherent structures. In addition to these quantities, amplitudes of the extracted dynamic modes can be computed. To this end, an eigenvalue decomposition of F_{dmd}

$$F_{\text{dmd}}\underbrace{\begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}}_{Y}\begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_N \end{bmatrix}$$

can be used to determine the matrix of DMD modes (Schmid 2010)

$$\Phi \;=\; \begin{bmatrix} \phi_1 & \cdots & \phi_N \end{bmatrix} \;=\; U\,Y \;\in\; \mathbb{C}^{M\times N}.$$

With this expression, we can represent experimental or numerical snapshots as a linear combination of DMD modes according to

$$\psi_t \;=\; \sum_{i\,=\,1}^{N} \alpha_i\,\mu_i^t\,\phi_i, \quad t \;\in\; \{0,\ldots,N-1\}\,,$$

or, equivalently, in matrix form,

$$\overbrace{\begin{bmatrix} \psi_0 & \psi_1 & \cdots & \psi_{N-1} \end{bmatrix}}^{\Psi_0^{N-1}} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_N \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_N \end{bmatrix}}_{D_\alpha = \operatorname{diag}\{\alpha\}} \underbrace{\begin{bmatrix} 1 & \mu_1 & \cdots & \mu_1^{N-1} \\ 1 & \mu_2 & \cdots & \mu_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mu_N & \cdots & \mu_N^{N-1} \end{bmatrix}}_{V_{\text{and}}}.$$

The temporal evolution of the dynamic modes in discrete time is governed by the Vandermonde matrix V_{and}; this matrix is determined by the N complex eigenvalues μ_i which contain information about underlying frequencies and growth/decay rates. Determination of the unknown vector of amplitudes $\alpha = \begin{bmatrix} \alpha_1 & \cdots & \alpha_N \end{bmatrix}^T$ then amounts to finding the solution to the following optimization problem

$$\underset{\alpha}{\text{minimize}} \quad \|\Psi_0^{N-1} \; - \; \Phi \, D_\alpha \, V_{\text{and}}\|_F^2.$$

Using the singular value decomposition (2.6) of Ψ_0^{N-1} and the definition of Φ, we bring this problem into the following form

$$\underset{\alpha}{\text{minimize}} \quad J(\alpha) \; = \; \|\Sigma \, V^* \; - \; Y \, D_\alpha \, V_{\text{and}}\|_F^2, \tag{2.8}$$

which is a convex optimization problem that can be solved using standard methods (Boyd & Vandenberghe 2004; Grant & Boyd 2012).

A superposition of all DMD modes, properly weighted by their amplitudes and advanced in time according to their complex frequency, optimally approximates the entire data sequence. The key challenge that this brief addresses is identification of a truncated representation in order to capture the most important dynamic structures (by eliminating features that contribute weakly to the data sequence).

3. Low-rank dynamic mode decomposition

As described above, the dynamic mode decomposition algorithm of Schmid (2010) provides the optimal rank-N solution to the optimization problem (2.7). In what follows, we provide a solution to (2.7) that has an a priori specified rank r with $r < N$. Under this constraint, we reduce the number of modes that influence the quality of approximation of experimental or numerical snapshots at non-zero times to r.

A simple change of variables,

$$H \; = \; U^* \, \Psi_1^N \, V, \quad F_\Sigma \; = \; F \, \Sigma,$$

transforms (2.7) into the following optimization problem

$$\underset{F_\Sigma}{\text{minimize}} \quad \|H \; - \; F_\Sigma\|_F^2. \tag{3.1}$$

It is well known that the singular value decomposition of the matrix H,

$$H \; = \; U_h \, \Sigma_h \, V_h^* \; = \; \sum_{i=1}^{N} \sigma_{hi} \, u_{hi} \, v_{hi}^*,$$

provides the optimal rank r solution F_r to the optimization problem (3.1). The rank-r approximation to F_{dmd} is thus determined by

$$F_r \;=\; \left(\sum_{i=1}^{r} \sigma_{hi}\, u_{hi}\, v_{hi}^{*}\right) \Sigma^{-1}. \tag{3.2}$$

Following the procedure of the full-rank dynamic mode decomposition, the eigenvalue decomposition of F_r,

$$F_r\, Y_r \;=\; \Lambda_r\, Y_r,$$

can now be used to determine the matrix Θ of *low-rank DMD modes*

$$\Theta \;=\; \left[\begin{array}{ccc} \theta_1 & \cdots & \theta_N \end{array}\right] \;=\; U\, Y_r \;\in\; \mathbb{C}^{M\times N}.$$

Analogous to the full-rank case, the original data sequence can be approximated using a linear combination of dynamic modes, properly weighted by their amplitude and their time-dependence

$$\psi_t \;=\; \sum_{i=1}^{N} \beta_i\, \lambda_i^t\, \theta_i, \quad t \;\in\; \{0,\ldots,N-1\}\,. \tag{3.3}$$

Since only r eigenvalues (e.g., $\{\lambda_1, \ldots, \lambda_r\}$) in (3.3) are non-zero, the quality of approximation for $t \in \{1,\ldots,N-1\}$ is determined by the amplitudes $\{\beta_1, \ldots, \beta_r\}$ associated with the r non-zero eigenvalues of the matrix F_r. The problem of determining optimal amplitudes can now be approached in a similar manner as in Section 2.2, namely by solving the following optimization problem

$$\underset{\beta}{\text{minimize}} \quad \|\Sigma\, V^{*} \;-\; Y_r\, D_\beta\, V_{\mathrm{and}}^{\lambda}\|_F^2, \tag{3.4}$$

where $\beta = \left[\begin{array}{ccc} \beta_1 & \cdots & \beta_N \end{array}\right]^T \in \mathbb{C}^N$ is the vector of the unknown amplitudes, $D_\beta = \mathrm{diag}\,\{\beta\}$, and $V_{\mathrm{and}}^{\lambda}$ is the $N \times N$ Vandermonde matrix corresponding to the eigenvalues $\{\lambda_1, \ldots, \lambda_N\}$ of F_r. Alternatively, the amplitudes $\{\beta_1, \ldots, \beta_r\}$ associated with the non-zero eigenvalues of the matrix F_r can be determined to optimally approximate $\{\psi_1, \ldots, \psi_{N-1}\}$, and the remaining amplitudes $\{\beta_{r+1}, \ldots, \beta_N\}$ can be determined to provide an optimal approximation of the initial condition ψ_0.

The low-rank DMD modes and the associated eigenvalues do not represent a subset of their standard DMD counterparts; in fact, they could be very different from each other.

4. Sparse dynamic mode decomposition

In this section, we direct our attention to the problem of selecting the subset of DMD modes that has the most profound influence on the quality of approximation of snapshot sequence. In this sense, we are interested in a hierarchical description of the data sequence in terms of a set of dynamic modes. Our approach consists of two steps. In the first step, a *sparsity structure* is sought which strikes a user-defined balance between the approximation error with respect to the full data sequence and the number of extracted modes; see (4.2) below. In the second step, the sparsity structure for the amplitudes (identified in the first step) is fixed and the optimal values of the non-zero amplitudes are determined; see (4.3) below.

We approach this problem by augmenting the objective function $J(\alpha)$ in (2.8) with an additional term, $\mathbf{card}\,(\alpha)$, that penalizes the number of non-zero elements in the vector

of unknown amplitudes α,

$$\underset{\alpha}{\text{minimize}} \quad J(\alpha) \; + \; \gamma \, \textbf{card}\,(\alpha)\,. \tag{4.1}$$

In the modified optimization problem (4.1), γ is a positive parameter that reflects our emphasis on sparsity of the vector α. Larger values of γ place stronger emphasis on the number of non-zero elements in the vector α (relative to the quality of the least-squares approximation, $J(\alpha)$), thereby encouraging sparser solutions to (4.1).

In general, finding solutions to the problem (4.1) amounts to a combinatorial search that quickly becomes intractable for any problem of interest. To circumvent this predicament we will consider a relaxed version of (4.1) that is obtained by replacing the cardinality function by the ℓ_1-norm of the vector α. We thus obtain the substitute optimization problem

$$\underset{\alpha}{\text{minimize}} \quad J(\alpha) \; + \; \gamma \sum_{i\,=\,1}^{N} |\alpha_i|. \tag{4.2}$$

Note that the sparsity-promoting DMD problem (4.2) is a convex optimization problem whose global solution, for small and medium sizes, can be found using standard optimization solvers (Boyd & Vandenberghe 2004; Grant & Boyd 2012). We have also developed efficient iterative algorithms to solve (4.2) for large-scale problems. The details of our algorithms will be reported elsewhere.

After a specified balance between the quality of approximation of our experimental or numerical snapshots and the number of DMD modes is achieved, we fix the sparsity structure of the unknown vector of amplitudes and determine only the non-zero amplitudes as a solution to the following constrained convex optimization problem:

$$\begin{aligned} \underset{\alpha}{\text{minimize}} \quad & \|\Sigma\, V^* \; - \; Y\, D_\alpha\, V_{\text{and}}\|_F^2 \\ \text{subject to} \quad & E^T \alpha \; = \; 0. \end{aligned} \tag{4.3}$$

In this expression, the matrix $E \in \mathbb{R}^{N\times m}$ encodes information about the sparsity structure of the vector α. The columns of E are the unit vectors in $\mathbb{R}^N$ whose non-zero elements correspond to zero components of α. For example, for $\alpha \in \mathbb{C}^4$ with

$$\alpha \; = \; \begin{bmatrix} \; \alpha_1 & 0 & \alpha_3 & 0 \; \end{bmatrix}^T,$$

the matrix E is given as

$$E \; = \; \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

5. Example: a screeching supersonic jet

In this section, we apply low-rank and sparse DMD to a database of snapshots obtained from a large eddy simulation of a screeching supersonic rectangular jet. Screech is a component of supersonic jet noise that is connected to the presence of a train of shock cells within the jet column (Tam 1995). For a turbulent jet, the unsteady shear layers interact with the shocks to create sound. While this process is generally broadbanded, screech is a special case of shock-noise that arises from creation of a feedback loop between the upstream-propagating part of the acoustic field and the generation of new disturbances at the nozzle lip. This self-sustaining feedback loop leads to an extremely

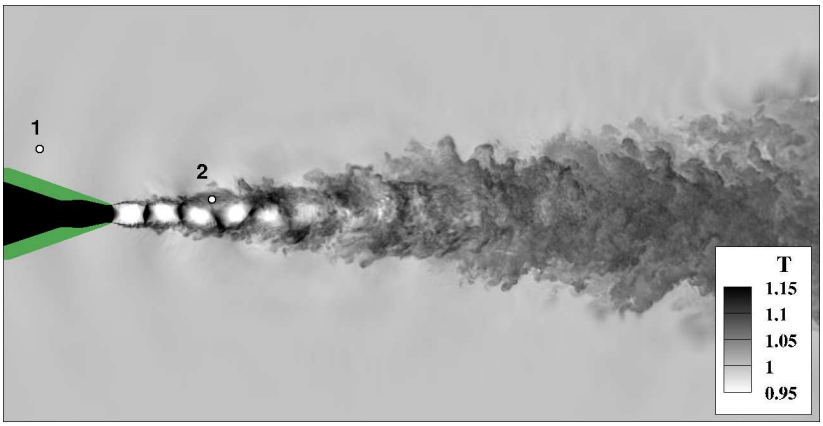

FIGURE 1. Temperature contours on a centerplane cross section taken from a snapshot of the screeching jet. Time histories of pressure were recorded at probe locations indicated by circles 1 and 2.

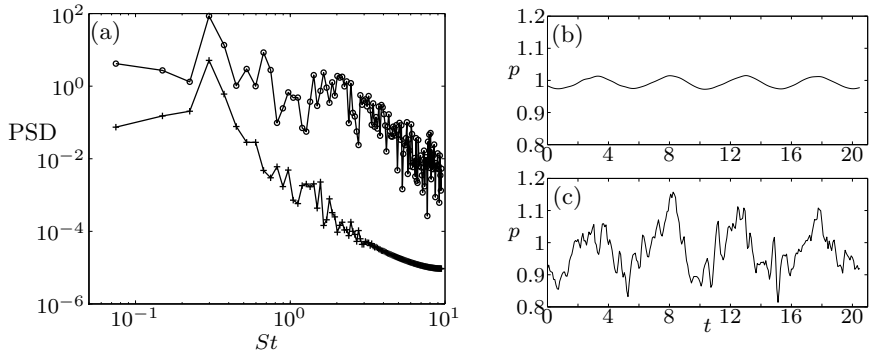

FIGURE 2. (a) Power spectra of pressure corresponding to the measurement locations indicated in Figure 1. Both spectra peak at $St \approx 0.3$, although the spectrum at location 2 (circles) is more broadbanded than at location 1 (crosses). Subfigures (b) and (c) show the corresponding time-dependence of pressure at locations 1 and 2, respectively.

loud (narrow-banded) screech tone at a specific fundamental frequency. The presence of a tonal process embedded in an otherwise broadbanded turbulent flow makes the screeching jet an excellent test case for the low-rank and sparse DMD methods. The objective of both methods is to extract the entire coherent screech feedback loop from the turbulent data and to describe the screech mechanism with as few modes as possible.

The supersonic jet used in this example was produced by a convergent rectangular nozzle of aspect ratio 4, precisely matching the geometry of an experimental nozzle (Frate & Bridges 2011). The entire flow inside, outside, and downstream of the nozzle was simulated using the low-dissipation low-dispersion LES solver `charles` on an unstructured mesh with about 45 million control volumes. This simulation was one of a sequence involving different mesh resolutions, and was validated against the experimental measurements (Nichols *et al.* 2011).

The stagnation pressure and temperature inside the nozzle were set so that the jet Mach number $M_j = 1.4$ and the fully expanded jet temperature matched the ambient temperature. Since the nozzle was purely convergent, however, the jet emerged from the nozzle in an underexpanded state, leading to a train of diamond-shaped shock cells as

shown in Figure 1. The figure shows contours of temperature on a centerplane cross section taken through the narrow dimension of the nozzle. From an animation of the jet (not shown), the first two shock cells are almost stationary, but the third and fourth shock cells undergo transverse oscillations along the narrow dimension of the jet (vertical in the figure). The transverse oscillation of the shock cells occurs precisely at the screech frequency of the jet, and is connected to a strong upstream-oriented acoustic tone.

Figure 2 shows the time histories and spectra of the pressure recorded at the locations indicated by the circles in Figure 1. Location 1 coincides with the center of the upstream-directed acoustic beam associated with the screech tone. At this location, we observe that the pressure spectrum is relatively narrow-banded. The pressure signal at location 2 contains a range of frequencies corresponding to the turbulence of the jet shear layers, resulting in the broadbanded spectrum. Both spectra show a strong peak at a Strouhal number of $St = fD_e/u_j = 0.3$, which is the screech tone frequency predicted for a Mach 1.4 jet from a 4:1 aspect ratio rectangular nozzle. Since the sample window included about 4 periods of the screech tone, in both cases this peak coincides with the fourth Fourier coefficient.

The database used for the DMD analysis consisted of 257 snapshots (so that $N = 256$) of the full three-dimensional pressure and velocity fields associated with the jet. The snapshots were equispaced in time with an interval of $\Delta t = 0.0528 D_e/u_j$, where D_e is the nozzle equivalent diameter (the diameter of the circle of same area as the nozzle exit) and u_j is the fully expanded jet velocity. Although the computational domain for the LES extended $\approx 32D_e$ downstream of the nozzle exit, DMD was applied to a subdomain focusing on the shock cells within the jet's potential core and the surrounding acoustic field (this domain extends to $10D_e$). This restriction reduced the number of cells from 45 million to 8 million. In spite of this reduction, each snapshot required 256Mb of storage in double precision format. To handle such large matrices, DMD was implemented in a MapReduce framework so that the matrix could be stored and processed across several storage discs. In particular, the algorithm relied upon a MapReduce QR-factorization of tall-and-skinny matrices developed by Constantine & Gleich (2011).

Figure 3(a) illustrates the dependence of the absolute value of the amplitudes of the DMD modes obtained by solving optimization problem (2.8) on the frequency of the corresponding eigenvalues. This type of plot is similar to the power-spectral density plots resulting from standard Fourier analysis.

It is not trivial to identify by mere inspection a subset of DMD modes that has the strongest influence on the quality of the least-squares approximation. As shown in Figure 3(b), the largest amplitudes appear to originate from the eigenvalues that are rather strongly damped. In what follows, we demonstrate that keeping only a subset of modes with largest amplitudes can lead to poor quality of approximation of numerically generated snapshots.

The sparsity level $\mathbf{card}\,(\alpha)$ and the Frobenius norm of the difference between Ψ_0^{N-1} and $\Phi\, D_\alpha\, V_{\mathrm{and}}$ for the optimal vector of amplitudes α resulting from the sparsity-promoting DMD algorithm are shown in Figure 4 as a function of the user-specified parameter γ (a measure of preference between approximation quality and solution sparsity). As expected, larger values of γ encourage sparser solutions, at the expense of compromising quality of the least-squares approximation. The values of γ in Figure 4 are selected in such a way that $\gamma_{\min}$ induces a dense vector α (with 256 non-zero elements), and $\gamma_{\max}$ induces α with a single non-zero element.

Eigenvalues resulting from the standard DMD algorithm (circles) along with the subset

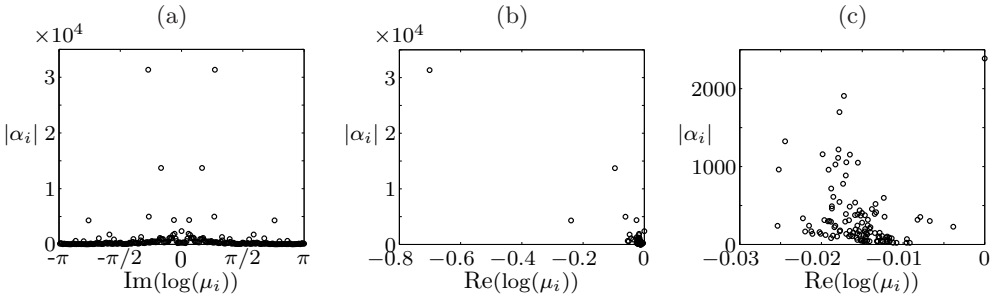

FIGURE 3. Dependence of the absolute value of the DMD amplitudes α_i on (a) the frequency and (b,c) the real part of the corresponding DMD eigenvalues μ_i. Subfigure (c) focuses on the amplitudes that correspond to lightly damped eigenvalues; it represents a zoomed version of subfigure (b).

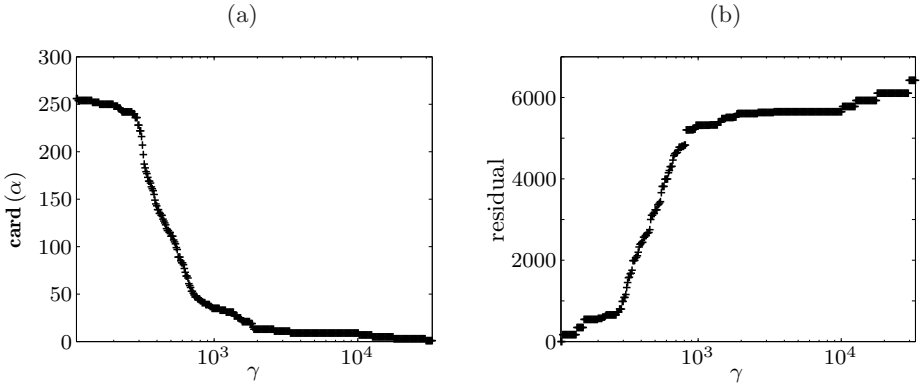

FIGURE 4. (a) The sparsity level **card** (α) and (b) the residual $\|\Psi_0^{N-1} - \Phi\, D_\alpha\, V_{\text{and}}\|_F$ of the optimal vector of amplitudes α resulting from the sparsity-promoting DMD algorithm.

of N_z eigenvalues selected by the sparsity-promoting DMD algorithm (crosses) are shown in Figure 5. Eigenvalues in the interior of the unit circle correspond to strong damping $\mathrm{Re}(\log(\mu_i))$. Because a strongly damped mode influences only the initial portion of the signal, its associated amplitude $|\alpha_i|$ can be large, as shown in Figure 3(b). For decreasing values of N_z, the sparsity-promoting DMD identifies eigenvalues that are responsible for the bulk of the dynamics contained in the data sequence. The selection of the retained eigenvalues is non-trivial, but increasingly concentrates on the low-frequency modes as N_z decreases. For $N_z = 3$, only the mean flow and one dominant frequency pair (that corresponds to the fundamental frequency of the screech tone) remain, while for $N_z = 5$ a second, lower frequency is identified. In this low-N_z limit we do see that the sparsity-promoting DMD approximates the processed data sequence using the most prevalent structures. The corresponding amplitudes for the various truncations N_z are displayed in Figure 6. Again, as N_z decreases a concentration on low frequencies is observed. We also note that the amplitudes of the original DMD do not provide sufficient guidance in reducing the full set of DMD modes to a few relevant ones.

A different picture emerges when probing the low-rank dynamic mode decomposition. As mentioned earlier, the eigenvalue distribution and modal shapes can deviate significantly from their standard counterparts. This is illustrated in Figure 7, where the eigenvalues μ_i are plotted in the complex plane (crosses) and compared with the eigenvalues resulting from full-rank DMD (circles). Except for the mean-flow eigenvalue, the

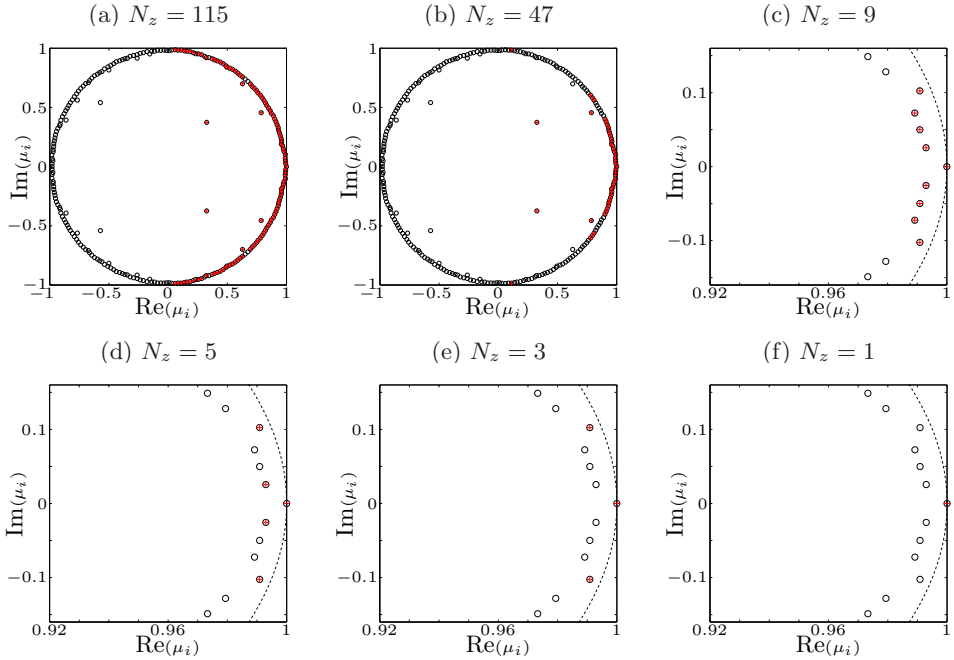

FIGURE 5. Eigenvalues resulting from the standard DMD algorithm (circles) along with the subset of N_z eigenvalues selected by the sparsity-promoting DMD algorithm (crosses). In subfigures (c-f), the dashed curves identify the unit circle.

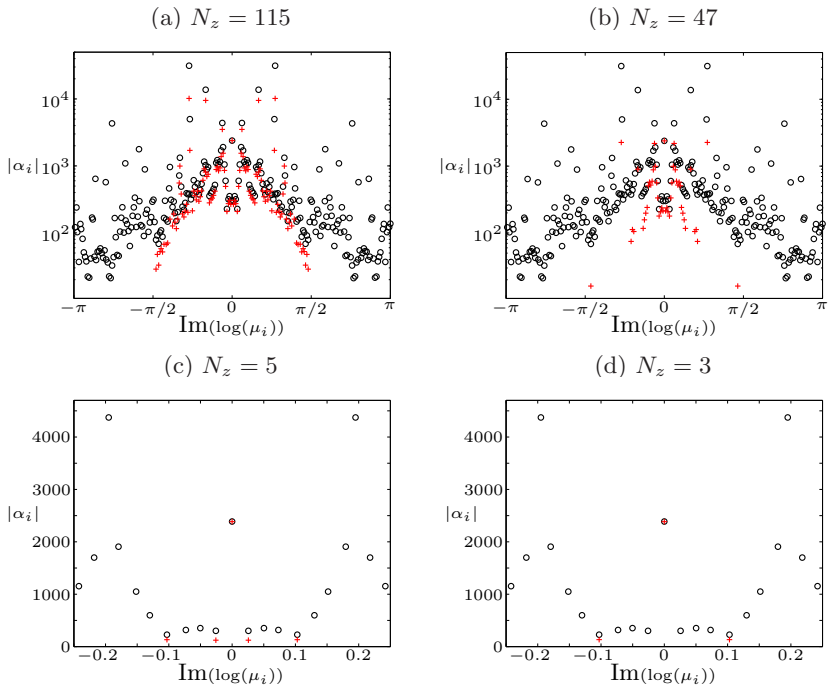

FIGURE 6. Dependence of the absolute value of the amplitudes α_i on the frequency (imaginary part) of the corresponding eigenvalues μ_i. The results are obtained using the standard DMD algorithm (circles) and the sparsity-promoting DMD algorithm (crosses) with N_z DMD modes.

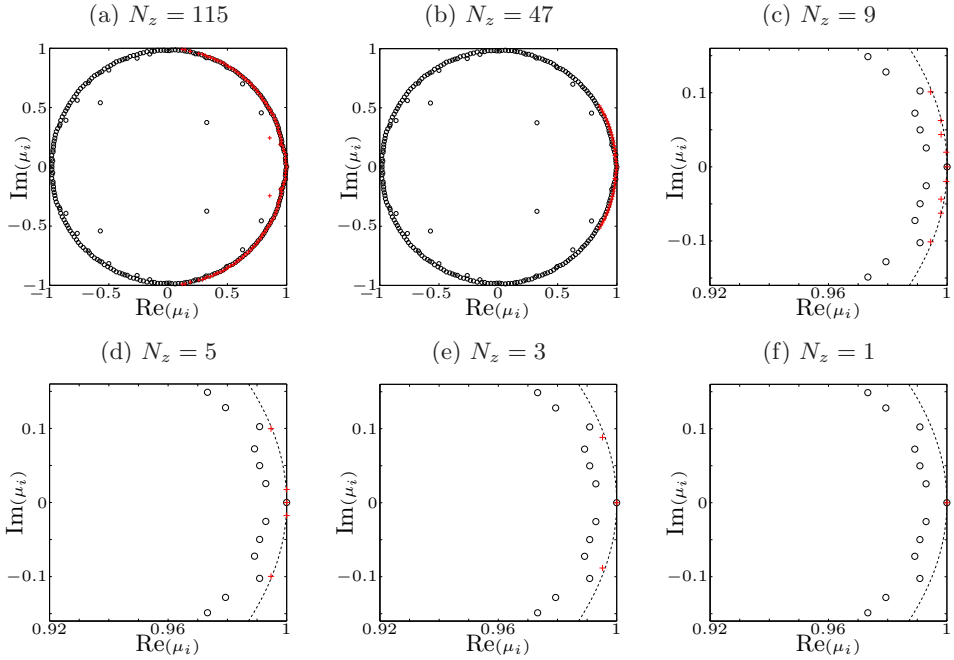

FIGURE 7. Eigenvalues resulting from the standard DMD algorithm (circles) and the low-rank DMD algorithm with r non-zero eigenvalues (crosses). In subfigures (c-f), the dashed curves identify the unit circle.

low-rank eigenvalues deviate markedly from the standard ones, even though we observe the same concentration on low-frequency structures as in the sparsity-promoting case. This deviation of the low-rank eigenvalues is such that they are driven towards the unit circle. Unlike a Fourier transform, the frequencies $\mathrm{Im}(\log(\mu_i))$ selected by the low-rank DMD are not uniform (see especially Figure 7(c)) and furthermore are not multiples of $1/T$ where T is the sample window length. We also note that low-rank DMD eliminates the high-amplitude, highly damped eigenvalues faster than sparse DMD as N_z is decreased.

Finally, Figure 8 juxtaposes the low-rank and sparsity-promoting DMD algorithms developed in this brief. The performance of each algorithm is quantified by the Frobenius norm of the approximation error between a low-dimensional representation and the full data sequence (in terms of the number of included structures). As expected, the performance of both algorithms deteriorates as the number of modes is reduced; however, Figure 8 demonstrates that the loss of performance for the low-rank DMD is far more precipitous than that for the sparsity-promoting DMD.

6. Concluding remarks

We have introduced two extensions of the standard DMD algorithm that address the reduction of the full rank-N decomposition to a low-dimensional one. The first algorithm (low-rank DMD) is based on the representation of the approximation problem by a rank-r solution (with $r < N$). The second algorithm (sparse DMD) uses ℓ_1-norm of the vector of the unknown amplitudes as a proxy for enforcing sparsity; a trade-off between the quality of approximation of the snapshot sequence and the number of DMD modes is achieved by augmenting the least-squares optimization problem by a term that penalizes

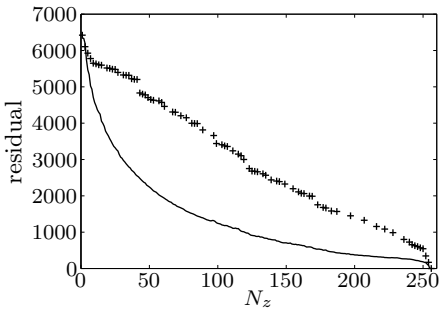

FIGURE 8. The optimal residual of the sparsity-promoting (symbols) and low-rank (solid line) DMD algorithms.

the ℓ_1-norm of the amplitude vector. This induces a sparse vector of amplitudes and consequently selects DMD modes and frequencies that have strongest influence on the quality of approximation on a given time interval.

Both algorithms have been tested on a data set of a screeching jet and have shown their value in quantitative data analysis. Even though the reported results are preliminary, they show a direction of research with great promise and a wealth of applications.

Acknowledgments

The authors gratefully acknowledge Prof. Parviz Moin for his encouragement to pursue this effort during the 2012 Center for Turbulence Research Summer Program.

REFERENCES

BOYD, S. & VANDENBERGHE, L. 2004 *Convex optimization*. Cambridge University Press.

CANDÈS, E. J., ROMBERG, J. & TAO, T. 2006 Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **52** (2), 489–509.

CANDÈS, E. J. & TAO, T. 2006 Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inf. Theory* **52** (12), 5406–5425.

CHEN, K., TU, J. & ROWLEY, C. 2012 Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses. *J. Nonlinear Sci.* DOI 10.1007/s00332-012-9130-9.

CONSTANTINE, P. G. & GLEICH, D. F. 2011 Tall and skinny QR factorizations in MapReduce architectures. In *Proceedings of the 2nd international workshop on MapReduce and its applications*, pp. 43–50.

DONOHO, D. L. 2006 Compressed sensing. *IEEE Trans. Inf. Theory* **52** (4), 1289–1306.

FRATE, F. C. & BRIDGES, J. E. 2011 Extensible rectangular nozzle model system. AIAA Paper 2011-975.

GRANT, M. & BOYD, S. 2012 CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`.

LUMLEY, J. 2007 *Stochastic Tools in Turbulence*. Dover Publications.

MOORE, B. 1981 Principal component analysis in linear systems: controllability, observability and model reduction. *IEEE Trans. Automat. Control* **AC-26(1)**.

Nichols, J. W., Ham, F. E. & Lele, S. K. 2011 High-fidelity large-eddy simulation for supersonic rectangular jet noise prediction. AIAA Paper 2011-2919.

Rowley, C. 2005 Model reduction for fluids using balanced proper orthogonal decomposition. *Int. J. Bifurcation Chaos* **15**, 997–1013.

Rowley, C., Mezic, I., Bagheri, S., Schlatter, P. & Henningson, D. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**, 115–127.

Schmid, P. 2010 Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28.

Sipp, D., Marquet, O., Meliga, P. & Barbagallo, A. 2010 Dynamics and control of global instabilities in open flows: a linearized approach. *Appl. Mech. Rev.* **63**, 030801.

Sirovich, L. 1987 Turbulence and the dynamics of coherent structures. Part I: Coherent structures. *Quart. Appl. Math.* **45(3)**, 561–571.

Tam, C. K. W. 1995 Supersonic jet noise. *Annu. Rev. Fluid Mech.* **27**, 17–43.