Efficiency of Sparse Dynamic Mode Decomposition in Mode Selection

Israa Fakih
Department of Mathematics, EPFL

1 Abstract

Dynamic mode decomposition (DMD) is a powerful tool for extracting the coherent structures and essential dynamics of complex flow fields. However, in many cases, traditional DMD may yield modes that contribute minimally to the overall approximation quality of the snapshot sequence. Sparse DMD addresses the issue of selecting the modes that have a significant influence and contribute the most by enforcing certain sparsity structure on the vector of amplitudes. This goal is achieved by augmenting the least square residuals with a regularization term. Utilizing the L_1 norm regularization facilitates this objective due to its inherent sparsity-inducing properties and transforming the optimization problem into convex one. Then the algorithm alternates between enforcing sparsity and minimizing the residual. Once a certain balance or trade-off is attained, the sparsity structure is fixed and the nonzero amplitudes are determined. Several examples from flow fields are used to evaluate the performance of sparse DMD and its efficiency in capturing the significant modes.

2 Introduction

In the study of fluid dynamics, the behavior of the fluid flows is often described by an infinite dimensional system governed by nonlinear equations. Despite the inherent complexity of these systems, they can often be approximated effectively by models of low complexity. And this gives the rise of some spatially organised patterns or features, coherent structures, that persist in the flow and play significant role in its dynamics. Understanding and identifying these coherent structures is at the core of unraveling the intricate dynamics and transport phenomena present within fluid flows.

In recent years, there has been a surge in developing techniques for extracting these structures. Among these techniques are the proper orthogonal decomposition (POD) and the dynamic mode decomposition (DMD). Both of these methods depend on a sequence of snapshots, which are numerical values observable in the flow field, such as vorticity, pressure, velocity, or temperature, to extract dominant spatial and temporal patterns inherent within the data. However, POD inforces orthogonality on the spatial modes whereas DMD modes may be non-normal, meaning they do not satisfy the orthogonality condition. But this non normality enables DMD to capture certain dynamical effects that may not be captured by POD.

However, not all modes contributes the same in the fluid flows and the challenge remains in capturing the ones that exert the most significant influence on flow dynamics . This limitation could introduce unwanted noise and give poor approximation. Some may propose to just truncate the modes with least amplitudes, however, such an approach may overlook modes with crucial effects on the flow dynamics. In response to this challenge, we introduce in this paper a variant of the DMD known as sparse dynamic mode decomposition (Sparse DMD). In the sparse DMD the computation of the modes remains unchanged while a novel approach is employed to compute their amplitudes. The goal is to enforce a certain sparsity structure on the vector of amplitudes by imposing a regularity term on the least squares problem between the matrix of snapshots and the linear combination of DMD modes. Through this regularity, l_1 is utilized which serves as a convex relaxation of the non-convex cardinality function. One of the sophisticated ways to solve this convex optimization problem is the usage of alternating direction of multipliers. This iterative method alternates between promoting sparsity and minimizing the least square residual. By striking the balance between them, a vector of amplitudes is obtained where modes with significant effects retain non-zero amplitudes, while less influential modes are effectively suppressed with zero amplitudes. Once the desired sparsity structure is achieved, we compute the optimal amplitudes of the dynamic modes to obtain a concise and interpretable representation of the flow dynamics.

Through this paper, in section 2, we provide a theoretical background on the koopman operator and DMD modes, elucidating their formulation and significance in the context of flow analysis. In section 3, we introduce the sparse DMD and present its formulation by including the derivation of the two optimization problems and their solution. In section 4, we present the evaluation

of the decomposition by checking the performance loss on a synthetic example. Furthermore, we validate it on two real world datasets: flow around a cylinder and flow over a heated cylinder with Boussinesq approximation. In section 6, we summarize the key findings and provide a detailed proof of the theoretical part in the appendix.

3 Theoretical Background and Problem formulation

3.1 Koopman Operator

The evolution of a fluid system represented by x with the state space $\mathcal{X} \in \mathbb{R}^M$ is given by the following system:

$$\begin{cases} \frac{dx}{dt} = F(x(t)) \\ x(t_0) = x_0 \end{cases}$$

where F is in general a nonlinear function that maps x_{i+1} to x_i over a time interval Δt. This function arises usually from partial differential equations which captures the essential characteristics of the flow from physical modeling and conservation principles.
In the case when F is linear, the solution of the system is expressed in terms of the spectral elements of F. Assume that Z and Λ are the eigenvectors and eigenvalues of F respectively (i.e. $FZ = \Lambda Z$) then the solution is $x_k = \sum_{i=1}^{M} b_i z_i e^{\lambda_i t}$ where b_i are the corresponding amplitudes calculated from the initial condition. These eigenvectors are called DMD modes. However as F is in general nonlinear we aim in transforming this system into a linear one.

Consider taking some observable of the flow which can be physical quantities such as temperature, pressure or energy of the inaccessible state x denoted by $\psi(x) : \mathcal{X} \to \mathbb{C}$. In this case the dynamics of the nonlinear system is projected into the space of observable by introducing Koopman operator $\mathcal{K}$ such that

$$\mathcal{K}(\psi(x_i)) = \psi(x_{i+1}) \tag{1}$$

It is clear from 1 that the Koopman operator is linear. Hence the koopman operator transforms the problem with a nonlinear function to one with linear operator. So the function ψ on which $\mathcal{K}$ operates can be written as a linear combination of the Koopman eigenfunctions $\{\phi_i\}$.

3.2 DMD Matrix

Each numerical value of a scalar or vector valued observable is a snapshot. Consider forming snapshot matrix S of M snapshots obtained in $N+1$ repeated

application of $\mathcal{K}$ with time lag Δt

$$S = \begin{bmatrix} \psi_1(x_0) & \psi_1(x_1) & ... & \psi_1(x_N) & \psi_1(x_{N+1}) \\ \psi_2(x_0) & \psi_2(x_1) & ... & \psi_2(x_N) & \psi_2(x_{N+1}) \\ : & : & : & : & : \\ : & : & : & : & : \\ \psi_M(x_0) & \psi_M(x_1) & ... & \psi_M(x_N) & \psi_M(x_{N+1}) \end{bmatrix}$$

Then it makes sense to consider a matrix $A \in \mathbb{C}^{M \times M}$ such that: $A\psi(x_t) = (\mathcal{K}\psi)(x_t) = \psi(x_{t+1})$.
In this Scenario, if we define $X = [\psi(x_0), ..., \psi(x_N)]$ and $Y = [\psi(x_1), ..., \psi(x_{N+1})]$ then the relationship between X and Y can be succinctly expressed as

$$Y = AX \tag{2}$$

So the matrix A solves the following least square problem $\min \|AX - Y\|_F^2$. In general this problem has infinite solution if X^T has a nontrivial null space. DMD aims for the minimal norm solution which is given explicitly by $A = YX^\dagger$ where $X^\dagger$ is the Moore–Penrose pseudoinverse of X. Now by considering the action of $\mathcal{K}$ on the space spanned by $\{\psi_1, \psi_2, ..., \psi_M\}$, we seek a matrix $\mathbf{U}$ that minimizes the residual of

$$(\mathcal{K}\psi_i)(x_k) = \sum_{j=1}^{M} u_{ji}\psi_j(x_k) + \rho_i(x_k), i = 1, ..., M, k = 0, ..., N \tag{3}$$

which is equivalent to solving the following least square problem

$$\min \|X^T\mathbf{U} - Y^T\|_F$$

It is clear from here that $\mathbf{U}^T = A$. Assume that $\mathbf{U} = Q\Lambda Q^{-1}$ then

$$(\mathcal{K}\phi_i)(x_k) \approx \lambda_i\phi_i(x_k)$$

where $\phi_i = \psi_i Q$. Therefore the Koopman mode decomposition

$$\begin{pmatrix} \mathcal{K}^k\psi_1(x) \\ \mathcal{K}^k\psi_2(x) \\ : \\ \mathcal{K}^k\psi_M(x) \end{pmatrix} = \sum_{i=1}^{M} z_i\phi_i\lambda_i^k \tag{4}$$

where $Z = Q^{-T}$. By using the fact that $\mathbf{U} = A^T$, the columns of Z are the right eigenvectors of A.

3.2.1 DMD modes

As discussed above to compute the Koopman or DMD modes we need to compute the eigenvectors of the matrix A. The eigenvalues and eigenvectors are computed using the Rayleigh quotient of A with respect to $Range(X)$. First we

start by taking the truncated SVD of $X = U\Sigma V^* \approx U_k\Sigma_k V_k^*$. The reason for considering a truncated SVD is that the smallest singular values can be computed with large errors due to the noise vector. In this case, the solution vector has an inaccurate component in the direction of the corresponding vector. Then using $Y = AU_k\Sigma_k V_k^*$, the Rayleigh quotient which is the DMD matrix has the form

$$F_{DMD} = U_k^* Y V_k \Sigma^{-1} \tag{5}$$

Each eigenpair (λ, w) of the matrix F_{DMD} gives the eigenpair $(\lambda, U_k w)$ of A.

3.2.2 Amplitudes of DMD modes

After computing the eigenvalues and eigenvectors of the DMD matrix, our next task is to approximate the numerical snapshots using linear combination of DMD mode, i.e.

$$\psi(x_i) = \psi^i = \sum_{j=1}^{l} \alpha_j z_j \lambda_j^{i-1}, i = 0, ..., N \tag{6}$$

where α_j are the amplitudes of the corresponding DMD modes. Determination of the unknown amplitudes corresponds to solving the following least square problem

$$\min_{\alpha} \|X - ZDiag(\alpha)V_{mode}\|_F^2$$

where $V_{mode} = \begin{pmatrix} 1 & \lambda_1 & ... & \lambda_1^{N-1} \\ 1 & \lambda_2 & ... & \lambda_2^{N-1} \\ : & : & : & : \\ 1 & \lambda_l & ... & \lambda_l^{N-1} \end{pmatrix}$ is the vandermonde matrix corresponding to the eigenvalues Λ. Using the SVD of the matrix X with $Z = U_k W$ and invariant of the norm under orthonormal matrices, the above problem could be rewritten as

$$\min_{\alpha} \|\Sigma_k V_k^* - WDiag(\alpha)V_{mode}\|_F^2 \tag{7}$$

Proposition 3.1. *The optimization problem 7 is equivalently represented in a quadratic minimization problem*

$$\min_{\alpha} H(\alpha) = \min_{\alpha}(\alpha^* P\alpha - q^*\alpha - \alpha^* q + s) \tag{8}$$

where

$$P = (W^*W) \circledast \overline{(V_{mode}V_{mode}^*)}, q = \overline{Diag(V_{mode}V\Sigma^* W)}, s = trace((\Sigma V^*)^*(\Sigma V^*))$$

where $\circledast$ represents the elementwise multiplication between two matrices and the overline signifies the complex conjugate.

By using 3.1, the solution to this quadratic minimization problem is

$$\alpha_{DMD} = P^{-1}q \tag{9}$$

The proof of this proposition is presented in appendix A.

4 Sparse DMD

Not all computed DMD modes have a great influence on the overall dynamics of the system. Including all the modes could introduce some noise and affect the accuracy of the data. To mitigate this issue, it is preferable to choose a subset of the DMD modes that contribute the most. One way is to ignore the modes with amplitudes less than a threshold value. However, sometimes the amplitudes computed by the traditional DMD are not sufficient to separate relevant from negligible modes especially when the data matrix contains statistical outliers or is otherwise compromised by high amplitude noise. For this reason, a better way in selecting subsets is the sparse DMD, which selects a subset of DMD modes that contributes the most to the quality of approximation. This goal is achieved in two steps. In the first step, the sparsity structure or pattern of the amplitudes α is identified. After establishing the sparsity pattern, the non-zero values of α are determined.

4.1 Sparsity Structure

To ascertain the sparsity pattern of the amplitude vector α, an additional term is added to the objective function in 8

$$\min_{\alpha} H(\alpha) + \gamma \sum_{i=1}^{l} |\alpha_i| \tag{10}$$

The reason for utilizing the L_1 norm goes back to its properties and ability to induce sparsity in the amplitude vector while preserving important features of the data and its robustness towards outliers. Unlike the L_2 norm which penalizes large values uniformly.

In order to solve the above optimization problem it is better to be written equivalently to constrained optimization problem with two variables

$$\min_{\alpha} \quad H(\alpha) + \gamma|\beta|_1 \text{ subject to} \qquad \alpha - \beta = 0 \tag{11}$$

In this formulation, an alternating method can be employed. This method alternates between having a sparsity structure and having a good approximation by minimzing the least square residual. First, we start with an initial β to calculate α. Subsequently, we fix α and calculate the updated β. This process iterates until we reach the optimal solution where the residuals are less than a certain tolerance.

$$res_{prim} = \|\alpha_{i+1} - \beta_{i+1}\|_2 \leq \epsilon \tag{12}$$

$$res_{dual} = \|\beta_{i+1} - \beta_i\|_2 \leq \delta \tag{13}$$

To solve the above constrained problem and get the iterates we start from the lagrangian formulation

$$\mathcal{L}_\rho(\alpha, \beta, \lambda) = H(\alpha) + \gamma\|\beta\|_1 + \frac{1}{2}(\lambda^*(\alpha - \beta) + (\alpha - \beta)^*\lambda + \rho\|\alpha - \beta\|_2^2)$$

Where $\lambda \in \mathbf{C}^M$ is the lagrange multiplier and ρ is a positive parameter that introduces a quadratic penalty on the deviation between α and β This gives the following iterates

$$\alpha_{i+1} = \arg\min_{\alpha} \mathcal{L}_\rho(\alpha, \beta_i, \lambda_i) \tag{14}$$

$$\beta_{i+1} = \arg\min_{\beta} \mathcal{L}_\rho(\alpha_{i+1}, \beta, \lambda_i) \tag{15}$$

$$\lambda_{i+1} = \lambda_i + \rho(\alpha_{i+1} - \beta_{i+1}) \tag{16}$$

In appendix A it is shown that solving these iterates results in

$$\alpha_{i+1} = (P + \frac{\rho}{2}I)^{-1}(q + \frac{\rho}{2}r_i) \tag{17}$$

$$\beta_{i+1}^j = S_\kappa(v_i^j) \tag{18}$$

with $r_i = \beta_i - \frac{\lambda_i}{\rho}$, $v_i = \alpha_{i+1} - \frac{\lambda_i}{\rho}$, $\kappa = \frac{\gamma}{\rho}$ and

$$S_\kappa(v_i^j) = \begin{cases} v_i^j - \kappa & \text{if } v_i^j > \kappa \\ 0 & \text{if } v_i^j \in [-\kappa, \kappa] \\ v_i^j + \kappa & \text{if } v_i^j < -\kappa \end{cases}$$

By the end of this, we get a vector β carrying the sparsity structure (number of non-zero amplitudes and the places or indices of these non-zero ones). Now, we move to the second step where this sparsity structure is fixed and the non-zero amplitudes are determined. This is equivalent to solving the following constrained convex optimization problem

$$\min_{\alpha} \quad H(\alpha) \text{ subject to } \quad E^T\alpha = 0 \tag{19}$$

where the matrix $E \in \mathbf{R}^{l \times s}$, s is the number of non-zero elements in beta, represents the sparsity structure. Each column of E is a unit vector with non-zero element corresponding to the non-zero amplitude of α.

The solution to 19, shown in appendix A, is

$$\alpha_{opt} = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} P & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} q \\ 0 \end{bmatrix} \tag{20}$$

Algorithm 1 Sparse DMD Algorithm

function SPARSEDMD(X, Y, γ, ρ, max_iter, ε_{abs}, ε_{rel}) ▷ Input parameters
  $[U, S, V] = \text{svd}(X)$ ▷ The SVD of matrix X.
  $U_r = U(:, 1:r), V_r = V(:, 1:r), S_r = S(:, 1:r)$ ▷ truncate to a rank r
  $F_{DMD} = U_r^{'} \times Y \times V_r \times s^{-1}$ ▷ Compute the DMD matrix
  $[W_r, D_r] = \text{eig}(F_{DMD})$ ▷ Compute the eigenvalues and eigenvectors of the DMD matrix
  $\phi = U_r \times W_r$ ▷ DMD modes
  Calculate P, q, s and V_{mode} ▷ Calculate auxiliary variables
  Initialize β and λ ▷ Initialize parameters
  for $i = 1$ to max_iter **do** ▷ Loop over maximum iterations
    Compute $\alpha_{i+1}, \beta_{i+1}, \lambda_{i+1}$ according to 17,18,16
    if $res_{prim} < \varepsilon_{prim}$ and $res_{dual} < \varepsilon_{dual}$ **then**
      break
    end if
  end for
  $ind\_zero = \text{find}(|\beta| < 1.e-12)$
  $E = I(:, ind\_zero)$ ▷ Compute E which resembles sparsity structure
  Compute α_{opt} according to 20
  Return a_opt, ϕ, λ
end function

5 Experimental Result

5.1 Synthetic Example

In this section, sparse DMD is applied on certain snapshots. We started testing it first on a synthetic example. A test matrix $A \in \mathbb{R}^{M \times N}$ with $M = 1000$ and $N = 400$ is generated with entries uniformly distributed in $[0, 1]$, and then $A = A/\|A\|_2$. Eigenvalues resulting from the standard DMD along with the subsets selected from the sparse DMD are shown in Figure 1. In Figure 2(a), sparsity structure (number of non-zero amplitudes) is determined as a function of the sparsity parameter γ. While the performance $Perf = 100 \times \frac{res}{\|X\|_F}$, where residual $res = \|X - ZDiag(\alpha_{opt})V_{monde}\|_F$, with respect to γ is represented in Figures 2(b),2(c). As expected, as the sparsity parameter γ increases, a sparser solution is obtained but this comes at the cost of compromising the quality of the approximation.

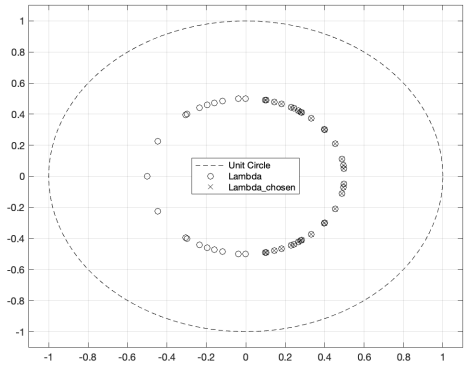
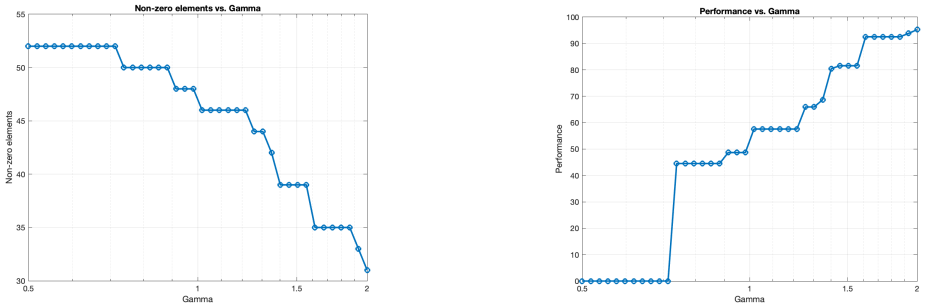


Figure 1: Eigenvalues resulting from the standard DMD algorithm represented by circles and subset of eigenvalues resulting from sparse DMD represented by crosses for the synthetic matrix A.



(a) (b)

Figure 2: Illustration of the decay of the number of nonzero amplitudes (resembling the sparsity structure) and the variation of the performance loss of the approximation with respect to the sparsity parameter γ for the synthetic example

5.2 Flow around a cylinder

In this example, the data snapshots taken from [4] are vorticity data of a flow around cylinder, discretized with dimension 89351. The simulation data with $\Delta t = 0.02$ are down-sampled, and the test case contains 151 snapshots. In Figure 3, the eigenvalues obtained from the DMD analysis are depicted, accompanied by the subset chosen through the sparsity-promoting DMD approach. In figure 4(a) and 4(b), consistent conclusions are drawn about the behavior of the non zero amplitudes and the performance with respect to sparsity parameter γ. In figure 5, the dynamic modes with nonzero amplitudes are depicted. By check-

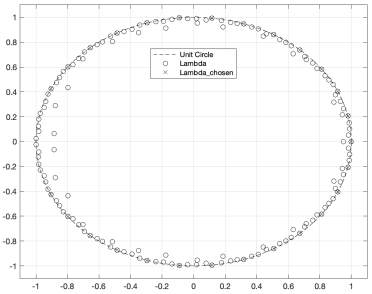
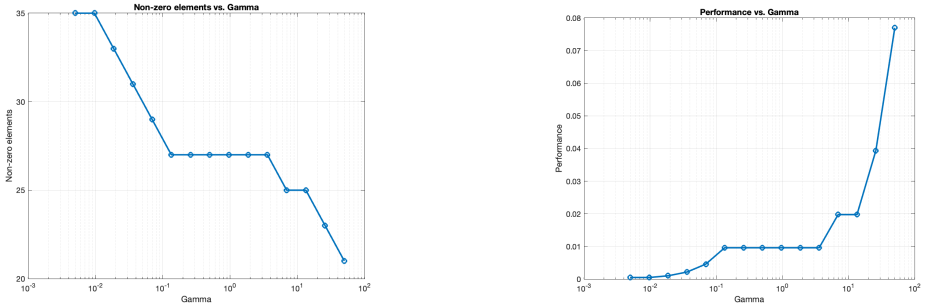


Figure 3: Eigenvalues resulting from the standard DMD algorithm represented by circles and the subset of eigenvalues resulting from sparse DMD represented by crosses for the cylinder flow.

ing the residuals of these modes, $residual(i) = \|YV_rS_R^{-1}W_r(:,i) - \lambda(i)\phi(:,i)\|_2$, it is noticed that only the modes with good residuals have nonzero amplitudes and are being selected.



(a) (b)

Figure 4: Illustration of the decay of the number of nonzero amplitudes (resembling the sparsity structure) and the variation of the performance loss of the approximation with respect to the sparsity parameter γ for the cylinder flow.

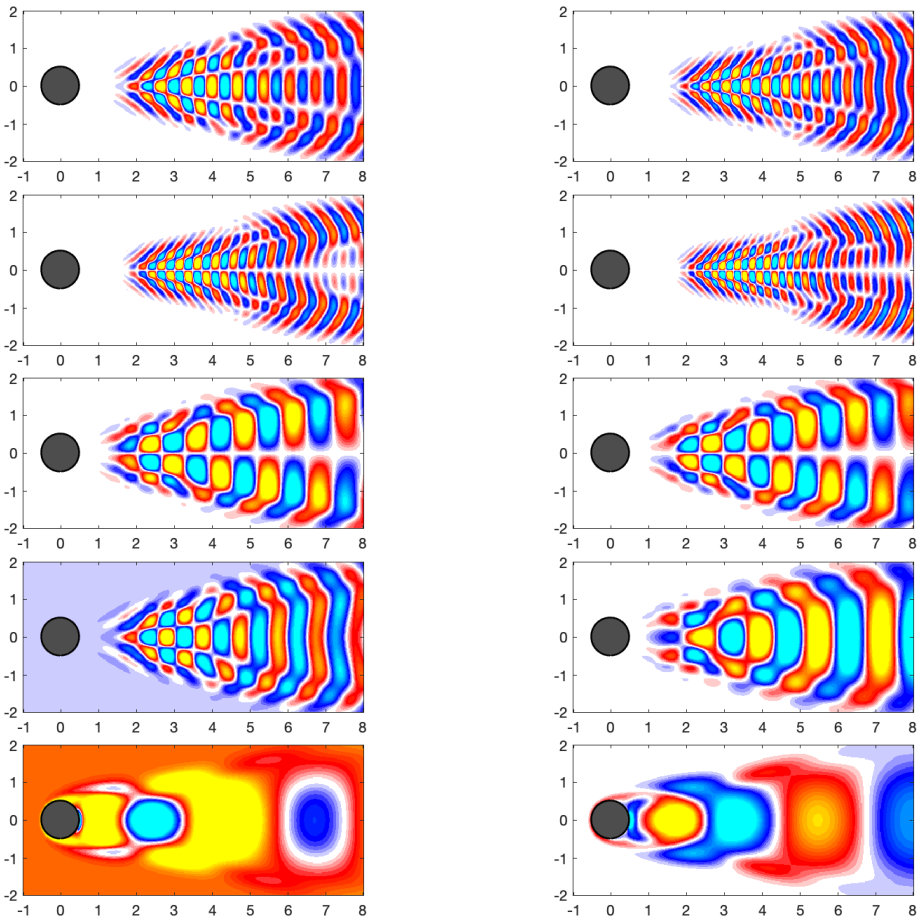


Figure 5: DMD modes with nonzero amplitudes resulting from sparsity promoting DMD for the cylinder flow case.

5.3 Heated Cylinder with Boussinesq Approximation

In this example, a simulation of a 2D flow is generated by a heated cylinder, and the fluid flow and heat transfer are analyzed using the Boussinesq approximation to model and solve the buoyancy problem. The data set is taken from [5][1] , where the simulation was done with Gerris flow solver and was resembled onto a regular grid. The data snapshots represents the v-velocity component of the flow, discretized with dimension 67500. The test case contains 151 snapshots. The set of lambda chosen through sparse DMD is shown in figure 6 and the performance loss and number of nonzero amplitudes with respect to gamma are illustrated in figure 7(a) and 7(b). In figure 8,9 and 10, It is shown from the calculated residuals that the modes chosen from the sparse DMD has the least residual compared to the ones from the traditional DMD. Which proves the

efficiency of sparse DMD in extracting the modes contributing the most.

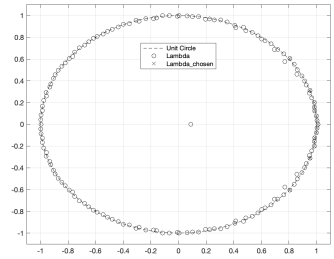


Figure 6: Eigenvalues resulting from the standard DMD algorithm represented by circles and the subset of eigenvalues resulting from sparse DMD represented by crosses for the heated cylinder flow.

(a) (b)

Figure 7: Illustration of the decay of the number of nonzero amplitudes (resembling the sparsity structure) and the variation of the performance loss of the approximation with respect to the sparsity parameter γ for the heated cylinder flow.

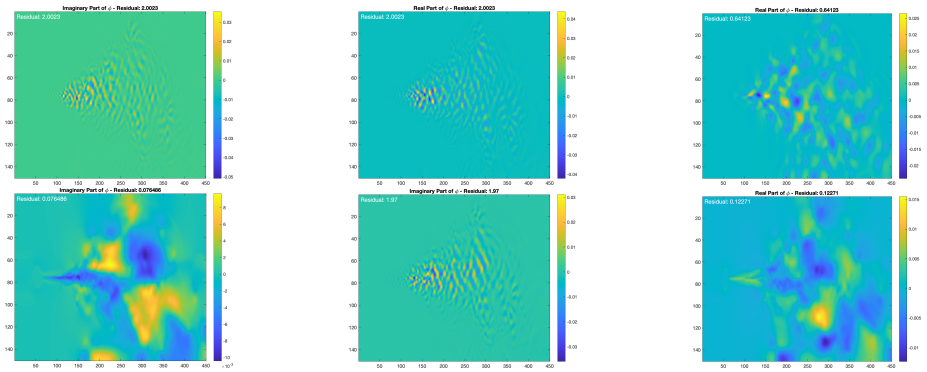


Figure 8: Some of the DMD modes with their residuals from the traditional DMD.

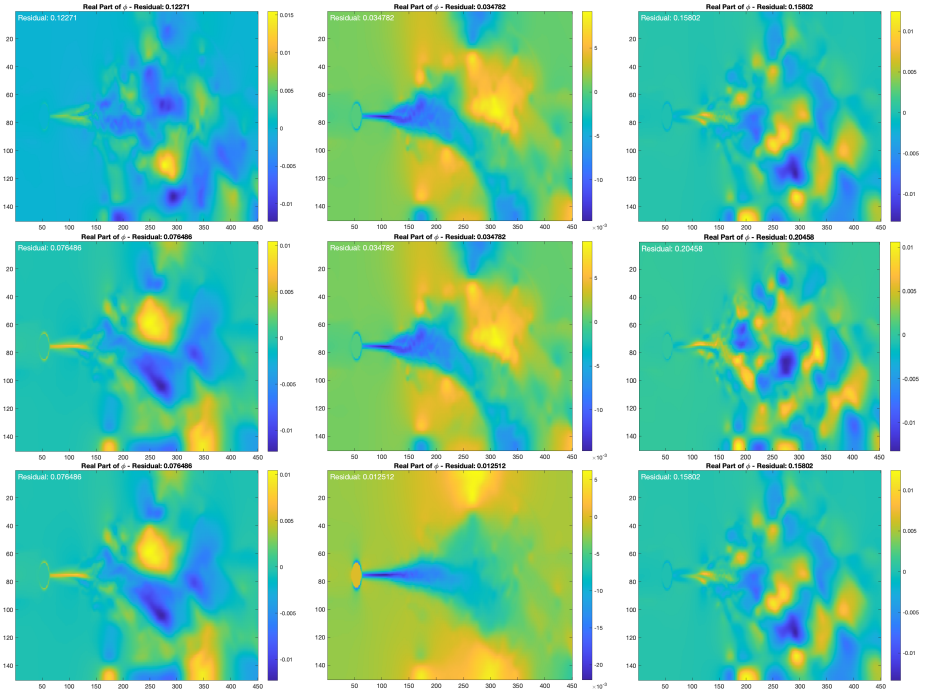


Figure 9: Real part of DMD modes with nonzero amplitudes resulting from sparsity promoting DMD for the cylinder flow case.

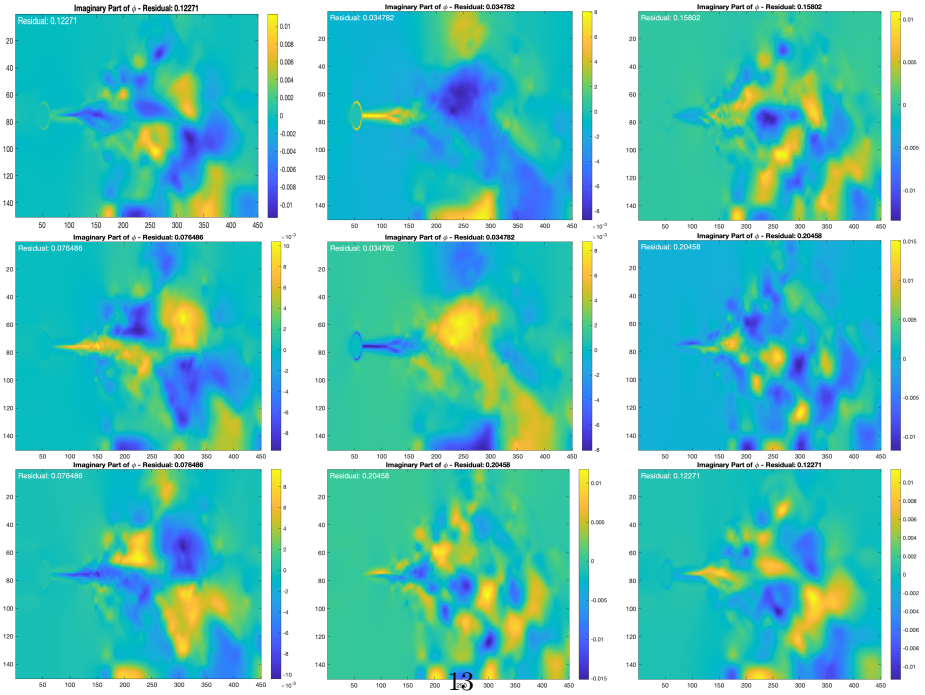


Figure 10: Imaginary part of DMD modes with nonzero amplitudes resulting from sparsity promoting DMD for the cylinder flow case.

6 Conclusion

In this paper, we addressed the challenge of selecting the modes that contribute the most in the fluid dynamics while discarding the ones that introduce noise to the data. This objective was accomplished by introducing the L_1 norm of the vector of amplitudes to the least square problem for determining the amplitudes of the modes. The L_1 norm promotes sparsity in the amplitude vector, transforming the optimization problem into a convex one, that could be efficiently solved by alternating between minimizing residual and promoting sparsity. Once the desired sparsity is achieved, it is fixed and the amplitudes are calculated accordingly. Through the implementation of the sparse DMD on the flow through cylinder and flow over a heated cylinder with Boussinesq approximation, we demonstrated its effectiveness in assigning nonzero amplitudes to the most influential modes while eliminating those that are less relevant. Furthermore, in the synthetic example, we illustrated that the degree of sparsity and the performance loss depends on the sparsity parameter chosen.

A Appendix A

Proof of proposition 3.1

$$
\begin{aligned}
&\|\Sigma V^* - W D_\alpha V_{mode}\|_F \\
&= trace((\Sigma V^* - W D_\alpha V_{mode})^*(\Sigma V^* - W D_\alpha V_{mode})) \\
&= trace((\Sigma V^*)^* - V_{mode}^* D_\alpha^* W^*)(\Sigma V^* - W D_\alpha V_{mode}) \\
&= trace((\Sigma V^*)^*(\Sigma V^*) - (\Sigma V^*)^* W D_\alpha V_{mode} - V_{mode}^* D_\alpha^* W^* \Sigma V^* + V_{mode}^* D_\alpha^* W^* W D_\alpha V_{mode}) \\
&= trace((\Sigma V^*)^*(\Sigma V^*) - \Sigma V^* W V_{mode} D_\alpha - V_{mode}^* W^* \Sigma V^* D_\alpha^* + D_\alpha^* W^* W D_\alpha V_{mode} V_{mode}^*) \\
&= trace((\Sigma V^*)^*(\Sigma V^*)) - trace(\Sigma V^* W V_{mode} D_\alpha) - trace(V_{mode}^* W^* \Sigma V^* D_\alpha^*) \\
&+ trace(D_\alpha^* W^* W D_\alpha V_{mode} V_{mode}^*)
\end{aligned}
$$

where the third equality is obtained by using $trace(AB) = trace(BA)$. The first term in the last equality is just s. For the second and third term we apply the fact that $trace(QD_\alpha) = \overline{diag(Q)}^* \alpha$ for any matrix Q, in particular for $Q = trace(V_{mode}^* W^* \Sigma V^)*$, which gives the term q. And for the last term, the equality $trace(D_\beta^* A D_\alpha \overline{B}^*) = \beta^*(A \circledast B)\alpha$ is utilized with $\beta = \alpha$, $A = W^*W$ and $B = \overline{V_{mode}^* V_{mode}}$ which gives $P = A \circledast B$.

proof of the lagrangian iterates

For 14, we have using completing the square that

$$
\begin{aligned}
\mathcal{L}_\rho &= H(\alpha) + \gamma\|\beta_i\|_1 + \frac{1}{2}(\lambda^*(\alpha - \beta_i) + (\alpha - \beta)^*\lambda_i + \rho\|\alpha - \beta_i\|_2^2 \\
&= H(\alpha) + \gamma\|\beta_i\|_1 + \|\frac{\lambda_i}{\rho}\|_2^2 + \frac{1}{2}(\lambda_i^*(\alpha - \beta_i) + (\alpha - \beta_i)^*\lambda_i + \rho\|\alpha - \beta_i\|_2^2 - \|\frac{\lambda_i}{\rho}\|_2^2) \\
&= H(\alpha) + \gamma\|\beta_i\|_1 + \|(\alpha - \beta_i) + \frac{\lambda_i}{\rho}\|_2^2 - \|\frac{\lambda_i}{\rho}\|_2^2
\end{aligned}
$$

as the terms $\gamma\|\beta_i\|_1$ and $\|\frac{\lambda_i}{\rho}\|_2^2$ are independent of α, then 14is equivalent to the following

$$\arg\min(H(\alpha) + \|(\alpha - \beta_i) + \frac{\lambda_i}{\rho}\|_2^2)$$

Substituting the expression of $H(\alpha)$ and denoting $r_i = \beta_i - \frac{\lambda_i}{\rho}$ gives the following

$$\begin{aligned}
&\arg\min(\alpha^* P\alpha - q^*\alpha - \alpha^* q + s + \frac{\rho}{2}\|\alpha - r_i\|_2^2) \\
&= \arg\min(\alpha^* P\alpha - q^*\alpha - \alpha^* q + s + \frac{\rho}{2}\alpha^*\alpha - \frac{\rho}{2}\alpha^* r_i - \frac{\rho}{2} r_i^*\alpha + \rho\|r_i\|_2^2) \\
&= \arg\min /(\alpha^*(P + \frac{\rho}{2}I)\alpha - (q^* + \frac{\rho}{2}r_i)\alpha - \alpha^*(q + \frac{\rho}{2}r_i) + s + \rho\|r_i\|_2^2)
\end{aligned}$$

By solving the above quadratic optimization problem, we get

$$\alpha_{i+1} = (P + \frac{\rho}{2}I)^{-1}(q + \frac{\rho}{2}r_i)$$

For 15, we have also using completing the square that $\mathcal{L}_\rho = H(\alpha_{i+1}) + \gamma\|\beta\|_1 + \|(\beta - \alpha_{i+1}) - \frac{\lambda_i}{\rho}\|_2^2 - \|\frac{\lambda_i}{\rho}\|_2^2$ and the terms $H(\alpha_{i+1})$ and $\|\frac{\lambda_i}{\rho}\|_2^2$ are indepenedent of β, then 15 is equivalent to the following

$$\arg\min(\gamma\|\beta\|_1 + \frac{\rho}{2}\|\beta - v_i\|_2^2)$$

The solution to the above optimization problem is

$$\beta_{i+1}^j = S_\kappa(v_i^j)$$

with $\kappa = \frac{\gamma}{\rho}$ and

$$S_\kappa(v_i^j) = \begin{cases} v_i^j - \kappa & \text{if } v_i^j > \kappa \\ 0 & \text{if } v_i^j \in [-\kappa, \kappa] \\ v_i^j + \kappa & \text{if } v_i^j < -\kappa \end{cases}$$

Proof of 19

The Lagrangian formulation of this constrained problem gives

$$\begin{aligned}
\mathcal{L}(\alpha, \lambda) &= H(\alpha)_\lambda^* E^T\alpha + (E^T\alpha)^*\lambda \\
&= \alpha^* P\alpha - q^*\alpha - \alpha^* q + s + \lambda^* E^T\alpha + (E^T\alpha)^*\lambda
\end{aligned}$$

which gives the rise of two equations. The first with respect to α

$$\begin{aligned}
&2P\alpha - 2q + E\lambda + E\lambda = 0 \\
&p\alpha + E\lambda = q
\end{aligned}$$

and the second with respect to λ

$$E^T\alpha = 0$$

Hence the solution for the above system is $\begin{bmatrix} P & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \lambda \end{bmatrix} = \begin{bmatrix} q \\ 0 \end{bmatrix}$

References

[1] Tobias Günther, Markus Gross, and Holger Theisel. Generic objective vortices for flow visualization. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 36(4):141:1–141:11, 2017.

[2] Mihailo R Jovanović, Peter J Schmid, and Joseph W Nichols. Sparsity-promoting dynamic mode decomposition. *Physics of Fluids*, 26(2), 2014.

[3] Mihailo R Jovanovic, Peter J Schmid, and JW Nichols. Low-rank and sparse dynamic mode decomposition. *Center for Turbulence Research Annual Research Briefs*, 2012:139–152, 2012.

[4] J Nathan Kutz, Steven L Brunton, Bingni W Brunton, and Joshua L Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

[5] S. Popinet. Free computational fluid dynamics. *ClusterWorld*, 2(6), 2004.

[6] Peter J Schmid. Dynamic mode decomposition and its variants. *Annual Review of Fluid Mechanics*, 54:225–254, 2022.