The 8th European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS Congress 2022
5-–9 June 2022, Oslo, Norway

FEATURE DETECTION ALGORITHMS AND MODAL DECOMPOSITION METHODS

B. BEGIASHVILI[1], J. GARICANO-MENA[1,2], S. LE CLAINCHE[1] AND E. VALERO[1,2]

[1] ETSI Aeronáutica y del Espacio – Universidad Politécnica de Madrid.
Plaza Cardenal Cisneros 3, Madrid E-28040, Spain

[2] Center for Computational Simulation (CCS)
Boadilla del Monte, E-28660, Spain

Key words: modal decompositions; machine learning; feature detection; low order algorithms; dynamic mode decomposition; proper orthogonal decomposition; matrix factorization.

Abstract. Various modal decomposition techniques have been developed in the last decade [1–11]. We focus on data-driven approches, and since data flow volume is increasing day by day, it is important to study the performance of order reduction and feature detection algorithms.

In this work we compare the performance and feature detection behaviour of energy and frequency based algorithms (Proper Orthogonal Decomposition [1–3] and Dynamic Mode Decomposition [4–6, 8–11]) on two data set testcases taken from fluid dynamics. The datasets considered (the velocity field of laminar wake around the mid-section of a very long cylinder at $Re_D = 100$ and the pressure field of turbulent jet (axisymetric) at $Re_D = 10^6$) represent different flow regimes.

The performance of these algorithms is thoroughly assessed concerning both the accuracy of the results retrieved and the computational performance.

From this assessment, those techniques that are potentially better suited for the applications are identified and after the possibility of parallelizing the algorithms will be studied with a final objective: to enable data-driven analysis of industrially relevant fluid mechanical problems.

1 INTRODUCTION

As the amount of available data increases, it is necessary to develop tools capable of extracting useful information from them and, either from numerical simulation or experiments of fluid flows, eventually reconstruct the system they represent with the least possible error. To achieve this need, data-based fluid flow decomposition algorithms have emerged in the last decade with the purpose of detecting and extracting coherent structures from high-dimensional datasets.
Data-driven algorithms can be classified in two classes, energy-based and frequency-based. Energy-based algorithms can optimally identify the most energetic contributions within a few modes and they can be designed to prioritize different physical phenomena by the appropriate use of mass matrix (kinetic energy, entropy, vorticity, etc). On the contrary, frequency-based methods infer the frequency of each mode from the dataset instead of defining it *a priori* from the time discretization (like in discrete Fourier transform). Also, frequency-based methods give

information about modes growth rate, making them suitable for data-driven stability analysis and feedback control for linear systems. However, frequency-based algorithms can be highly challenged for non-linear systems and temporally localized events.

This work is structured as follows. In Sect. 2 we will review modal decomposition techniques, from algorithmic point of view, used for dynamical systems order reduction, specifically Proper Orthogonal Decomposition (**POD**) and Dynamic Mode Decomposition (**DMD**). In Sect. 3 we will describe the datasets used in this work. In Sect. 4 we will apply these techniques to the datasets in order to evaluate the capacity of each of them to detect the dominant structures and frequency (or Strouhal number) of the system. In Sect. 5 we will analyze the results in order to draw conclusions about the studied algorithms, the relationship between them, and propose next steps.

2 Methodology

In order to standardize the data input for each algorithm, we are going to put the data in a snapshots matrix $\mathbf{X} \in \mathbb{R}^{m \times n_t}$, where $m \equiv n_u \times n_x \times n_y \times n_z$. Here, n_u is the number of analysed fluid flow variables, and n_x, n_y, n_z spatial dimensions. n_t is the number of snapshots and they will be sampled with time interval Δt.

2.1 Proper Orthogonal Decomposition

POD originates from the turbulence field. Lumley introduced it to the fluid dynamics community in 1967 [12] as an attempt to decompose a turbulent fluid motion into a set of deterministic functions that capture a portion of the total fluctuating kinetic energy in the flow. In other words, **POD** helps us to find coherent structures in a flow that can describe different fluid flow phenomena. **POD** is related to Singular Value Decomposition (**SVD**) of the snapshots matrix **X**:

$$\mathbf{X} \underset{SVD}{\equiv} \mathbf{L}\,\boldsymbol{\Sigma}\,\mathbf{R}', \tag{1}$$

where $(\cdot)'$ denotes transpose of a matrix. $\mathbf{L} = [\mathbf{l}_1, \mathbf{l}_2, ..., \mathbf{l}_{n_t}] \in \mathbb{R}^{m \times n_t}$ are left singular vectors, or **POD** spatial modes/structures of the flow. $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_{n_t}] \in \mathbb{R}^{n_t \times r}$ are right singular vectors, or **POD** temporal modes and gives information about how corresponding spatial modes evolve in time. $\boldsymbol{\Sigma}$ is a diagonal matrix with $[\sigma_1, \sigma_2, ..., \sigma_{n_t}] \in \mathbb{R}^{n_t \times n_t}$ non-zero entries and are **POD** singular values, it is organized hierarchically and gives information about the importance of the corresponding modes. Computationally, **SVD** is implemented into many scientific packages, such as `MATLAB` [13] or `NumPy` [14].

One way to reduce the dimensionality of the original matrix **X** is using a subset of **POD** basis. We define ϵ as the residual of the energy after choosing the size of the subset of **POD**, r, as it follows:

$$\epsilon = \frac{\sum_{i=r}^{n_t} \sigma_i^2}{\sum_{i=1}^{n_t} \sigma_i^2}. \tag{2}$$

This will yield the r-rank optimal representation of the dataset. On the one hand, we can choose the size of subset, r, or we can choose ϵ_k and calculate r so that $\epsilon_k \leq \epsilon$. Then $\mathbf{L}$, $\mathbf{\Sigma}$ and $\mathbf{R}$ are truncated: $\mathbf{L}_r \in \mathbb{R}^{m\times r}$, $\mathbf{\Sigma}_r \in \mathbb{R}^{r\times r}$ and $\mathbf{R}_r \in \mathbb{R}^{r\times n_t}$.

In order to extract the dominant Strouhal number of the flow, Discrete Fourier Transform (**DFT**) will be applied on $\mathbf{R}$ matrix columns, this is, right singular vectors.

2.2 Dynamic Mode Decomposition

In the original definition of **DMD** [5], input data was provided as a single time series. However a generalization of **DMD**, that produces the same results and requires a dataset of shanpshot pairs, was introduced in [6]. **DMD** assumes a linear relationship between consecutive snapshots:

$$\mathbf{x}_{t+1} = \mathcal{A}\,\mathbf{x}_t. \tag{3}$$

DMD requires a dataset of snapshot pairs, or a pair of datasets as follows: $\mathbf{X}_1^{n_t-1} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{n_t-1}]$ and $\mathbf{X}_2^{n_t} = [\mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_{n_t}]$. Next, using truncated **SVD**, reduced $\mathcal{A}$ is computed:

$$\bar{\mathcal{A}} \equiv \mathbf{L}_r'\,\mathcal{A}\,\mathbf{L}_r = \mathbf{L}_r'\,\mathbf{X}_2^{n_t}\,\mathbf{R}_r\,\mathbf{\Sigma}_r^{-1}, \tag{4}$$

where $\bar{\mathcal{A}}$ is a projection of $\mathcal{A}$ into **POD** modes.

Then right eigendecomposition is calculated:

$$\bar{\mathcal{A}}\,\mathbf{\Psi} = \mathbf{\Psi}\,\mathbf{\Lambda}_\mu, \tag{5}$$

where the columns of $\mathbf{\Psi}$ are projected **DMD** modes with the corresponding **DMD** eigenvalues μ_i (the diagonal entries of $\mathbf{\Lambda}_\mu$). Temporal growth rates can be calculated ($\alpha_i = \Re(\log \mu_i/\Delta t)$) as well as angular pulsation ($\omega_i = \Im(\log \mu_i/\Delta t)$). Exact **DMD** modes will be recovered as $\mathbf{\Phi} = \mathbf{X}_2^{n_t}\,\mathbf{R}_r\,\mathbf{\Sigma}^{-1}\,\mathbf{\Psi}\,\mathbf{\Lambda}_\mu^{-1}$. We calculate amplitude a_i corresponding to each **DMD** mode as in [15]:

$$\min_a \|\mathbf{X}_1^{n_t-1} - \mathbf{\Phi}\,\mathbf{D}_a\,\mathbf{V}_{and}(\mu)\|, \tag{6}$$

where $\mathbf{D}_a$ is diagonal matrix with amplitudes as entries, and $\mathbf{V}_{and}(\mu)$ Vandermonde matrix formed with **DMD** singular values.

3 Testcase description

We have considered two publicly available testcases, encompassing laminar and turbulent flows.

The first testcase considered is the flow field around the mid-section of a very long cylinder. The Reynolds number is $Re_D = 100$, defined as $Re_D = UD/\nu$, where U is the free stream velocity, ν is the kinematic viscosity and $D = 1$ is the diameter of the cylinder. At these conditions, the flow is laminar and two-dimensional. The flow past a circular cylinder is a benchmark problem in fluid dynamics, generally used in a large number of applications, i.e., to validate novel methodologies, algorithms, etc. This makes it a suitable database to test the performance of all the methods analysed in this article. The flow dynamics in the two-dimensional flow past a circular cylinder is already known, studied since the past and presented in detail by Barkley & Henderson [16] (among others). More specifically, at $Re_D = 100$ this two-dimensional flow field is periodic with leading non-dimensional frequency Strouhal number $St = 0.16$, where $St = fD/U$, being f the frequency in Hertz.

The database analysed is available from Ref. [17]. This database consists of 151 snapshots, equiseparated in time with $\Delta t = 0.2$, representing the saturated flow regime of the streamwise and normal velocity components. The spatial domain is limited to the area surrounding the cylinder wake $x \in [-1, 8]$ and $y \in [-2, 2]$ for the streamwise and normal spatial components, where the center of the cylinder is located at the point $(x, y) = (0, 0)$. The spatial domain is formed by 450×200 grid points, also equidistant in space. Figure 1 shows two representative snapshots of the stramwise velocity component in the database analysed. The figure also shows the temporal evolution of the streamwise and normal velocities in a representative point of the computational domain, where it is possible see the periodic character of the studied solution.

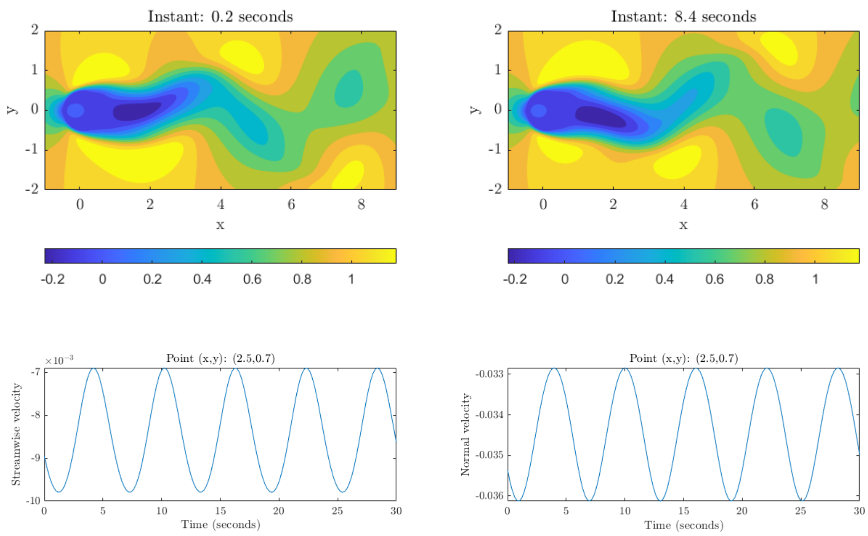

Figure 1: Flow past a circular cylinder at $Re_D = 100$. Top: two representative snapshots of the streamwise velocity field. Bottom: time history of the database analysed represented by the streamwise (left) and normal (right) velocity components. Point extracted at $(x, y) = (2.5, 0.7)$.

The second testcase, taken from Ref. [18], considers the pressure field in a turbulent jet at $Re_D = 10^6$ (based on the nozzle diameter D) and Mach number $M = 0.4$, defined as $M = U/c$, with U the free stream velocity and c the sound speed, taken in an azimuthal plane. This testcase is a good example of a DNS generated turbulent flow, consisting of multiple interacting spatio-temporal scales and suitable to show the performance of the methods in complex flows. Although it is notorious that, in contrast to wall bounded turbulent flows [11], the frequency spectrum in jet flows presents selected high-amplitude frequencies, which are in charge of driving the flow dynamics. These high-amplitude frequencies are in many cases connected with flow instabilities occurring in the shear layer of the jet, which is in good agreement with the own nature of this type of flows.

The database studied is conformed by 5000 snapshots representing the pressure field in a solution statistically converged. Each snapshot is formed by 39×175 grid points equidistant in space, and they are organized equiseparated in time with $\Delta t = 0.2$. The flow is axi-symmetric. In the case studied, the jet nozzle is located at $(x, r) = (0, 1)$ in the spatial domain defined in $x \in [0, 20]$ and $r \in [0, 4]$, being r the radial component, non-dimensionalized with the jet diameter $D = 1$. At the present flow conditions, the flow is characterized by a high-amplitude frequency driving the flow dynamics: $St \simeq 0.6$, based on the jet nozzle diameter (see more details about this flow characterization in Ref. [18]). The high complexity and the turbulent character of the database analysed is presented in Fig. 2, which shows two representative snapshots of the flow field and the evolution of the pressure field in two representative points of the spatial domain.

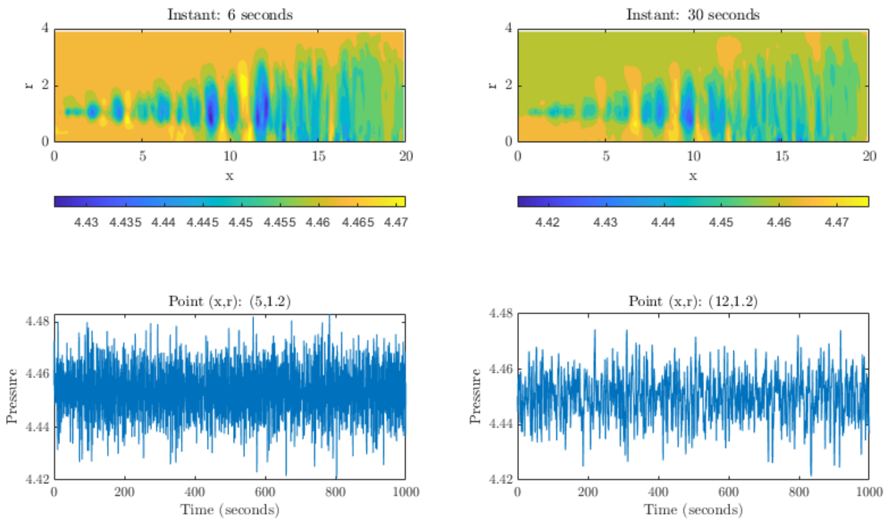

Figure 2: Turbulent jet flow at $Re_D = 10^6$ and $M = 0.4$. Top: two representative snapshots of the pressure field. Bottom: time history of the database analysed represented by the pressure field extracted at $(x, r) = (5, 1.2)$ (left) and $(x, y) = (12, 1.2)$ (right).

All the computations carried out have been performed on a computer equipped with an 4-core Intel(R) Core(TM) `i7-10750H` CPU at 2.6 GHz, a cache memory of 6144 kB and 16.0 GB of `RAM`.

4 Results

This section introduces the results of the two testcases for well known **POD** and **DMD** algorithms. Fig. 3 shows **POD** singular values (left) and frequency content for the first temporal mode (right) for the first testcase, laminar wake around cylinder. Singular values are found in pairs as expected from a periodic flow. As expected, there is only one dominant frequency since it is a saturated regime and with no noise. Fig. 4 shows **DMD** spectrum (left) and **DMD** growth rates (right) of corresponding modes. The **DMD** spectrum shows a clear periodic solution with leading $St = 0.16$ and its harmonics. The growth rates are nearly zero, as expected in a saturated regime. In noisy data or numerical simulations still not in saturated regime, **DMD** growth rates will never be zero.

Fig. 5 shows that **POD** and **DMD** modes, or spatial structures, are very similar. **POD**

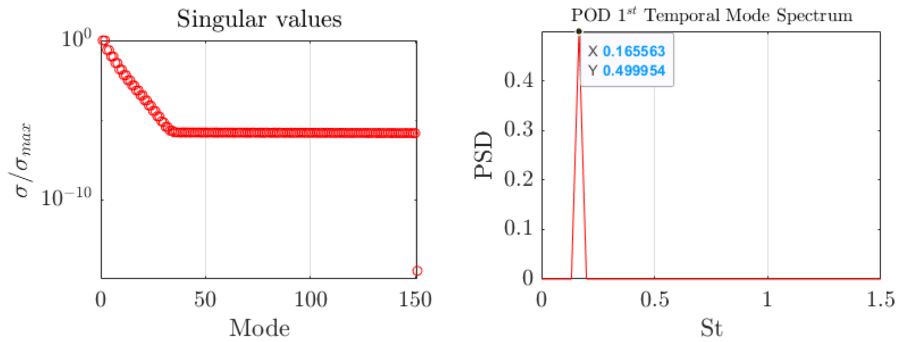

Figure 3: **POD** in the two-dimensional cylinder wake. Left: singular values. Right: first temporal mode's frequency content.

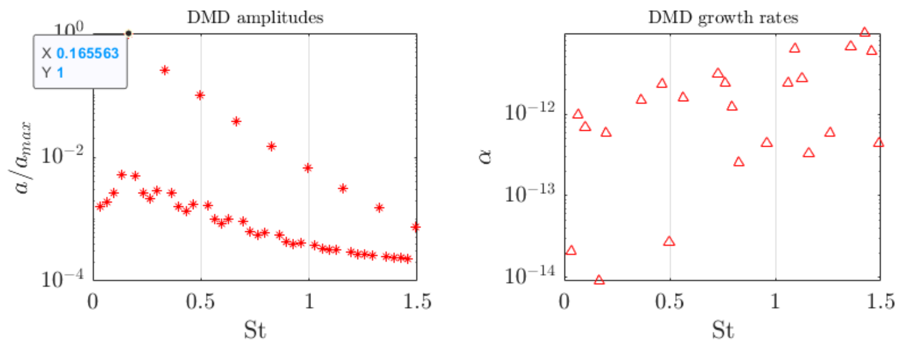

Figure 4: **DMD** in the two-dimensional cylinder wake. Left: amplitude as a function of Strouhal number for each **DMD** mode. Right: growth rate as a function of Strouhal number for each **DMD** mode.

modes are real, while **DMD** modes are complex, revealing travelling character of the modes by phase shifting. In this case, since the dataset comes from a saturated flow regime and with no noise, **POD** modes are also travelling with a clear dominant frequency.

Fig. 6 shows singular values (left) and first **POD** temporal mode's frequency content (right) for turbulent jet. Unlike the previous case, the singular values do not reach a certain number of modes and stop decreasing, but decrease continuously. This is due to the fact that it is a turbulent flow and multi-scale phenomena take place. For the same reason, the **POD** mode contains several frequencies.

Fig. 7 shows **DMD** amplitudes (left) and growth rates (right) in the turbulent jet. Unlike the laminar wake around cylinder, growth rates are not zero. They are negative, this means that the corresponding modes, representing different phenomena, are decreasing. As seen, **DMD** is sensitive to the tolerance choice for **SVD** part of the algorithm.

The **POD** and **DMD** modes are presented in Fig. 8. The results show that both algorithms, **POD** and **DMD** are able to detect the dominant spatial sturcture. All the modes take a form of wavepacket structures. These wavepackets have been linked to the Kelvin-Helmoltz instability of the annular jet shear layer for $St \geq 0.3$ [19]. The fact that there is a link between a real turbulent jet and these simple linear stability concepts provides the possibility for non-empirical and reduced-order models.

Fig. 9 shows the capacity of detecting dominant Strouhal number, presenting Strouhal number as a function of number of snapshots considered by the algorithm. As seen in Fig. 9, in the laminar wake around cylinder, **DMD** converges before **POD**. This is due to the **DFT** that is calculating on the temporal **POD** modes. When calculating **DFT**, frequency resolution is determined by the time domain considered. In the turbulent jet **POD** and **DMD** present very similar behaviour.

5 Conclusions and future steps

In this contribution we have illustrated how these modes can be leveraged to identify flow patterns. These algorithms can also be used to derive lower-order models of the dynamics investigated [20,21]. Parallelized versions of these algorithms allow to tackle large-scale industrial problems. Both avenues are currently under investigation.

REFERENCES

[1] S. Volkwein, Proper orthogonal decomposition: Theory and reduced-order modelling, Lecture Notes, University of Konstanz (01 2012).

[2] G. Berkooz, P. Holmes, J. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, Annual Review of Fluid Mechanics 25 (2003) 539–575. doi:10.1146/annurev.fl.25.010193.002543.

[3] L. Grinberg, A. Yakhot, G. Karniadakis, Analyzing transient turbulence in a stenosed carotid artery by proper orthogonal decomposition, Annals of biomedical engineering 37 (2009) 2200–17. doi:10.1007/s10439-009-9769-z.

[4] C. Rowley, I. Mezic, S. Bagheri, P. Schlatter, D. Henningson, Spectral analysis of nonlinear flows, Journal of Fluid Mechanics 641 (2009) 115 – 127. doi:10.1017/S0022112009992059.

[5] P. Schmid, Dynamic mode decomposition of numerical and experimental data, Journal of Fluid Mechanics 656 (11 2008). doi:10.1017/S0022112010001217.

[6] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, J. N. Kutz, On dynamic mode decomposition: Theory and applications, Journal of Computational Dynamics 1 (2) (2014) 391–421.

[7] J. Garicano-Mena, B. Li, E. Ferrer, E. Valero, A composite dynamic mode decomposition analysis of turbulent channel flows, Physics of Fluids 31 (2019) 115102. doi:10.1063/1.5119342.

[8] M. Jovanovic, P. Schmid, J. Nichols, Sparsity-promoting dynamic mode decomposition, Physics of Fluids 26 (09 2013). doi:10.1063/1.4863670.

[9] M. R. Jovanović, From bypass transition to flow control and data-driven turbulence modeling: An input–output viewpoint, Annual Review of Fluid Mechanics 53 (1) (2021) 311–345. arXiv:https://doi.org/10.1146/annurev-fluid-010719-060244, doi:10.1146/annurev-

fluid-010719-060244.
URL `https://doi.org/10.1146/annurev-fluid-010719-060244`

[10] S. Le Clainche, J. Vega, Higher order dynamic mode decomposition to identify and extrapolate flow patterns, Physics of Fluids 29 (08 2017). doi:10.1063/1.4997206.

[11] S. Le Clainche, D. Izbassarov, M. Rosti, L. Brandt, O. Tammisola, Coherent structures in the turbulent channel flow of an elastoviscoplastic fluid, Journal of Fluid Mechanics 888 (2020) A5. doi:10.1017/jfm.2020.31.

[12] J. L. Lumley, The structure of inhomogeneous turbulent flows, Atmospheric Turbulence and Radio Wave Propagation (1967).
URL `https://cir.nii.ac.jp/crid/1571980075051475712`

[13] MATLAB, https://es.mathworks.com/.

[14] NumPy, https://numpy.org/.

[15] M. R. Jovanović, P. J. Schmid, J. W. Nichols, Sparsity-promoting Dynamic Mode Decomposition, Phys. Fluids 26 (2) (2014). doi:10.1063/1.4863670.

[16] D. Barkley, R. Henderson, Three-dimensional floquet stability analysis of the wake of a circular cylinder, J. Fluid Mech. 322 (1996) 215–241.

[17] J. N. Kutz, S. L. Brunton, B. W. Brunton, J. L. Proctor, Dynamic mode decomposition: data-driven modeling of complex systems, SIAM–Society for Industrial and Applied Mathematics, 2016.

[18] A. Towne, O. T. Schmidt, T. Colonius, Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis, Journal of Fluid Mechanics 847 (2018) 821–867. doi:10.1017/jfm.2018.283.

[19] T. Suzuki, T. Colonius, Instability waves in a subsonic round jet detected using a near-field phased microphone array, Journal of Fluid Mechanics 565 (2006) 197–226. doi:10.1017/S0022112006001613.

[20] S. Le Clainche, J. Vega, A Review on Reduced Order Modeling using DMD-Based Methods, 2020, pp. 55–66. doi:10.1007/978-3-030-21013-7_4.

[21] J. Du, F. Fang, C. Pain, I. Navon, J. Zhu, D. Ham, Pod reduced-order unstructured mesh modeling applied to 2d and 3d fluid flow, Computers Mathematics with Applications 65 (3) (2013). doi:doi.org/10.1016/j.camwa.2012.06.009.

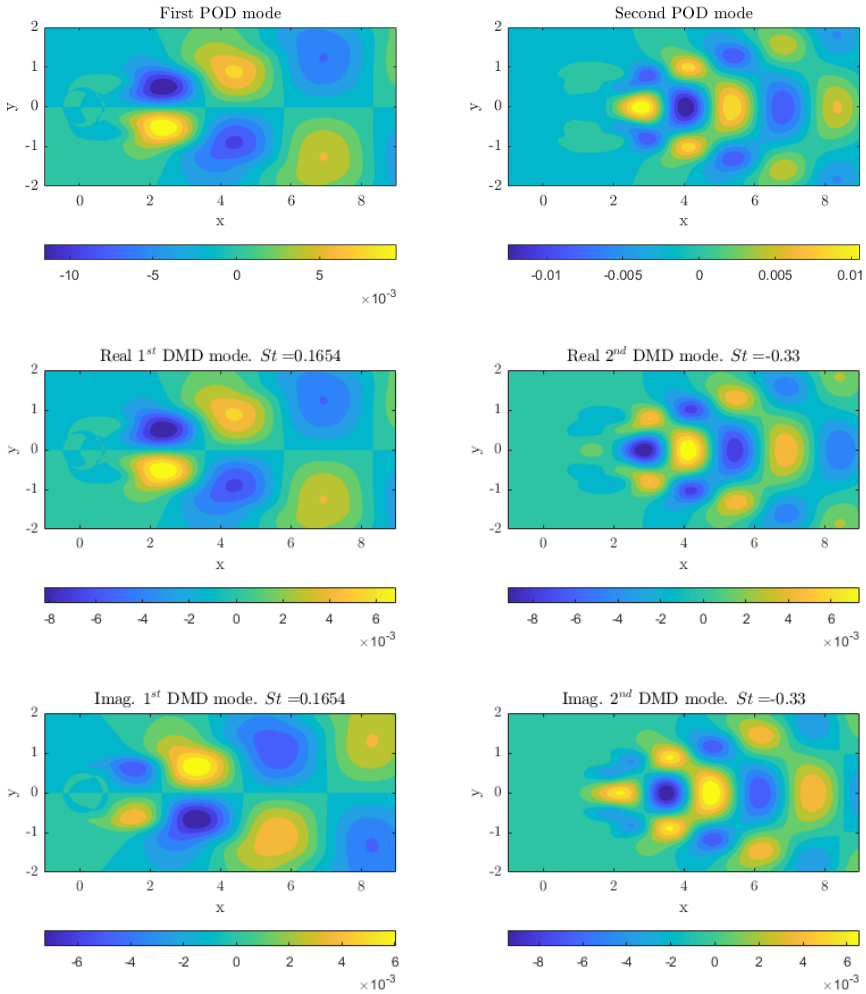

Figure 5: The two-dimensional cylinder wake. **POD** (top), **DMD** real (mid) and **DMD** imaginary (bottom) dominant modes. From left to right: first and second highest energy/amplitude modes.

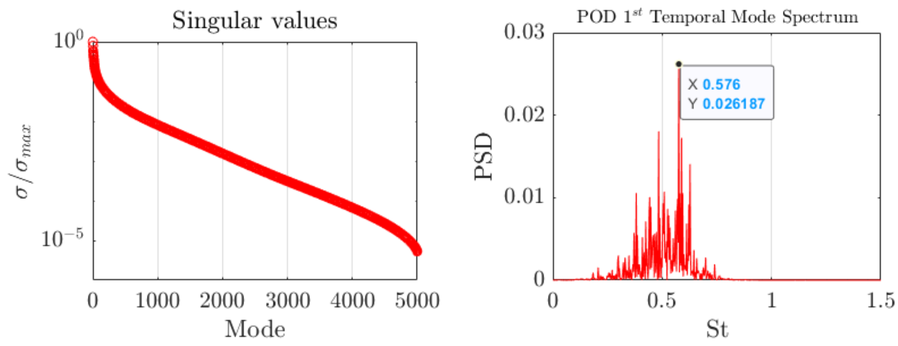

Figure 6: POD in the turbulent jet. Left: singular values. Right: first temporal mode's frequency content.

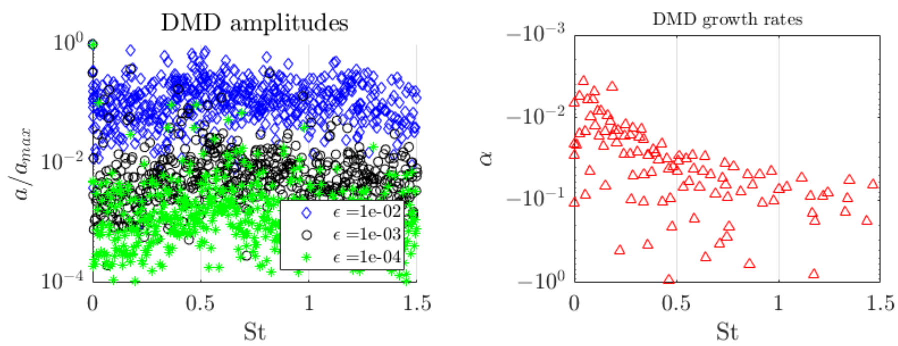

Figure 7: **DMD** in the turbulent jet. Left: amplitude as a function of ϵ and Strouhal number for each **DMD** mode. Right: growth rate as a function of Strouhal number for each **DMD** mode.

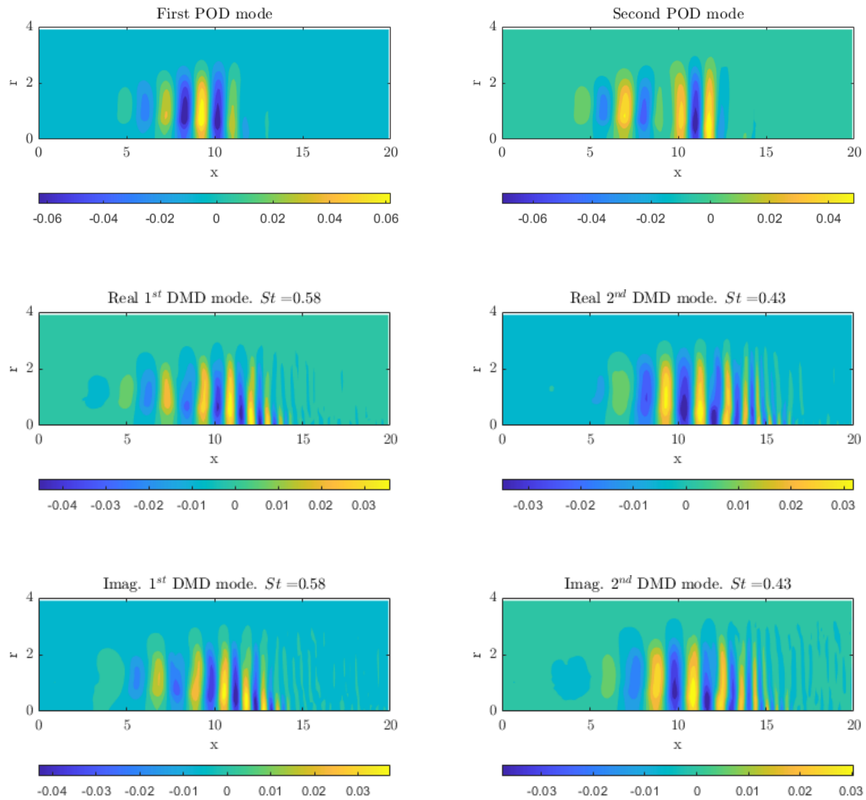

Figure 8: The turbulent jet. **POD** (top), **DMD** real (mid) and **DMD** imaginary (bottom) dominant modes. From left to right: first and second highest energy/amplitude modes.

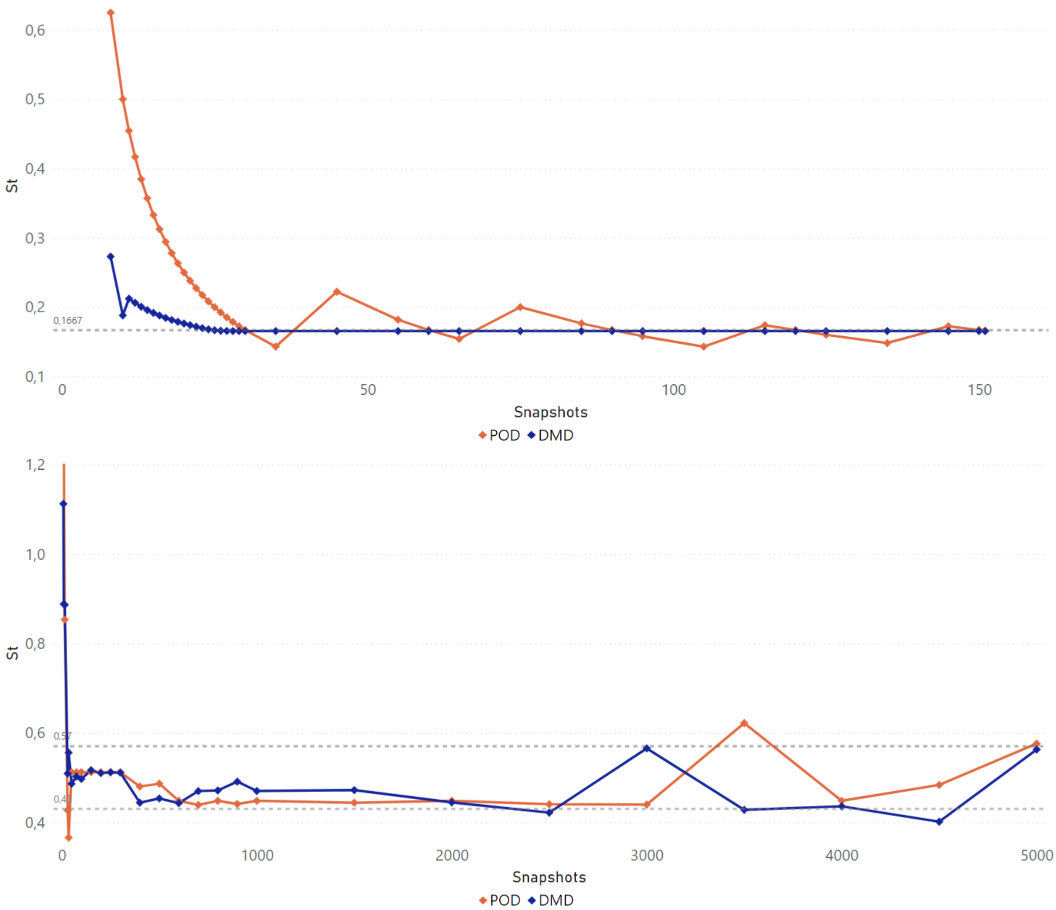

Figure 9: Dominant St detected vs the number of snapshots considered for **POD** and **DMD** in the laminar wake around cylinder (top); and turbulent jet (bottom).