Feature Extraction from Turbulent Channel Flow Databases via Composite DMD Analysis

B. Li\textsuperscript{1,2}, J. Garicano-Mena\textsuperscript{1}, E. Valero\textsuperscript{1}

\textsuperscript{1}School of Aeronautics, Universidad Politécnica de Madrid, 28040 Madrid, Spain
\textsuperscript{2}Center for Engineering and Scientific Computation, and School of Aeronautics and Astronautics, Zhejiang University, 310027 Zhejiang, PR China

E-mail: binghua.li@alumnos.upm.es, jesus.garicano.mena@upm.es

Abstract. In this contribution we consider the Dynamic Mode Decomposition (DMD) framework as a purely data-driven tool to investigate a \( Re \approx 950 \) turbulent channel database. Specifically, composite-based DMD analyses are conducted, with hybrid snapshots composed by skin friction and Reynolds stress. A small number of dynamic modes (less than 1\% of the number of snapshots) is found to be able to recover accurately the DNS Reynolds stresses near the wall, with a weighted factor as an indicator for the modes selections. As a possibility of analysis large turbulent database, we conclude that composite DMD is an attractive, purely data-driven, feature extraction tool to study turbulent flows.

1. Introduction

Turbulent flows involve a wide ranges of spatial and temporal scales, which impose stringent constrains on the experimental tools and/or numerical techniques employed to describe such flows. Dense grids of computational/data-acquisition points and high time resolution are required to obtain a well-resolved, statistically independent description of a turbulent flow. Modern data-acquisition systems (\textit{e.g.} particle image velocimetry, PIV [1]) fulfill the requirements to gain an accurate description. Advances in high-performance computing systems and development of stable and accurate numerical schemes, designed to exploiting parallel architectures, make possible to compute and analyze the complex real world processes with higher resolution and wider flow conditions (\textit{e.g.} via direct numerical simulation, DNS [2]). However, the application of such techniques to the turbulent flow lead to an enormous size of data, and the computational grids even scales roughly as \( Re^{9/4} \), where \( Re \) is the Reynolds number of the flow. Then the analyses of those data become a challenging task.

Feature extraction algorithms assist on the classification of the information buried in the data from experiments or DNS computations. In recent years, Proper Orthogonal Decomposition (POD, [3, 4]) and Dynamic Mode Decomposition (DMD, [5, 6]) have arisen as probably the most wide spread techniques.
Proper Orthogonal Decomposition techniques, which work on the sequences of snapshots acquired at successive time interval either from experimental measurements or numerical solutions, provide an optimal representation of these sequences. Features identified by POD are orthogonal to each other [4], and are classified by a decreasing energy content. Applications of POD techniques to the analysis of turbulent flow are reported in [7, 8].

POD algorithms can also be leveraged to build Reduced Order Models (ROMs) for the turbulent flows. Recently, [9] introduced a low-order model through a two-step Galerkin projection of the Navier-Stokes equations, which simulated regeneration of the streaks by the lift-up effect and reproduced the periodic spatial-temporal flow patterns. The link to the near-wall dynamics of turbulent boundary layers also attracted our attention on the features in the inner region of the turbulent boundary layers.

Dynamic mode decomposition (DMD) techniques, which emerged about ten years ago with the contributions of Schmid [6] and Rowley and collaborators [5], have arisen as prominent feature identification methods in the field of fluid dynamics. DMD also operates on a sequence of snapshots, but structures retrieved oscillate harmonically at specific frequencies. Many variants of the DMD method exist, allowing to retrieve meaningful flow structures from either experimental or numerical flow data in a purely data-driven (i.e., equation free) manner. Examples of successful application of variants of the method to a great variety of flow data can be found in [10, 11, 12, 13, 14].

Recently, Le Clainche and her collaborators [15, 16] introduced a High Order version of the DMD algorithm. Reference [17] describes an Optimized variant of DMD which is less sensitive to noise in the input data. Contribution [18] describes a Sparsity-Promoting variant of DMD, which assists in the selection of a subset of dynamic modes retained. The Non-Uniform DMD strategy of [19], standing on compressed-sensing principles, is capable of handling snapshots sampled at varying temporal separation. It also offers the opportunity to reduce the size of the snapshots considered by having recourse to K-means algorithms. By combining both effects, an effective reduction of the computational effort needed to perform the DMD is achieved. Alternatively, [20] offers a distributed-memory implementation of DMD, based upon a parallelized QR decomposition. This same algorithm was later on applied on composite snapshots -i.e., snapshots formed by considering two or more different magnitudes (velocity components, skin friction, $\lambda_2$ invariant, ...) - to compute laminar-to-turbulent transition in [13].

However, no matter which specific implementation of POD/DMD is considered, handling large turbulent databases is invariably problematic. Indeed, since a wide range of temporal and spatial scales interacting in complex fashion is involved, it is never trivial to select a few of them which can reproduce the flow behavior accurately enough. However, the attempt of reproducing a specific feature or functional (e.g., drag) rather than the complete system behavior may ease the work at hand, as done e.g. in aerodynamic design [21, 22].

In this contribution we investigate the performance of composite DMD techniques when applied to a turbulent channel flow database at a moderate friction Reynolds
number \( Re_\tau \approx 950 \). In particular, we are interested in the flow physics in the near-wall region. Considering the work about the near-wall dynamics in the boundary layers [9, 10, 23], we will focus on the region \( y^+ \in [0, 50] \).

In a recent work [24], we conducted a composite DMD analysis of a turbulent channel flow database obtained at the relatively low \( Re_\tau \approx 200 \). By considering a temporal sequence of composite snapshots assembled by concatenating Reynolds stress field \( u'v'(\vec{x}, t) \) and wall skin friction \( C_f(t) \) we were capable of establishing an informed classification of the DMD modes retrieved. This classification ultimately allowed us to identify a few modes that reconstruct accurately the Reynolds stress distribution responsible of drag generation.

Our goal in this work is to assess whether the methodology set up in [24] is still applicable to turbulent channel flows at moderate friction Reynolds number. We thus consider the \( Re_\tau \approx 950 \) database, provided by Prof. Jiménez’ group [25]. The long term goal is to reveal whether flow features linked to drag exist by comparing the it with previous works [24] and -if that is the case- learn how those structures could be modified to better understand the near-wall dynamics with the change of the Reynolds number.

This contribution is organized as follows: next section describes the DNS solver employed to generate the turbulent databases and the specific implementations of DMD strategies applied to analyze them. Section 3 discusses the results obtained. Finally, section 4 presents the conclusions of our work.

2. Numerical Methodology

2.1. Databases description

The work conducted pertains to the \( Re_\tau \approx 950 \) database in [25]. However, whenever necessary, we refer to prior results on a \( Re_\tau \approx 200 \) database [24]. Accordingly, we will indicate with either R950 or R200 to which specific database we are referring to. The parameters of the DNS database studied are summarized in Table 1. The R200 database was generated by the incompressible DNS solver described in [26]. The code follows the paradigm introduced in [27]: it solves for the wall-normal components of velocity \( v \) and vorticity \( \eta \). This quantities are Fourier-transformed (de-aliased using the \( 2/3 \) rule) along the homogeneous directions, and discretized using explicit compact finite-differences along the wall normal direction. Both the streamwise \( u \) and spanwise \( w \) velocity components are retrieved using the continuity equation with the relation \( \eta = \frac{\partial w}{\partial x} - \frac{\partial u}{\partial z} \). Time integration is accomplished by an explicit third order, low-storage Runge–Kutta method, combined with an implicit second-order Crank–Nicolson scheme. The simulations have been conducted under the assumption of constant mass flux, and more detailed see [24].

The case R950 is described in more detail in Lozano-Durán & Jiménez [28, 29]. The code employed also follows [27]. The spatial discretization employed dealiased Fourier
Table 1. Parameters of the simulation.

| Case | $Re$ | $L_x/h$ | $L_z/h$ | $\Delta x^+$ | $\Delta z^+$ | $\Delta y_{\text{max}}^+$ | $N_x, N_z$ | $N_y$ | $\Delta t_s^+$ | $u_\tau$ |
|------|------|---------|---------|--------------|--------------|----------------|-----------|------|----------------|--------|
| R200 | 200  | $2\pi$  | $\pi$   | 6.54         | 3.27         | 5.18           | 192       | 129  | 0.1            | 0.04198|
| R950 | 932  | $2\pi$  | $\pi$   | 11           | 5.7          | 7.6            | 768       | 385  | 0.8            | 0.04539|

in the two wall-parallel directions, and either Chebychev or seven-point compact finite differences in the wall-normal one.

The streamwise and spanwise coordinates are $x$ and $z$, and the wall-normal coordinate, $y$, is zero at the wall. Throughout this paper, $u'$, $v'$ and $w'$ are streamwise, wall-normal and spanwise velocity fluctuations, measured with respect to their mean, which is defined over the two homogeneous directions and time.

The channel half-height is $h$, and ‘$+$’ superscripts denote wall units defined in terms of the friction velocity $u_\tau$ and of the kinematic viscosity $\nu$. The Kármán number is $Re_\tau = u_\tau h/\nu$. The Kolmogorov length and time scales are $\mu = (\nu^3/\varepsilon)^{1/4}$ and $t_\mu = (\mu/\varepsilon)^{1/2}$, respectively, where $\varepsilon$ is the mean dissipation rate of the kinetic energy.

The analysis of the temporal evolution of the flow requires storing approximately $10^3 \sim 10^4$ snapshots for each simulation, implying hundreds of gigabytes (GBs) to tens of terabytes (TBs) for each channel in Table 1. For the channel dimensions, they had been shown in [30, 28] that this box size is the minimum needed to accommodate the widest flow structures and results in correct one-point statistics. For the analyses in this work, we may try several efforts to reduce the burden of applying the DMD techniques, which make it possible to analyses the turbulent flow of multiscales in considerable time and memory requirements.

2.2. Feature detection algorithms: composite DMD

We present here a brief summary of the DMD technique, as proposed in [6]. Given a sequence of instantaneous flow fields numbered from 1 to $n_s$ (e.g., taking one or all recorded variables), the following data matrix can be constructed:

$$V^{n_s}_{1} = \{v(t_1), v(t_2), \ldots, v(t_{n_s})\},$$

where the subindex and superindex identify, respectively, the first and last time instants of the sequence. The data is ordered in time, and separated by a constant sampling time interval $\Delta t^a$ such that: $t_{k+1} = t_k + \Delta t^a$ for all $k = 1, \ldots, n_s - 1$. In the case of linear stability analysis and within the exponential growth region, it is possible to define a linear operator $A$ (i.e., a numerical approximation of the linearized Navier–Stokes operator) such that $v(t_{k+1}) = A v(t_k)$. For non-linear systems, $A$ represents the Koopman operator. Eq. 1 can then be rewritten as a Krylov sequence (see [31]):

$$V^{n_s}_{1} = \{v(t_1), A v(t_1), \ldots, A^{n_s-1} v(t_1)\}.$$
For an ordered sequence, Eq. 2 can be equated to Eq. 1:

$$A\{v(t_1), v(t_2), ..., v(t_{n_s-1})\} = \{v(t_2), v(t_3), ..., v(t_{n_s})\},$$

which can alternatively be written in matrix form as:

$$AV_1^{n_s-1} = V_2^{n_s}. \quad (4)$$

Next, the Singular Value Decomposition (SVD) of the matrix $V_1^{n_s-1} = U\Sigma W^H$ is obtained; the superscript $H$ denotes conjugate transposition. Matrix $\Sigma$ is a diagonal matrix with entries $\sigma_i$ the singular values. The left singular vectors – the columns of $U$ – can be related to the POD modes of the input data sequence [32]; the DMD algorithm of Schmid [6] offers the POD modes as a by-product.

The SVD of the snapshot matrix is then inserted into Eq. 4, which yields $AU\Sigma W^H = V_2^{n_s}$. The reduced matrix $\tilde{A} = U^H A U$ associated with the initial system described by $A$, can be rewritten using the former equality as:

$$\tilde{A} = U^H A U = U^H V_2^{n_s} W \Sigma^{-1}. \quad (5)$$

The reduced matrix $\tilde{A}$ is the projection of the matrix $A$ onto the space contained in $U$, and previously obtained through the SVD operation [6]. The DMD operates under the assumption that the projected matrix $\tilde{A}$ conveys most of the information codified into operator $A$.

Once the reduced matrix $\tilde{A}$ has been calculated, the reduced DMD modes $y_i$ can be obtained, as well as the associated eigenvalues $\mu_i$ (i.e., growth rates $\Re(\mu_i)$ and frequencies $\Im(\mu_i)$ mapped to the unit circle) of the reduced system by solving the eigenvalue problem $\tilde{A}y_i = \mu_i y_i$. The approximated eigenmodes of the matrix $A$ can then be recovered via a projection onto the original space, using relation $\phi_i = U y_i$. Eventually, the growth rates and frequencies in the complex half-plane can be recovered from the eigenvalues as $\lambda_i = \log(\mu_i)/\Delta t_s$ (do not mistake the $i$-th eigenvalue $\lambda_i$ with the $\lambda_2$ invariant).

Finally, note that the DMD decomposition allows to reconstruct the original data sequence as:

$$v(t) = \sum_{i=1}^{n_s-1} \alpha_i \phi_i e^{\lambda_i t}. \quad (6)$$

In this contribution, the amplitudes $\alpha_i$ are computed following the formulation in [18]. That is, the $\alpha_i$’s stem from the minimization problem in the Fröhbenius norm:

$$\min_{\alpha_i} \|\Sigma W^H - YD_\alpha V\|_F^2, \quad (7)$$

where the columns in matrix $Y$ are the eigenvectors $y_i$, diagonal matrix $D_\alpha$ contains the unknown amplitudes $\alpha_i$ and $V$ is a Vandermonde matrix whose columns are generated by the successive powers of the column vector $[\mu_1^k, \ldots, \mu_{n_s-1}^k]^T$, with $k = 0, \ldots, n_s - 1$. 


2.3. Composite DMD analysis of the turbulent databases

The DNS solver provides the flow state at every $\Delta t^*$ time instant, which we obtained the velocity field from the R950 database.

In this work two types of DMD analyses have been considered: classical DMD analysis –performed on snapshots of instantaneous Reynolds stress distribution, $u'v'(\vec{x}, t_k)$– and composite DMD analysis performed on snapshots obtained by concatenating instantaneous skin friction at the wall $C_f(t_k)$ and Reynolds stress $u'v'(\vec{x}, t_k)$. The reason why we do this choice could source back to the link between both magnitudes, as the FIK identity confirms [33]:

$$C_f = \frac{12}{Re_b} \text{Laminar} + 12 \int_{0}^{1} 2 (1 - y) < -u'v'> dy \text{ Turbulent}.$$  

(8)

In equation above, operator $< \circ >$ represents the temporal and spatial (along the homogeneous directions $x$ and $z$) averaging.

It should be noted that if the upper limit in expansion Eq. 6 is truncated -i.e., if a smaller number $n_r < n_s - 1$ is considered- a reduced order model for the process is obtained. Different criteria exist to discriminate which dynamic modes are to be retained, e.g., the cardinality-penalization-based criterion introduced in [18] or by the criterion of significant time-integrated contributions on the amplitude in [12].

In this work we have considered instead two simple criteria. The first criterion consists in retaining only those modes fulfilling $|\alpha_i| \geq 10\%$, where the $\alpha_i$’s are the amplitudes in Eq. 6. The second criterion retains in the expansion those modes that contribute most to the skin friction at the wall. This can be accomplished by defining the quantities $\beta_i \equiv (\phi_i \cdot e_{C_f})\alpha_i$. If $e_{C_f}$ is the unit vector along the component of the skin friction, then factor $(\phi_i \cdot e_{C_f})$ extracts the $C_f$-related component from the dynamic mode $\phi_i$ obtained in a composite DMD analysis. This second weighted criterion, shows the links between Reynolds stresses and skin friction. As we shall see in section 3.2, retaining those modes with $|\beta_i| \geq 10\%$ allows to recover with sufficient accuracy the Reynolds stresses distribution for $n_r \ll n_s$.

Finally, whatever the upper limit $n_r$ taken in expansion Eq. 6, note that the Reynolds stress profile extracted from the DMD analysis, are obtained as:

$$< u'v' >^{DMD} (y) = \frac{1}{n_s \Delta t^*} \sum_{i=1}^{n_r} \alpha_i < \phi_i - (\phi_i \cdot e_{C_f})e_{C_f} > \int_{0}^{n_s \Delta t^*} e^{\lambda_i t} dt.$$  

(9)

It is those profiles which, averaged by the $n_r$ reconstructed modes selected, will be compared against those retrieved directly from the DNS simulation in section 3.
3. Results and discussion

In this section, we firstly identify a subset of temporal snapshots that could relatively accurately represent the turbulent flow statistics at section 3.1. The subdomain taken into consideration is also discussed. Next, the results obtained by DMD with or without composite data will be discussed in section 3.2.

3.1. Data sequence definition and verification

First, since achieving a statistically converged representation of turbulent process requires averaging on large domains and over long times, processing the whole database with the DMD method is, quite often, not affordable. Accordingly, we investigate what spatial resolution and number of snapshots can be considered that is adequate to approximate the turbulent physics we are attempting to analyze.

To address this question, we show the second-order moments obtained from shorter snapshot subsequences \( n_s = 400, 500 \) and 800 applying different spatial decimation along the homogeneous directions \( x \) and \( z \), e.g., retaining every other three point (indicated as \( s_x = s_z = 4 \)). As visible in Fig. 1(a), considering spatially decimated snapshots does not impact second-order quantities. And for the number of snapshots retained, Fig. 1(b) also shows that at least in the near-wall region \( y^+ < 50 \), the error incurred is relatively small. Therefore, the data sequence considered in the rest of this work is the one consisting of the first \( n_s = 500 \) snapshots spatially decimated along both \( x \) and \( z \) directions with \( s_x = s_z = 4 \). Invoking vertical symmetry, one can also consider a half of the domain. In this manner, the data to be processed by DMD is reduced from the original 15754 snapshots of \( n_x \times n_y \times n_z \approx 2.27 \times 10^8 \) points to 500 snapshots of size \( \frac{n_x}{s_x} \times (\text{floor}(\frac{n_y}{s_y}) + 1) \times \frac{n_z}{s_z} \approx 3.16 \times 10^6 \). We considered as well sequences of snapshots in the range \( y^+ \in [0, y^+_{\text{max}}] \) for \( y^+_{\text{max}} = 50 \), which include the viscous and buffer layers, see Fig. 2.

Since then, the whole over time of the instantaneous flow-fields taken in this analysis are from \( t_1 \approx 806.044 \) to \( t_{500} \approx 814.204 \) with a time gap of \( \Delta T \approx 8.160 \) (note that \( \Delta T \neq \Delta t_x \)). One thing should be mentioned is that, as the sampling times of the snapshots provided in R950 database are slightly non-uniform, we did a linear interpolation to these sequences we are going to study, which helps to promote the analyses by DMD techniques.

3.2. DMD analysis

We have applied DMD analysis to \( n_s = 500 \) data sequences, considering snapshots formed either by just \( u'v'(\vec{x}, t_k) \) fields (classical DMD) or snapshots joining \( u'v'(\vec{x}, t_k) \) fields and \( C_f(t_k) \) values (composite DMD). From the eigenvalues distribution, Fig. 3, we find that the most of the modes lie near the locus \( |\mu| = 1 \), which respects the statistically stationary nature of this turbulent flow. Still some modes located inside the unit circle. From the eigenvalues distributions shown in Fig. 3, we also observe that both for the...
R200 and the R950 cases, classical and composite DMD eigenvalues practically coincide (see Figs. 3(a) and 3(c)).

As we said, classical and composite DMD give the practically the same eigenvalues distribution (see Fig. 3). However, by sorting according to the $\beta_i$ factor introduced in section 2.3, a smaller number of modes of a certain relevance is retrieved. Indeed, we observe how -no matter the Reynolds number considered - the number of modes associated either with large values of $|\beta_i|$ is always smaller than the number of modes associated with large $|\alpha_i|$, see Figs. 3(b) and 3(d).

Figure 4, in turn, establishes how the superposition of those few (less than 1%) composite DMD modes associated with large values of $|\beta_i|$ reconstructs (Eq. 9) quite accurately the Reynolds stress profile. On the contrary, still a fair amount of classical DMD modes associated with large values of $|\alpha_i|$ is not capable of providing an accurate

---

Figure 1. Reynolds shear stress $\langle u'v' \rangle$: data sequence sensitivity to length $n_s$ and spatial decimation $s_x$, $s_z$, based on the R950 database. In (a), sensitivity to length $n_s$; (b), sensitivity to spatial decimation.

Figure 2. The friction velocity normalized wall normal streamwise velocity profile.
reconstruction of the Reynolds stress. This seems to indicate that modes selected by large values of $|\beta|$ are more relevant to the physics investigated than those associated with large values of $|\alpha|$.

Finally, let us visualize the Dynamic modes from the composite DMD analyses. Fig. 5(a) shows the reconstructed $u'v'$ field for the R200 case using the $n_r = 6$ modes with $|\beta_i| / |\beta_{max}| > 10\%$. Fig. 5(b) does the same for single mode fulfilling previous condition. In both cases, features staying $\Delta z^+ \approx 100 - 120$ apart from each other are visible, which is consistent with the findings in e.g. [10, 23].
4. Conclusion

In this contribution, we presented a strategy to perform a composite DMD analysis on a turbulent channel flow at moderate Reynolds number ($Re \approx 950$). On this turbulent flow with more than $2.27 \times 10^8$ grid points per snapshot, we conducted the composite based DMD analyses on sequence of snapshots formed by skin friction $C_f(t_k)$ and the Reynolds shear stress field $u'v'$. Since the database memory size is far beyond the capacity of our machine, the first step was to find a proper scheme to reduce the data size while keeping the turbulent flow physics. The reduction of both the spatial resolution and temporal extent have been undertaken. Finally, we have retained a snapshot sequence of $n_s = 500$ snapshots and the resolution decimation $s_x = s_z = 4$ as an acceptable input for our investigation; also half channel was considered. We found that the first DMD mode (out of 499) can
As for the composite DMD analyses, and the classification of the modes according to factor $\beta_i$, we can identify a small number of dynamic modes (less than 1% of the original sequence length) and still reconstruct the shear stress properly.

To summarize, this work has described the application of the composite DMD technique to a higher Reynolds number turbulent channel flow, showing the possibility of investigating large database by the DMD techniques. Future work will consist of composite DMD analyses on either $C_f(t_k) - u'v'(\vec{x}, t_k)$ or $C_f(t_k) - u'(\vec{x}, t_k)$ on longer data sequences for which the strategies described in [34] will be leveraged.

**Acknowledgments**

The authors are grateful to Professor W. Li for his valuable comments on an early version of this work.

The support provided by China Scholarship Council (CSC, No.201806320222) during a visit of Mr. Binghua Li to UPM-ETSIAE is acknowledged. This work is funded by Projects: Drag Reduction via Turbulent Boundary Layer Flow Control (DRAGY, GA-690623), 2016-2019, China-EU Aeronautical Cooperation project, co-funded by Ministry of Industry and Information Technology (MIIT), China, and Directorate -General for Research and Innovation (DG RTD), European Commission. The first author also would like to thank Prof. Javier Jiménez for his kind invitation to the fourth Madrid Turbulence Workshop, which was supported by the Coturb program of the European Research Council.
References

[1] R Theunissen, F Scarano, and M L Riethmüller. Spatially adaptive piv interrogation based on data ensemble. *Exp. Fluids*, 48(5):875–887, 2010.

[2] D GE Grigoriadis, B J Geurts, H Kuerten, J Fröhlich, and V Armenio. *Direct and large-eddy simulation*. X. Springer, 2018.

[3] G Berkooz, P Holmes, and J L Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.*, 25(1):539–575, 1993.

[4] S Volkwein. Proper orthogonal decomposition: Theory and reduced-order modelling. *Lecture Notes, University of Konstanz*, 4(4):1–29, 2013.

[5] C W Rowley, J Mezić, S Bagheri, P Schlatter, and D S Henningson. Spectral analysis of nonlinear flows. *J. Fluid Mech.*, 641:115–127, 2009.

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

[7] P Moin and R D Moser. Characteristic-eddy decomposition of turbulence in a channel. *J. Fluid Mech.*, 200:471–509, 1989.

[8] B Podvin. A proper-orthogonal-decomposition–based model for the wall layer of a turbulent channel flow. *Phys. Fluids*, 21(1):015111, 2009.

[9] M Lagha. A comprehensible low-order model for wall turbulence dynamics. *Phys. Fluids*, 26(8):085111, 2014.

[10] A Cassinelli, M de Giovannetti, and Y Hwang. Streak instability in near-wall turbulence revisited. *J. Turbul.*, 18(5):443–464, 2017.

[11] S Le Clainche, J M Vega, and J Soria. Higher order dynamic mode decomposition of noisy experimental data: The flow structure of a zero-net-mass-flux jet. *Exp. Therm. Fluid Sci.*, 88:336–353, 2017.

[12] J Kou and W Zhang. An improved criterion to select dominant modes from dynamic mode decomposition. *Eur. J. Mech. B Fluids*, 62:109–129, 2017.

[13] T Sayadi, P J Schmid, J W Nichols, and P Moin. Reduced-order representation of near-wall structures in the late transitional boundary layer. *J. Fluid Mech.*, 748:278–301, 2014.

[14] P J Schmid, D Violato, and F Scarano. Decomposition of time-resolved tomographic piv. *Exp. Fluids*, 52(6):1567–1579, 2012.

[15] S Le Clainche and J M Vega. Higher order dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.*, 16(2):882–925, 2017.

[16] S Le Clainche, Z Han, and E Ferrer. An alternative method to study cross-flow instabilities based on high order dynamic mode decomposition. *Phys. Fluids*, 31(9):094101, 2019.

[17] K K Chen, J H Tu, and C W Rowley. Variants of dynamic mode decomposition: boundary condition, koopman, and fourier analyses. *J. Nonlinear Sci.*, 22(6):887–915, 2012.

[18] M R Jovanović, P J Schmid, and J W Nichols. Sparsity-promoting dynamic mode decomposition. *Phys. Fluids*, 26(2):024103, 2014.

[19] F Guéniat, L Mathelin, and L R Pastur. A dynamic mode decomposition approach for large and arbitrarily sampled systems. *Phys. Fluids*, 27(2):025113, 2015.

[20] T Sayadi and P J Schmid. Parallel data-driven decomposition algorithm for large-scale datasets: with application to transitional boundary layers. *Theor. Comput. Fluid Dyn.*, 30(5):415–428, 2016.

[21] T Bui-Thanh, M Damodaran, and K E Willcox. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA J.*, 42(8):1505–1516, 2004.

[22] T Brachonnier, M Ferrer, J C Jouhaud, M Montagnac, and P Sagaut. Towards an adaptive pod/svd surrogate model for aeronautical design. *Comput. Fluids*, 40(1):195–209, 2011.

[23] M Lagha, J Kim, J D Eldredge, and X Zhong. A numerical study of compressible turbulent boundary layers. *Phys. Fluids*, 23(1):015106, 2011.

[24] J Garicano-Mena, B Li, E Ferrer, and E Valero. A composite dynamic mode decomposition
analysis of turbulent channel flows. *Phys. Fluids*, 31(11):115102, 2019.

[25] UPM turbulent database. [https://torroja.dmt.upm.es/turbdata](https://torroja.dmt.upm.es/turbdata). Accessed: 2019-06-30.

[26] P Luchini and M Quadrio. A low-cost parallel implementation of direct numerical simulation of wall turbulence. *J. Comput. Phys.*, 211(2):551–571, 2006.

[27] J Kim, P Moin, and R Moser. Turbulence statistics in fully developed channel flow at low reynolds number. *J. Fluid Mech.*, 177:133–166, 1987.

[28] A Lozano-Durán and J Jiménez. Effect of the computational domain on direct simulations of turbulent channels up to \( re_x = 4200 \). *Phys. Fluids*, 26(1):011702, 2014.

[29] A Lozano-Durán and J Jiménez. Time-resolved evolution of coherent structures in turbulent channels: characterization of eddies and cascades. *J. Fluid Mech.*, 759:432–471, 2014.

[30] O Flores and J Jiménez. Hierarchy of minimal flow units in the logarithmic layer. *Phys. Fluids*, 22(7):071704, 2010.

[31] Y Saad. *Numerical methods for large eigenvalue problems*. Manchester University Press, 1992.

[32] C W Rowley and S TM Dawson. Model reduction for flow analysis and control. *Annu. Rev. Fluid Mech.*, 49:387–417, 2017.

[33] K Fukagata, K Iwamoto, and N Kasagi. Contribution of reynolds stress distribution to the skin friction in wall-bounded flows. *Phys. Fluids*, 14(11):L73–L76, 2002.

[34] B Li, J Garicano-Mena, E Ferrer, and E Valero. Feature extraction algorithms applied to turbulent channel flow databases: The effect of agglomeration strategies. In *Proceedings 8th European Conference for Aeronautics and Aerospace Sciences*, EUCASS, 2019.