A sparse optimal closure for a reduced-order model of wall-bounded turbulence

Zhao Chua Khoo, Chi Hin Chan†, and Yongyun Hwang

Department of Aeronautics, Imperial College London, South Kensington, London SW7 2AZ, UK

(Received xx; revised xx; accepted xx)

In the present study, a set of physics-informed and data-driven approaches are examined towards the development of an accurate reduced-order model for a turbulent plane Couette flow. Based on the utilisation of the proper orthogonal decomposition (POD), a particular focus is given to the development of a reduced-order model where the number of POD modes are not large enough to cover the full dynamics of the given turbulent state, the situation directly relevant to the reduced-order modelling for turbulent flows. Starting from the conventional Galerkin projection approach ignoring the truncation error, three approaches enhanced by both physics and data are examined: 1) sparse regression of the POD-Galerkin dynamics; 2) Galerkin projection with an empirical eddy viscosity model; 3) Galerkin projection with an optimal eddy viscosity obtained from a newly-proposed sparse regression - an idea applying the Sparse Identification of Nonlinear Dynamics (SINDy) framework to an eddy-viscosity model. The sparse regression of the POD-Galerkin dynamics does not perform well, as the number of POD modes for the given chaotic dynamics appears to be too small. While the unsatisfactory performance of the Galerkin projection model with an empirical eddy viscosity is observed, the newly proposed approach, which combines the concept of an optimal eddy-viscosity closure with a sparse regression, more accurately approximates the chaotic dynamics than the other reduced-order models considered. This is demonstrated with the mean and time scales of the POD mode amplitudes as well as the first- and second-order turbulence statistics.

Key words: Low-dimensional models, Turbulent boundary layers

1. Introduction

1.1. *Dynamical systems approach for wall-bounded turbulence*

Coherent structures in turbulent flows have been studied for many decades. These highly organised fluid motions in a chaotic flow field often carry significant amount of turbulent kinetic energy and momentum. Understanding and modelling of their dynamics have been a central challenge in turbulence research. In wall-bounded shear flows, a coherent structure was first discovered in the near-wall region (Kline *et al.* 1967) and many different kinds of coherent structures were subsequently observed over the past half century (e.g. Kovasznay

† Email address for correspondence: chi.chan19@imperial.ac.uk

et al. 1970; Falco 1977; Head & Bandyopadhay 1981; Jeong *et al.* 1997; Kim & Adrian 1999; del Álamo & Jiménez 2003; Hutchins & Marusic 2007, and many others). The growing evidence suggests that these coherent structures are organised in the form of the so-called 'attached eddies' originally hypothesised by Townsend (1956, 1976) for the logarithmic layer: coherent structures in the logarithmic layer emerge in the form of a self-similar hierarchy and their characteristic length scale is proportional to the distance between the eddy centre and the wall (see Hwang & Lee 2020 for a mathematical proof). The attached eddy hypothesis can be extended to include the near-wall and outer regions in a broad sense (Hwang 2015), and there has been a growing body of evidence supporting this idea over the past decade (see the recent review of Marusic & Monty 2019, and the references therein). In particular, each of these attached eddies was found to retain a sustaining mechanism independent of the others (Flores & Jiménez 2010; Hwang & Cossu 2010*c*, 2011; Hwang & Bengana 2016). This mechanism has been referred to as the 'self-sustaining process' and is based on a quasi-cyclic interaction of streaks and quasi-streamwise vortices (Hamilton *et al.* 1995; Waleffe 1997): 1) amplification of streamwise-elogated streaks by streamwise vortices through the lift-up effect (Butler & Farrell 1993; del Álamo & Jiménez 2006; Cossu *et al.* 2009; Hwang & Cossu 2010*a*,*b*); 2) breakdown of the amplified streaks via an instability and/or transient growth (Hamilton *et al.* 1995; Schoppa & Hussain 2002; Park *et al.* 2011; Alizard 2015; Cassinelli *et al.* 2017; de Giovanetti *et al.* 2017; Lozano-Durán *et al.* 2021); 3) nonlinear regeneration of streamwise vortices (Hamilton *et al.* 1995; Schoppa & Hussain 2002; Hwang & Bengana 2016). A key on-going challenge is to understand the interactions between the self-sustaining processes at multiple length scales, and recent studies have suggested that the interaction dynamics appear to be dauntingly complex (Cho *et al.* 2018; Lee & Moser 2019; Doohan *et al.* 2021*b*).

The discovery of the self-sustaining process has played a central role in advancing the notions of dynamical systems for turbulence research. In particular, it physically underpins the existence of non-trivial unstable equilibrium and time-periodic solutions in wall-bounded shear flows (Nagata 1990; Waleffe 2001; Kawahara & Kida 2001; Jiménez & Simens 2001; Waleffe 2003; Faisst & Eckhardt 2003; Wedin & Kerswell 2004; Gibson *et al.* 2008, 2009; Hall & Sherwin 2010; Park & Graham 2015; Hwang *et al.* 2016; Yang *et al.* 2019; Doohan *et al.* 2019, and many others). These solutions form a state-space skeleton for the birth of turbulence through a sequence of local and global bifurcations (Eckhardt *et al.* 2007; Kawahara *et al.* 2012; Graham & Floryan 2020), and their use has been central to the description for the temporal dynamics of transition to turbulence (e.g. Kreilos & Eckhardt 2012) and for the local behaviour in the spatio-temporal dynamics (see the review by Barkley 2016). Furthermore, given that they are exact solutions to the Navier-Stokes equations, they provide precise understanding for turbulence dynamics in an interpretable and mathematically analysable manner especially compared to the structures obtained with conventional conditional average.

Despite these advances, the computation of the unstable equilibrium and periodic solutions are increasingly infeasible as Reynolds number increases. A key reason to this is that the typical algorithms used for the search of these solutions are designed to iteratively find the initial condition that leads to the same flow field after a given time interval (an arbitrary small time interval for equilibrium and a time period for periodic orbits; e.g. Viswanath 2007; Willis *et al.* 2013; Farazmand 2016). Apart from the computational cost required for the repeated simulations, the numerical convergence of such an algorithm depends on the leading Lyapunov exponent of the related chaotic state – it becomes increasingly difficult to find an initial flow field which converges with a sufficiently small residual when the leading Lyapunov exponent is very large. An approach employed to bypass this difficulty was to approximate the complex multi-scale dynamics at high Reynolds numbers with a closure model (e.g. Rawat *et al.* 2015; Hwang *et al.* 2016; Yang *et al.* 2019). Yet, the unstable equilibrium and periodic solutions obtained with this strategy are often obtained by continuing the existing solutions at low Reynolds numbers, thereby not retaining the key multi-scale processes of interest at high Reynolds numbers (e.g. scale interactions). Indeed, a very recent work by Doohan *et al.* (2021*a*) showed that when turbulence exhibits multi-scale behaviours explicitly, most of the equilibrium solutions obtained in this manner do not sit anywhere near turbulent state in the physically relevant phase portraits. An obvious way to resolve this issue is to compute unstable periodic orbits with sufficiently long time periods, as they should be able to capture the key periodic and/or quasi-periodic multi-scale dynamics. However, in practice, their computation is much more difficult than that of the equilibrium solutions and is practically almost impossible due to the rapidly vanishing convergence of the solutions by the increasing Reynolds number. This poses an important challenge in extending the notions of dynamical systems for the description of multi-scale behaviours of turbulent flows.

Given this challenge, it is worth mentioning that many of the unstable equilibrium and periodic solutions have previously been computed in a highly reduced system even at low Reynolds numbers. Indeed, a large number of such solutions have been found in highly confined computational domains, in which the full spatio-temporal dynamics in a large computational domain would be drastically reduced (Nagata 1990; Waleffe 1995; Kawahara & Kida 2001; Jiménez & Simens 2001; Faisst & Eckhardt 2003; Wedin & Kerswell 2004; Gibson *et al.* 2008; Park & Graham 2015; Hwang *et al.* 2016; Doohan *et al.* 2019, 2021*a*). For the same rationale, perhaps, the key to tackling the multi-scale dynamics of turbulence using the dynamical systems notions may lie in a suitable dimensionality reduction of the given turbulent state without losing the core dynamics of interest. The equilibrium and periodic solutions to the reduced-order dynamical system can then be obtained much more easily using existing techniques, with hope that they can subsequently be used as proxies and/or a symbolic description for the multi-scale dynamics of a turbulent system.

1.2. *Reduced-order modelling*

Reduced-order modelling has been a long-standing topic in fluid mechanics, and wall-bounded turbulence is a flow configuration to which the earliest modelling efforts were made (Aubry *et al.* 1988; Rempfer & Fasel 1994*a*). The previous efforts may be classified into two categories. One is based on the proper orthogonal decomposition (e.g. Lumley 1967, 1981; Sirovich 1987; Holmes *et al.* 1996), and the reduced-order model is subsequently obtained by projecting the Navier-Stokes equations onto the space spanned by the POD modes (i.e. Galerkin projection; Aubry *et al.* 1988; Rempfer & Fasel 1994*a*,*b*; Smith *et al.* 2005). The other is equivalent to a heavily truncated spectral approximation to the system of interest (Waleffe 1997; Moehlis *et al.* 2004, 2005; Lagha & Manneville 2007; Chantry *et al.* 2017; Cavalieri 2021), which is often used to study low-dimensional dynamics of bifurcation and transition to turbulence. Given the scope of the present study, here we shall employ the former approach based on POD modes, as they provide the best orthonormal basis in terms of capturing the energy of the given flow fields.

One of the key challenges in the development of a reduced-order model using the POD modes are often associated with the number of POD modes considered and the resulting truncation. This problem becomes significant especially for high Reynolds number flows, where the small-scale motions, not captured by the chosen POD modes, play a crucial role in the energy cascade and turbulent dissipation. Indeed, one of the most fundamental features of turbulence is the exact balance between the production at large scale and the dissipation at the small (Kolmogorov) scale. Therefore, not accounting for such small-scale motions leads to an excess of energy and/or an erroneous behaviour in the reduced-order model.

This issue has often been resolved by incorporating an additional eddy-viscosity model that removes the excess in the energy of the reduced-order model (Aubry *et al.* 1988; Rempfer & Fasel 1994*b*; Couplet *et al.* 2005; Smith *et al.* 2005; Noack *et al.* 2011; Östh *et al.* 2014; Protas *et al.* 2015). Given the scope of the present study, the approach of Protas *et al.* (2015) is particularly appealing, in which an 'optimal' eddy viscosity was proposed to minimise the difference between the data from the measurement and from the reduced-order model. In this approach, the gradient of the given objective functional is computed over multiple time intervals using the adjoint-based formulation. The best-performing eddy viscosity is subsequently obtained by updating its value at every time interval, an approach reminiscent of the 'sub-optimal' control in flow control problem (Choi *et al.* 1993; Lee *et al.* 1999). It is worth mentioning that such an optimisation problem is ideally formulated with an objective functional considering a long time interval. However, the resulting adjoint equation for the Lagrange multiplier is mathematically unstable around the chaotic trajectory due to the positive leading Lyapunov exponent. This inherently limits the size of the time interval that can be used for the gradient calculation. Furthermore, the leading Lyapunov exponent rapidly grows with the Reynolds number, forming an important challenge in the application of such an approach for high Reynolds numbers – indeed, a similar issue of the time interval in the adjoint-based optimisation problem was previously discussed in the context of an optimal control problem (e.g. Bewley *et al.* 2001). Recently, there have been a rapidly growing number of studies aiming to identify the governing dynamics solely from simulation and experimental data available (e.g. Schmidt & Lipson 2009; Brunton *et al.* 2016; Loiseau & Brunton 2018). In the context of fluid mechanics, a novel approach was proposed by Brunton *et al.* (2016), who accurately identified the governing low-dimensional dynamics of Lorenz chaos and the two-dimensional laminar cylinder wake from data. In their study, a reduced-order model based on POD modes is identified in a data-driven manner by formulating an l_1-regularisation-based optimisation which minimises the difference between the full and the reduced-order dynamics – the approach referred to as 'sparse identification of nonlinear dynamics (SINDy)'. Furthermore, the SINDy framework has been extended by Loiseau & Brunton (2018), who introduced a constraint enforcing energy-preserving nonlinearity by reformulating the sparse regression problem with the Karush-Kuhn-Tucker conditions. There is also on-going effort to improve low-dimensional representation of nonlinear dynamics by nonlinear correlations in the temporal POD coefficients (Callaham *et al.* 2021*a*). For high Reynolds number turbulent flows, a recent study by Callaham *et al.* (2021*b*), the statistical-state dynamics of turbulence in three-dimensional bluff-body wake was modelled by combining a normal form amplitude equation with stochastic noise determined using a similar sparse regression. This approach based on a simple normal form amplitude equation is, however, not attractive for wall-bounded shear flows, as the key coherent structure dynamics in this case involves a much more sophisticated global bifurcation (e.g. the crisis bifurcation; see Kreilos & Eckhardt 2012) as well as a rich spatio-temporal dynamics that can be described with the concept of thermodynamic phase transition (e.g. Avila *et al.* 2011; Barkley 2016). It is finally worth mentioning that, in the context of turbulence modelling, there have been several successful recent efforts made with the use of SINDy for a closure of the Reynolds-Averaged Navier-Stokes (RANS) equations (e.g. Beetham & Capecelatro 2020; Schmelzer *et al.* 2020; Beetham *et al.* 2021). A comprehensive review on the utilisation of machine learning approaches for turbulence modelling in RANS and LES (large eddy simulation) can also be found in Duraisamy (2021), where the need to augment the given turbulent model using machine learning methods with suitable physical constraints was discussed.

1.3. *Objective of the present study*

The objective of the present study is to develop a reduced-order model more accurate than conventional ones for a turbulent state in wall-bounded shear flows. As the first step towards this task, we will consider a relatively simple but turbulent system: i.e. the minimal flow unit of a plane Couette flow (Hamilton *et al.* 1995), where the chaotic dynamics is well described in terms of the self-sustaining process. Based on the utilisation of POD modes, four different approaches will be considered and compared to devise a best performing reduced-order model: 1) Galerkin projection with a simple truncation; 2) sparse Galerkin regression (i.e. application of SINDy Brunton *et al.* 2016); 3) Galerkin projection with a simple eddy viscosity closure (Smith *et al.* 2005); 4) Galerkin projection with a SINDy closure for the truncation error. In particular, the last approach mentioned here is new and proposed by the present study, in which the concept of the optimal eddy viscosity (Protas *et al.* 2015) will be combined with a sparse regression. The idea of calibration of a given reduced-order model with an eddy-viscosity closure was previously proposed by several studies (e.g. Couplet *et al.* 2005; Cordier *et al.* 2010) with different types of regularisations. A similar idea was also recently explored in the recent work by Mohebujjaman *et al.* (2019), who applied a data-driven correction to a two-dimensional cylinder flow without an explicit physical closure model. In this context, it is worth mentioning that an important benefit of the l_1-regularisation used in this study over the common l_2-regularisation is that it prevents possible overfit of the eddy viscosity model which might yield an overdamped reduced-order model. The l_1-regularisation will also offer a more parsimonious low-dimensional description. We will see that this approach enables us to effectively determine an 'optimal' eddy viscosity, significantly improving the accuracy of the low-dimensional model based on POD modes.

This paper is organised as follows. In §2, the equations of motion, the flow geometry and the POD modes are introduced. The four reduced-order models are subsequently introduced and formulated in §3 and their performance will be examined and mutually compared in §4. This paper concludes in §5.

2. Background

2.1. *Plane Couette Flow (PCF)*

We consider a plane Couette flow (PCF) of an incompressible fluid, where the two walls move in opposite directions with the streamwise velocity, $\pm U_w$. The wall-normal distance between the two walls is given by $2h$. The kinematic viscosity of the fluid is ν, and the Reynolds number is defined by $Re = U_w h/\nu$. The streamwise, wall-normal and spanwise coordinates are made dimensionless with h, and they are denoted by $\mathbf{x} = (x, y, z)$. The two walls are set to be located at $y = \pm 1$. The equations of motion are then given by

$$\frac{\partial \mathbf{u}}{\partial t} = -(\mathbf{u} \cdot \nabla)\mathbf{u} - y\frac{\partial \mathbf{u}}{\partial x} - \mathrm{v}\mathbf{e}_x - \nabla p + \frac{1}{Re}\nabla^2\mathbf{u}, \tag{2.1}$$

where t is the time, $\mathbf{u} = (\mathrm{u}, \mathrm{v}, \mathrm{w})$ is the perturbation velocity from the laminar base flow, $\mathbf{U} = (y, 0, 0)$, p the pressure and $\mathbf{e}_x$ the unit vector in the streamwise direction.

To build a reduced-order model, we first perform a DNS confined to a minimal flow unit (MFU) (Jiménez & Moin 1991; Hamilton *et al.* 1995). Following Hamilton *et al.* (1995), the computational domain of $(L_x, L_y, L_z) = (1.75\pi, 2, 1.2\pi)$ is considered at $Re = 400$. The simulation was performed using `channelflow2.0`, an open-source DNS code (`https://www.channelflow.ch/`). The code uses the Fourier-Galerkin discretisation in the streamwise and spanwise directions and the Chebyshev-tau discretisation in the wall-normal

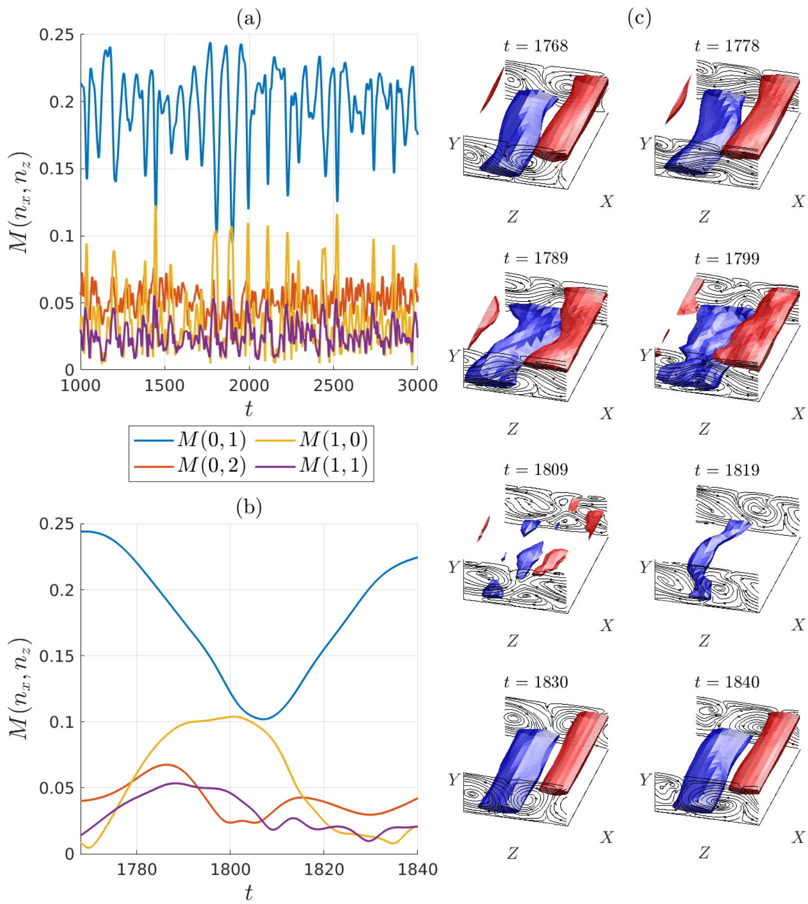


Figure 1: Time trace of $M(0, 1)$, $M(0, 2)$, $M(1, 0)$, $M(1, 1)$ (a) for $t = 1000 - 3000$ and (b) for $t = 1768 - 1840$. (c) Flow snapshots at $t = 1768, 1778, 1789, 1799, 1809, 1819, 1830, 1840$, where the blue and red iso-surfaces indicate $u = \pm 0.65$, respectively.

direction. The number of grid points in each spatial direction is given by $(N_x, N_y, N_z) = (16, 33, 16)$ (after dealiasing). A third-order semi-implicit backward difference formula (SDBF3) is used for the time-integration scheme. The simulation has been performed by setting a zero pressure gradient. The domain size normalised by viscous inner units (denoted by the superscript $(\cdot)^+$) is obtained as $(L_x^+, L_y^+, L_z^+) \simeq (186, 68, 127)$, in good agreement with that in Hamilton *et al.* (1995).

Figure 1 shows the DNS results, which exhibits the self-sustaining process (SSP). To examine the time evolution of the flow fields, the square root of energy of each plane Fourier mode is introduced:

$$M(n_x, n_z) = \left(\int_{-1}^{1} [\hat{\mathrm{u}}^2(n_x, n_z, y) + \hat{\mathrm{v}}^2(n_x, n_z, y) + \hat{\mathrm{w}}^2(n_x, n_z, y)] \, \mathrm{d}y \right)^{\frac{1}{2}}, \tag{2.2}$$

where $\hat{(\cdot)}$ denotes the Fourier coefficients for the perturbation velocity, and n_x and n_z define the streamwise and spanwise wavenumbers such that $n_x\alpha$ and $n_z\beta$ ($\alpha = 2\pi/L_x$ and

$\beta = 2\pi/L_z$). Figure 1(a) shows the time trace featuring the quasi-periodic oscillation at the SSP time scale, $T_{SSP} \approx 80 - 90$. A sequence of flow field snapshots for a single SSP cycle, which correspond to the time trace in figure 1(b), are shown in figure 1(c). The initial flow field is featured with an amplified state of the high- and low-speed streamwise velocity streaks ($t = 1768$) (note that the time evolution of the streaks is depicted by $M(0, 1)$ in figure 1b). The streaks subsequently become unstable (Hamilton *et al.* 1995) or experience the related transient growth (Schoppa & Hussain 2002), leading to their breakdown in a sinuously meandering manner. The streak breakdown emerges at $t = 1809$ where the streaks evidently disappear. Nonlinear processes subsequently regenerate streamwise vortices, leading to an increase in $M(1, 1)$. Finally, for $t = 1819 - 1840$, the regenerated streamwise vortices ($y - z$ cut planes) redistribute the momentum from mean shear, resulting in the formation of new streaks especially near the lower wall region. This mechanism is known as the 'lift-up' effect.

2.2. *Proper orthogonal decomposition*

The proper orthogonal decomposition (POD) seeks a set of orthonormal functions that maximises the ensemble-averaged projection of the velocity perturbation, $\mathbf{u}$ (e.g. Holmes *et al.* 1996). The optimisation is performed by solving the following eigenvalue problem:

$$\int_\Omega \langle \mathbf{u}(\mathbf{x}, t)\mathbf{u}^H(\mathbf{x}', t)\rangle\, \boldsymbol{\Phi}^{(n)}_{n_x,n_z}(\mathbf{x}')d^3\mathbf{x}' = \lambda^{(n)}_{n_x,n_z} \boldsymbol{\Phi}^{(n)}_{n_x,n_z}(\mathbf{x}), \tag{2.3}$$

where $(\cdot)^H$ is the complex conjugate transpose, $\langle\cdot\rangle$ an ensemble average, $\lambda^{(n)}_{n_x,n_z}$, the eigenvalue representing the average kinetic energy contained in each POD mode and $\boldsymbol{\Phi}^{(n)}_{n_x,n_z}$ are the POD modes. Here, the eigenvalue and the corresponding POD modes are indexed by a positive integer n in decreasing order of the eigenvalue for each pair of the streamwise and spanwise wavenumber indices, (n_x, n_z). The perturbation velocities are then represented in terms of a linear superposition of the POD modes:

$$\mathbf{u}(\mathbf{x}, t) = \sum_{n_x=-\infty}^{\infty} \sum_{n_z=-\infty}^{\infty} \sum_{n=1}^{\infty} a^{(n)}_{n_x,n_z}(t)\boldsymbol{\Phi}^{(n)}_{n_x,n_z}(\mathbf{x}), \tag{2.4a}$$

where $a^{(n)}_{n_x,n_z}(t)$ is the time-dependent amplitude of each POD mode. Given the translational invariance in the streamwise and spanwise directions, Fourier expansions are *optimal* and the POD mode is further written as

$$\boldsymbol{\Phi}^{(n)}_{n_x,n_z}(\mathbf{x}) = \frac{1}{\sqrt{L_x L_z}} e^{i(\alpha n_x x + \beta n_z z)} \boldsymbol{\phi}^{(n)}_{n_x,n_z}(y), \tag{2.4b}$$

where $\boldsymbol{\phi}^{(n)}_{n_x,n_z}(y)$ describes the wall-normal structure of each POD mode. Since the velocities in physical space are real-valued, the following conjugate symmetries are also satisfied:

$$a^{(n)}_{n_x,n_z}(t) = a^{(n)*}_{-n_x,-n_z}(t) \quad \text{and} \quad \boldsymbol{\Phi}^{(n)}_{n_x,n_z} = \boldsymbol{\Phi}^{(n)*}_{-n_x,-n_z}, \tag{2.4c}$$

where the superscript $(\cdot)^*$ indicates the complex conjugate. Substituting (2.4a) and (2.4b) into (2.3), the eigenvalue problem is simplified to

$$\int_{-1}^{1} \underbrace{\langle \hat{\mathbf{u}}(n_x, n_z, y, t)\hat{\mathbf{u}}^H(n_x, n_z, y', t)\rangle}_{\equiv \hat{\mathbf{R}}} \boldsymbol{\phi}^{(n)}_{n_x,n_z}(y')\, \mathrm{d}y' = \lambda^{(n)}_{n_x,n_z} \boldsymbol{\phi}^{(n)}_{n_x,n_z}(y). \tag{2.5}$$

The ensemble-averaged covariance $\hat{\mathbf{R}}$ is computed by enforcing the discrete symmetries

(n_x, n_z, n)	$(0,0,1)$	$(0,\pm1,1)$	$(0,\pm2,1)$	$(\pm1,0,1)$	$(0,\pm3,1)$	$(0,0,2)$	$(0,\pm1,2)$
λ	4.481	0.767	0.0574	0.0388	0.0201	0.0159	0.0113
E [%]	68.4	23.4	1.75	1.18	0.614	0.243	0.345

Table 1: Eigenvalues of the first 7 POD modes, ranked in terms of the eigenvalue of the POD mode λ. Here, E [%] is the total energy content of both (n_x, n_z, n) and $(-n_x, -n_z, n)$ wavenumber pairs.

of the PCF:

$$\begin{aligned} \mathcal{I} &: [\mathrm{u}, \mathrm{v}, \mathrm{w}, p](x, y, z, t) = [\mathrm{u}, \mathrm{v}, \mathrm{w}, p](x, y, z, t), \\ \mathcal{P} &: [\mathrm{u}, \mathrm{v}, \mathrm{w}, p](x, y, z, t) = [\text{-u}, \text{-v}, \text{-w}, p](-x, -y, -z, t), \\ \mathcal{R} &: [\mathrm{u}, \mathrm{v}, \mathrm{w}, p](x, y, z, t) = [\mathrm{u}, \mathrm{v}, \text{-w}, p](x, y, -z, t), \\ \mathcal{RP} &: [\mathrm{u}, \mathrm{v}, \mathrm{w}, p](x, y, z, t) = [\text{-u}, \text{-v}, \mathrm{w}, p](-x, -y, z, t), \end{aligned} \tag{2.6}$$

where $\mathcal{I}$ is the identify transformation, $\mathcal{P}$, a point reflection about the origin, $\mathcal{R}$, a reflection about the z-plane, and $\mathcal{RP}$, a 180° rotation about the z-axis. Given the statistically stationary nature of the turbulent state and the invariance of the PCF under the discrete group transformation in (2.6), the covariance operator is obtained as

$$\hat{\mathbf{R}}(n_x, n_z, y, y') = \frac{1}{4T} \int_{t_0}^{t_0+T} \sum_{\mathcal{T} \in \mathcal{D}_2} (\widehat{\mathcal{T} : \mathbf{u}})(n_x, n_z, y, t)(\widehat{\mathcal{T} : \mathbf{u}})^H(n_x, n_z, y', t)\, \mathrm{d}t, \tag{2.7}$$

where $\mathcal{D}_2 = \{\mathcal{I}, \mathcal{P}, \mathcal{R}, \mathcal{RP}\}$.

For the computation of the POD modes, the ensemble-averaged covariance $\hat{\mathbf{R}}$ is first obtained and the eigenvalue problem (2.5) is subsequently solved. Since the ensemble average is equivalent to time average for a statistically stationary flow, $\hat{\mathbf{R}}$ is obtained by averaging in time over an interval $t \in [-10000, 0]$ with a sampling time interval $\Delta t = 1$ (i.e. $t_0 = -10000$ and $T = 10000$). Within this time interval, the turbulent state is chosen to be statistically stationary. We also note that the typical time period of the SSP is about $T_{SSP} \approx 80-90$, implying that more than 100 cycles are considered for the construction of the POD modes. Table 1 shows the leading eigenvalues obtained in the present study, and they are found to match closely with those reported in Smith *et al.* (2005). The structures of the 9 leading POD modes are also visualised in figure 2.

3. Reduced-order models

3.1. *POD-Galerkin projection*

To build a reduced-order model, we consider the velocity given by (2.4*a*) with a finite number of POD modes:

$$\mathbf{u}(\mathbf{x}, t) = \sum_{n_x=-M_x}^{M_x} \sum_{n_z=-M_z}^{M_z} \sum_{n=1}^{N_p} a_{n_x,n_z}^{(n)}(t) \boldsymbol{\Phi}_{n_x,n_z}^{(n)}(\mathbf{x}) + \mathbf{u}_R(\mathbf{x}, t), \tag{3.1}$$

where M_x, M_z and N_p are the numbers of streamwise, spanwise Fourier modes and $\boldsymbol{\phi}_{n_x,n_z}^{(n)}$, respectively. $\mathbf{u}_R(\mathbf{x}, t)$ is the residual velocity field that will not be resolved by the reduced-order model. After substituting (3.1) into (2.1), the projection onto each POD basis yields the following system of ordinary different equations (ODE):

$$\dot{\mathbf{a}} = \mathbf{L}\mathbf{a} + \mathbf{N}(\mathbf{a}, \mathbf{a}) + \mathbf{T}, \tag{3.2}$$

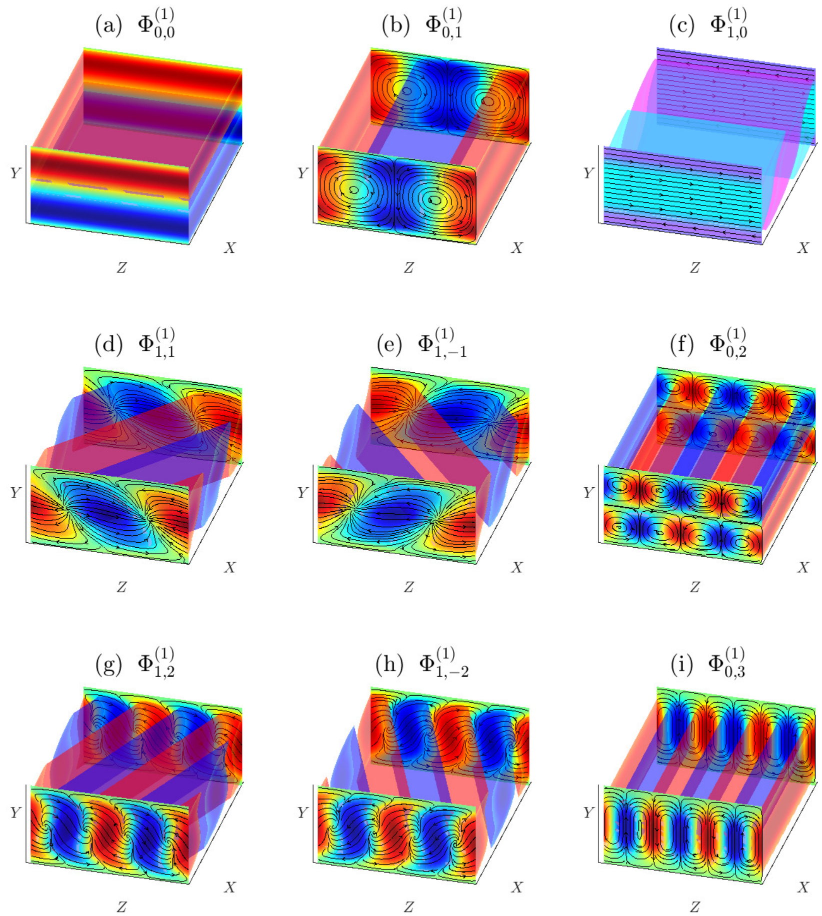


Figure 2: Visualisation of 9 leading POD modes (n_x, n_z, n): (a) $(0,0,1)$; (b) $(0,1,1)$; (c) $(1,0,1)$; (d) $(1,1,1)$; (e) $(1,-1,1)$; (f) $(0,2,1)$; (g) $(1,2,1)$; (h) $(1,-2,1)$; (i) $(0,3,1)$. Iso-surfaces of (a,b,d–i) denote $u=-0.15$ (red) and $u=0.15$ (blue). Iso-surfaces of (c) denote $w=0.15$ (teal) and $w=-0.15$ (purple).

where $\dot{(\cdot)} \equiv d/dt$, $\mathbf{a} \in \mathbb{C}^r$ with $r = [(2M_x+1)(2M_z+1)N_P+1]/2$ as the column vector, the element of which is given by $a^{(n)}_{n_x,n_z}$ for $-M_x \leqslant n_x \leqslant M_x$, $-M_z \leqslant n_z \leqslant M_z$ and $1 \leqslant n \leqslant N_p$, $\mathbf{L}$ and $\mathbf{N}$ are from the projection of the linear and quadratic nonlinear parts of (2.1) onto the finite number of POD modes (for further details on their definitions, see Appendix A), and $\mathbf{T}$ is the residual term originating from $\mathbf{u}_R(\mathbf{x},t)$. From (3.2), the simplest low-dimensional model is obtained by ignoring the residual term: i.e.

$$\mathbf{T} = \mathbf{0}. \tag{3.3}$$

This case shall be referred to as the *POD-Galerkin model*.

3.2. *Sparse POD-Galerkin regression*

Several recent studies have proposed to identify $\mathbf{L}$ and $\mathbf{N}$ directly from the data (Brunton *et al.* 2016; Loiseau & Brunton 2018). If the dimension of the given nonlinear oscillation is sufficiently low (e.g. two-dimensional laminar cylinder wake), $\mathbf{T}$ can be ignored with the use of a small number of POD modes. In this case, $\mathbf{L}$ and $\mathbf{N}$ can directly be obtained from the snapshots of $\mathbf{a}$ using a sparse regression technique (i.e. SINDy; e.g. Brunton *et al.* 2016). SINDy is formulated by first collecting a set of time snapshots of the POD amplitudes from DNS (e.g. $\mathbf{a}_{dns}(t)$) into the following data matrices:

$$\mathbf{X} = \begin{bmatrix}\mathbf{a}_{dns}(t_1) & \mathbf{a}_{dns}(t_2) & \ldots & \mathbf{a}_{dns}(t_m)\end{bmatrix}^T, \tag{3.4a}$$

and

$$\dot{\mathbf{X}} = \begin{bmatrix}\dot{\mathbf{a}}_{dns}(t_1) & \dot{\mathbf{a}}_{dns}(t_2) & \ldots & \dot{\mathbf{a}}_{dns}(t_m)\end{bmatrix}^T. \tag{3.4b}$$

A set of candidate library functions, $\Theta(\mathbf{X})$, is subsequently constructed. In the present study, we restrict the library functions to be

$$\Theta(\mathbf{X}) = \begin{bmatrix} \mathbb{P}_\mathbf{L}(\mathbf{X}) & \mathbb{P}_\mathbf{N}(\mathbf{X}) \end{bmatrix} = \begin{bmatrix} \mathbb{P}_\mathbf{L}(\mathbf{a}_{dns}(t_1)) & \mathbb{P}_\mathbf{N}(\mathbf{a}_{dns}(t_1)) \\ \mathbb{P}_\mathbf{L}(\mathbf{a}_{dns}(t_2)) & \mathbb{P}_\mathbf{N}(\mathbf{a}_{dns}(t_2)) \\ \vdots & \vdots \\ \mathbb{P}_\mathbf{L}(\mathbf{a}_{dns}(t_m)) & \mathbb{P}_\mathbf{N}(\mathbf{a}_{dns}(t_m)) \end{bmatrix}, \tag{3.5a}$$

where $\mathbb{P}_\mathbf{L}(\mathbf{X})$ and $\mathbb{P}_\mathbf{N}(\mathbf{X})$ denote the linear and quadratic combinations of the state vector admitted by the form of $\mathbf{L}$ and $\mathbf{N}$ in (3.2), respectively. We also introduce a coefficient matrix

$$\boldsymbol{\Xi} = \begin{bmatrix}\boldsymbol{\Xi}_\mathbf{L} & \boldsymbol{\Xi}_\mathbf{N}\end{bmatrix}^T, \tag{3.5b}$$

where $\boldsymbol{\Xi}_\mathbf{L}$ and $\boldsymbol{\Xi}_\mathbf{N}$ contain the (unknown) coefficients for $\mathbb{P}_\mathbf{L}(\mathbf{X})$ and $\mathbb{P}_\mathbf{N}(\mathbf{X})$. The coefficient matrix, $\boldsymbol{\Xi}$, is subsequently determined by solving the following convex least squares regression problem:

$$\min_{\boldsymbol{\Xi}} \|\dot{\mathbf{X}} - \Theta(\mathbf{X})\,\boldsymbol{\Xi}\|_2 + \gamma\|\boldsymbol{\Xi}\|_1, \tag{3.6}$$

where $\|\cdot\|_2$ and $\|\cdot\|_1$ denote the standard ℓ_2 and ℓ_1 norms, respectively, and γ is the penalty introduced for the sparsity promoting ℓ_1-regulariser. An advantage of using an ℓ_1-regularisation compared to an ℓ_2-regularisation is that they tend to prevent data overfit by promoting a model with the least complexity (sparse) required to model the dynamics (Brunton & Kutz 2019). The optimsation problem in (3.6) can be solved with the well-known LASSO (least absolute shrinkage and selection operator) algorithm. For large datasets, an alternative based on sequential threshold least squares was recommended instead – i.e. the SINDy approach (Brunton *et al.* 2016). This approach is used in the present study. We note that if the given dataset is even larger, the approach proposed by Gelß *et al.* (2019) may further be considered (i.e. multidimensional approximation of nonlinear dynamical systems (MANDy)). The sparse regression using a template given by the Galerkin projection will be referred to as the *POD-SINDy model*.

It is worth mentioning that the optimisation (3.6) was recently proposed to be solved with an equality constraint which explicitly enforces the energy conservation in the nonlinear operator $\mathbf{N}$ (Loiseau & Brunton 2018): i.e. $\mathbf{a}^H\mathbf{N}(\mathbf{a},\mathbf{a}) = 0$. However, this approach will not be considered in the present study, where the number of POD modes for the construction of a reduced-order model will not necessarily be large enough to fully cover the dimension of the given chaotic state. In such a case, the sparse regression in (3.6) implies that the residual term $\mathbf{T}$ in (3.2) is modelled as

$$\mathbf{T} = \mathbf{L}_{\Delta}\mathbf{a} + \mathbf{N}_{\Delta}(\mathbf{a}, \mathbf{a}), \tag{3.7}$$

where $\mathbf{L}_{\Delta}$ and $\mathbf{N}_{\Delta}$ are the difference in $\mathbf{L}$ and $\mathbf{N}$ obtained by the Galerkin projection and the sparse regression. Therefore, the energy-conservation constraint for the nonlinear operator proposed by Loiseau & Brunton (2018) would enforce $\mathbf{a}^H(\mathbf{N}(\mathbf{a},\mathbf{a})+\mathbf{N}_{\Delta}(\mathbf{a},\mathbf{a})) = 0$. However, this is not necessarily desirable in the present case, as will be discussed in the following subsection. Hence in §4.2, the library terms for the set of POD modes from the same wavenumber space, $\mathbf{a}_{n_x,n_z}$, are made dependent on the wavenumber of the POD modes as is suggestive from a Galerkin projection. The library function will consist of first-order polynomials of POD modes from the same wavenumber space. It attempts to model the $\mathbf{L}_{\Delta}$ terms of the dynamics. The modelling of the $\mathbf{N}_{\Delta}$ terms are accounted for by including the combination of POD modes that form a set of second-order polynomials which reflect the triadic wavenumber interactions, where $(n_x, n_z) = (k_x, k_z)+(m_x, m_z)$, from a POD-Galerkin model.

3.3. *The needs for an eddy-viscosity closure*

Now, we consider a physical model for $\mathbf{T}$. To rationalise the need of such a model, we first define $\mathbf{a}_\infty$ as the solution to (3.2) for infinitely large M_x, M_z and N_p. In this case, the residual $\mathbf{T}$ in (3.2) should vanish, thus $\mathbf{a}_\infty$ would be identical to those obtained by projecting $\mathbf{u}$ from DNS onto the POD modes. Given the energy-conserving nature of the nonlinear term in (2.1) (e.g. Joseph 1976), the contribution of the resulting nonlinear term to the change rate of the perturbation kinetic energy ($\mathbf{a}_\infty^H\mathbf{a}_\infty$) should be zero for every time instance, t, i.e.

$$\mathbf{a}_\infty^H \mathbf{N}_\infty(\mathbf{a}_\infty, \mathbf{a}_\infty) = 0, \tag{3.8}$$

where $\mathbf{N}_\infty$ is the quadratic nonlinear term obtained by considering infinitely large M_x, M_z and N_p. We note that (3.8) must also be true even if $\mathbf{a}_\infty$ is replaced by any arbitrary vector. This observation motivated Loiseau & Brunton (2018) to impose $\mathbf{a}^H\mathbf{N}(\mathbf{a},\mathbf{a}) = 0$ as an equality constraint into the optimisation problem in (3.6).

Let us now consider small values of M_x, M_z and N_p which define the size of the reduced-order model. In particular, we will assume that M_x, M_z and N_p are not large enough to cover the full energy cascade dynamics of the given turbulent state. We define a projection operator $\mathcal{P}_l$ onto the subspace defined by the small values of M_x, M_z and N_p. Then, $\mathbf{a}_\infty$ can be decomposed into $\mathbf{a}_\infty = \mathbf{a}_{\infty,l} + \mathbf{a}_{\infty,h}$, where $\mathbf{a}_{\infty,l} = \mathcal{P}_l[\mathbf{a}_\infty]$ and $\mathbf{a}_{\infty,h} = \mathcal{P}_h[\mathbf{a}_\infty]$ with $\mathcal{I}_0[\cdot] = \mathcal{P}_l[\cdot] + \mathcal{P}_h[\cdot]$ ($\mathcal{I}_0[\cdot]$ is the identity operator). Using this decomposition and the quadratic nature of $\mathbf{N}_\infty$, (3.8) can be written as

$$\begin{aligned} &\mathbf{a}_{\infty,l}^H\left[\mathbf{N}_\infty(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,h}) + \mathbf{N}_\infty(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,l}) + \mathbf{N}_\infty(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,h})\right] \\ &+ \mathbf{a}_{\infty,h}^H\left[\mathbf{N}_\infty(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,l}) + \mathbf{N}_\infty(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,h}) + \mathbf{N}_\infty(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,l})\right] = 0, \end{aligned} \tag{3.9}$$

where the top and bottom lines describe the nonlinear energy transport of the perturbation kinetic energy in the $\mathcal{P}_l$ and $\mathcal{P}_h$ subspaces, respectively. Here, we note that the term $\mathbf{a}_{\infty,l}^H\mathbf{N}_\infty(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,l})$, equivalent to $\mathbf{a}^H\mathbf{N}(\mathbf{a},\mathbf{a})$ expected from (3.2), vanishes due to (3.8), and so does $\mathbf{a}_{\infty,h}^H\mathbf{N}_\infty(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,h})$ (note that the only difference between $\mathbf{a}_{\infty,l}^H\mathbf{N}_\infty(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,l})$ and $\mathbf{a}^H\mathbf{N}(\mathbf{a},\mathbf{a})$ are their dimension). The top line in (3.9) should indicate the rate of perturbation energy transferred from the $\mathcal{P}_l$ to the $\mathcal{P}_h$ subspace. Importantly, the energy cascade from large to small scales in the three-dimensional Navier-Stokes equations implies

$$\left\langle \mathbf{a}_{\infty,l}^H\left[\mathbf{N}_\infty(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,h}) + \mathbf{N}_\infty(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,l}) + \mathbf{N}_\infty(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,h})\right]\right\rangle < 0 \tag{3.10a}$$

and

$$\left\langle \mathbf{a}_{\infty,h}^{H} \left[\mathbf{N}_{\infty}(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,h}) + \mathbf{N}_{\infty}(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,l}) + \mathbf{N}_{\infty}(\mathbf{a}_{\infty,h}, \mathbf{a}_{\infty,h}) \right] \right\rangle > 0 \tag{3.10b}$$

for a statistically stationary flow where the ensemble average is equivalent to a time average. Indeed, this has been observed in a number of previous studies where the inter-scale energy transfer is analysed in detail (e.g. Cho *et al.* 2018; Lee & Moser 2019; Hwang & Lee 2020; Doohan *et al.* 2021*b*). Given the equivalence of $\mathbf{a}_{\infty,l}^{H}\mathbf{N}_{\infty}(\mathbf{a}_{\infty,l}, \mathbf{a}_{\infty,l})$ to $\mathbf{a}^{H}\mathbf{N}(\mathbf{a}, \mathbf{a})$, a condition for $\mathbf{T}$ to meet from a physical viewpoint would then be

$$\left\langle \mathbf{a}^{H}\mathbf{T} \right\rangle < 0, \tag{3.11}$$

indicating that the residual term, $\mathbf{T}$, in (3.2) must contain an energy-removal mechanism.

The discussion above evidently justifies the use of an eddy viscosity model in many previous studies (e.g. Aubry *et al.* 1988; Rempfer & Fasel 1994*a*,*b*; Smith *et al.* 2005; Östh *et al.* 2014; Protas *et al.* 2015): i.e.

$$\mathbf{T} = \nu_t \mathbf{D}\mathbf{a}, \tag{3.12}$$

where ν_t is a scalar-valued eddy viscosity and $\mathbf{D}$ is the Laplacian operator for the reduced-order model defined in Appendix A. For $\nu_t > 0$, (3.12) yields a sufficient condition for (3.11), as it ensures $\mathbf{a}^{H}\mathbf{T}(= \nu_t \mathbf{a}^{H}\mathbf{D})\mathbf{a} < \mathbf{0}$ at every time for non-zero $\mathbf{a}$ with $(n_x, n_z) \neq (0, 0)$ due to the negative semi-definite nature of $\mathbf{D}$ (see Appendix A). In the previous studies, various form of eddy viscosity have been introduced and examined. In the present study, we will first consider a simple spectral eddy viscosity model similar to the one in Smith *et al.* (2005) where an empirical real-valued constant of ν_t is employed. It is evident that the performance of this simple single-valued eddy viscosity model is expected to be limited, as there is evidence supporting the complex nature of the inter-modal energy transfer (e.g. Couplet *et al.* 2003; Podvin 2009, see also §4.3). Therefore, this approach here is considered for the purpose of comparing with the other models. This approach shall be referred to as *POD-Galerkin-E model* (the '*E*' stands for an 'empirical' eddy viscosity).

3.4. *Sparse optimal eddy-viscosity closure*

Although the eddy viscosity model in (3.12) ensures the physical property that originates from the energy cascade of turbulent state, (3.11), it is evidently too crude. Indeed, $\mathbf{a}^{H}\mathbf{T}$ does not have to be negative for every point in time like the one ensured by (3.12) – only its time average needs to be negative. Furthermore, there is no physical reason that different POD modes should feel the 'same' eddy viscosity: for example, the POD modes for large n_x, n_z and n are not expected to experience a large amount of energy removal from their dynamics by $\mathbf{T}$. More flexible forms of eddy viscosity have therefore been proposed previously (for a review, see Östh *et al.* 2014, where various forms of eddy viscosity have been examined for a three-dimensional turbulent bluff-body wake). In particular, Protas *et al.* (2015) introduced the concept of 'optimal' eddy viscosity by formulating an adjoint-based optimisation problem that minimises the difference between the POD amplitudes from the measurement and the reduced-order model. However, as discussed in §1, the application of the adjoint-based optimisation to a turbulent flow does not always allow for a sufficient long optimisation time interval due to the unstable nature of the adjoint system around a chaotic state. Also, the approach of Protas *et al.* (2015) is still based on a scalar-valued eddy viscosity, although its generalisation to a sophisticated form of eddy viscosity is easily possible.

In the present study, we take an alternative formulation which enables us to consider a long time horizon for a similar optimisation problem. In particular, we determine an optimal eddy viscosity with the sparse regression in §3.2. For the demonstrative purpose, we consider a nonlinear closure model for $\mathbf{T}$ (Östh *et al.* 2014; Protas *et al.* 2015), in which a complex-valued eddy viscosity is set to vary with POD modes:

$$\mathbf{T} = \mathbf{V}_t\mathbf{D}\mathbf{a} \tag{3.13a}$$

with

$$\mathbf{V}_t(t) = e(t)diag[\mathbf{c}], \tag{3.13b}$$

where $e(t)$ is a scalar-valued function of $\mathbf{a}$ that can be chosen for a nonlinear eddy viscosity model and $\mathbf{c}$ the constant vector to be determined for each n_x, n_z and n. We note that the choice of e usually ensures $\mathbf{V}_t$ to vanish as $\mathbf{a}$ becomes zero. In previous studies (e.g. Östh *et al.* 2014; Protas *et al.* 2015), the perturbation kinetic energy, $e(t) = (\mathbf{a}^H\mathbf{a})^{1/2}$, has often been considered. In the present study, a simpler form, $e(t) = |a_{0,0}^{(1)}|$, is chosen, given that the corresponding mode contains approximately 70% of the total perturbation kinetic energy (see table 1). A preliminary test also reveals that this choice makes the optimised reduced-order model perform slightly better than that of $e(t) = (\mathbf{a}^H\mathbf{a})^{1/2}$ (see Appendix B).

Now, we formulate an optimisation problem determining $\mathbf{V}_t$ in (3.13*b*). Using the POD mode amplitudes obtained from DNS (i.e. $\mathbf{a}_{dns}$) and (3.2), the desired residual term of the low-dimensional system is given by

$$\mathbf{T}_{dns}(t) \equiv \dot{\mathbf{a}}_{dns}(t) - \mathbf{L}\mathbf{a}_{dns} - \mathbf{N}(\mathbf{a}_{dns}, \mathbf{a}_{dns}), \tag{3.14}$$

where $\mathbf{T}_{dns}$ indicates the residual term calculated with $\mathbf{a}_{dns}$. Similarly to the sparse regression in §3.2, a set of time snapshots of $\mathbf{T}_{dns}(t)$ is introduced into a data matrix using (3.14):

$$\mathbf{Y} = \begin{bmatrix}\mathbf{T}_{dns}(t_1) & \mathbf{T}_{dns}(t_2) & \dots & \mathbf{T}_{dns}(t_m)\end{bmatrix}^T. \tag{3.15a}$$

Given (3.13*a*), the related library function for the optimisation is subsequently formed to be

$$\Theta_e(\mathbf{Y}) = \begin{bmatrix}e_{dns}(t_1)\mathbf{a}_{dns}(t_1)\mathbf{D} & e_{dns}(t_2)\mathbf{a}_{dns}(t_2)\mathbf{D} & \dots & e_{dns}(t_m)\mathbf{a}_{dns}(t_m)\mathbf{D}\end{bmatrix}^T, \tag{3.15b}$$

where $e_{dns}(t)$ is e obtained from DNS. Since we seek a complex-valued constant vector $\mathbf{c}$ that minimises the difference between $\mathbf{T}_{dns}$ given by (3.14) and $\mathbf{T}$ from the residual model in (3.13*a*), the following optimisation is defined:

$$\min_{\mathbf{c}} \|\mathbf{Y} - \Theta_e(\mathbf{Y})\, diag[\mathbf{c}]\|_2 + \gamma_e\|\mathbf{c}\|_1, \tag{3.16}$$

where γ_e is the parameter for the sparsity-promoting ℓ_1-regulariser. Like the optimisation problem in (3.6), (3.16) is solved by applying the SINDy approach (Brunton *et al.* 2016). This approach will be referred to as *POD-Galerkin-R model* (the '*R*' stands the determination of an eddy viscosity with a sparse 'regression'). As mentioned in §1, several previous studies (e.g. Couplet *et al.* 2005; Cordier *et al.* 2010) proposed a similar idea of calibrating the residual term, $\mathbf{T}$, using Tikhonov regularisation, an optimisation based on the ℓ_2-regularisation. Here, the ℓ_1-regularisation has a benefit over the ℓ_2-regularisation, as it is designed to prevent data overfit, offering a parsimonious low-dimensional model. Lastly, we note that setting $\gamma_e \to \infty$ results in the POD-Galerkin model, while $\gamma_e \to 0$ yields the least-square eddy viscosity (for a further discussion on the effect of γ_e, see Appendix C).

It is also useful to make some remarks on the least-square sparse regression in (3.16). First, the regression (3.16) can now consider a very long sampling time, as it simply relies on the POD mode amplitudes, $\mathbf{a}_{dns}$, taken from DNS. Therefore, it no longer suffers from the finite optimisation time-interval issue that one may face in the conventional adjoint-based formulation. Second, the regression problem (3.16) can flexibly be formulated by accounting for various forms of eddy viscosity, and this can be achieved by adding more library functions in (3.15*b*). Third, the regression (3.16) allows for negative elements of $\mathbf{V}_t$ (or $\mathbf{c}$), indicating that the 'backwards scattering' in the energy cascade can be taken into account. Although

Model	Physics	Model for $\mathbf{T}$	Technique for $\mathbf{T}$
POD-Galerkin	-	$\mathbf{0}$	-
POD-SINDy	-	$\mathbf{L}_\Delta \mathbf{a} + \mathbf{N}_\Delta(\mathbf{a}, \mathbf{a})$	Sparse Regression
POD-Galerkin-E	Energy Cascade	$\nu_t \mathbf{D}\mathbf{a}$	Empirical
POD-Galerkin-R	Energy Cascade	$\mathbf{V}_t^T \mathbf{D}\mathbf{a}$	Sparse Regression

Table 2: Summary of the reduced-order models.

this issue may cause a potential numerical instability of the resulting reduced-order model, it may well be fixed by imposing an additional inequality constraint (e.g. all the element of $\mathbf{c}$ is greater than or equal to zero). Finally, $\mathbf{V}_t$ (or $\mathbf{c}$) in (3.13*b*) can also be a complex vector, given that $\mathbf{a}$ is complex. In other words, some 'dispersive' effect in the dynamics of the reduced-order model can also be added with non-zero imaginary part of $\mathbf{V}_t$.

4. Results and discussions

We examine the reduced-order models obtained by applying the approaches in §3 to the PCF introduced in §2. In table 2, the extent that each model utilises physical information and data from DNS is summarised with the form of the closure. The approach relying on the data most is the POD-SINDy model, as it determines all the elements of the operator $\mathbf{L}_\Delta$ and $\mathbf{N}_\Delta$ by solving the regression problem in (3.6). On the other hand, the approach accounting for both physics and data to the largest extent would be the POD-Galerkin-R model, as a physics-informed and flexible form of $\mathbf{T}$ is determined in a data-driven manner.

4.1. *Dimension of the reduced-order model: POD-Galerkin model*

The POD-Galerkin model introduced in §3.1 is first studied to determine a few reference reduced-order models. For simplicity, only the most energetic POD mode is taken for each n_x and n_z, and the dimension of the reduced-order model is varied with the number of the plane Fourier modes. Three different cases are considered: i) 6-modes with $\mathbf{a} = [a^{(1)}_{0,0}, a^{(1)}_{0,\pm1}, a^{(1)}_{0,\pm2}, a^{(1)}_{\pm1,0}, a^{(1)}_{\pm1,\pm1}, a^{(1)}_{\pm1,\mp1}]$ examined in Smith *et al.* (2005); ii) 25-modes with $\mathbf{n}(\equiv [N_p, M_x, M_z]) = [1, 3, 3]$ (see (3.1)); iii) 41-modes with $\mathbf{n} = [1, 4, 4]$. Figure 3 shows the time trace of the Fourier-mode energy, $M(n_x, n_z)$, for the three cases. In all cases, the initial condition for $\mathbf{a}$ is obtained by projecting a DNS field onto the corresponding POD mode subspace. For the 6-mode case, the state vector, $\mathbf{a}$, reaches a non-trivial equilibrium state after an oscillatory transient, consistent with the result of Smith *et al.* (2005) (figure 3a). Given that the 6-mode model exhibits a stable non-trivial equilibrium (i.e. $\mathbf{a} \neq \mathbf{0}$), this would not be a good reference case to build a low-dimensional model exhibiting chaotic dynamics. Considering a larger number of plane Fourier modes (i.e. 25- and 41-modes), the state vector, $\mathbf{a}$, of the reduced-order model exhibits a chaotic trajectory (figures 3b,c). However, the values of $M(n_x, n_z)$ appear to be far off from those in DNS. In particular, the second most energetic mode of the POD-Galerkin model, $M(0, 1)$, which would represent the time evolution of streaks, significantly deviates from DNS. Despite this issue, the presence of the chaotic dynamics in the POD-Galerkin models with 25- and 41-modes indicates that they would be good reference cases which the other modelling approaches in §3 can further be employed. Therefore, the remaining part of the present study will only consider the 25- and 41-mode cases.

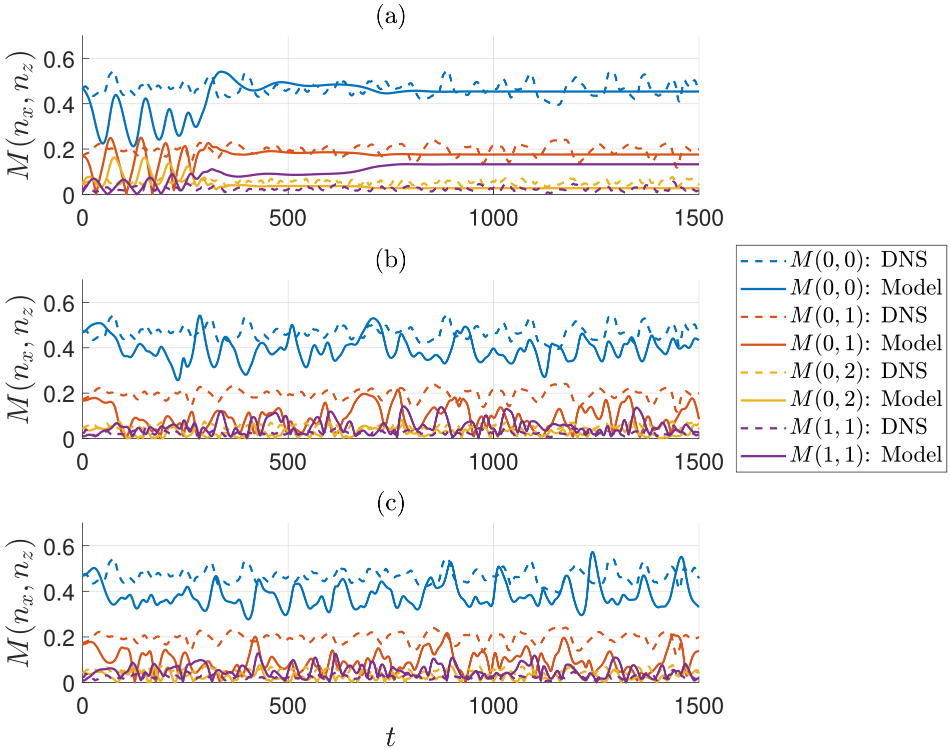


Figure 3: Time trace of $M(n_x, n_z)$ of the POD-Galerkin model: (a) 6 modes (Smith *et al.* 2005); (b) 25 modes; (c) 41 modes.

4.2. *Sparse POD-Galerkin regression: POD-SINDy model*

Based on the 25- and 41-mode cases in §4.1, the optimisation problem in (3.6) is solved for $t \in [-10000, 0]$ with an sampling interval of $\Delta t = 0.5$, in order to obtain the corresponding POD-SINDy model. The model constructed is subsequently examined by considering an additional time interval $t \in [0, 5000]$. The relative error of the least-squares regression defined in (3.6) is reported in table 3 for the 25-mode model. Here, the relative error is defined as

$$\mathcal{E}_{PS} = \frac{1}{N_{\text{mode}}} \sum_{n_x=-M_x}^{M_x} \sum_{n_z=-M_z}^{M_z} \sum_{n=1}^{N_p} \frac{\left\| \left[\dot{\mathbf{X}} - \Theta(\mathbf{X})\Xi\right]_{n_x,n_z}^{(n)} \right\|_2}{\left\| \left[\dot{\mathbf{X}}\right]_{n_x,n_z}^{(n)} \right\|_2}, \tag{4.1}$$

where $[\cdot]_{n_x,n_z}^{(n)}$ indicates the component defined by the POD mode indices given in (2.4*a*). N_{mode} is defined as the number of POD modes used in the sparse regression. We note that the relative error is normalised by N_{mode} such that $\mathcal{E}_{PS} \in [0, 1]$. As expected, the relative error reaches the minimum when γ is zero (least-squares regression). When γ is increased, the relative error becomes larger due to the increased sparsification penalty as defined in equation (3.6).

The 6-mode POD-SINDy model obtained and the 25-mode POD-SINDy model obtained in this way is subsequently simulated. It is found that the POD-SINDy model rapidly blows up for $t \in [0, 4]$, and this behaviour remains unchanged for relatively low values of $\gamma \in [0, 0.1]$. For $\gamma = 1$, only $M(0, 2)$ blows up while the rest remains relatively stable, although all the

γ	0	0.0001	0.001	0.01	0.1	1	10
$\mathcal{E}_{PS}$ (training)	0.69	0.69	0.69	0.69	0.75	0.98	0.99
$\mathcal{E}_{PS}$ (validation)	0.68	0.68	0.68	0.68	0.73	0.98	0.99
N_0 (**L**)	0	0	0	0	31	49	49
N_0 (**N**)	0	0	0	2	78	534	664

Table 3: Relative error of the least-squares regression in (3.6) for the POD-SINDy model with 25-modes. Here, $\mathcal{E}_{PS}$ is defined in (4.1), which is the summation of the the relative error of each mode when compared to DNS data, normalised by the total number of POD modes. The 'training' and 'validation' in second and third lines imply $\mathcal{E}_{PS}$ from $t \in [-10000, 0]$ and $t \in [0, 5000]$, respectively. Also, $N_0(\cdot)$ indicates the number of zero terms in the linear operator **L** and in the nonlinear operator **N**, except the Reynolds shear-stress term in the mean equation, in the model. Note that the number of the POD modes used for the 25-mode POD-SINDy model is $N_{\mathrm{mode}} = 49$ due to the conjugate symmetry (i.e. 1 mode for the mean equation and 2×24 modes for the fluctuation equations.

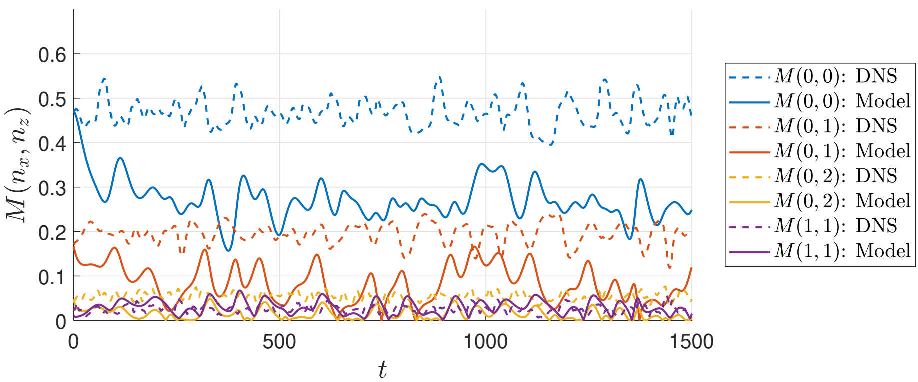


Figure 4: Time trace of $M(n_x, n_z)$ from DNS and the POD-Galerkin-E model with 25 modes ($\nu_t = 0.003$).

linear terms are zero in this case (see table 3). For $\gamma = 10$, all the terms in the regression from the GP-template are zero, except for the nonlinear terms for the equation for $\dot{a}_{0,0}^{(1)}$. The application of this approach to the 41-mode case also exhibits a similar behaviour, thus it is not pursued any more. The blow-up of the POD-SINDy model was not reported in previous studies (e.g. Brunton *et al.* 2016; Loiseau & Brunton 2018; Rubini *et al.* 2021). However, it is worth mentioning that the dimension of nonlinear oscillations in such cases are low (e.g. Lorenz chaos, two-dimensional laminar vortex shedding, and two-dimensional cavity flows). As such, the number of POD modes considered in those studies appear to be large enough to cover the full nonlinear dynamics using sparse regression. This suggests that the blow-up of the POD-SINDy model shown here are presumably caused by the number of POD modes that is not large enough to cover the full chaotic dynamics of interest. As discussed in §3.2, in this case, the POD-SINDy model takes the residual term **T** in the form of (3.7), which does not necessarily ensure (3.11).

γ_e	0.0003	0.0005	0.001	0.005	0.1
$\mathcal{E}_{PGR}$ (training)	0.93	0.93	0.96	1	1
$\mathcal{E}_{PGR}$ (validation)	0.92	0.92	0.95	1	1
N_0	11	15	35	49	49

Table 4: Relative error of the least-squares regression in (3.16) for the POD-Galerkin-R model with 25 modes. Here, $\mathcal{E}_{PGR}$ is defined in (4.2). The 'training' and 'validation' in second and third lines imply $\mathcal{E}_{PS}$ from $t \in [-10000, 0]$ and $t \in [0, 5000]$, respectively. As in Table 3, N_0 indicates the number of zero valued eddy-viscosity terms for $N_{\text{mode}} = 49$.

4.3. *Utilisation of an empirical eddy viscosity: POD-Galerkin-E model*

The POD-Galerkin-E model, which utilises the simple spectral eddy viscosity of Smith *et al.* (2005), is examined by considering a few values of ν_t defined in (3.12): $\nu_t = (0.001, 0.003, 0.005)$. It is, however, found that the introduction of such a simple eddy viscosity closure does not significantly improve the accuracy of the reduced-order model compared to the original POD-Galerkin model. Indeed, $\nu_t = 0.001$ is found to be too small to influence the original POD-Galerkin model, while $\nu_t = 0.005$ is too large and stabilises the chaotic dynamics into a stable non-trivial equilibrium. Here, we present the results for $\nu_t = 0.003$, which was determined by accounting for this observation like Smith *et al.* (2005). Figure 4 compares the time trace of $M(n_x, n_z)$ from DNS with that from the POD-Galerkin-E model with 25-modes for $\nu_t = 0.003$. The mean component, $M(0, 0)$, which contains the largest perturbation energy, exhibits a large difference from that of DNS. In fact, this deviation is greater than that of the POD-Galerkin model which does not employ any model of $\mathbf{T}$ (compare figure 4 with figure 3b). However, it should be mentioned that not all of the mode amplitudes exhibit such a deterioration. The POD-Galerkin-E model is also found to exhibit a much more improved $M(1, 1)$ (compare figure 4 with figure 3b; see also table 5), indicating that the utilisation of a suitable eddy viscosity closure would improve the performance of a reduced-order model. This will be seen in §4.5.

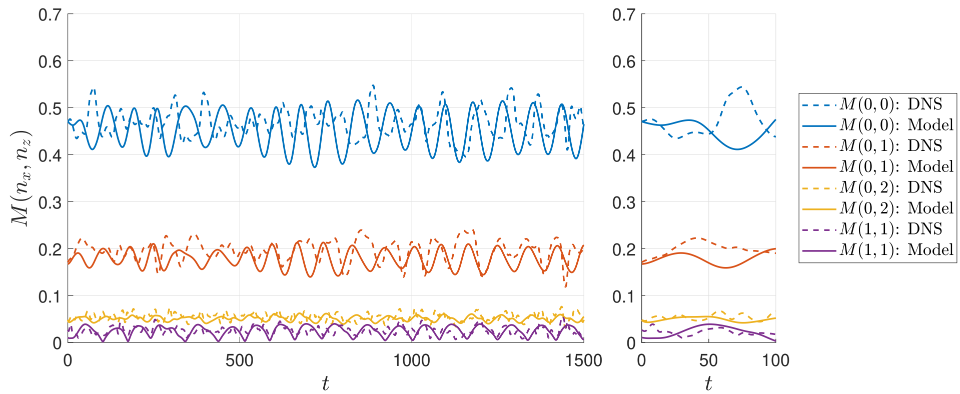
4.4. *Sparse optimal closure: POD-Galerkin-R model*

We now consider the POD-Galerkin-R model where a flexible form of eddy viscosity is determined by solving the least-squares regression problem in (3.16). The regression is performed with the data taken for $t \in [-10000, 0]$ with an sampling interval of $\Delta t = 0.5$, and the model is subsequently examined by considering an additional time interval $t \in [0, 5000]$. Similarly to (4.2), the relative error of the regression is defined as

$$\mathcal{E}_{PGR} = \frac{1}{N_{\text{mode}} - 1} \sum_{n_x=-M_x}^{M_x} \sum_{n_z=-M_z}^{M_z} \sum_{n=1}^{N_p} \frac{\left\| [\mathbf{Y} - \Theta_e(\mathbf{Y}) diag[\mathbf{c}]]^{(n)}_{n_x,n_z} \right\|_2}{\left\| [\mathbf{Y}]^{(n)}_{n_x,n_z} \right\|_2} \tag{4.2}$$

for $(n_x, n_z) \neq (0, 0)$, and it is reported in table 4 for the 25-mode model. Note that the mode with $(n_x, n_z) = (0, 0)$ is excluded from the relative error statistics as no residual term was applied to this mode. As expected, the relative error becomes larger as the ℓ_1-regularisation penalty, γ_e, increases. For $\gamma_e > 0.005$, $\mathcal{E}_{PGR}$ remains unchanged, indicating that the regression would not make any improvement. In the present study, we have chosen to present the result for $\gamma_e = 0.0005$ which renders the proposed regression sufficiently effective, while not allowing for too small values in $\mathbf{c}$ that could well be from some numerical issues (e.g. sampling time interval). We also ensure that our residual model does not overfit

(a)



(b)

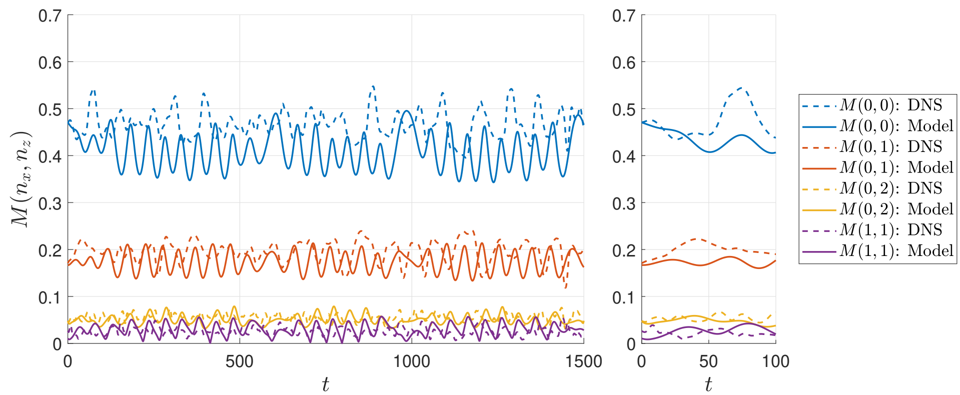


Figure 5: Time trace of $M(n_x, n_z)$ from DNS and the POD-Galerkin-R model: (a) 25-modes ($\gamma_e = 0.0005$); (b) 41-modes ($\gamma_e = 0.0001$). The misalignment of the initial condition at $t = 0$ for $M(1,1)$ between the model and DNS data are due to the residual flow-field not being capture by the POD basis.

our training dataset by comparing the training and validation set error quantitatively as seen in table 4.

Figure 5 shows time trace of $M(n_x, n_z)$ from the POD-Galerkin-R model utilising 25- and 41-modes, and it is compared with that from DNS. The time traces of $M(n_x, n_z)$ from the POD-Galerkin-R model are now quite close to those from DNS, including the initial time evolution of the three most energetic modes ($M(0,0)$, $M(0,1)$ and $M(0,2)$ for $t < 50$) – note that the Lyapunov time would be expected to be at the order of the smallest time scale of the flow (i.e. the Kolmogorov time scale), which is given by $t \sim O(10)$ in the present case (e.g. Ruelle 1979; Crisanti *et al.* 1993). This indicates that the performance of this model is evidently far superior to that of the POD-Galerkin and the POD-Galerkin-E models. Both of the 25- and 41-mode cases of the POD-Galekrin-R model also exhibit a chaotic oscillation with the time scale close to that of DNS, observed in figure 8, and which will be discussed in §4.5.

The constant vector $\mathbf{c}(=c^{(n)}_{n_x,n_z})$ used for the eddy viscosity in (3.13*b*) is also visualised in figure 6. We first consider the 25-mode case (figure 6a). Here, we note that the constant

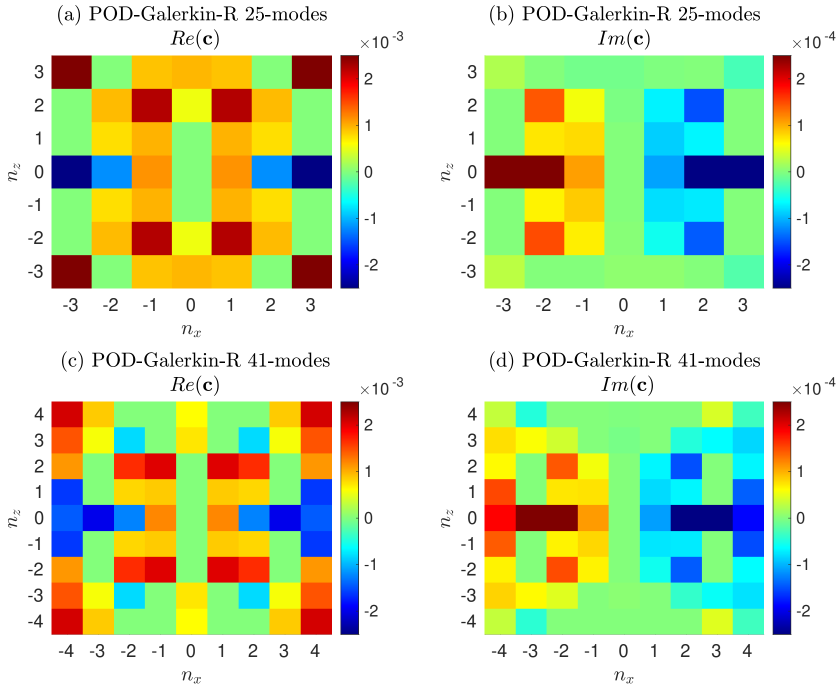


Figure 6: Distribution of real (left column) and imaginary (right column) part of $c^{(n)}_{n_x,n_z}$ in the $n_x - n_z$ plane : (a,b) 25-modes ($\gamma_e = 0.0005$); (c,d) 41-modes ($\gamma_e = 0.0001$).

vector, $\mathbf{c}$, exhibits a conjugate symmetry because the velocity in this study is real-valued. It also appears to be distributed highly symmetrically in the n_x-n_z plane. This originates from the $\mathcal{R}$ symmetry in (2.6) – the $\mathcal{R}$ symmetry imposes the streamwise and wall-normal components of $\boldsymbol{\phi}^{(n)}_{n_x,n_z}$ and $\boldsymbol{\phi}^{(n)}_{n_x,-n_z}$ to be identical and their spanwise component to have the opposite signs, while the energy of the streamwise component dominates over the other two components (see also figure 7). The real part of $\mathbf{c}$, proportional to the strength of dissipation (i.e. the removal of the energy from the given dynamics), is positive for most pairs of (n_x, n_z) in the $n_x - n_z$ plane, as expected from (3.11). However, there are some $c^{(n)}_{n_x,n_z}$ which exhibit negative values, and this is particularly pronounced around $(n_x, n_z) = (\pm 3, 0)$. This indicates that allowing for negative elements of $\mathbf{c}$ does help to improve the performance of the reduced-order model as discussed in §3.4. It is interesting to note that the higher streamwise and spanwise independent modes are energised (inverse energy transfer) and dissipated (forward energy cascade) respectively, consistent with the findings from Podvin (2009). Furthermore, some $c^{(n)}_{n_x,n_z}$ also exhibit non-zero imaginary values. However, their amplitudes are overall an order-of-magnitude smaller than those of the real counter part, implying that the dispersive effect imposed by the imaginary part of $c^{(n)}_{n_x,n_z}$ would not be significant. Finally, the distribution of $c^{(n)}_{n_x,n_z}$ for the 41-mode case (figures 6c,d) shows that increasing the number of POD modes do not significantly change the overall distribution and strength of $c^{(n)}_{n_x,n_z}$ in the $n_x - n_z$ plane (compare figures 6a,b with 6c,d). We note that as the number of POD modes is increased, the eddy viscosity also needs to vanish. However, this is not observed in figure 6. Instead, the similar distribution of $c^{(n)}_{n_x,n_z}$ for the 25- and 41-mode cases suggests that the eddy viscosity obtained with (3.16) is coherently compensating for

Cases	# of POD modes	$\langle\lvert a^{(1)}_{0,0}\rvert\rangle$	$\langle\lvert a^{(1)}_{0,1}\rvert\rangle$	$\langle\lvert a^{(1)}_{0,2}\rvert\rangle$	$\langle\lvert a^{(1)}_{1,1}\rvert\rangle$
DNS	N/A	2.110	0.881	0.230	0.079
POD-Galerkin	25	1.831 (13%)	0.453 (49%)	0.126 (45%)	0.227 (187%)
	41	1.772 (16%)	0.421 (52%)	0.127 (45%)	0.239 (203%)
POD-Galerkin-E	25	1.239 (41%)	0.389 (56%)	0.072 (69%)	0.132 (67%)
	41	1.268 (40%)	0.210 (76%)	0.055 (76%)	0.085 (8%)
POD-Galerkin-R	25	2.055 (5%)	0.809 (8%)	0.231 (0.4%)	0.111 (41%)
	41	1.820 (14.5%)	0.744 (16%)	0.235 (2%)	0.144 (82%)

Table 5: Time-averaged amplitudes of POD modes from DNS and the reduced-order models. Here, the numbers in the parenthesis indicate the relative error to the values from DNS.

some physical processes which are not simply resolved by the increase in the number of plane Fourier modes (i.e. M_x and M_z). In this respect, it is worth reminding that the reduced-order models in the present study only consider the leading POD mode for each (n_x, n_z). It is therefore presumable that the compensation made by the eddy viscosity model is associated with the lack of the higher-order POD modes for each (n_x, n_z).

4.5. *Comparison of the reduced-order models*

Having examined all of the reduced-order models introduced in §3, their performance is compared in this subsection. Table 5 shows the time-averaged amplitudes of the four leading POD modes from DNS and the reduced-order models, except the POD-SINDy model whose solution was found to blow up (see §4.2). We note that the four POD modes contain approximately 95% of total perturbation kinetic energy (see table 1). The POD-Galerkin model performs sensibly only for the mean component, $a^{(1)}_{0,0}$, while the rest of the components with $(n_x, n_z) \neq (0, 0)$ exhibit considerable errors ranging from 50% to 200%. The addition of an empirical eddy viscosity does not improve the POD-Galerkin model greatly (i.e. POD-Galerkin-E model), since the model still shows errors of 50%-80% across all the four leading POD modes. This model may also be viewed to perform most poorly, given the largest errors for $a^{(1)}_{0,0}$ that contains the largest amount of perturbation energy. Finally, the POD-Galerkin-R model shows the best performance and it has only a maximum 16% error for the first three leading POD modes. Although this model still shows a relatively large error for $a^{(1)}_{1,1}$, the energy contained by this mode in DNS is only about 1% (see table 1). Therefore, this error would be relatively insignificant.

The mean and turbulent velocity fluctuations from DNS and the reduced-order models are compared in figure 7. As expected from table 5, the mean velocity from the POD-Galerkin model and the POD-Galerkin-R model shows the best agreement with that from DNS. However, the POD-Galerkin model exhibits large differences in the velocity fluctuation profiles, while the statistics from the POD-Galerkin-E model are overall damped. A closer inspection reveals that the cross-stream and wall-normal turbulence fluctuations are predicted better by the 41-mode model, with a slightly poorer prediction in the streamwise mean velocity as compared to the 25 mode model. In any case, the level of agreement of the POD-Galerkin-R models in turbulence statistics has not been observed in any of the previous reduced-order models in plane Couette flow (e.g. Smith *et al.* 2005; Cavalieri 2021).

Next, to assess the dynamical behaviour of the leading POD modes from the reduced-order

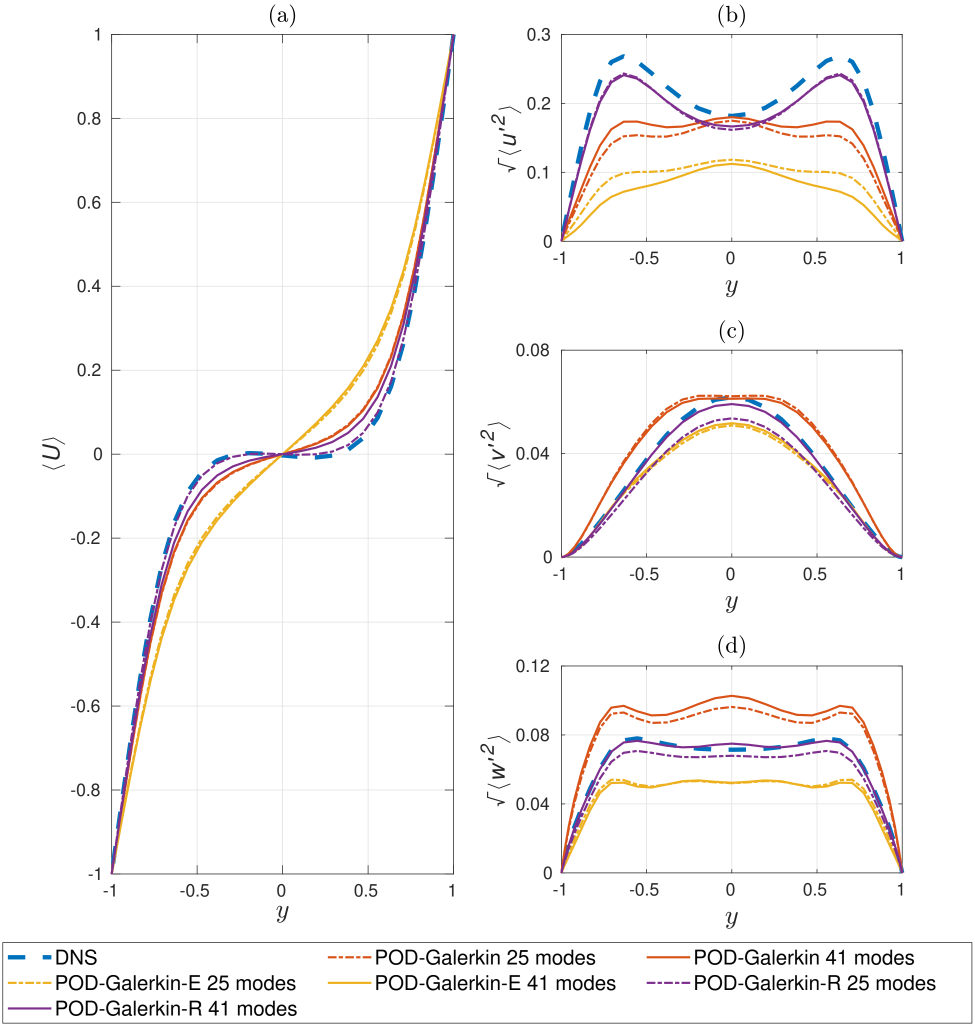


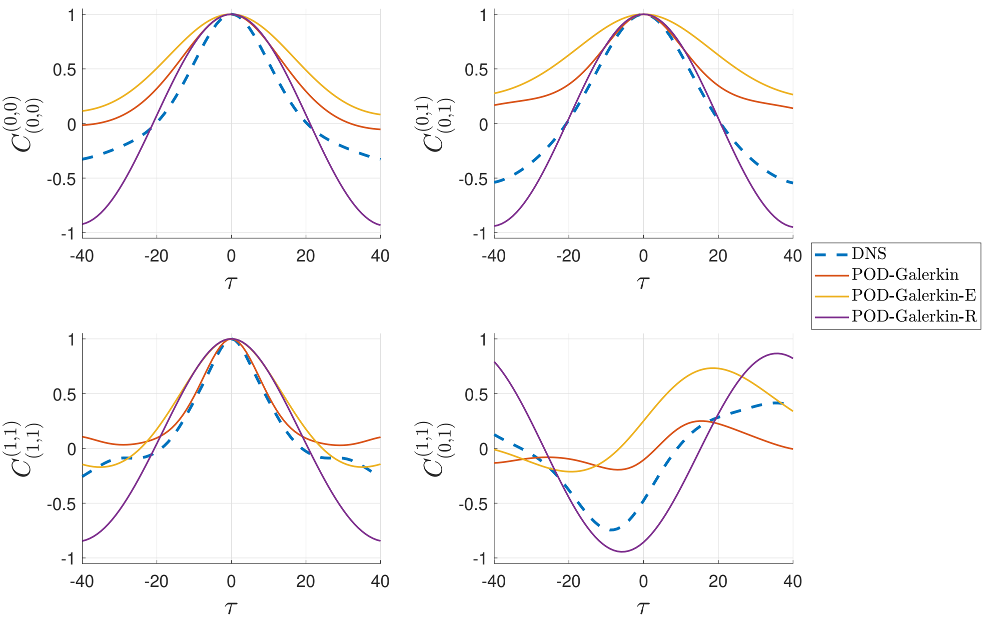
Figure 7: Turbulence statistics from DNS and the reduced-order models: (a) streamwise mean velocity; (b,c,d) root-mean-squared velocity fluctuations.

models, the following temporal auto- and cross-correlations of the main observables defined in (2.2) are computed:

$$C^{(n_x,n_z)}_{(m_x,m_z)}(\tau) = \frac{\langle \widetilde{M}(t+\tau;n_x,n_z)\widetilde{M}(t;m_x,m_z)\rangle}{\sqrt{\widetilde{M}^2(t;n_x,n_z)}\sqrt{\widetilde{M}^2(t;m_x,m_z)}}, \tag{4.3}$$

where $\widetilde{M}(t;n_x,n_z) = M(t;n_x,n_z) - \langle M(t;n_x,n_z)\rangle$. Figure 8 compares the correlation functions of the 25-mode and 41-mode models considered with those of DNS. In general, for the temporal correlations of $C^{(0,1)}_{(0,1)}$, the POD-Galerkin-R model have a closer match to the DNS data when compared to POD-Galerkin and POD-Galerkin-E models. The inclusion of more POD modes have the effect of improving the correlations of $C^{(1,1)}_{(1,1)}$ especially for $\tau = [-15, 15]$. By doing so, we observe a notable improvement of the $C^{(1,1)}_{(0,1)}$ correlations, an

(a)



(b)

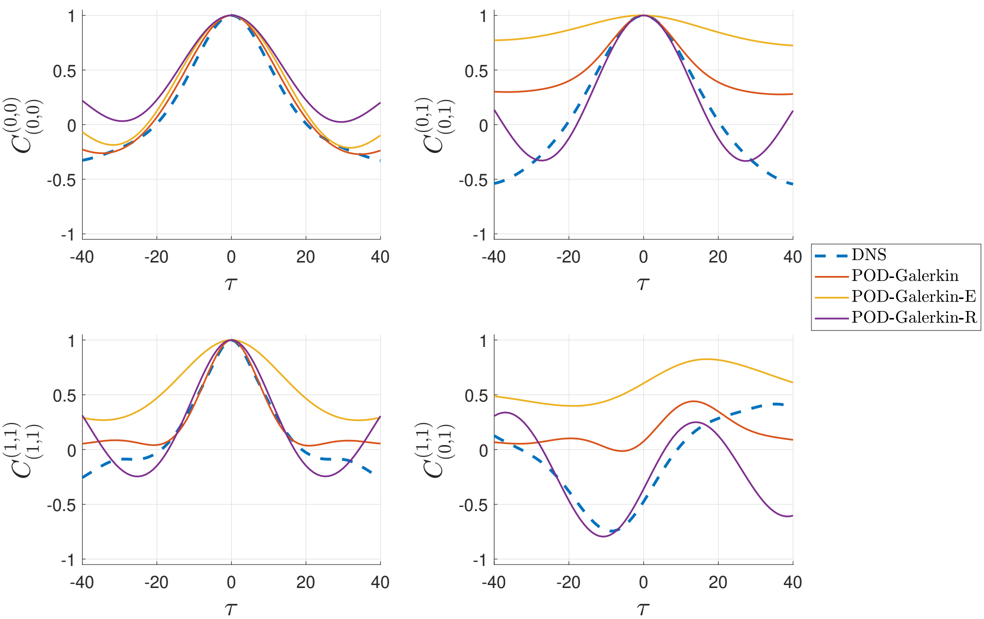


Figure 8: Temporal auto- and cross-correlations of: (a) 25-modes (γ_e = 0.0005 for the POD-Galerkin-R model); (b) 41-modes (γ_e = 0.0001 for the POD-Galerkin-R model).

important behaviour expected from the self-sustaining process, as it captures the breakdown of a streak structure and the regeneration of the streamwise vortices. Neither the POD-Galerkin or the POD-Galerkin-E model was able to closely replicate this behaviour with the inclusion of more POD modes.

Finally, figure 9 shows a time trace of the observables defined in (2.2) and a set of flow-field snapshots visualising a self-sustaining process generated by the POD-Galerkin-R 41-mode model. The strong streaky motions are apparent in $t = 1840 - 1857$, shown as a peak in $M(0,1)$ in figure 9b. The streaks breakdown into a wavy-behaviour from $t = 1874 - 1891$, accompanied by a decrease in $M(0,1)$ and an increase in $M(1,0)$ in figure 9b. The streaks breakdown completely from $t = 1909 - 1926$ while the quasi-streamwise vortices are regenerated, leading to an increase in $M(1,1)$. Finally, the quasi-streamwise vortices feed energy to the streaks from $t = 1943 - 1960$, known as the 'lift-up' effect. We note that the self-sustaining process from the POD-Galerkin-R 41-mode model is qualitatively similar to that of figure 1, supporting the good agreements in temporal auto- and cross-correlations of figure 8.

5. Concluding remarks

In the present study, we have examined a set of physics-informed and data-driven approaches towards the development of a low-dimensional model more accurate than the conventional ones for turbulent wall-bounded shear flows. Based on the utilisation of POD modes, a particular focus is given to the case where the number of the POD modes is not necessarily large enough to cover the full dynamics of the given chaotic state. Starting from the conventional POD-Galerkin model, three additional approaches have been examined: 1) sparse regression of the POD-Galerkin dynamics (POD-SINDy model); 2) POD-Galerkin projection with an empirical eddy viscosity model (POD-Galerkin-E model; Smith *et al.* 2005); 3) a newly-proposed POD-Galerkin projection with an optimal eddy viscosity determined using a spare regression (POD-Galerkin-R model). The sparse regression of the POD-Galerkin dynamics has been found to be unsuccessful presumably due to the small number of POD modes considered, although this might be able to be improved by incorporating the energy-preserving nonlinearity constraint into the model (Loiseau & Brunton 2018). In the present study, this issue can be tackled by introducing a data-driven eddy-viscosity model for a highly turbulent flow, as the POD-Galerkin projection with a sparse optimal viscosity has been found to well approximate the given chaotic dynamics. It should be mentioned that this eddy-viscosity model was introduced to have a better nonlinear energy balance (3.8) only at large scale spanned by the POD modes of interest (see also discussion in §3.3). In this respect, the data-driven eddy viscosity model here may be viewed to be a pragmatic alternative of the the energy-preserving nonlinearity constraint in Loiseau & Brunton (2018) for highly turbulent flows.

The key reason to the success of the POD-Galerkin-R model is that it considers the largest amount of physical information: i.e. Galerkin projection and energy cascade. It is important to emphasise that the Galerkin projection allows the reduced-order model to inherit the mathematical structure of the Navier-Stokes equations. In other words, this feature makes the reduced-order model analysable, as it contains all the mathematical elements previously utilised to study the flow physics: e.g. linearised dynamics and production/dissipation, etc. Having said this, the energy cascade via nonlinear and non-local interactions modelled here is still an active and challenging research topic (e.g. Vassilicos 2015), and it may take years to gain the full physical understanding, if not possible. The eddy-viscosity model utilised in the present study is still very minimal to incorporate the full energy cascade dynamics into a reduced-order model. However, a notable point of doing so is that a data-driven approach

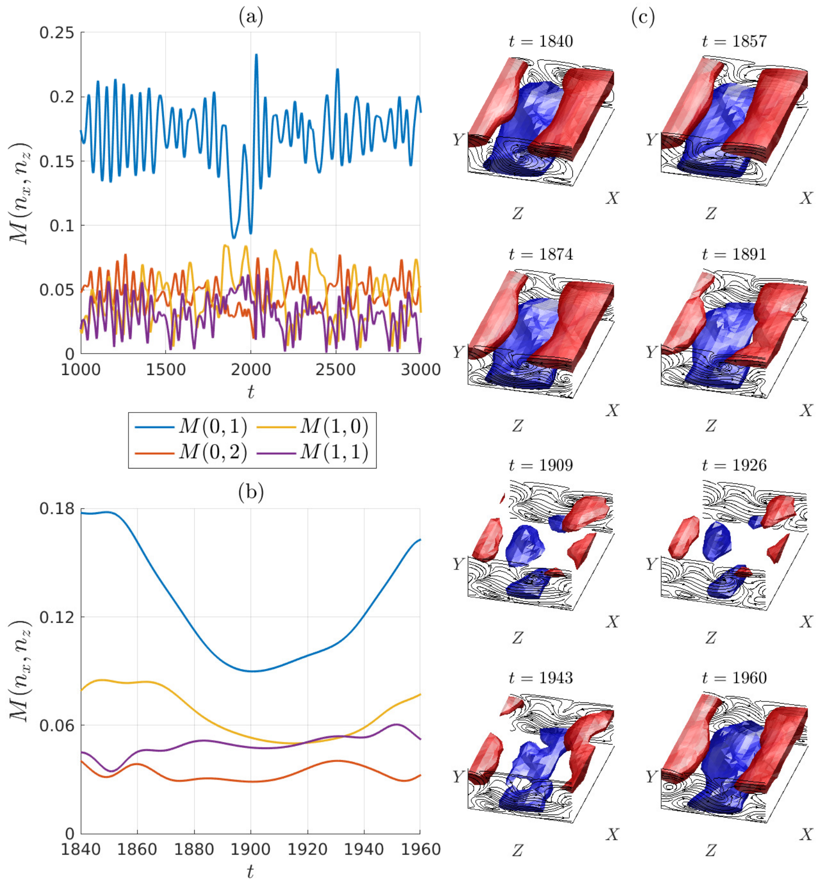


Figure 9: Time trace of $M(0, 1)$, $M(0, 2)$, $M(1, 0)$, $M(1, 1)$ obtained from the POD-Galerkin-R 41 mode model (a) for $t = 1000 - 3000$ and (b) for $t = 1840 - 1960$. (c) Flow snapshots at $t = 1840, 1857, 1874, 1891, 1909, 1926, 1943, 1960$, where the blue and red iso-surfaces indicate $u = \pm 0.38$, respectively.

(i.e. sparse regression), which itself does not provide any insight into the given flow physics, was applied to model the flow physics which is not fully understood. We have shown that classical physics-based reduced-order modelling (i.e POD-Galerkin) of a complex process is limited, and data-driven approaches can be exploited to improve the reduced-order models.

It should also be mentioned that there have recently been a surge of data-driven flow modelling approaches using optimisation and machine learning (see the recent review by Brunton *et al.* 2020). In the context of reduced-order modelling, utilisation of some machine learning algorithms (e.g. reservoir computing) was proposed for the prediction of a chaotic dynamical system (e.g. Pathak *et al.* 2018 and the other recent studies). While such an approach may well be practically useful for the prediction of extreme events relevant to weather forecasting, it does not offer insights into the flow physics required for modelling in a wider context. Indeed, how one would smartly incorporate the known flow physics into a data-driven modelling approach has been a central issue of many current investigations, especially when the equations of motion (e.g. Navier-Stokes equations) are fully available. In this respect, the utilisation of Galerkin projection in the present study may perhaps provide a new opportunity as it directly offers a mathematical structure from the governing equations. Indeed, instead of utilising a model given by (3.13), a highly flexible form of model for **T** may well be considered with a machine learning algorithm.

It is also worth mentioning about the extrapolation capability of the model obtained at a given set of parameters to the others. This issue has often been regarded to be generally challenging for a model reduction problem. In the present study, the optimal eddy viscosity obtained here is, in fact, intricately linked to the physical processes of the given system. The optimal value would vary with the change of system parameters (e.g. the Reynolds-number dependent role of small scales modelled with the eddy viscosity here). Therefore, further efforts need to be made to address this issue in the future.

Finally, given the original scope of the present paper discussed in §1.1, the natural next step of the present study is to apply the approach proposed here to flows at higher Reynolds numbers where coherent structures begin to emerge at multiple integral length scales as in the attached eddy hypothesis of Townsend (1956, 1976). An obvious issue for this next step would lie in the determination of the number of POD modes that capture the core interaction dynamics at integral length scales, while effectively excluding the dissipative dynamics that can be modelled using the data-driven eddy-viscosity approach here. Once this process is completed with an appropriate validation using DNS data, invariant solutions (e.g. unstable periodic orbits) of the reduced-order model can subsequently be computed to study the multi-scale dynamics. The current hope is that the total degree of freedom of the reduced-order model remains at $O(10^2-10^3)$ at a sufficiently high Reynolds number (e.g. $Re_\tau \simeq 500-1000$) to tackle this challenge.

Funding

Y.H. gratefully acknowledges the financial support of the Leverhulme Trust (RPG-2019-123) and the Engineering and Physical Sciences Research Council (EPSRC; EP/T009365/1) in the UK.

Declaration of interest

The authors report no conflict of interest.

Appendix A. Galerkin projection

The projection of (2.4*a*) onto the Navier-Stokes equations (2.1) leads to the following system of ordinary differential equations:

$$\dot{a}^{(n)}_{n_x,n_z} = \sum_{m=1}^{N_p} L^{(n,m)}_{n_x,n_z} a^{(m)}_{n_x,n_z} + \mathcal{N}^{(n)}_{n_x,n_z}, \tag{A 1a}$$

where

$$\mathcal{N}^{(n)}_{n_x,n_z} = \sum_{k_x=-N_x}^{N_x} \sum_{k_z=-N_z}^{N_z} \sum_{k=1}^{N_p} \sum_{m=1}^{N_p} N^{(n,k,m)}_{\substack{n_x,n_z\\k_x,k_z}} \; a^{(k)}_{k_x,k_z} a^{(m)}_{\substack{m_x=n_x-k_x\\m_z=n_z-k_z}}, \tag{A 1b}$$

with

$$L_{n_x,n_z}^{(n,m)} = -\frac{1}{Re}\left[\left(\frac{2\pi n_x}{L_x}\right)^2 + \left(\frac{2\pi n_z}{L_z}\right)^2\right]\delta_{nm} - \frac{1}{Re}\int_y \left(\frac{\mathrm{d}\boldsymbol{\phi}_{n_x,n_z}^{(m)}}{\mathrm{d}y}\right)^H \frac{\mathrm{d}\boldsymbol{\phi}_{n_x,n_z}^{(n)}}{\mathrm{d}y}\ \mathrm{d}y$$
$$-\left(\frac{2\pi i n_x}{L_x}\right)\int_y y(\boldsymbol{\phi}_{n_x,n_z}^{(m)})^H \boldsymbol{\phi}_{n_x,n_z}^{(n)}\ \mathrm{d}y - \int_y (\phi_{2,n_x,n_z}^{(m)})^H \phi_{1,n_x,n_z}^{(n)}\ \mathrm{d}y, \tag{A 1c}$$

and

$$N_{\substack{n_x,n_z\\k_x,k_z}}^{(n,k,m)} = -\frac{1}{\sqrt{L_x L_z}}\int_y (\boldsymbol{\phi}_{n_x,n_z}^{(n)})^H \left[\frac{2\pi i k_x}{L_x}\boldsymbol{\phi}_{k_x,k_z}^{(k)} \quad \frac{\mathrm{d}\boldsymbol{\phi}_{k_x,k_z}^{(k)}}{\mathrm{d}y} \quad \frac{2\pi i k_z}{L_z}\boldsymbol{\phi}_{k_x,k_z}^{(k)}\right]\boldsymbol{\phi}_{\substack{m_x=n_x-k_x\\m_z=n_z-k_z}}^{(m)}\ \mathrm{d}y, \tag{A 1d}$$

Therefore, (A 1) may be written as the following quadratic nonlinear dynamical system form:

$$\dot{\mathbf{a}} = \mathbf{La} + \mathbf{N}(\mathbf{a},\mathbf{a}), \tag{A 2a}$$

where $\mathbf{a}$ is defined as a column vector, each element of which is given by $a_{n_x,n_z}^{(n)}$,

$$\mathbf{La} \equiv \sum_{m=1}^{N_p} L_{n_x,n_z}^{(n,m)} a_{n_x,n_z}^{(m)}, \tag{A 2b}$$

and

$$\mathbf{N}(\mathbf{a},\mathbf{a}) \equiv \mathcal{N}_{n_x,n_z}^{(n)}. \tag{A 2c}$$

Similarly, the diffusion operator used for the eddy-viscosity closure in §3.3 and §3.4 is defined as

$$\mathbf{Da} \equiv \sum_{m=1}^{N_p} D_{n_x,n_z}^{(n,m)} a_{n_x,n_z}^{(m)}, \tag{A 3a}$$

where

$$D_{n_x,n_z}^{(n,m)} = -\left[\left(\frac{2\pi n_x}{L_x}\right)^2 + \left(\frac{2\pi n_z}{L_z}\right)^2\right]\delta_{nm} - \int_y \left(\frac{d\boldsymbol{\phi}_{n_x,n_z}^{(n)}}{dy}\right)^H \frac{d\boldsymbol{\phi}_{n_x,n_z}^{(m)}}{dy}\ dy \tag{A 3b}$$

for $(n_x,n_z) \neq (0,0)$ and

$$D_{n_x,n_z}^{(n,m)} = 0. \tag{A 3c}$$

for $(n_x,n_z) = (0,0)$, so that the eddy viscosity is not applied to the mean equation. Using the diffusion operator above, the eddy viscosity model defined in (3.13) is finally written as

$$\mathbf{V}_t\mathbf{Da} \equiv c_{n_x,n_z}^{(n)} e(t) \sum_{m=1}^{N_p} D_{n_x,n_z}^{(n,m)} a_{n_x,n_z}^{(m)}, \tag{A 4a}$$

where $c_{n_x,n_z}^{(n)}$ forms each element of $\mathbf{c}$ in (3.13).

Appendix B. The choice of e in §3.4

Here, we report a POD-Galerkin-R model, in which $e(t) = \mathbf{a}^H\mathbf{a}$ is considered instead of $e(t) = a_{0,0}^{(1)}$. The sparse regression in (3.16) is performed with the DNS data for $t \in [-10000, 0]$ and the resulting model is subsequently examined for $t \in [0, 5000]$. The time trace of $M(n_x,n_z)$ from the reduced-order model and from DNS is shown in figure 10.

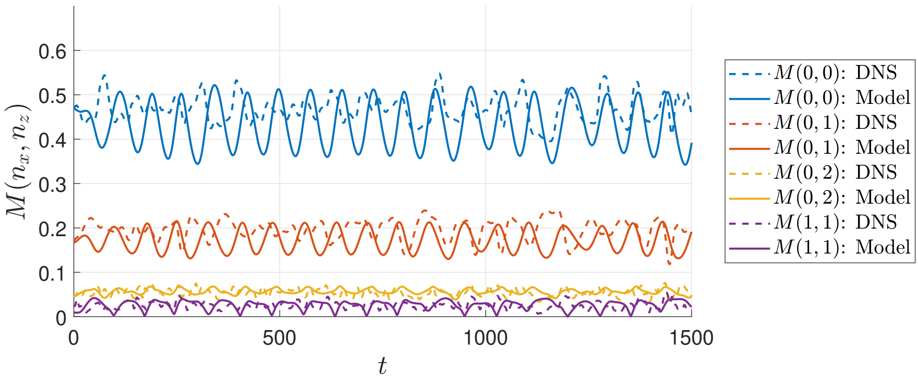


Figure 10: Time trace of $M(n_x, n_z)$ from DNS and the POD-Galerkin-R model ($\gamma_e = 0.0005$ and $e(t) = \mathbf{a}^H\mathbf{a}$) with 25 modes.

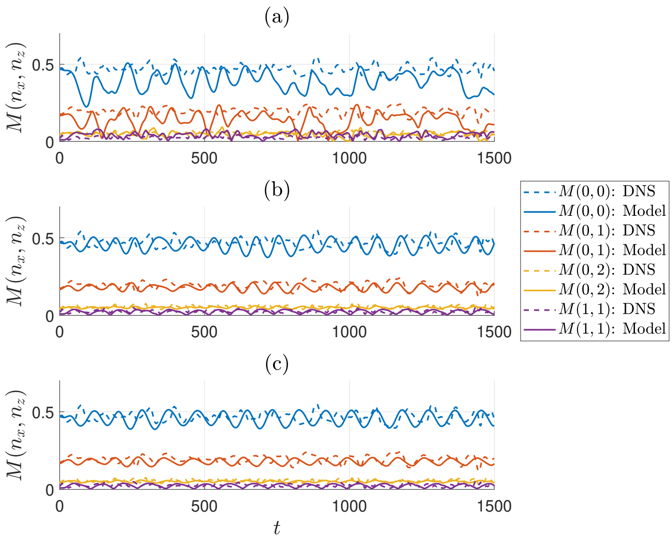


Figure 11: Time trace of $M(n_x, n_z)$ from DNS and the POD-Galerkin-R of the 25-mode model : (a) $\gamma_e = 0.001$; (b) $\gamma_e = 0.0005$; (c) $\gamma_e = 0.0003$.

Overall, the mean of $M(n_x, n_z)$ and its oscillation time scale from the reduced-order model compare fairly well with those from DNS. The oscillation magnitude of $M(0,0)$ in this case is slightly stronger than that from the POD-Galerkin-R model with $e(t) = a_{0,0}^{(1)}$ (figure 5a), and the oscillation appears to be slightly less chaotic.

Appendix C. The effect of γ_e on model dynamics.

The sparsity-promoting ℓ_1-regulariser acts as a control parameter balancing between the effect of the GP model and the residual model on the overall dynamics. In Figure 11, we observe that for $\gamma_e = 0.001$, the effect of the residual model adversely affects the temporal dynamics as only certain POD modes are being selectively damped. For $\gamma_e = 0.0003$, the residual model dominates and the POD modes are strongly coupled to the mean POD mode where we obtain oscillatory behaviour due to excessive damping.

REFERENCES

del Álamo, J.C. & Jiménez, J. 2003 Spectra of the very large anisotropic scales in turbulent channels. *Phys. Fluids* **15**, L41.

del Álamo, J. C. & Jiménez, J. 2006 Linear energy amplification in turbulent channels. *J. Fluid Mech.* **559**, 205–213.

Alizard, F. 2015 Linear stability of optimal streaks in the log-layer of turbulent channel flows. *Phys. Fluids* **27**, 105103.

Aubry, N., Holmes, P., Lumley, J. L. & Stone, E. 1988 The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.* **192**, 115–173.

Avila, K., Moxey, D., de Lozar, A., Avila, M., Barkley, D. & Hof, Bjorn 2011 The onset of turbulence in pipe flow. *Science* **333**, 192–196.

Barkley, D. 2016 Theoretical perspective on the route to turbulence in a pipe. *Journal of Fluid Mechanics* **803**.

Beetham, S. & Capecelatro, J. 2020 Formulating turbulence closures using sparse regression with embedded form invariance. *Physical Review Fluids* **5** (8).

Beetham, S., Fox, R.O. & Capecelatro, J. 2021 Sparse identification of multiphase turbulence closures for coupled fluid–particle flows. *Journal of Fluid Mechanics* **914**, A11.

Bewley, T. R., Moin, P. & Temam, R. 2001 Dns-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *J. Fluid Mech.* **447**, 179–225.

Brunton, Steven L & Kutz, J Nathan 2019 *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press.

Brunton, S. L., Noack, B. R. & Koumoutsakos, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. of Fluid Mec.* **52**, 477–508.

Brunton, Steven L., Proctor, Joshua L. & Kutz, J. Nathan 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **113** (15), 3932–3937, arXiv: https://www.pnas.org/content/113/15/3932.full.pdf.

Butler, K. M. & Farrell, B. F. 1993 Optimal perturbations and streak spacing in wall-bounded turbulent shear flow. *Phys. Fluids* **5**, 774–777.

Callaham, J. L., Brunton, S. L. & Loiseau, J.-C. 2021*a* On the role of nonlinear correlations in reduced-order modeling. *arXiv preprint arXiv:2106.02409* ArXiv:2106.02409.

Callaham, J. L., Loiseau, J.-C., Rigas, G. & Brunton, S. L. 2021*b* Nonlinear stochastic modelling with langevin regression. *Pro. Roy. Soc. Lond. A* **477** (2250), 487–505.

Cassinelli, A., de Giovanetti, M. & Hwang, Y. 2017 Streak instability in near-wall turbulence revisited. *J. Turbul.* **18** (5), 443–464.

Cavalieri, A. V. G. 2021 Structure interactions in a reduced-order model for wall-bounded turbulence. *Phys. Rev. Fluids* **6**, 034610.

Chantry, M., Tuckerman, L. S. & Barkley, D. 2017 Universal continuous transition to turbulence in a planar shear flow. *J. Fluid Mech.* **824**, R1.

Cho, M., Hwang, Y. & Choi, H. 2018 Scale interactions and spectral energy transfer in turbulent channel flow. *J. Fluid Mech.* **854**, 474–504.

Choi, H., Temam, R., Moin, P. & Kim, J 1993 Feedback control for unsteady flow and its application to the stochastic burgers equation. *J. Fluid Mech.* **245**, 509–543.

Cordier, Laurent, El Majd, B Abou & Favier, Julien 2010 Calibration of pod reduced-order models using tikhonov regularization. *International journal for numerical methods in fluids* **63** (2), 269–296.

Cossu, C., Pujals, G. & Depardon, S. 2009 Optimal transient growth and very large scale structures in turbulent boundary layers. *J. Fluid Mech.* **619**, 79–94.

Couplet, M, Basdevant, C & Sagaut, P 2005 Calibrated reduced-order pod-galerkin system for fluid flow modelling. *Journal of Computational Physics* **207** (1), 192–220.

Couplet, M, Sagaut, P & Basdevant, C 2003 Intermodal energy transfers in a proper orthogonal decomposition–galerkin representation of a turbulent separated flow. *Journal of fluid mechanics* **491** (491), 275–284.

Crisanti, A., Jensen, M. H., Paladin, G. & Vulpiani, A. 1993 Predictability of velocity and temperature fields in intermittent turbulence. *J. Phys. A* **26**, 6943.

Doohan, P., Bengana, Y., Yang, Q., Willis, A. P. & Hwang, Y. 2021*a* Multi-scale state space and travelling waves in wall-bounded turbulence. *J. Fluid Mech.* Under review.

Doohan, P., Willis, A. P. & Hwang, Y. 2019 Shear stress-driven flow: the state space of near-wall turbulence as $Re_\tau \to \infty$. *J. Fluid Mech.* **874**, 606–638.

Doohan, P., Willis, A. P. & Hwang, Y. 2021*b* Minimal multi-scale dynamics of near-wall turbulence. *J. Fluid Mech.* **913**, A8.

Duraisamy, Karthik 2021 Perspectives on machine learning-augmented reynolds-averaged and large eddy simulation models of turbulence. *Phys. Rev. Fluids* **6**, 050504.

Eckhardt, B., Schneider, T.M., Hof, B. & Westerweel, J. 2007 Turbulence Transition in Pipe Flow. *Ann. Rev. Fluid Mech.* **39**, 447.

Faisst, H. & Eckhardt, B. 2003 Travelling waves in pipe flow. *Phys. Rev. Lett.* **91**, 224502.

Falco, R. E. 1977 Coherent motions in the outer region of turbulent boundary layers. *Phys. Fluids* **20**, S124–S132.

Farazmand, M. 2016 An adjoint-based approach for finding invariant solutions of navier–stokes equations. *J. Fluid Mech.* **795**, 278–312.

Flores, O. & Jiménez, J 2010 Hierarchy of minimal flow units in the logarithmic layer. *Phys. Fluids* **22**, 071704.

Gelss, Patrick, Klus, Stefan, Eisert, Jens & Schütte, Christof 2019 Multidimensional approximation of nonlinear dynamical systems. *Journal of Computational and Nonlinear Dynamics* **14** (6).

Gibson, J. F., Halcrow, J. & Cvitanovic, P. 2008 Visualizing the geometry of state space in plane Couette flow. *J. Fluid Mech.* **611**, 107–130.

Gibson, J. F., Halcrow, J. & Cvitanovic, P. 2009 Equilibrium and traveling-wave solutions of plane Couette flow. *J. Fluid Mech.* **638**, 243–266.

de Giovanetti, M., Sung, H. J. & Hwang, Y. 2017 Streak instability in turbulent channel flow: the seeding mechanism of large-scale motions. *J. Fluid Mech.* **832**, 483–513.

Graham, M. D. & Floryan, D. 2020 Exact coherent states and the nonlinear dynamics of wall-bounded turbulent flows. *Annu. Rev. Fluid. Mech.* **53**, 227–253.

Hall, P. & Sherwin, S. J. 2010 Streamwise vortices in shear flows: harbingers of transition and the skeleton of coherent structures. *J. Fluid Mech.* **661**, 178–205.

Hamilton, James M, Kim, John & Waleffe, Fabian 1995 Regeneration mechanisms of near-wall turbulence structures. *Journal of Fluid Mechanics* **287** (1), 317–348.

Head, M. R. & Bandyopadhay, P. 1981 New aspects of turbulent boundary-layer structure. *J. Fluid Mech* **107**, 297–338.

Holmes, Philip, Lumley, John L. & Berkooz, Gal 1996 *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. *Cambridge Monographs on Mechanics* . Cambridge University Press.

Hutchins, N. & Marusic, I. 2007 Evidence of very long meandering features in the logarithmic region of turbulent boundary layers. *J. Fluid Mech.* **579**, 1–28.

Hwang, Y. 2015 Statistical structure of self-sustaining attached eddies in turbulent channel flow. *J. Fluid Mech.* **723**, 264–288.

Hwang, Y. & Bengana, Y. 2016 Self-sustaining process of minimal attached eddies in turbulent channel flow. *J. Fluid Mech.* **795**, 708–738.

Hwang, Y. & Cossu, C. 2010*a* Amplification of coherent streaks in the turbulent Couette flow: an input-output analysis at low Reynolds number. *J. Fluid Mech.* **643**, 333–348.

Hwang, Y. & Cossu, C. 2010*b* Linear non-normal energy amplification of harmonic and stochastic forcing in the turbulent channel flow. *J. Fluid Mech.* **664**, 51–73.

Hwang, Y. & Cossu, C. 2010*c* Self-sustained process at large scales in turbulent channel flow. *Phys. Rev. Lett.* **105**, 044505.

Hwang, Y. & Cossu, C. 2011 Self-sustained processes in the logarithmic layer of turbulent channel flows. *Phys. Fluid* **23**, 061702.

Hwang, Y. & Lee, M. 2020 The mean logarithm emerges with self-similar energy balance. *J. Fluid Mech.* **903**, R6.

Hwang, Y., Willis, A. P. & Cossu, C. 2016 Invariant solutions of minimal large-scale structures in turbulent channel flow for Re_τ up to 1000. *J. Fluid Mech.* **802**, R1.

Jeong, J., Benney, F., Schoppa, W. & Kim, J. 1997 Coherent structures near the wall in a turbulent channel flow. *J. Fluid Mech.* **332**, 185–214.

Jiménez, Javier & Moin, Parviz 1991 The minimal flow unit in near-wall turbulence. *Journal of Fluid Mechanics* **225**, 213–240.

Jiménez, J. & Simens, M. P. 2001 Low-dimensional dynamics of a turbulent wall flow. *J. Fluid Mech.* **435**, 81–91.

Joseph, D. D. 1976 *Stability of Fluid Motions*. *Springer Tracts in Natural Philosophy, vol. 27* . New York: Springer.

Kawahara, G. & Kida, S. 2001 Periodic motion embedded in plane Couette turbulence: regeneration cycle and burst. *J. Fluid Mech.* **449**, 291–300.

Kawahara, G., Uhlmann, M. & van Veen, L. 2012 The significance of simple invariant solutions in turbulent flows. *Annu. Rev. Fluid Mech.* **44**, 203–225.

Kim, K. C. & Adrian, R. 1999 Very large-scale motion in the outer layer. *Phys. Fluids* **11** (2), 417–422.

Kline, S. J., Reynolds, W. C., Schraub, F. A. & Runstadler, P. W. 1967 The structure of turbulent boundary layers. *J. Fluid Mech.* **30**, 741–773.

Kovasznay, L. S. G., Kibens, V. & Blackwelder, R. F. 1970 Large-scale motion in the intermittent region of a turbulent boundary layer. *J. Fluid Mech.* **41**, 283–325.

Kreilos, T. & Eckhardt, B. 2012 Periodic orbits near onset of chaos in plane Couette flow. *Chaos* **22**, 047505.

Lagha, M. & Manneville, P. 2007 Modelling transitional plane Couette flow. *Eur. Phys. J.* **58**, 433–447.

Lee, C., Kim, J. & Choi, H. 1999 Suboptimal control of turbulent channel flow for drag reduction. *J. Fluid Mech.* **358**, 245–258.

Lee, M. K. & Moser, R. D. 2019 Spectral analysis of the budget equation in turbulent channel flows at high reynolds number. *J. Fluid Mech.* **860**, 886–938.

Loiseau, Jean-Christophe & Brunton, Steven L 2018 Constrained sparse galerkin regression. *Journal of Fluid Mechanics* **838**, 42–67.

Lozano-Durán, A., Nikolaidis, M.-A., Constantinou, N. C. & Karp, M. 2021 Cause-and-effect of linear mechanisms sustaining wall turbulence. *J. Fluid Mech.* **914**, A8.

Lumley, J. L. 1967 The structure of inhomogeneous turbulent flows. In *Atmospheric Turbulence and Radio Wave Propagation* (ed. A. M. Yaglom & V. I. Tatarski), *Nonlinear Problems of Fluid Dynamics* . Nakua, Moscow.

Lumley, John L 1981 Coherent structures in turbulence. In *Transition and turbulence*, pp. 215–242. Elsevier.

Marusic, I. & Monty, J. P. 2019 Attached eddy model of wall turbulence. *Annu. Rev. Fluid Mech.* **51**, 49–74.

Moehlis, J., Faisst, H. & Eckhardt, B. 2004 A low-dimensional model for turbulent shear flows. *New Journal of Physics* **6**, 56.

Moehlis, J., Faisst, H. & Eckhardt, B. 2005 Periodic orbits and chaotic sets in a low-dimensional model for shear flows. *SIAM J. Appl. Dyn. Syst.* **4**, 352–376.

Mohebujjaman, Muhammad, Rebholz, Leo G & Iliescu, Traian 2019 Physically constrained data-driven correction for reduced-order modeling of fluid flows. *International Journal for Numerical Methods in Fluids* **89** (3), 103–122.

Nagata, M. 1990 Three-dimensional finite-amplitude solutions in plane Couette flow: bifurcation from infinity. *J. Fluid Mech.* **217**, 519–527.

Noack, B. R., MORZýnski, M. & Tadmor, G. 2011 *Reduced-Order Modelling for Flow Control (CISM Courses and Lectures)*. Springer.

Östh, J., Noack, B. R., Krajnović, S., Barros, D. & Borée, J. 2014 On the need for a nonlinear subscale turbulence term in pod models as exemplified for a high-reynolds-number flow over an ahmed body. *J. Fluid Mech.* **747**, 518–544.

Park, J., Hwang, Y. & Cossu, C. 2011 On the stability of large-scale streaks in the turbulent Couette and Poiseuille flows. *C. R. Mècanique* **339** (1), 1–5.

Park, J. S. & Graham, M. D. 2015 Exact coherent states and connections to turbulent dynamics in minimal channel flow. *J. Fluid Mech.* **782**, 430–454.

Pathak, J., Hunt, B., Girvan, M., Lu, Z. & Ott, E. 2018 Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. *Phys. Rev. Lett.* **120**, 024102.

Podvin, Bérengère 2009 A proper-orthogonal-decomposition–based model for the wall layer of a turbulent channel flow. *Physics of Fluids* **21** (1), 015111, arXiv: https://doi.org/10.1063/1.3068759.

Protas, B., Noack, B. R. & Östh, J. 2015 Optimal nonlinear eddy viscosity in galerkin models of turbulent flows. *J. Fluid Mech.* **766**, 337–367.

Rawat, S., Cossu, C., Hwang, Y. & Rincon, F. 2015 On the self-sustained nature of large-scale motions in turbulent Couette flow. *J. Fluid Mech.* **782**, 515–540.

Rempfer, D. & Fasel, F. H. 1994*a* Dynamics of three-dimensional coherent structures in a flat-plate boundary-layer. *J. Fluid Mech.* **275**, 257–283.

Rempfer, D. & Fasel, F. H. 1994*b* Evolution of three-dimensional coherent structures in a flat-plate boundary-layer. *J. Fluid Mech.* **260**, 351–375.

Rubini, R., Lasagna, D. & Ronch, A. Da 2021 The l_1-based sparsification of energy interactions in unsteady lid-driven cavity flow. *J. Fluid Mech.* **905**, A18.

Ruelle, D. 1979 Microscopic fluctuations and turbulence. *Phys. Lett. A* **72**, 81.

Schmelzer, Martin, Dwight, Richard P & Cinnella, Paola 2020 Discovery of algebraic reynolds-stress models using sparse symbolic regression. *Flow, Turbulence and Combustion* **104** (2), 579–603.

Schmidt, M. & Lipson, H. 2009 Distilling free-form natural laws from experimental data. *Science* **264**, 255–275.

Schoppa, W. & Hussain, F. 2002 Coherent structure generation in near-wall turbulence. *J. Fluid Mech.* **453**, 57–108.

Sirovich, L. 1987 Turbulence and the dynamics of coherent structures; part I: Coherent structures. *Q. Appl. Math* **XLV** (3), 561–571.

Smith, TR, Moehlis, J & Holmes, Ph 2005 Low-dimensional models for turbulent plane couette flow in a minimal flow unit. *Journal of Fluid Mechanics* **538**, 71.

Townsend, A. A. 1956 *The structure of turbulent shear flow*, 1st edn. Cambridge U. Press.

Townsend, A. A. 1976 *The structure of turbulent shear flow*, 2nd edn. Cambridge U. Press.

Vassilicos, J. C. 2015 Dissipation in turbulent flows. *Annu. Rev. Fluid Mech.* **47**, 95–114.

Viswanath, D. 2007 Recurrent motions within plane Couette turbulence. *J. Fluid Mech.* **580**, 339–358.

Waleffe, F. 1995 Hydrodynamic stability and turbulence: Beyond transients to a self-sustaining process. *Stud. Appl. Math.* **95**, 319–343.

Waleffe, F. 1997 On a self-sustaining process in shear flows. *Phys. Fluids* **9**, 883–900.

Waleffe, F. 2001 Exact coherent structures in channel flow. *J. Fluid Mech.* **435**, 93–102.

Waleffe, F. 2003 Homotopy of exact coherent structures in plane shear flows. *Phys. Fluids* **15**, 1517–1534.

Wedin, H. & Kerswell, R. R. 2004 Exact coherent structures in pipe flow: travelling wave solutions. *J. Fluid Mech.* **508**, 333–371.

Willis, A. P., Cvitanovic, P. & Avila, M. 2013 Revealing the state space of turbulent pipe flow by symmetry reduction. *J. Fluid Mech.* **721**, 514–540.

Yang, Q., Willis, A. P. & Hwang, Y. 2019 Exact coherent states of attached eddies in channel flow. *J. Fluid Mech.* **862**, 1029–1059.