JM PAPERS

Informative and non-informative decomposition of turbulent flow fields

Gonzalo Arranz[1],† **and Adrián Lozano-Durán**[1,2]

[1]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA 91125, USA

Not all the information in a turbulent field is relevant for understanding particular regions or variables in the flow. Here, we present a method for decomposing a source field into its informative $\boldsymbol{\Phi}_I(\boldsymbol{x}, t)$ and residual $\boldsymbol{\Phi}_R(\boldsymbol{x}, t)$ components relative to another target field. The method is referred to as informative and non-informative decomposition (IND). All the necessary information for physical understanding, reduced-order modelling and control of the target variable is contained in $\boldsymbol{\Phi}_I(\boldsymbol{x}, t)$, whereas $\boldsymbol{\Phi}_R(\boldsymbol{x}, t)$ offers no substantial utility in these contexts. The decomposition is formulated as an optimisation problem that seeks to maximise the time-lagged mutual information of the informative component with the target variable while minimising the mutual information with the residual component. The method is applied to extract the informative and residual components of the velocity field in a turbulent channel flow, using the wall shear stress as the target variable. We demonstrate the utility of IND in three scenarios: (i) physical insight into the effect of the velocity fluctuations on the wall shear stress; (ii) prediction of the wall shear stress using velocities far from the wall; and (iii) development of control strategies for drag reduction in a turbulent channel flow using opposition control. In case (i), IND reveals that the informative velocity related to wall shear stress consists of wall-attached high- and low-velocity streaks, collocated with regions of vertical motions and weak spanwise velocity. This informative structure is embedded within a larger-scale streak–roll structure of residual velocity, which bears no information about the wall shear stress. In case (ii), the best-performing model for predicting wall shear stress is a convolutional neural network that uses the informative component of the velocity as input, while the residual velocity component provides no predictive capabilities. Finally, in case (iii), we demonstrate that the informative component of the wall-normal velocity is closely linked

† Email address for correspondence: garranz@mit.edu

to the observability of the target variable and holds the essential information needed to develop successful control strategies.

Key words: turbulence simulation, turbulence theory, turbulent boundary layers

1. Introduction

Since the early days of turbulence research, there have been multiple attempts to decompose the flow into different components to facilitate its physical understanding, control its behaviour and devise reduced-order models. One of the earliest examples is the Reynolds decomposition (Reynolds 1895), which divides the velocity field into its mean and fluctuating components. More sophisticated approaches rapidly emerged, aiming at extracting the coherent structure of the flow through correlations and structure identification (Robinson 1991; Panton 2001; Adrian 2007; Smits, McKeon & Marusic 2011; McKeon 2017; Jiménez 2018). This interest is justified by the hope that insights into the dynamics can be gained by analysing a subset of the entire flow, while the remaining incoherent flow plays only a secondary role in understanding the overall dynamics. In this work, we introduce a method to decompose turbulent flow fields into informative and non-informative components, referred to as informative and non-informative decomposition (IND), such that the informative component contains all the useful information for physical understanding, modelling and control with respect to a given quantity of interest.

The quest to divide turbulent flows in terms of coherent and incoherent motions has a long history, tracing back to the work of Theodorsen (1952), and has been a subject of active research since the pioneering experimental visualisations of Kline *et al.* (1967) and the identification of large-scale coherent regions in mixing layers by Brown & Roshko (1974). Despite this rich history, the field still lacks consensus about the definition of a coherent structure due to the variety of interpretations proposed by different researchers. One of the initial approaches to distinguish turbulent regions was the turbulent/non-turbulent discriminator circuits introduced by Corrsin & Kistler (1954). Since then, single- and two-point correlations have become conventional tools for identifying coherent regions within the flow (e.g. Sillero, Jiménez & Moser 2014). The development of more sophisticated correlation techniques, such as the linear stochastic estimation (Adrian & Moin 1988) – together with its extensions (Tinney *et al.* 2006; Baars & Tinney 2014; Encinar & Jiménez 2019) – and the characteristic-eddy approach (Moin & Moser 1989), has further improved our understanding of the coherent structure of turbulence. An alternative set of methods focuses on decomposing the flow into localised regions where certain quantities of interest are particularly intense. The first attempts, dating back to the 1970s, include the variable-interval time average method (Blackwelder & Kaplan 1976) for obtaining temporal structures of bursting events, and its modified version, the variable-interval space average method (Kim 1985) for characterising spatial rather than temporal structures. With the advent of larger databases and computational resources, more refined techniques have emerged to extract three-dimensional, spatially localised flow structures. These include investigations into regions of rotating fluid (e.g. vortices; Moisy & Jiménez 2004; Del Álamo *et al.* 2006), motions carrying most of the kinetic energy (e.g. regions of high- and low-velocity streaks; Hwang & Sung 2018; Bae & Lee 2021), and those responsible for most of the momentum transfer in wall turbulence (e.g. quadrant events and uniform momentum zones; Meinhart & Adrian 1995; Adrian,

Meinhart & Tomkins 2000; Lozano-Durán, Flores & Jiménez 2012; Lozano-Durán & Jiménez 2014; de Silva, Hutchins & Marusic 2016; Wallace 2016).

The methods described above offer a local-in-space characterisation of coherent structures, in contrast to the global-in-space modal decompositions of turbulent flows (Taira *et al.* 2017, 2020). One of the first established global-in-space methods is the proper orthogonal decomposition (POD) (Lumley 1967), wherein the flow is decomposed into a series of eigenmodes that optimally reconstruct the energy of the field. This method has evolved in different directions, such as space-only POD (Sirovich 1987), spectral POD (Towne, Schmidt & Colonius 2018) and conditional POD (Schmidt & Schmid 2019), to name a few. Another popular approach is dynamic mode decomposition (DMD) (Schmid 2010; Schmid *et al.* 2011), along with decompositions based on the spectral analysis of the Koopman operator (Rowley *et al.* 2009; Mezić 2013). Similar to POD, various modifications of DMD have been developed, e.g. the extended DMD (Williams, Kevrekidis & Rowley 2015), the multi-resolution DMD (Kutz, Fu & Brunton 2016), and the high-order DMD (Le Clainche & Vega 2017) (see Schmid (2022) for a review). The POD and DMD methods do not explicitly account for nonlinear interactions. To overcome this, extensions to detect quadratic nonlinear interactions based on the bispectrum have also been developed (Baars & Tinney 2014; Schmidt 2020). Another noteworthy modal decomposition approach is empirical mode decomposition, first proposed by Huang *et al.* (1998) and recently used in the field of fluid mechanics (e.g. Cheng *et al.* 2019). While the methods listed above are purely data-driven, other modal decompositions, such as resolvent analysis and input–output analysis, are grounded in the linearised Navier–Stokes equations (Trefethen *et al.* 1993; Jovanović & Bamieh 2005; McKeon & Sharma 2010). It has been shown that POD, DMD and resolvent analysis are equivalent under certain conditions (Towne *et al.* 2018). Recently, machine learning has opened new opportunities for nonlinear modal decompositions of turbulent flows (Brunton, Noack & Koumoutsakos 2020).

The flow decomposition approaches presented above, either local or global in space, have greatly contributed to advancing our knowledge about the coherent structure of turbulence. Nonetheless, there are still open questions, especially regarding the dynamics of turbulence, that cannot be answered easily by current methodologies. Part of these limitations stems from the linearity of most methods, yet turbulence is a nonlinear system. A more salient issue perhaps lies in the fact that current methods (with exceptions, such as the extended POD; Borée 2003) tend to focus on decomposing source variables without accounting for other target variables of interest. In general, it is expected that different target variables would require different decomposition approaches of the source variable. For example, we might be interested in a decomposition of the velocity that is useful for understanding the wall shear stress. Hence the viewpoint adopted here aims to answer the question: what part of the flow is relevant to understanding the dynamics of another variable? In this context, coherent structures are defined as those containing the useful information needed to understand the evolution of a target variable.

The concept of information alluded to above refers to the Shannon information (Shannon 1948; Cover & Thomas 2006), i.e. the average unpredictability in a random variable. The systematic use of information-theoretic tools for causality, modelling and control in fluid mechanics has been discussed recently by Lozano-Durán & Arranz (2022). Betchov (1964) was one of the first authors to propose an information-theoretic metric to quantify the complexity of turbulence. Some works have leveraged Shannon information to analyse different aspects of two-dimensional turbulence and energy cascade models (Cerbus & Goldburg 2013; Materassi *et al.* 2014; Granero-Belinchon 2018; Shavit & Falkovich
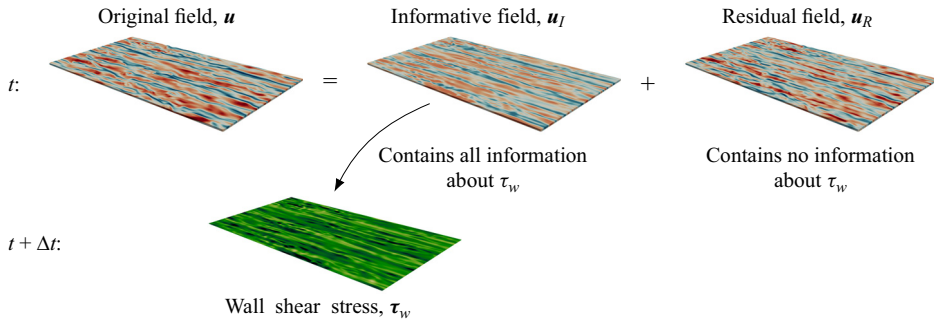
Figure 1. Schematic of IND applied to wall-bounded turbulent flow. The source variable is the velocity fluctuation in the fluid volume $x \in \Omega_u$ at t, and the target variable is the wall shear stress vector at the wall at $t + \Delta t$. For the sake of visualisation, only the streamwise component of the velocity fluctuations and the wall shear stress is displayed. The velocity fluctuation at time t are decomposed into their informative and residual components to the wall shear stress in the future, $t + \Delta t$.

2020; Lee 2021; Tanogami & Araki 2024). Information theory has also been used for causal inference in turbulent flows (Liang & Lozano-Durán 2016; Lozano-Durán, Bae & Encinar 2019; Wang *et al.* 2021; Lozano-Durán & Arranz 2022; Martínez-Sánchez *et al.* 2023), and reduced-order modelling (Lozano-Durán *et al.* 2019). The reader is referred to Lozano-Durán & Arranz (2022) for a more detailed account of the applications of information-theoretic tools in fluid mechanics.

This work is organised as follows. The formulation of the flow decomposition into informative and non-informative components is introduced in § 2: we first discuss the exact formulation of IND in §§ 2.1 and 2.2, followed by its numerically tractable approximation, aIND, in § 2.3. Section 3 demonstrates the application of the method to the decomposition of the velocity field, using wall shear stress in a turbulent channel flow as the target variable. This decomposition is leveraged for physical understanding, prediction of the wall shear stress using velocities away from the wall via convolutional neural networks, and drag reduction through opposition control. Finally, conclusions are presented in § 4.

2. Methodology

2.1. *The IND of the source variable*

Let us denote the source variable by $\boldsymbol{\Phi}(x, t)$, with $x \in \Omega_{\boldsymbol{\Phi}}$, and the target variable by $\boldsymbol{\Psi}(x, t)$, with $x \in \Omega_{\boldsymbol{\Psi}}$, where x and t represent the spatial and time coordinates, respectively. For example, in the case of a turbulent channel flow, the source variable could be the velocity fluctuations defined over the entire domain, $\boldsymbol{\Phi}(x, t) = u(x, t)$, and the target variable could be the shear stress vector at every point over one of the walls, $\boldsymbol{\Psi}(x, t) = \boldsymbol{\tau}_w(x, t)$, as shown in figure 1. We seek to decompose $\boldsymbol{\Phi}(x, t)$ into two independent contributions: an informative contribution to the target variable in the future, $\boldsymbol{\Psi}_+ = \boldsymbol{\Psi}(x, t + \Delta T)$ with $\Delta T \geq 0$, and a residual term that conveys no information about $\boldsymbol{\Psi}_+$ (i.e. the non-informative component):

$$\boldsymbol{\Phi}(x, t) = \boldsymbol{\Phi}_I(x, t) + \boldsymbol{\Phi}_R(x, t), \tag{2.1}$$

where $\boldsymbol{\Phi}_I$ and $\boldsymbol{\Phi}_R$ are the informative and residual contributions, respectively. The decomposition is referred to as IND.

To find a decomposition of the form shown in (2.1), we need to introduce a definition of information. We rely on the concept of Shannon information (Shannon 1948), which

quantifies the average information in the variable $\boldsymbol{\Psi}_+$ as

$$H(\boldsymbol{\Psi}_+) = -\sum_{S \in \mathcal{S}} p_{\boldsymbol{\Psi}_+}(\boldsymbol{\Psi}_+ = S) \log p_{\boldsymbol{\Psi}_+}(\boldsymbol{\Psi}_+ = S) \geq 0, \qquad (2.2)$$

where $H(\boldsymbol{\Psi}_+)$ is referred to as the Shannon entropy or information of $\boldsymbol{\Psi}_+$, $p_{\boldsymbol{\Psi}_+}(\boldsymbol{\Psi}_+ = S)$ denotes the probability of $\boldsymbol{\Psi}_+$ being in the state S, and $\mathcal{S}$ represents the set of all possible states of $\boldsymbol{\Psi}_+$. The remaining information in $\boldsymbol{\Psi}_+$, after discounting for the information in $\boldsymbol{\Phi}$, is measured by the conditional Shannon information:

$$H(\boldsymbol{\Psi}_+ \mid \boldsymbol{\Phi}) = -\sum_{S \in \mathcal{S}} \sum_{R \in \mathcal{R}} p_{\boldsymbol{\Psi}_+,\boldsymbol{\Phi}}(S, R) \log \frac{p_{\boldsymbol{\Psi}_+,\boldsymbol{\Phi}}(S, R)}{p_{\boldsymbol{\Phi}}(R)} \geq 0, \qquad (2.3)$$

where $p_{\boldsymbol{\Psi}_+,\boldsymbol{\Phi}}$ is the joint probability distribution of $\boldsymbol{\Psi}_+$ and $\boldsymbol{\Phi}$, R is a particular state of $\boldsymbol{\Phi}$, and $\mathcal{R}$ is the set of all possible states of $\boldsymbol{\Phi}$. The difference between (2.2) and (2.3) quantifies the amount of shared information between the variables,

$$I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi}) = H(\boldsymbol{\Psi}_+) - H(\boldsymbol{\Psi}_+ \mid \boldsymbol{\Phi}), \qquad (2.4)$$

and is referred to as the mutual information between $\boldsymbol{\Psi}_+$ and $\boldsymbol{\Phi}$. The condition $H(\boldsymbol{\Psi}_+) \geq H(\boldsymbol{\Psi}_+ \mid \boldsymbol{\Phi})$ – known as 'information can't hurt' (Cover & Thomas 2006) – guarantees that $I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi})$ is always non-negative. The mutual information is equal to 0 only when the variables are independent, i.e. $p_{\boldsymbol{\Psi}_+,\boldsymbol{\Phi}}(S, R) = p_{\boldsymbol{\Psi}_+}(S) \, p_{\boldsymbol{\Phi}}(R)$ for all possible states $S \in \mathcal{S}$ and $R \in \mathcal{R}$.

We are now in a position to define the conditions that $\boldsymbol{\Phi}_I$ and $\boldsymbol{\Phi}_R$ must satisfy. First, the informative contribution should maximise $I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi}_I)$ from (2.4), which is achieved when

$$I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi}_I) = H(\boldsymbol{\Psi}_+), \qquad (2.5)$$

namely, $\boldsymbol{\Phi}_I$ contains all the information in $\boldsymbol{\Psi}_+$. Equation (2.5) can be rewritten using (2.4) as

$$H(\boldsymbol{\Psi}_+ \mid \boldsymbol{\Phi}_I) = 0, \qquad (2.6)$$

which is mathematically equivalent to expressing $\boldsymbol{\Psi}_+$ as a function of $\boldsymbol{\Phi}_I$, namely, $\boldsymbol{\Psi}_+ = \mathcal{F}(\boldsymbol{\Phi}_I)$. Second, the residual term $\boldsymbol{\Phi}_R$ and the informative term $\boldsymbol{\Phi}_I$ must be independent, which requires

$$I(\boldsymbol{\Phi}_R; \boldsymbol{\Phi}_I) = 0. \qquad (2.7)$$

This also ensures that the residual component has no information about $\boldsymbol{\Psi}_+$, namely $I(\boldsymbol{\Phi}_R; \boldsymbol{\Psi}_+) = 0$, since $I(\boldsymbol{\Phi}_R; \boldsymbol{\Psi}_+) \leq I(\boldsymbol{\Phi}_R; \boldsymbol{\Phi}_I)$. The previous inequality is known as the data-processing inequality, and states that no transformation of a variable can increase its information content, which can only remain the same or decrease (Cover & Thomas 2006, Theorem 2.8.1). In addition, since $\boldsymbol{\Phi}_R$ and $\boldsymbol{\Phi}_I$ are statistically independent from (2.7), the equality

$$\|\boldsymbol{\Phi}\|^2 = \|\boldsymbol{\Phi}_I\|^2 + \|\boldsymbol{\Phi}_R\|^2 \qquad (2.8)$$

is satisfied. If $\boldsymbol{\Phi}$ contains no information about $\boldsymbol{\Psi}_+$, then $\|\boldsymbol{\Phi}_I\|^2/\|\boldsymbol{\Phi}\|^2 \approx 0$ and $\|\boldsymbol{\Phi}_R\|^2/\|\boldsymbol{\Phi}\|^2 \approx 1$. Conversely, if $\boldsymbol{\Phi}$ exclusively contains all the information necessary to understand $\boldsymbol{\Psi}_+$, then $\|\boldsymbol{\Phi}_I\|^2/\|\boldsymbol{\Phi}\|^2 = 1$. Note that, in general, $\boldsymbol{\Phi}_I$, $\boldsymbol{\Phi}_R$ and $\mathcal{F}$ are functions of ΔT, which has been omitted here for the sake of simplicity in the notation.

Since the Shannon information is based on the joint probability distribution of the variables, rather than their specific values, there may exist many functions that satisfy

(2.5) and (2.7). To identify a unique solution, we impose that the informative field $\boldsymbol{\Phi}_I(\boldsymbol{x}, t)$ is smooth. Note that, assuming $\boldsymbol{\Phi}(\boldsymbol{x}, t)$ is smooth, the previous condition also implies that the residual field must be smooth.

In summary, the necessary conditions that IND satisfies are as follows.

(i) The source variable is decomposed as the sum of the informative and residual contributions: $\boldsymbol{\Phi} = \boldsymbol{\Phi}_I + \boldsymbol{\Phi}_R$ (2.1).

(ii) The informative field contains all the information about the target variable in the future: $I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi}_I) = H(\boldsymbol{\Psi}_+)$ (2.5).

(iii) The informative and residual components share no information: $I(\boldsymbol{\Phi}_R; \boldsymbol{\Phi}_I) = 0$ (2.7).

(iv) The informative field is smooth.

2.2. *The IND of the target variable*

Alternatively, we can seek to decompose the target variable as $\boldsymbol{\Psi} = \boldsymbol{\Psi}_I + \boldsymbol{\Psi}_R$, where $\boldsymbol{\Psi}_I$ and $\boldsymbol{\Psi}_R$ are, respectively, the informative and residual components of $\boldsymbol{\Psi}$ with respect to $\boldsymbol{\Phi}_- = \boldsymbol{\Phi}(\boldsymbol{x}, t - \Delta T)$, with $\Delta T > 0$. The constraints to be satisfied are

$$I(\boldsymbol{\Phi}_-; \boldsymbol{\Psi}_I) = H(\boldsymbol{\Phi}_-), \quad I(\boldsymbol{\Psi}_R; \boldsymbol{\Psi}_I) = 0, \tag{2.9a,b}$$

together with the smoothness of $\boldsymbol{\Psi}_I$. In this case, $\boldsymbol{\Psi}_I$ corresponds to the part of $\boldsymbol{\Psi}$ that can explain the source variable $\boldsymbol{\Phi}$ in the past, while $\boldsymbol{\Psi}_R$ is the remaining term, which is agnostic to the information in the source variable.

2.3. *Approximate IND*

We frame the conditions of IND described in § 2.1 as a minimisation problem. To that end, several assumptions are adopted. First, (2.5) and (2.7) require calculating high-dimensional joint probability distributions, which might be impractical due to limited data and computational resources. The curse of high dimensionality comes from both the high dimensionality of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$, and the large number of points in $\boldsymbol{x}$. To make the problem tractable, we introduce the approximate IND, or aIND for short. First, the source and target variables are restricted to be scalars, Φ and Ψ, respectively. Second, we consider only two points in space: $\Phi(\boldsymbol{x}, t)$ and $\Psi_+(\boldsymbol{x} - \Delta \boldsymbol{x}, t + \Delta T)$, where $\boldsymbol{x}$ and $\Delta \boldsymbol{x}$ are fixed. This reduces the problem to the computation of two-dimensional joint probability distributions, which is trivially affordable in most cases, even enabling the use of experimental data.

Another difficulty arises from the constraint in (2.7), which depends on the unknown probability distribution of the variable $\Phi_R = \Phi - \Phi_I$, which adds to the complexity of the optimisation problem. To alleviate this issue, we seek to minimise $I(\Phi_R; \Phi_I)$ rather than include it as a hard constraint.

Finally, provided that Φ and Ψ_+ are smooth, minimising $\|\Phi - \Phi_I\|^2$ ensures that Φ_I is smooth too. Therefore, we include the mean square error as a penalisation term in the minimisation problem. Thus the formulation of the aIND is posed as

$$\arg \min_{\Phi_I, \mathcal{F}} I(\Phi_R; \Phi_I) + \gamma \|\Phi - \Phi_I\|^2, \quad \text{s.t.} \ \Psi_+ = \mathcal{F}(\Phi_I), \tag{2.10}$$

where $\gamma \geq 0$ is a regularisation constant, and $\Phi_R = \Phi - \Phi_I$. Equation (2.10) is solved by assuming that the mapping $\mathcal{F}$ is invertible over a given interval. This allows replacing $\Phi_I(t) = \mathcal{F}^{-1}(\Psi_+(t))$ over that interval in (2.10) and solving for $\mathcal{F}^{-1}$ using standard optimisation techniques. More details about the solution of (2.10) are provided in § A.1.

Equation (2.10) yields the informative and residual components for a given $\boldsymbol{x}$, $\Delta\boldsymbol{x}$ and t, denoted as $\Phi_{I,\Delta}(\boldsymbol{x}, t; \Delta\boldsymbol{x})$ and $\Phi_{R,\Delta}(\boldsymbol{x}, t; \Delta\boldsymbol{x})$, together with the mapping $\mathcal{F}$. We can find the best approximation to IND by selecting the value of $\Delta\boldsymbol{x}$ that maximises the informative component. To that end, we introduce the relative energy of $\Phi_{I,\Delta}$ as

$$E_I(\Delta\boldsymbol{x}; \boldsymbol{x}, \Delta T) = \frac{\|\Phi_{I,\Delta}\|^2}{\|\Phi\|^2}. \tag{2.11}$$

High values of E_I define the informative region of $\Phi_{I,\Delta}$ over Ψ_+, and constitute the information-theoretic generalisation of the two-point linear correlation (see Appendix C). We define $\Delta\boldsymbol{x}^{max}$ as the shift $\Delta\boldsymbol{x}$ that maximises E_I for a given $\boldsymbol{x}$ and ΔT. Hence we use $\Delta\boldsymbol{x} = \Delta\boldsymbol{x}^{max}$ for aIND, and simply refer to the variables in this case as Φ_I and Φ_R. During the optimisation, we ensure that $2\,I(\Phi_I; \Phi_R) < 0.03\,H(\Phi_I, \Phi_R)$ to guarantee that Φ_I and Φ_R are independent, and that (2.8) holds. We also assess *a posteriori* that $I(\Phi_R; \Psi_+)$ remains small for all $\boldsymbol{x}$ (see Appendix E).

Finally, we list below the main simplifications of aIND with respect to the general IND framework.

  (i) The source and the target variable are restricted to be scalars.

  (ii) The constraint in (2.7) is cast as the minimisation term in (2.10).

 (iii) The minimisation problem in (2.10) is computed for two points in space. The closest approximation to IND is achieved by selecting the value of $\Delta\boldsymbol{x}$ that maximises the magnitude of the informative component.

 (iv) Equation (2.10) is solved by assuming that the mapping $\mathcal{F}$ is invertible over a given interval.

Despite the simplifications above, aIND still successfully recovers the exact analytical solution in the validation cases presented in Appendix B, even outperforming correlation-based methods such as linear stochastic estimation (LSE) and extended POD (EPOD).

2.4. *Validation*

The methodology presented in § 2.1 and its numerical implementation (§ A.1) have been validated with several analytical examples. In this subsection, we discuss one of these examples that also illustrates the use and interpretation of the IND.

Consider the source and target fields:

$$\text{source} \quad \Phi(\boldsymbol{x}, t) = f(\boldsymbol{x}, t) + g(\boldsymbol{x}, t), \tag{2.12}$$

$$\text{target} \quad \Psi_+(\boldsymbol{x}, t) = \Psi(\boldsymbol{x}, t+1) = 0.5 f(\boldsymbol{x}, t)^2 - 0.2 f(\boldsymbol{x}, t) + \epsilon(\boldsymbol{x}, t), \tag{2.13}$$

where

$$f(\boldsymbol{x}, t) = 2\sin(2\pi x - 2t)\sin(2\pi y), \tag{2.14}$$

$$g(\boldsymbol{x}, t) = \tfrac{1}{5}\sin(7\sqrt{2}\,\pi x - 0.1t)\sin(8\sqrt{3}\,\pi y - 0.5t). \tag{2.15}$$

The source field is a combination of the streamwise travelling wave f and the lower amplitude, higher wavenumber travelling wave g. The target is a function of f and ϵ, where the latter is a random variable that follows the pointwise normal distribution with zero mean and standard deviation (σ) equal to 0.1: $\epsilon(\boldsymbol{x}, t) \sim \mathcal{N}(0, \sigma)$. Snapshots of Φ and Ψ are shown in figures 2(*a,b*), respectively.
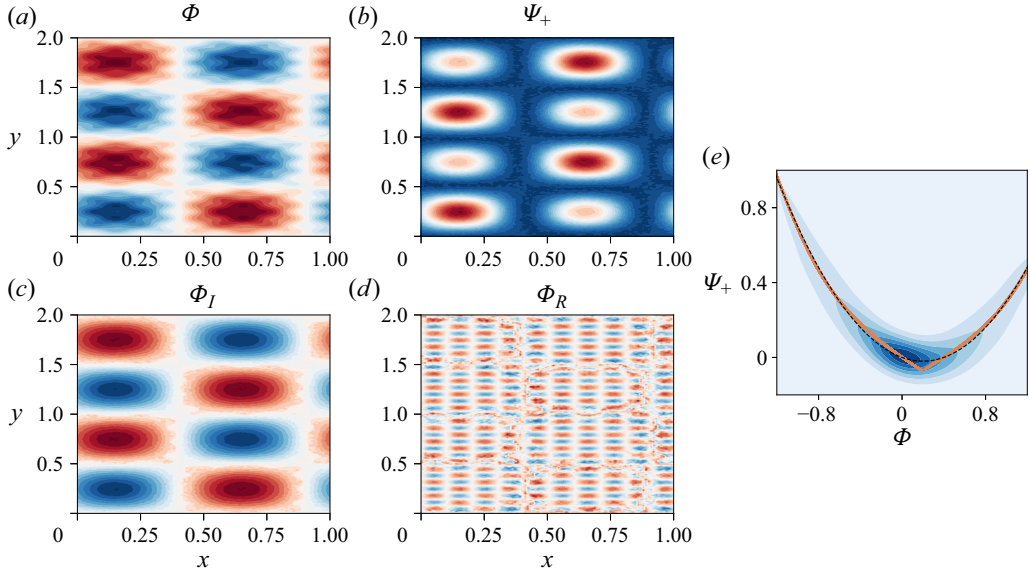
Figure 2. Validation of aIND for the system in (2.12). (*a–d*) Snapshots of Φ, Ψ_+, Φ_I and Φ_R, respectively. The contours range from the minimum value (dark blue) to the maximum value (dark red) for each quantity; these correspond to Φ, $\Phi_I \approx [-2, 2]$, $\Psi_+ \approx [-1.1, 1.4]$ and $\Phi_R \approx [-0.35, 0.35]$. (*e*) Contours of the joint probability (Φ, Ψ_+) from lower (white) to higher (blue) probability. The analytical solution is $\mathcal{F}^{exact}(\Phi_I) = 0.5\Phi_I^2 - 0.2\Phi_I$ (dashed black), and the numerical solution is $\mathcal{F}(\Phi_I)$ (orange).

For $\Delta T = 1$ and values of $\sigma \to 0$, the analytical solution of the IND is

$$\Phi_I^{exact} = f, \qquad \Phi_R^{exact} = g, \qquad (2.16a,b)$$

where the mapping to comply with $H(\Psi_+ \mid \Phi_I^{exact}) = 0$ is $\mathcal{F}^{exact}(\Phi_I) = 0.5\Phi_I^2 - 0.2\Phi_I$, and the residual term satisfies the condition $I(\Phi_I^{exact}; \Phi_R^{exact}) = 0$, since the variables are independent.

The results of solving the optimisation problem using aIND, denoted by Φ_I, Φ_R and $\mathcal{F}$, are displayed in figures 2(*c–e*). It can be observed that Φ_I approximates well the travelling wave represented by $\Phi_I^{exact} = f$. The small differences between Φ_I and Φ_I^{exact}, also appreciable in Φ_R, are localised at values $f \approx 0.2$ and can be explained by the small discrepancies between $\mathcal{F}$ and $\mathcal{F}^{exact}$ at the minimum, as seen in figure 2(*e*). These are mostly a consequence of ϵ and the numerical implementation (see § A.1), and they diminish as $\sigma \to 0$. Additional validation cases, together with a comparison of aIND with EPOD and LSE, can be found in Appendix B.

3. Results

We study the aIND of the streamwise (u), wall-normal (v) and spanwise (w) velocity fluctuations in a turbulent channel flow using as target the streamwise component of the shear stress at the wall, $\tau_x(x, z, t) = \rho \nu \, \partial U(x, 0, z, t)/\partial y$, where ρ is the fluid density, ν is the kinematic viscosity, U is the instantaneous streamwise velocity, and x, y and z are the streamwise, wall-normal and spanwise directions, respectively. The wall is located at $y = 0$. The data are obtained from direct numerical simulations in a computational domain of size $8\pi h \times 2h \times 4\pi h$ in the streamwise, wall-normal and spanwise directions, respectively, where h represents the channel half-height. The flow is driven by a constant

mass flux imposed in the streamwise direction. The Reynolds number, based on the friction velocity u_τ, is $Re_\tau = u_\tau h/\nu \approx 180$. Viscous units, defined in terms of ν and u_τ, are denoted by superscript $*$. The time step is fixed at $\Delta t^* = 5 \times 10^{-3}$, and snapshots are stored every $\Delta t_s^* = 0.5$. A description of the numerical solver and computational details can be found in Lozano-Durán *et al.* (2020).

The source and target variables for aIND are

$$\text{source} \quad u(\boldsymbol{x}, t),\ v(\boldsymbol{x}, t) \text{ or } w(\boldsymbol{x}, t), \tag{3.1}$$

$$\text{target} \quad \tau_{x,+} = \tau_x(x - \Delta x_\square^{max}, z - \Delta z_\square^{max}, t + \Delta T), \tag{3.2}$$

where $\square = u$, v or w. The aIND gives

$$u(\boldsymbol{x}, t) = u_I(\boldsymbol{x}, t) + u_R(\boldsymbol{x}, t), \tag{3.3}$$

$$v(\boldsymbol{x}, t) = v_I(\boldsymbol{x}, t) + v_R(\boldsymbol{x}, t), \tag{3.4}$$

$$w(\boldsymbol{x}, t) = w_I(\boldsymbol{x}, t) + w_R(\boldsymbol{x}, t), \tag{3.5}$$

where the informative and residual components are also functions of ΔT. We focus our analysis on $\Delta T^* \approx 25$ unless otherwise specified. This value corresponds to the time shift at which $H(\tau_{x,+} \mid \tau_x)/H(\tau_{x,+}) \gtrsim 0.97$, meaning that $\tau_{x,+}$ shares no significant information with its past. For $\Delta T^* > 25$, the value of $H(\tau_{x,+} \mid \tau_x)$ gradually increases towards $H(\tau_{x,+})$ asymptotically. This value is similar to that reported by Zaki & Wang (2021), who found using adjoint methods that wall observations at $\Delta T^* \approx 20$ are the most sensitive to upstream and near-wall velocity perturbations. The shifts $\Delta \boldsymbol{x}_\square^{max} = [\Delta x_\square^{max}, \Delta z_\square^{max}]$ for $\square = u$, v or w are computed by a parametric sweep performed in Appendix D. Their values are functions of y, but can be approximated as $\Delta \boldsymbol{x}_u^{max}/h \approx [-1, 0]$, $\Delta \boldsymbol{x}_v^{max}/h \approx [-1.2, 0]$ and $\Delta \boldsymbol{x}_w^{max}/h \approx [-0.8, \pm 0.15]$. Due to the homogeneity and statistical stationarity of the flow, the mapping $\mathcal{F}$ is a function of only y and ΔT. The validity of the approximations made in the aIND is discussed in Appendix E, where it is shown that the residual component of u contains almost no information about the future wall shear stress. For the interested reader, we also include the relative energy field $E_I(\Delta \boldsymbol{x}; \boldsymbol{x}, \Delta T^* = 25)$ of the three velocity components in Appendix D.

3.1. *Coherent structure of the informative and residual components of u, v and w to τ_x*

We start by visualising the instantaneous informative and residual components of the flow. We focus on the streamwise component, as it turns out to be the most informative to τ_x, as detailed below. Figure 3(*a*) displays iso-surfaces of $u(\boldsymbol{x}, t)$, revealing the alternating high- and low-velocity streaks attached to the wall, along with smaller detached regions. The informative and residual components $u_I(\boldsymbol{x}, t)$ and $u_R(\boldsymbol{x}, t)$ are shown in figures 3(*b,c*), respectively. The structures in u_I exhibit an alternating pattern similar to that in the original field, with the high- and low-velocity streaks located approximately in the same positions as $u(\boldsymbol{x}, t)$. These structures are also attached to the wall, but do not extend as far as the streaks in the original field, especially for $u_I(\boldsymbol{x}, t) > 0$. In contrast, the residual field $u_R(\boldsymbol{x}, t)$ lacks most of the elongated streaks close to the wall, but resembles $u(\boldsymbol{x}, t)$ far away, once the flow bears barely no information about $\tau_{x,+}$.

Figure 4 displays the root mean squared turbulence intensities as functions of the wall distance. Note that from the minimised term in (2.10), $\langle u^2 \rangle(y) = \langle u_I^2 \rangle(y) + \langle u_R^2 \rangle(y)$ (and similarly for the other components). From figure 4(*a*), we observe that $\langle u_I^2 \rangle^{1/2}$ is predominantly located within the region $y^* \leq 50$. This finding aligns with our earlier
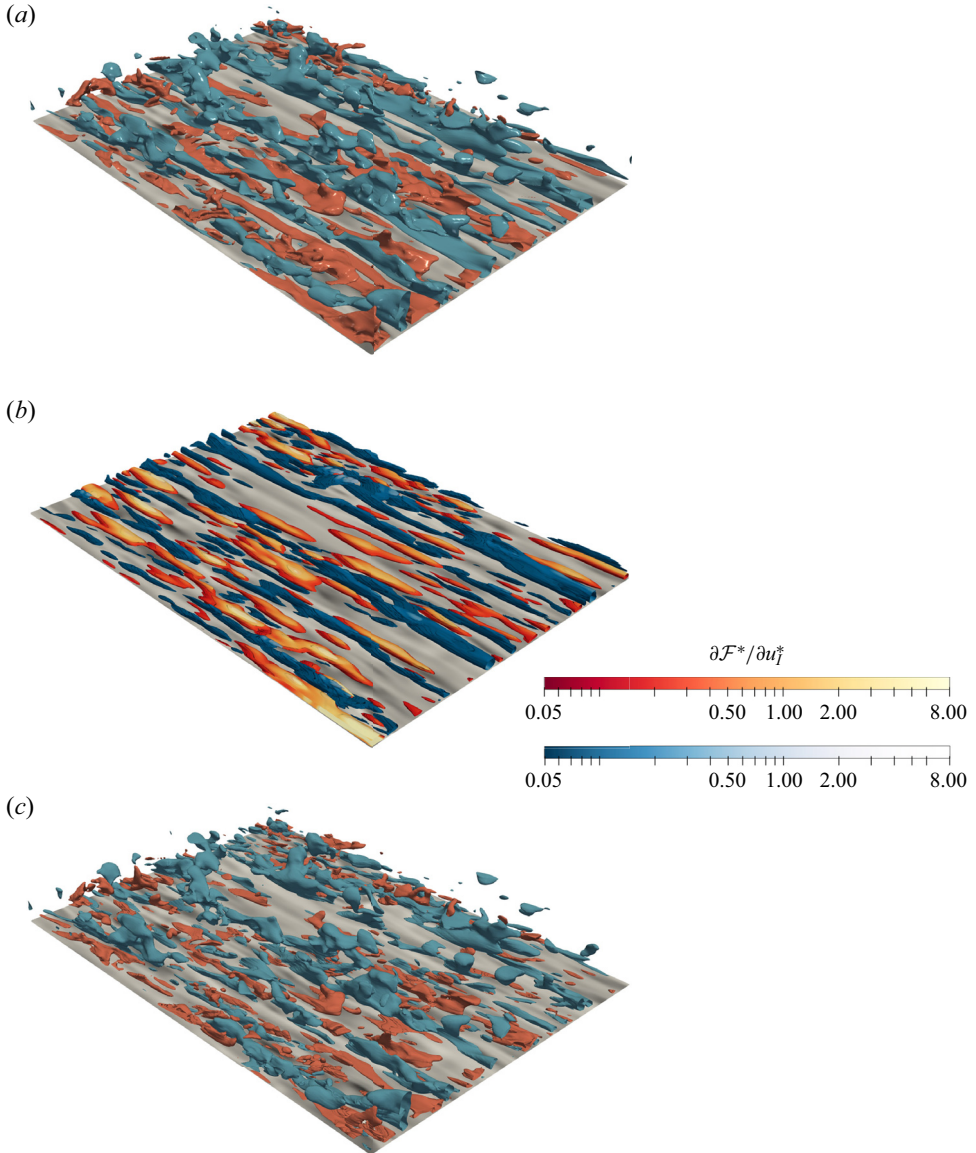
(*a*)

(*b*)

$\partial \mathcal{F}^* / \partial u_I^*$

(*c*)

Figure 3. Instantaneous flow structures: (*a*) iso-contours of the streamwise velocity fluctuations *u*; (*b*) iso-contours of the informative streamwise velocity fluctuations u_I; and (*c*) iso-contours of the residual streamwise velocity fluctuations u_R. In (*a*) and (*c*), the iso-contours correspond to blue for $u^* \approx -2.7$, red for $u^* \approx 2.7$; and (*b*) uses blue for $u_I^* \approx -1.8$, and orange for $u_I^* \approx 1.8$. The wall is coloured by the instantaneous wall shear stress at ΔT, from $\tau_x^* \approx 0.5$ (white) to $\tau_x^* \approx 2$ (black).

visual assessments from figure 3. The residual component $\langle u_R^2 \rangle^{1/2}$ also has a strong presence close to the wall, although it is shifted towards larger values of *y*. Interestingly, about half of the streamwise kinetic energy in the near-wall region originates from $\langle u_R^2 \rangle$, despite its lack of information about $\tau_{x,+}$. This phenomenon is akin to the inactive motions in wall turbulence (e.g. Townsend 1961; Jiménez & Hoyas 2008; Deshpande, Monty & Marusic 2021), with the difference that here inactive structures are interpreted as those
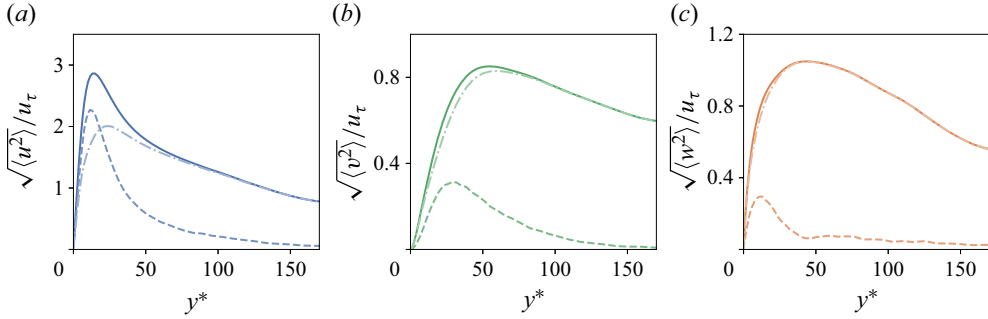
Figure 4. Root mean squared turbulence intensities of the (*a*) streamwise, (*b*) wall-normal and (*c*) spanwise velocity components: solid lines indicate the original flow field, dashed lines indicate the informative flow field, and dash-dotted lines indicate the residual field.

that do not reflect time variations of the wall shear stress. Another interesting observation is that $\langle u_I^2 \rangle^{1/2}$ peaks at $y^* \approx 10$, which is slightly below the well-known peak for $\langle u^2 \rangle^{1/2}$, whereas $\langle u_R^2 \rangle^{1/2}$ peaks at $y^* \approx 30$. This suggests that the near-wall peak of $\langle u^2 \rangle^{1/2}$ is controlled by a combination of active and inactive motions as defined above.

The root mean squared velocities for the cross-flow are shown in figures 4(*b*,*c*). The informative component of the wall-normal velocity $\langle v_I^2 \rangle^{1/2}$ is predominantly confined within the region $y^* \leq 70$, although its magnitude is small. The residual component $\langle v_R^2 \rangle^{1/2}$ is the major contributor to the wall-normal fluctuations across the channel height. The dominance of $\langle v_R^2 \rangle^{1/2}$ has important implications for control strategies in drag reduction, which are investigated in § 3.3. A similar observation is made for $\langle w^2 \rangle^{1/2}$, with $\langle w_I^2 \rangle^{1/2}$ being negligible except close to the wall for $y^* < 40$.

The statistical coherence of the informative and residual velocities in the wall-parallel plane is quantified with the two-point autocorrelation

$$C_{\phi\phi}(\Delta x, \Delta z; y_{ref}) = \frac{\langle \phi(x, y_{ref}, z, t)\, \phi(x + \Delta x, y_{ref}, z + \Delta z, t) \rangle}{\langle \phi(x, y_{ref}, z, t)^2 \rangle}, \qquad (3.6)$$

where ϕ is any component of the velocity field, and $y_{ref}^* = 15$. The autocorrelations are shown in figure 5 for the total, informative and residual components of the three velocities. The shape of the informative structure is elongated along the streamwise direction for the three correlations $C_{u_I u_I}$, $C_{v_I v_I}$ and $C_{w_I w_I}$. The results for *u*, shown in figure 5(*a*), reveal that u_I closely resembles the streaky structures of *u* in terms of streamwise and spanwise lengths. On the other hand, u_R consists of more compact and isotropic eddies in the (x, z)-plane. Figure 5(*b*) shows that v_I captures the elongated motions in *v*, which represents a small fraction of its total energy, whereas the shorter motions in *v* are contained in v_R. A similar conclusion is drawn for *w*, as shown in figure 5(*c*), where both *w* and w_R share a similar structure, differing from the elongated motions of w_I. The emerging picture from the correlations is that informative velocities tend to comprise streamwise elongated motions, whereas the remaining residual components are shorter and more isotropic. The differences between the structures of *v* and *w* and their informative counterparts are consistent with the lower intensities of v_I and w_I discussed in figure 4. It should be noted that the shape of the structures depends on the target variable, and they may differ for a different target quantity. For example, wall pressure fluctuations have been linked to more isotropic structures in the streamwise direction by several authors (Schewe
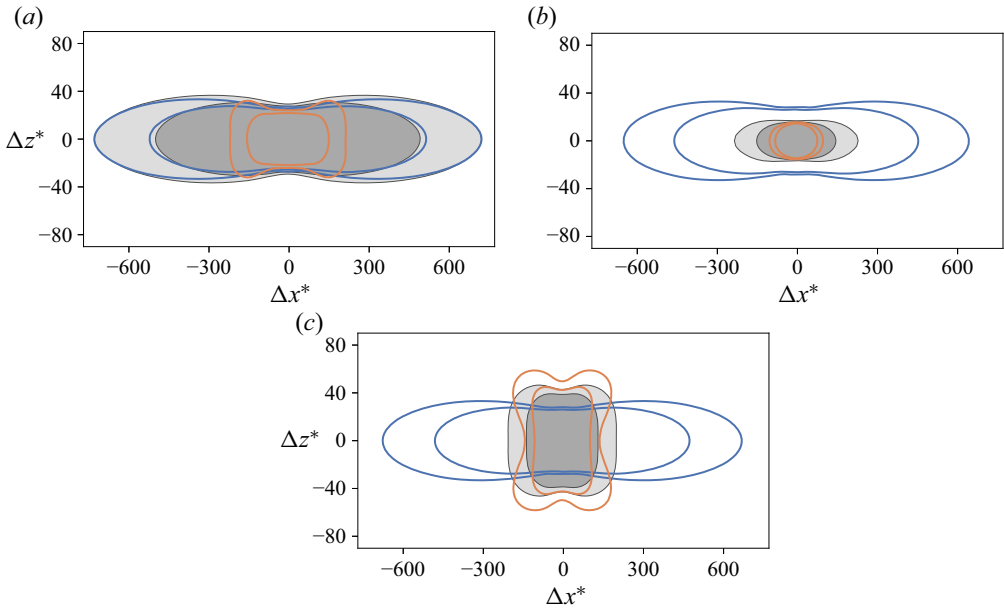
Figure 5. Autocorrelation coefficient of the velocity fluctuations in the $y^* = 15$ plane: (*a*) streamwise component, (*b*) wall normal component, and (*c*) spanwise component. Grey indicates the original field, blue indicates the informative field, and orange indicates the residual field. The contours correspond to $C_{\square\square} = [0.05, 0.1]$.

1983; Johansson, Her & Haritonidis 1987; Kim, Moin & Moser 1987; Ghaemi & Scarano 2013). The aIND may provide insights in this regard, as it has been noted in the literature that at least quadratic terms are needed to capture the interaction between the velocity and the wall pressure (Naguib, Wark & Juckenhöfel 2001; Murray & Ukeiley 2003).

We now analyse the average coherent structure of the flow in the (y, z)-plane. It is widely recognised in the literature that the most dynamically relevant energy-containing structure in wall turbulence comprises a low-velocity streak accompanied by a collocated roll (e.g. Kline *et al.* 1967; Kim *et al.* 1987; Farrell & Ioannou 2012; Lozano-Durán *et al.* 2012). A statistical description of this structure can be obtained by conditionally averaging the flow around low-velocity streaks. To this end, low-velocity streaks were identified by finding local minima of u at $y^* = 15$. For each streak, a local frame of reference was introduced with axes parallel to the original x, y and z coordinates. The origin of this local frame of reference is at the wall, such that its y-axis is aligned with the local minimum of u. The z-axis, denoted by Δz, points towards the nearest local maximum of u. This orientation ensures that any nearby high-speed streak is located in the region $\Delta z > 0$. Then the conditional average flow was computed by averaging $[u, v, w]$ over a window of size $\pm h$. The resulting conditionally averaged flow in the (y, z)-plane is shown in figure 6(*a*). This process was repeated for the informative and residual velocity fields using the same streaks identified previously for u. The conditionally averaged informative and residual velocities are shown in figures 6(*b,c*), respectively.

The conditional average velocity is shown in figure 6(*a*), which captures the structure of the low-/high-velocity streak pair and the accompanying roll characteristic of wall-bounded turbulence. The informative velocity (figure 6*b*) is dominated by streak motions, although these are smaller than the streaks of the entire field. The informative wall-normal velocity is present mostly within the streaks, while the informative spanwise
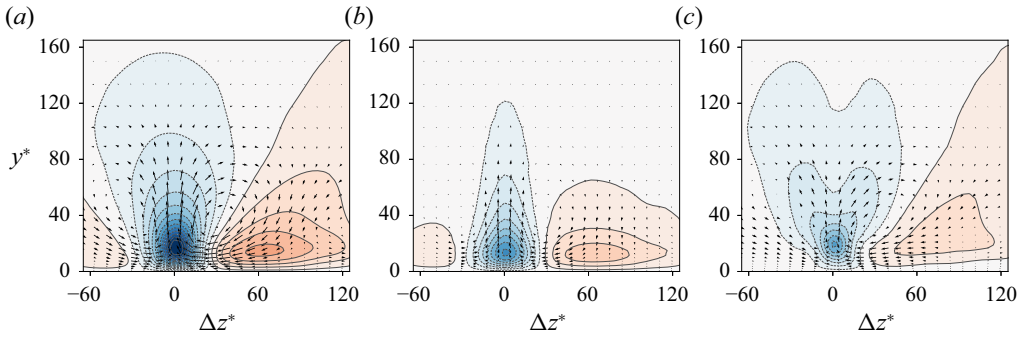
Figure 6. Conditionally averaged flow in the (y, z)-plane centred about low-velocity streaks: (*a*) original field, (*b*) informative field, and (*c*) residual field. The contours range from blue for $-5u_\tau$ to red for $5u_\tau$. The arrows display the in-plane velocity components ($v_\square$ and $w_\square$).

component is active close to the wall in the interface of the streak. Conversely, figure 6(*c*) shows that the residual velocity contains the large-scale streaks and the remaining spanwise motions. The emerging picture is that the informative component of the velocity contributing to the wall shear stress consists of smaller near-wall streaks collocated with vertical motions (i.e. sweeps and ejections), and spanwise velocity at the near-wall root of the roll. This informative structure is embedded within a larger-scale streak–roll structure of residual velocity, which bears no information about the wall shear stress.

We close this subsection by analysing the mappings $\tau_{x,+} = \mathcal{F}_u(u_I)$, $\tau_{x,+} = \mathcal{F}_v(v_I)$, $\tau_{x,+} = \mathcal{F}_w(w_I)$ obtained from the constraints $H(\tau_{x,+} \,|\, u_I) = 0$, $H(\tau_{x,+} \,|\, v_I) = 0$, $H(\tau_{x,+} \,|\, w_I) = 0$, respectively. The mapping are depicted in figure 7 at the wall-normal position where the energy for u_I, v_I and w_I is maximum, namely, $y^* \approx 8$, 19 and 6, respectively (see Appendix D). Figure 7(*a*) reveals an almost linear relationship between u_I and $\tau_{x,+}$ within the range $0 \le \tau_{x,+}^* \le 2$. Negative values of u_I align with $\tau_{x,+}^* < 1$, while positive values of u_I correspond to $\tau_{x,+}^* > 1$. This is clearly a manifestation of the proportionality between streak intensity and τ_x, such that higher streamwise velocities translate into higher wall shear stress by increasing $\partial U / \partial y$. However, the process saturates, and a noticeable change in the slope occurs for larger values of $\tau_{x,+}$, leading to u_I values that are relatively independent of $\tau_{x,+}$. This finding indicates that u_I provides limited information about high values of $\tau_{x,+}$ at the time scale $\Delta T^* = 25$. In other words, minor uncertainties in u_I result in significant uncertainties in $\tau_{x,+}$ after ΔT.

The effect of ΔT on $\mathcal{F}_u(u_I)$ is also analysed in figure 7(*a*). The main effect of decreasing ΔT^* is to decrease the slope of $\mathcal{F}_u(u_I)$ for $u_I^* > 5$. This result reveals that there exists a time horizon beyond which it is not possible to predict extreme events of wall shear stress from local fluctuations. Hence extreme values of the wall shear stress can be attributed to almost instantaneous high fluctuations of the streamwise velocity. The latter is in agreement with Guerrero, Lambert & Chin (2020), who linked extreme positive wall shear stresses with the presence of high-momentum regions created by quasi-streamwise vortices.

The mapping of v_I is shown in figure 7(*b*), which demonstrates again a nearly linear, albeit negative, relationship between v_I and $\tau_{x,+}$ in the range $0 \le \tau_{x,+}^* \le 2$. Positive values of v_I are indicative of $\tau_{x,+}^* < 1$, whereas negative values imply $\tau_{x,+}^* > 1$. Note that changes in the value of $\tau_{x,+}$ encompass either $u_I > 0$ and $v_I < 0$, or $u_I < 0$ and $v_I > 0$, revealing a connection between the dynamics of $\tau_{x,+}$ and the well-known sweep and ejection motions in wall-bounded turbulence (Wallace, Eckelman & Brodkey 1972; Wallace 2016). The
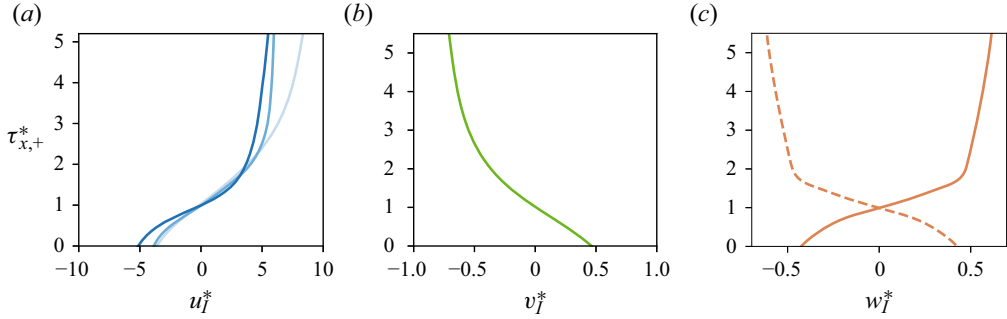
Figure 7. Mapping functions of the informative contributions of (*a*) streamwise, (*b*) wall-normal, and (*c*) spanwise velocity fluctuations to the streamwise wall shear stress for $\Delta T^* = 25$. Plot (*a*) also shows the effect of the time lag ΔT in the mapping $\tau_x = \mathcal{F}(u_I)$. Line colours correspond as follows: dark blue for $\Delta T^* \approx 25$, cyan for $\Delta T^* \approx 10$, and light blue for $\Delta T^* \approx 5$. In (*c*), the solid line corresponds to $+\Delta z_w^{max}$, and the dashed line to $-\Delta z_w^{max}$.

mappings also show that excursions into large wall shear stresses are caused by sweeps. Analogous to u_I, the value of v_I remains approximately constant for $\tau_{x,+}^* > 2$. Beyond that threshold, v_I provides no information about $\tau_{x,+}$.

The mapping of w_I presents two maxima ($\pm\Delta z_w^{max}$) due to the spanwise symmetry of the flow. The results for each maximum, shown in figure 7(*c*), are antisymmetric with respect to w_I. Similarly to u_I and v_I, there is an almost linear relationship between w_I and $\tau_{x,+}$ in the range $0 \leq \tau_{x,+}^* \leq 2$. For $+\Delta z_w^{max}$, negative values of w_I indicate $\tau_{x,+}^* < 1$, whereas positive values are linked to $\tau_{x,+}^* > 1$. The opposite is true for $-\Delta z_w^{max}$. Low values of $\tau_{x,+}$ are connected to low u_I, and positive (negative) values of w_I for $+\Delta z_w^{max}$ ($-\Delta z_w^{max}$). This outcome is consistent with the conditional average flow from figure 6, where it was shown that the information transfer between w_I and $\tau_{x,+}$ is mediated through the bottom part of the roll structure that accompanies high-/low-velocity streaks. The saturation of the influence of w_I to intense values of the wall shear stress is again observed for $\tau_{x,+}^* \gtrsim 2$.

The information provided by the mappings can be embedded into the instantaneous coherent structures. In figure 3(*b*), the $u_I(\boldsymbol{x}, t)$ structures are coloured by the local value of $\partial\mathcal{F}/\partial u_I$. This metric serves as a measure of the uncertainty in the wall shear stress as a function of u_I. Low values of $\partial\mathcal{F}/\partial u_I$ are associated with low uncertainty in $\tau_{x,+}$. This implies that small changes in u_I result in small changes in $\tau_{x,+}$. On the other hand, high values of $\partial\mathcal{F}/\partial u_I$ are associated with high uncertainty in $\tau_{x,+}$, such that small variations in u_I result in large changes in $\tau_{x,+}$. Interestingly, figure 3(*b*) shows that low-speed streaks – associated with ejections – are connected to low uncertainty values for τ_x along their entire wall-normal extent. On the contrary, the high-speed streaks of u_I, linked to extreme events, carry increasing uncertainty in τ_x (indicated by the light yellow colour) as they move further away from the wall.

3.2. *Reduced-order modelling: reconstruction of the wall shear stress from u*

We evaluate the predictive capabilities of the informative and residual components of the streamwise velocity fluctuations to reconstruct the wall shear stress in the future. The main aim of this subsection is to illustrate that when u is used as the input for developing a model, the resulting model exclusively utilises information from u_I, while u_R is disregarded.
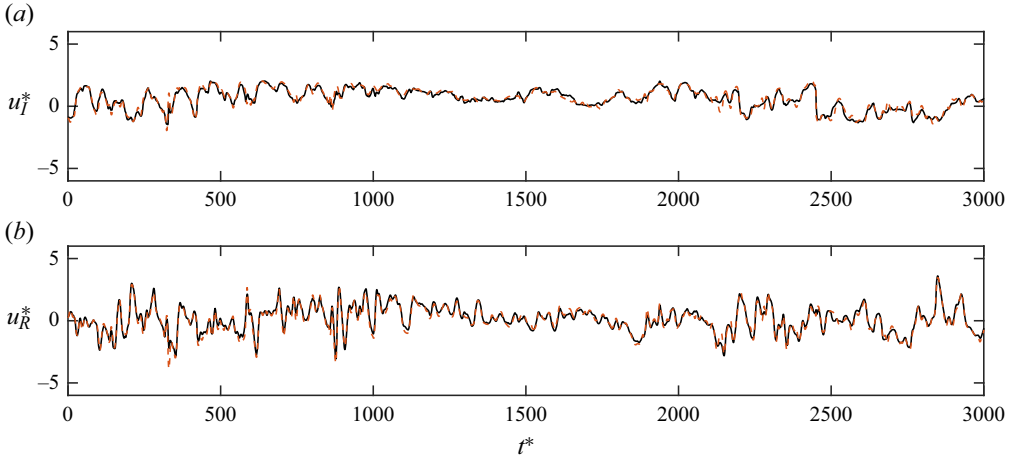
1000 A95-14

Figure 8. Example of approximate decomposition of u into (a) informative $\tilde{u}_I$ and (b) residual $\tilde{u}_R$ components using the ANN from (3.7) without need of $\tau_{x,+}$. Lines correspond to solid black for the actual u_I and u_R, and orange dashed for the predicted values.

Two scenarios are considered. In the first case, we devise a model for the pointwise, temporal forecasting of $\tau_{x,+}$ using pointwise data of u. In the second scenario, the spatially two-dimensional wall shear stress is reconstructed using u data from a wall-parallel plane located at given distance from the wall.

First, we discuss the pointwise forecasting of $\tau_{x,+}$ using pointwise data of u. We aim to predict the future of the wall shear stress at one point at the wall, $\tau_{x,+} = \tau_x(x_0, z_0, t + \Delta T)$, where x_0 and z_0 are fixed, and the time lag is $\Delta T^* = 25$. Three models are considered, using as input $u(x_0, t)$, $u_I(x_0, t)$ and $u_R(x_0, t)$, respectively, where $x_0 = [x_0 + \Delta x_u^{max}, y_{ref}, z_0]$ and $y_{ref}^* \approx 10$. The data are extracted from a simulation with the same set-up and friction Reynolds number as in § 3.1 but in a smaller computational domain ($\pi h \times 2h \times \pi/2h$). Note that all the points $[x_0, z_0]$ are statistically equivalent and can be used to train the model.

As a preliminary step to developing the forecasting models, we use a feedforward artificial neural network (ANN) to separate u into u_I and u_R without need of $\tau_{x,+}$. This step is required to make the models predictive, as in a practical case, the future of τ_x is unknown and cannot be used to obtained the informative and residual components. The model is given by

$$[\tilde{u}_I(x_0, t), \tilde{u}_R(x_0, t)] = \text{ANN}_{I,R}(u(x_0, t), u(x_0, t - \delta t), \dots, u(x_0, t - p\,\delta t)), \qquad (3.7)$$

where the tilde in $\tilde{u}_I$ and $\tilde{u}_R$ denotes estimated quantities, $\delta t^* = 0.5$, and $p = 1000$ is the number of time lags considered. Multiple time lags are required for predicting $\tilde{u}_I$ and $\tilde{u}_R$, in the same manner as time series of u and $\tau_{x,+}$ were used to compute u_I. The function $\text{ANN}_{I,R}$ comprises 6 hidden layers with 50 neurons per layer and ReLU activation functions. The approximately 700 000 samples are divided into 80 % for training and 20 % for validation. The Adam algorithm (Kingma & Ba 2017) is used to find the optimum solution. An example of the approximate decomposition from (3.7) is shown in figure 8.

The three ANN models trained to forecast $\tau_{x,+}$ are

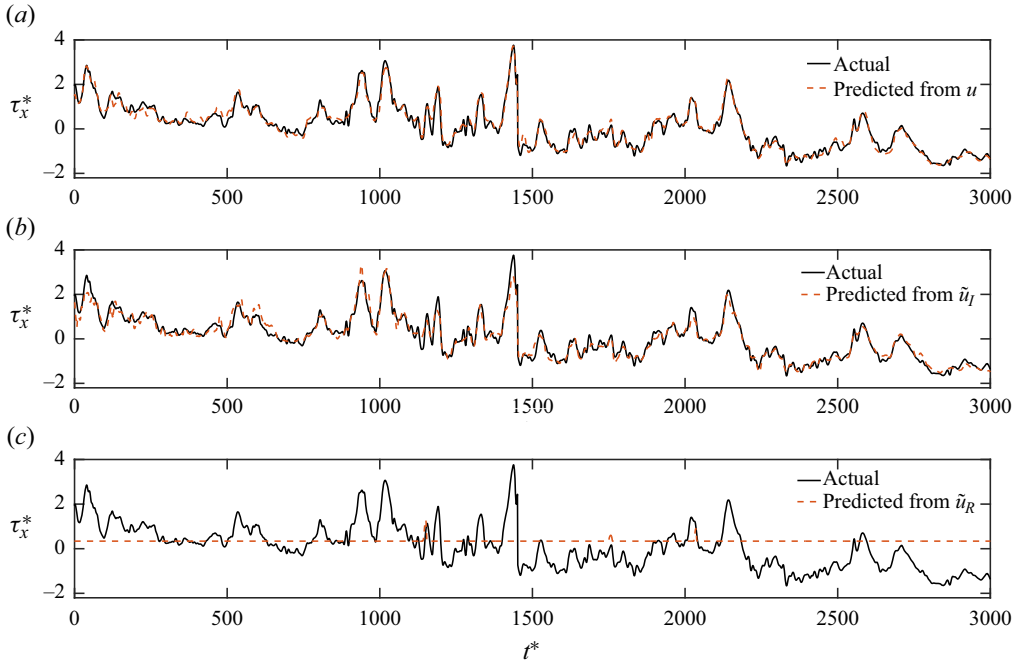$$\tilde{\tau}_x^I(x_0, z_0, t + \Delta T) = \text{ANN}_I(\tilde{u}_I(x_0, t)), \qquad (3.8a)$$

1000 A95-15

Figure 9. Temporal reconstruction of the wall shear stress using ANNs in (3.8) trained with (*a*) u, (*b*) $\tilde{u}_I$, and (*c*) $\tilde{u}_R$ as inputs, respectively. Solid black lines correspond to the actual wall shear stress; dashed orange lines correspond to the ANN reconstruction.

$$\tilde{\tau}_x^R(x_0, z_0, t + \Delta T) = \text{ANN}_R(\tilde{u}_R(x_0, t)), \tag{3.8b}$$

$$\tilde{\tau}_x^U(x_0, z_0, t + \Delta T) = \text{ANN}_U(u(x_0, t - \delta t), \dots, u(x_0, t - p\,\delta t)). \tag{3.8c}$$

Note that (3.8*a*) and (3.8*b*) use only one time step of $\tilde{u}_I$ and $\tilde{u}_R$, respectively, while (3.8*c*) incorporates multiple time lags of u. This approach is chosen because (3.7) (used to predict $\tilde{u}_I$ and $\tilde{u}_R$) also depends on multiple time lags of u. By training (3.8*c*) using the same time lags as (3.7), the predictions for $\tilde{\tau}_x^U$ rely on a model that accesses an equivalent amount of information about past states of the flow as do the models for predicting $\tilde{\tau}_x^I$ and $\tilde{\tau}_x^R$. This ensures a fair comparison among models.

The forecasting of the wall shear stress by the three models is illustrated in figure 9. The results indicate that the predictions based on u and $\tilde{u}_I$ are comparable, with relative mean squared errors 18 % and 22 %, respectively. The marginally larger error from the model using $\tilde{u}_I$ as input arises from inaccuracies within the ANN responsible for decomposing u into $\tilde{u}_I$ and $\tilde{u}_R$. In a perfect scenario, the forecasting errors using either u or $\tilde{u}_I$ as input would be identical, implying that $\tilde{u}_I$ contains all the information in u to make predictions. In contrast, the model that utilises the residual component $\tilde{u}_R$ fails to accurately predict the wall shear stress (approximately by 100 % error), yielding values that are nearly constant and close to the time average of $\tilde{\tau}_x$. These findings demonstrate that when u is used as input, the model extracts predictive information from $\tilde{u}_I$, while $\tilde{u}_R$ provides no predictive value.

It is important to clarify that we are not advocating for the separation of inputs into informative and residual components as a standard practice for training models. Instead, our goal is to illustrate that the training process of a model implicitly discriminates
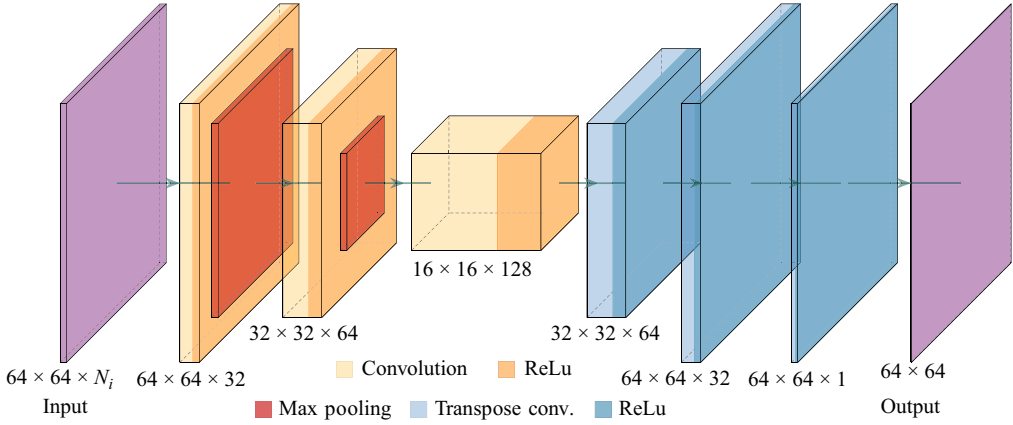
1000 A95-16

Figure 10. Schematic of the architecture for the (temporal) CNNs in (3.9) and (3.10). The numbers below the convolution and transpose convolution blocks correspond to the size of the filter and the number of channels applied. For $\mathrm{CNN}_{I,R}$ in (3.9), and CNN_U in (3.10c), $N_i = 500$; for CNN_I and CNN_R from (3.10a) and (3.10b), $N_i = 1$.

between these components, supporting our claim that all the necessary information for reduced-order modelling is encapsulated in u_I. An interesting consequence of this property is that the characteristics and structure of u_R are not useful for understanding the predictive capabilities of the model; instead, they help to discern which factors are irrelevant. For further discussion on the role of information in predictive modelling, the reader is referred to Lozano-Durán & Arranz (2022) and Yuan & Lozano-Durán (2024).

Next, we reconstruct the spatially varying wall shear stress $\tau_x(x, z, t + \Delta T)$ using $u(\mathbf{x}_{ref}, t)$, where $\mathbf{x}_{ref} = [x, y_{ref}, z]$ and $y_{ref}^* = 10$. The steps followed are analogous to those described above for the time signal prediction. First, we train a model to approximately decompose $u(\mathbf{x}_{ref}, t)$ into its informative and residual parts without requiring information about $\tau_x(x, z, t + \Delta T)$. To that end, we use a temporal convolutional neural network (CNN) (Long, Shelhamer & Darrell 2015; Guastoni *et al.* 2021) of the form

$$[\tilde{u}_I(\mathbf{x}_{ref}, t), \tilde{u}_R(\mathbf{x}_{ref}, t)] = \mathrm{CNN}_{I,R}(u(\mathbf{x}_{ref}, t), u(\mathbf{x}_{ref}, t - \delta t), \ldots, u(\mathbf{x}_{ref}, t - p\,\delta t)), \quad (3.9)$$

where $p = 500$ and $\delta t^* = 0.5$. The CNN is designed to process input data shaped as three-dimensional arrays, where dimensions represent spatial coordinates and temporal slices. The CNN comprises an image input layer, followed by three blocks consisting each of a convolutional layer, batch normalisation, and a ReLU activation function. Spatial dimensions are reduced through successive max pooling layers, while feature maps are subsequently upscaled back to original dimensions via transposed convolutional layers with ReLU activations. Further details of the CNN are provided in figure 10. A total of 12 000 snapshots are used, split into training (80 %) and validation (20 %). An example of the approximate decomposition from (3.9) is shown in figure 11.

The three models to predict the two-dimensional wall shear stress are

$$\tilde{\tau}_x^I(x, z, t + \Delta T) = \mathrm{CNN}_I(\tilde{u}_I(\mathbf{x}_{ref}, t)), \quad (3.10a)$$

$$\tilde{\tau}_x^R(x, z, t + \Delta T) = \mathrm{CNN}_R(\tilde{u}_R(\mathbf{x}_{ref}, t)), \quad (3.10b)$$

$$\tilde{\tau}_x^U(x, z, t + \Delta T) = \mathrm{CNN}_U(u(\mathbf{x}_{ref}, t), u(\mathbf{x}_{ref}, t - \delta t), \ldots, u(\mathbf{x}_{ref}, t - p\,\delta t)). \quad (3.10c)$$
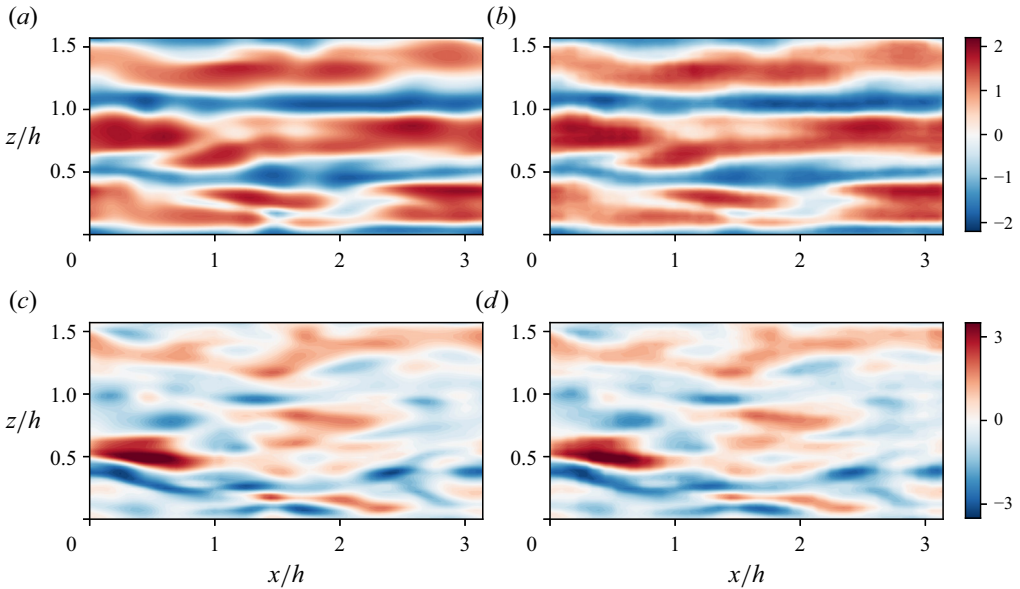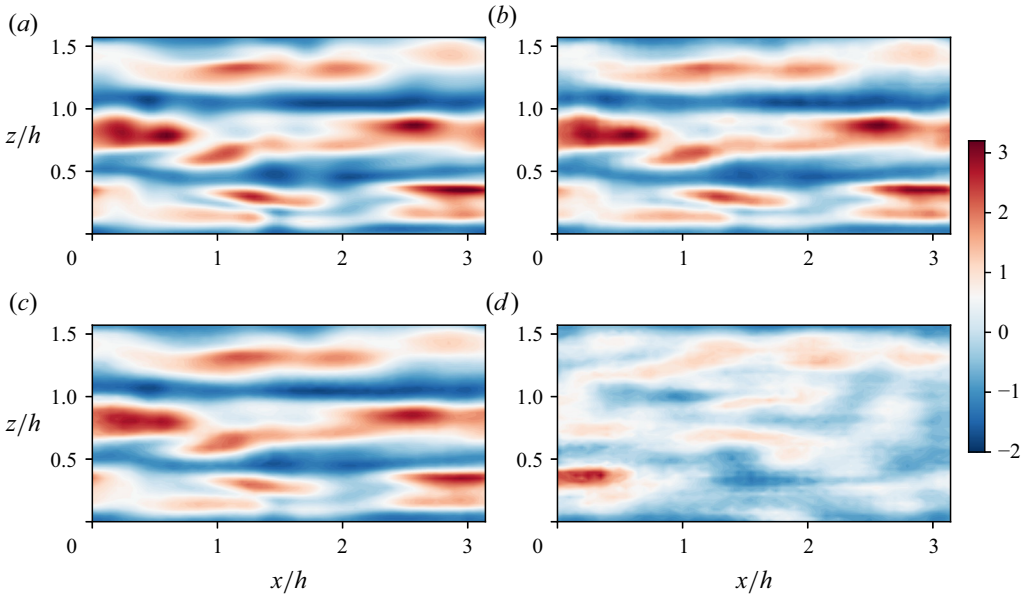
1000 A95-17

Figure 12. Example of spatial reconstruction of the instantaneous wall shear stress (a) using CNNs trained with (b) u, (c) $\tilde{u}_I$, and (d) $\tilde{u}_R$ as inputs, respectively.

the observability of a variable ($\tau_{x,+}$) in terms of the knowledge gained from another variable (v) as

$$O_{v \to \tau_{x,+}} = \frac{I(\tau_{x,+}; v)}{H(\tau_{x,+})}. \tag{3.11}$$

The variable $\tau_{x,+}$ is said to be perfectly observable with respect to v when $O_{v \to \tau_{x,+}} = 1$, i.e. there is no uncertainty in the state to be controlled conditioned to knowing the state of the sensor. Conversely, $\tau_{x,+}$ is completely unobservable when $O_{u \to \tau_{x,+}} = 0$, i.e. the sensor does not have access to any information about $\tau_{x,+}$. The greater the observability, the more information is available for controlling the system. By substituting (2.5) and (2.7) into (3.11), it is easy to show that $\tau_{x,+}$ is unobservable with respect to the residual component ($O_{v_R \to \tau_{x,+}} = 0$), and perfectly observable from the perspective of the informative component ($O_{v_I \to \tau_{x,+}} = 1$).

Figure 13 shows a schematic of the problem set-up for opposition control in a turbulent channel flow. The channel is as in § 3.2, but the wall-normal velocity at the wall is replaced by $v(x, 0, z, t) = f(v(x, y_s, z, t))$, where y_s is the distance to the sensing plane, and f is a user-defined function. In the original formulation by Choi *et al.* (1994), $f \equiv -v(x, y_s, z, t)$, hence the name opposition control. Here, we set $y_s^* \approx 14$, which is the optimum wall distance reported in previous works (Chung & Talha 2011; Lozano-Durán & Arranz 2022). Two Reynolds numbers are considered, $Re_\tau = 180$ and 395.

We split $v(x, y_s, z, t)$ into its informative (v_I) and residual (v_R) components to $\tau_x(x, z, t)$. Three controllers are investigated. In the first case, the function of the controller f is such that it uses only the informative component of $v(x, y_s, z, t)$, namely $f(v(x, y_s, z, t)) \equiv -v_I(x, y_s, z, t)$. In the second case, the controller uses the residual component $f(v(x, y_s, z, t)) \equiv -v_R(x, y_s, z, t)$. Finally, the third controller follows the original formulation $f(v(x, y_s, z, t)) \equiv -v(x, y_s, z, t)$.
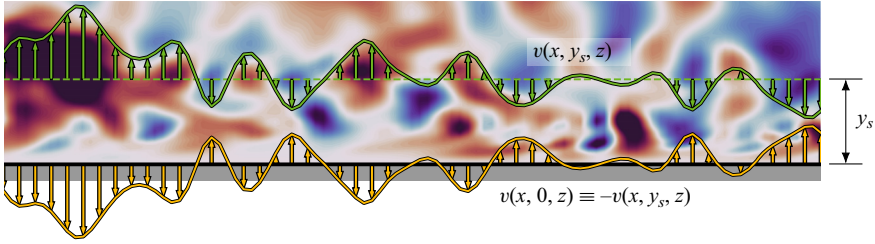
Figure 13. Schematic of the opposition control. The contour corresponds to the instantaneous vertical velocity on an (x, y)-plane: the green line depicts the instantaneous velocity at the sensing plane y_s, and the yellow line depicts the velocity at the wall imposed by the classical opposition control technique. The colour map ranges from blue for $v^* = -1.8$ to red for $v^* = 1.8$.

This is a more challenging application of the IND due to the dynamic nature of the control problem. When the flow is actuated, the dynamics of the system changes, and the controller should re-compute v_I (or v_R) for the newly actuated flow. This problem is computationally expensive, and we resort to calculating an approximation. The control strategy is implemented as follows.

   (i) A simulation is performed with $f \equiv -v(x, y_s, z, t)$, corresponding to the original version of opposition control.
  (ii) The informative term (v_I) of $v(x, y_s, z, t)$ related to the wall shear stress $\tau_x(x, z, t)$ is extracted for $\Delta T = 0$.
 (iii) We find an approximation of the controller, such that $\tilde{v}_I = f(v) \approx -v_I$. To obtain this approximation, we solve the minimisation problem

$$\arg \min_{\tilde{v}_I} \|v_I - \tilde{v}_I\|^2 + \gamma \, \frac{I(\tau_x; \tilde{v}_R)}{H(\tau_x)}, \qquad (3.12)$$

where $\gamma = 0.75$. The approximated informative term is modelled as a feedforward ANN with 3 layers and 8 neurons per layer.
 (iv) Two new simulations are conducted, using either $\tilde{v}_I$ or $\tilde{v}_R = v - \tilde{v}_I$ for opposition control.

Note that the devised controller can be applied in real time (i.e. during simulation runtime), since the estimated information component $\tilde{v}_I(t)$ is computed using only information from the present time instant, $v(t)$.

Figure 14 summarises the drag reduction for the three scenarios, namely $f \equiv -v(x, y_s, z, t), f \equiv -\tilde{v}_I(x, y_s, z, t)$ and $f \equiv -\tilde{v}_R(x, y_s, z, t)$. The original opposition control achieves drag reductions approximately 22 % and 24 % for $Re_\tau = 180$ and $Re_\tau = 395$, respectively. Similar reductions in drag using the same controller have been documented in the literature (Chung & Talha 2011; Luhar, Sharma & McKeon 2014). The values show a marginal dependency on Re_τ, in agreement with previous studies (Iwamoto, Suzuki & Kasagi 2002). Opposition control based on $\tilde{v}_I$ yields a moderate increase in drag reduction with a 24 % and 26 % drop for each Re_τ, respectively. Conversely, the drag reduction is only up to 7 % for the control based on the estimated residual velocity, $\tilde{v}_R$. Note that v_I is the component of v with the highest potential to modify the drag. Whether the drag increases or decreases depends on the specifics of the controller. On the other hand, the residual component v_R is expected to have a minor impact on the drag. As such, one might anticipate a 0 % drag reduction by using v_R. However, the approximation $\tilde{v}_R$ retains
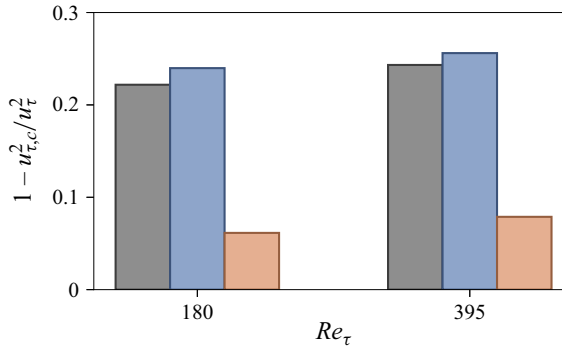
Figure 14. Drag reduction, computed as $1 - u_{\tau,c}^2/u_\tau^2$, where $u_{\tau,c}$ is the friction velocity for the controlled case. Each colour corresponds to a different controller: black for $f \equiv -v$, blue for $f \equiv -\tilde{v}_I$, and orange for $f \equiv -\tilde{v}_R$.

some information from the original velocity for intense values of the latter, which seems to reduce the drag on some occasions. Simulations using $f \equiv -k\tilde{v}_R$ – with k adjusted to $f \sim \|v(x, y_s, z, t)\|^2$ – were also conducted, yielding no additional improvements in the drag reduction beyond 8 %. It is also interesting to note that after performing steps (i)–(iv) of the control strategy, the informative content in v increases substantially (from $E_I^v \approx 0.1$ to $E_I^v \approx 0.8$). This phenomenon exposes the dynamic nature of the control problem highlighted above.

Figures 15(a,c) show the wall-normal velocity in the sensing plane for the controlled cases at $Re_\tau = 180$ with $f \equiv -\tilde{v}_I$ and $f \equiv -\tilde{v}_R$, respectively. Larger velocity amplitudes are observed in figure 15(c) compared to figure 15(a), indicating that higher Reynolds stresses are expected, which aligns with a larger average wall shear stress. On the other hand, figures 15(b,d) display the negative wall-normal velocity imposed at the boundary for the cases with $f \equiv -\tilde{v}_I$ and $f \equiv -\tilde{v}_R$, respectively. The informative component $\tilde{v}_I$ closely resembles the original velocity but with smaller amplitudes at extreme events of v. This appears to play a slightly beneficial role in drag reduction. Conversely, figure 15(d) shows that the estimated residual component is negligible except for large values of v. This is responsible for the smaller reduction in the mean drag. Although not shown, similar flow structures are observed for $Re_\tau = 395$, and the same discussion applies. In summary, we have utilised an example of opposition control in a turbulent channel to demonstrate the utility of IND. However, it is important to emphasise that the primary focus of this section is not on the real-time applicability or the performance of the control in this specific case. Instead, the main message that we aim to convey is more fundamental: the informative component of the variable measured by the sensor holds the essential information needed to develop successful control strategies, while the residual component is not useful in this regard.

4. Conclusions

We have presented informative and non-informative decomposition (IND), a method for decomposing a flow field into its informative and residual components relative to a target field. The informative field contains all the information necessary to explain the target variable, contrasting with the residual component, which holds no relevance to the target variable. The decomposition of the source field is formulated as an optimisation problem based on mutual information. To alleviate the computational cost and data requirements
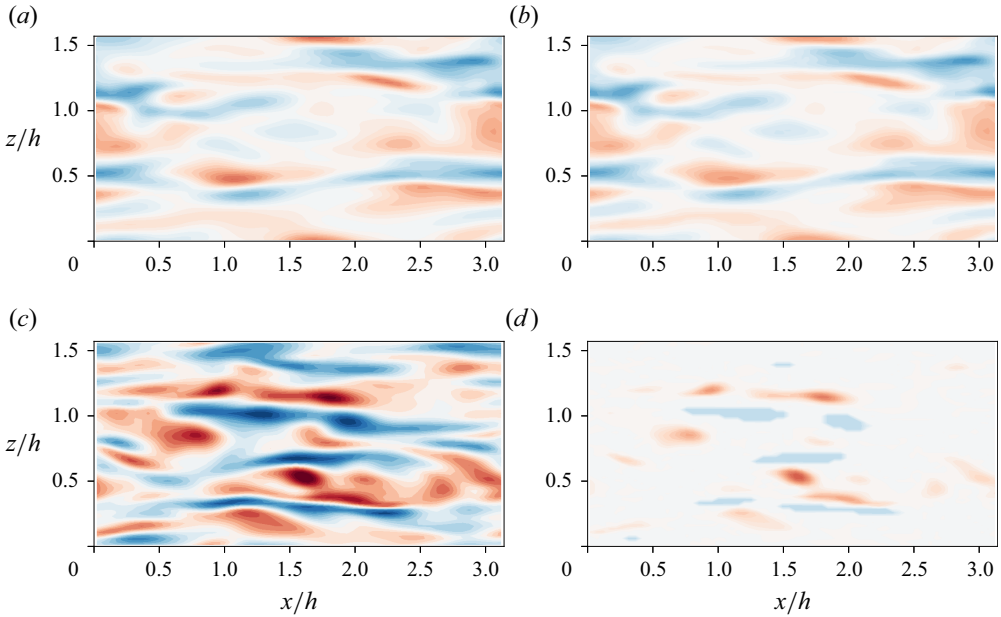
Figure 15. (*a*) Wall-normal velocity in the sensing plane, $v(x, y_s, z)$ and (*b*) the minus velocity imposed at the wall for the case $f \equiv -\tilde{v}_I$. (*c,d*) Same as (*a,b*), but for the case $f \equiv -\tilde{v}_R$. Contours range from blue for $v^* = -0.9$ to red for $v^* = 0.9$.

of IND, we have introduced an approximate solution, referred to as aIND. This approach still ensures that the informative component retains the information about the target, by minimising the mutual information between the residual and the target in a pointwise manner.

The IND is grounded in the fundamental principles of information theory, offering key advantages over other methods. As such, it is invariant under shifting, rescaling, and, in general, nonlinear $\mathcal{C}^1$-diffeomorphism transformations of the source and target variables (Kaiser & Schreiber 2002). The method is also fully nonlinear, and does not rely on simplifications such as the Gaussianity of the variables. This makes IND a suitable tool for studying turbulent phenomena, which are intrinsically nonlinear. In contrast, other linear correlation-based methods, such as LSE and EPOD, are not well equipped to capture nonlinearities in the flows. Additionally, we have shown that the pointwise formulation of the method (aIND) represents a cost-effective and memory-efficient implementation of IND without sacrificing performance compared to correlation-based methods. This approach also allows for the assimilation of experimental data.

The method has been applied to study the information content of the velocity fluctuations in relation to the wall shear stress in a turbulent channel flow at $Re_\tau = 180$. Our findings have revealed that streamwise fluctuations contain more information about the future wall shear stress than the cross-flow velocities. The energy of the informative streamwise velocity peaks at $y^* \approx 10$, slightly below the well-known peak for total velocity, while the residual component peaks at $y^* \approx 30$. This suggests that the peak observed in the total velocity fluctuations results from both active and inactive velocities, with 'active' referring to motions connected to changes in the wall shear stress. Further investigation of the coherent structure of the flow showed that the informative velocity consists of smaller near-wall high- and low-velocity streaks collocated with vertical motions (i.e. sweeps and ejections). The spanwise informative

velocity is weak, except close to the wall within the bottom root of the streamwise rolls. This informative streak–roll structure is embedded within a larger-scale streak–roll structure from the residual velocity, which bears no information about the wall shear stress for the considered time scale. We have also shown that ejections propagate information about the wall stress further from the wall than sweeps, while extreme values of the wall shear stress are attributed to sweeps in close proximity to the wall.

The utility of IND for reduced-order modelling was demonstrated in the prediction of the wall shear stress in a turbulent channel flow. The objective was to estimate the two-dimensional wall shear stress in the future, after $\Delta T^* = 25$, by measuring the streamwise velocity in a wall-parallel plane at $y^* \approx 10$ as input. The approach was implemented using a fully convolutional neural network as the predictor. Two cases were considered, using either the informative or the residual velocity component as input, respectively. The main discrepancies were localised in regions with high wall shear stress values. This outcome aligns with our prior analysis, which indicated that extreme wall shear stress events are produced by short-time near-wall sweeps not captured in the input plane. In contrast, the residual velocity component offers no predictive power for wall shear stress, as it has no observability of the wall shear stress, meaning that it lacks any information relevant to the latter. This example in reduced-order modelling reveals that models achieving the highest performance are those that utilise input variables with the maximum amount of information about the output.

Finally, we have investigated the application of IND for drag reduction in turbulent channel flows at $Re_\tau = 180$ and 395. The strategy implemented involved blowing/suction via opposition control. To this end, the no-transpiration boundary condition at the wall was replaced with the wall-normal velocity measured in the wall-parallel plane at $y^* = 14$. We explored the use of three wall-normal velocities: the total velocity (i.e. as originally formulated in opposition control), its informative component, and its residual component. The largest reduction in drag was achieved using the informative component of v, which performed slightly better than the total velocity for both Reynolds numbers. The residual component was shown to yield the poorest results. The application to drag reduction demonstrated here illustrates that the informative component of v contains the essential information needed for effective flow control. This paves the way for using IND to devise enhanced control strategies by isolating the relevant information from the input variables while disregarding the irrelevant contributions.

We conclude this work by highlighting the potential of IND as a post-processing tool for gaining physical insight into the interactions among variables in turbulent flows. Nonetheless, it is also worth noting that the approach relies on the mutual information between variables, which requires estimating joint probability density functions. This entails a data-intensive process that could become a constraint in cases where the amount of numerical or experimental data available is limited. Future efforts will be devoted to reducing the data requirements of aIND and extending its capabilities to account for multi-variable and multi-scale interactions among variables.

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.
- Gonzalo Arranz https://orcid.org/0000-0001-6579-3791;
- Adrián Lozano-Durán https://orcid.org/0000-0001-9306-0261.

Appendix A. Numerical implementation

A.1. *Solution for scalar variables using bijective functions*

Here, we provide the methodology to tackle the minimisation problem posed in (2.10). For convenience, we write (2.10) again:

$$\arg \min_{\Phi_I, \mathcal{F}} I(\Phi_R; \Phi_I) + \gamma \|\Phi - \Phi_I\|^2, \quad \text{s.t. } \Psi_+ = \mathcal{F}(\Phi_I). \tag{A1}$$

To solve (A1), we note that there are two unknowns: Φ_I and the function $\mathcal{F}$. If we assume that $\mathcal{F}$ is invertible, namely

$$\Phi_I(t) = \mathcal{F}^{-1}(\Psi_+(t)) \equiv \mathcal{B}(\Psi_+(t)), \tag{A2}$$

then (A1) can be recast as

$$\arg \min_{\mathcal{B}} I(\Phi - \mathcal{B}(q_+); \mathcal{B}(q_+)) + \gamma \|\Phi - \mathcal{B}(\Psi_+)\|^2, \tag{A3}$$

which can be solved by standard optimisation techniques upon the parametrisation of the function $\mathcal{B}$.

However, by imposing bijectivity, we constrain the feasible $\Phi_I(t)$ solutions that satisfy $H(\Psi_+ \mid \Phi_I) = 0$ and could lead to lower values of the loss function than in the more lenient case, where $\mathcal{F}$ needs only to be surjective. To circumvent this limitation, we recall that a surjective function with $N - 1$ local extrema points (points where the slope changes sign) can be split into N bijective functions (see figure 16a). In particular, we define

$$\Phi_I(t) = \mathcal{B}_i(\Psi_+(t)) \text{ s.t. } \Phi(t) \in [r_{i-1}, r_i), \tag{A4}$$

where r_i is the ith local extremum, such that $r_i > r_{i-1}$, $r_0 \to -\infty$ and $r_N \to \infty$.

Therefore, the final form of the minimisation equation is

$$\arg \min_{\mathcal{B}_i} I(\Phi - \mathcal{B}^{\cup}(\Psi_+); \mathcal{B}^{\cup}(\Psi_+)) + \gamma \sum_i^N \|\Phi - \mathcal{B}_i^0(\Psi_+; \Phi)\|^2, \tag{A5}$$

with

$$\mathcal{B}_i^0(\Psi_+; \Phi) = \begin{cases} \mathcal{B}_i & \text{if } \Phi(t) \in [r_{i-1}, r_i), \\ 0 & \text{otherwise,} \end{cases} \quad \mathcal{B}^{\cup}(\Psi_+) = \sum_i^N \mathcal{B}_i^0(\Psi_+; \Phi), \tag{A6a,b}$$

where the extrema (r_i) are unknowns to be determined in the minimisation problem, and γ and N are the only free parameters. Once the functions $\mathcal{B}_i$ are computed, the informative component is obtained from

$$\Phi_I(t) = \mathcal{B}_j(\Psi_+(t)) \text{ s.t. } j = \arg \min_i (\Phi(t) - \mathcal{B}_i(\Psi_+(t)))^2 \tag{A7}$$
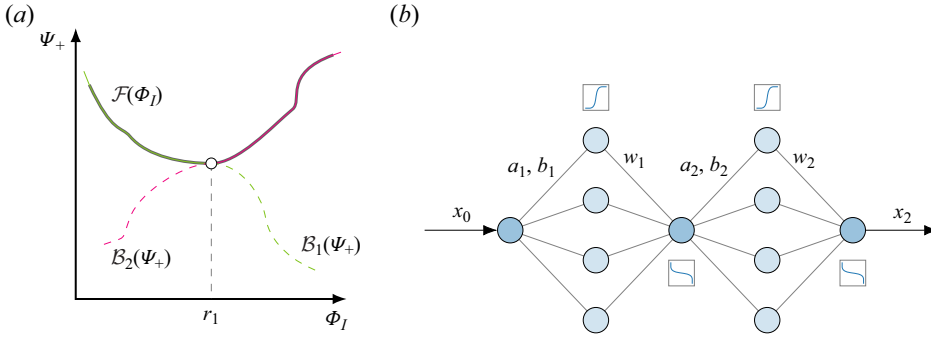
at every time step.

1000 A95-24

Figure 16. (*a*) Illustration of a surjective function $\mathcal{F}(\Phi_I)$ and its decomposition into two bijective functions, $\mathcal{B}_1(\Psi_+)$ and $\mathcal{B}_2(\Psi_+)$. (*b*) Example of a DSF architecture with 2 hidden layers and 4 neurons per hidden layer. The functions plotted within boxes are the activation functions acting on the neurons. Adapted from Huang *et al.* (2018).

We use feedforward networks to find $\mathcal{B}_i$, as they are able to approximate any Borel-measurable function on a compact domain (Hornik, Stinchcombe & White 1989). In particular, we use the deep sigmoidal flow (DSF) proposed by Huang *et al.* (2018), who proved that a feedforward ANN is a bijective transformation if the activation functions are bijective and all the weights are positive. The details of the DSF architecture and the optimisation can be found in § A.2.

One must emphasise that the current minimisation problem posed in (2.10) differs from the classical flow reconstruction problem (e.g. Erichson *et al.* 2020) where the maximum reconstruction of Φ is sought. In those cases, we look for a function $\mathcal{G}(\Psi_+)$ that minimises $\|\Phi - \mathcal{G}(\Psi_+)\|^2$. If the result is a non-bijective function, then the constraint $H(\Psi_+ \mid \Phi_I) = 0$ will not be satisfied.

A.2. *Networks architecture and optimisation details*

The present algorithm uses DSF networks to approximate bijective functions. This network architecture is depicted in figure 16(*b*). The DSF is composed of L stacked sigmoidal transformations. Each transformation produces the output,

$$x_l = \sigma^{-1}(w_l^T \times \sigma(a_l \times x_{l-1} + b_l)), \quad l = 1, \ldots, L, \tag{A8}$$

where x_{l-1} is the input, $\sigma(y) = 1/(1 + e^{-y})$ is the logistic function, σ^{-1} is the inverse of σ, a_l and b_l are vectors with the weights and biases of the decoder part of the l-layer, and w_l is a vector with the weights of the encoder part of the l-layer (see figure 16*b*). In addition, the weights for each layer have to fulfil $0 < w_{l,i} < 1$, $\sum_i w_{l,i} = 1$ and $a_{l,i} > 0$, $i = 1, \ldots, M$, where M is the number of neurons per layer. These constraints are enforced via the softmax and exponential activation functions for w_l and a_l, respectively, namely

$$w_{l,i} = \frac{\exp(w'_{l,i})}{\displaystyle\sum_{i=1}^{N} \exp(w'_{l,i})}, \quad a_{l,i} = \exp(a'_{l,i}). \tag{A9a,b}$$

More details on the DSF architecture can be found in Huang *et al.* (2018).

To compute the optimal weights and biases that yield the optimal $\mathcal{B}_i$ that minimise (A5), we use the Adam algorithm (Kingma & Ba 2017). This minimisation process requires all

operations to be continuous and differentiable. To achieve that, we compute the mutual information using a kernel density estimator, and the piecewise-defined functions $\mathcal{B}_i^0$ are made $\mathcal{C}^1$ continuous by applying the logistic function

$$\mathcal{B}_i^0(\Psi_+; \Phi) = \mathcal{B}_i(\Psi_+) \, \sigma(k(\Phi - \tilde{r}_{i-1})) \, \sigma(k(\tilde{r}_i - \Phi)), \tag{A10}$$

where the parameter $k > 0$ can be chosen to control the steepness of the function, and $\tilde{r}_j = r_j \pm \log(p/(p-1))/k$, which ensures $\mathcal{B}_i^0 = p\mathcal{B}_i$ at the boundaries.

In the present study, the first term in (2.10) is normalised with $\|\Phi\|^2$, and the second term is normalised with $H(\Phi_I, \Phi_R)/2$. Under this normalisation, free parameters $p = 0.99$ and $k = 500$ were determined to be adequate for the optimisation process. The number of bijective functions N was selected to minimise (A5) while producing a continuous mapping, as illustrated in figures 2 and 16(*a*). In the study presented in § 3, an N value of 1 was found to be optimal. We also explored different values for the regularisation constant γ. For $N = 1$, similar mappings were achieved for $0.5 \geq \gamma \geq 2$, and the results discussed in § 3 were calculated with $\gamma = 1$. In cases with $N \geq 1$, starting with a high γ value, approximately 10, during initial iterations proved beneficial for converging the solution. Subsequently, γ was gradually decreased to emphasise the minimisation of the first term in (A5). Currently, this adjustment is performed manually, but future developments in aIND could automate this process (Groenendijk *et al.* 2021). Finally, the DSF architecture was set to 3 layers with 12 neurons per layer.

Appendix B. Validation of aIND and comparison with EPOD and LSE

We include two additional validation cases of aIND applied to two-dimensional fields in a plane $\boldsymbol{x} = (x, z)$. These synthetic examples have an exact analytic solution that enables us to quantify the error produced by the different methods. We consider the system

$$\text{source} \quad \Phi(\boldsymbol{x}, t) = \Phi_I(\boldsymbol{x}, t) + \Phi_R(\boldsymbol{x}, t), \tag{B1}$$

$$\text{target} \quad \Psi_+(\boldsymbol{x}, t) = F(\Phi_I(\boldsymbol{x}, t)), \tag{B2}$$

where the fields Φ_I and Φ_R and the function F are given. In particular, Φ_I and Φ_R are the velocity fluctuations in the planes $y^* \approx 5$ and 40, respectively, of a turbulent channel flow with $Re_\tau = 180$ in a domain $8\pi h \times 2h \times 4\pi h$ in the streamwise, wall-normal and spanwise directions, respectively. Instantaneous snapshots of the fields are shown in figure 17. To ensure that the fields are independent (i.e. $I(\Phi_I, \Phi_R) = 0$), the informative field is extracted at $y^* \approx 5$ from the bottom wall, whereas the residual field is extracted at $y^* \approx 40$ from the top wall at a shifted time step.

We compare aIND with the extended POD method (EPOD) proposed by Borée (2003) and the spectral in space version of the LSE presented by Encinar & Jiménez (2019). In the following subsections, we provide a small overview of each method.

B.1. *Extended POD*

The EPOD offers a linear decomposition of a source field $\Phi(\boldsymbol{x}, t)$ into its correlated (*C*) and decorrelated (*D*) contributions to a given target field such that

$$\Phi_C(\boldsymbol{x}, t) = \sum_n a_\Psi^n(t) \, U_\Phi^n(\boldsymbol{x}), \tag{B3}$$

$$\Phi_D(\boldsymbol{x}, t) = \Phi(\boldsymbol{x}, t) - \Phi_C(\boldsymbol{x}, t), \tag{B4}$$
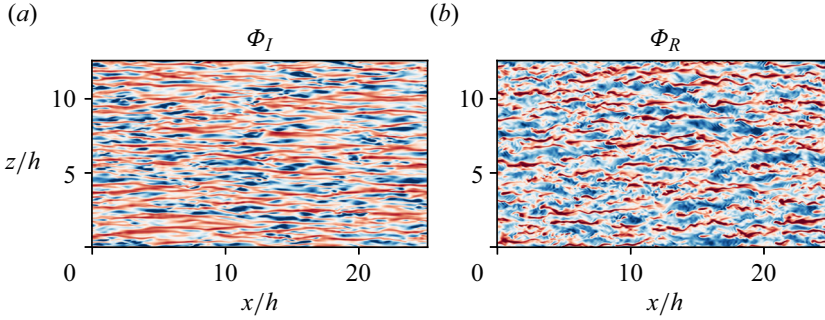
Figure 17. Instantaneous velocities used to create synthetic examples: (*a*) informative field and (*b*) residual field. The velocities are extracted from direct numerical simulations of a turbulent channel flow at $Re_\tau = 180$ at $y^+ \approx 5$ and 40, respectively.

where n is the number of modes, and a_Ψ^n is the temporal coefficient of the nth POD mode of the target field Ψ_+ and U_Φ^n in the nth spatial mode. The latter is computed as

$$U_\Phi^n(x) = \frac{\langle a_\Psi^n \, \Phi(x, t)\rangle_t}{\langle a_\Psi^n a_\Psi^n\rangle}, \tag{B5}$$

where $\langle \cdot \rangle_t$ denotes temporal average. The EPOD decomposition has the following properties (Borée 2003).

  (i) The correlation between the original source field and the target field is the same as the correlation between the correlated field and the target field, namely

$$\langle \Phi \Psi \rangle = \langle \Phi_C \Psi \rangle. \tag{B6}$$

  (ii) The decorrelated field is uncorrelated with the target field, i.e.

$$\langle \Phi_D \Psi \rangle = 0. \tag{B7}$$

Therefore, we define the EPOD informative component as the correlated field ($\Phi_I^{EPOD} \equiv \Phi_C$) and the EPOD residual component as the decorrelated field ($\Phi_R^{EPOD} \equiv \Phi_D$). In the following examples, the POD of the target field is obtained using 300 snapshots and the informative field is reconstructed using the 50 more energetic modes.

B.2. *Spectral LSE*

The LSE, proposed by Adrian & Moin (1988), provides the best mean square linear estimate of the 'response' field $\Phi(x, t)$ given the 'predictor' $\Psi_+(x, t)$ (Tinney *et al.* 2006). Considering a collection of discrete spatial locations x_i, the best linear estimate that minimises

$$\arg \min_{\tilde\Phi} \langle (\Phi(x_i, t) - \tilde\Phi(x_i, t))^2 \rangle_t \tag{B8}$$

is given by

$$\tilde\Phi(x_i, t) = L_{ij} \, \Psi_+(x_j, t), \tag{B9}$$

where repeated indices imply summation. The entries of the matrix L take the form (Adrian & Moin 1988)

$$L_{ij}(x_m) = \frac{\langle \Phi(x_i, t) \, \Psi(x_m, t)\rangle_t}{\langle \Psi(x_j, t) \, \Psi(x_m, t)\rangle_t}. \tag{B10}$$

From (B8), we define the LSE informative and residual components as $\Phi_I^{LSE}(\boldsymbol{x}, t) \equiv \tilde{\Phi}(\boldsymbol{x}, t)$ and $\Phi_R^{LSE}(\boldsymbol{x}, t) \equiv \Phi - \tilde{\Phi}(\boldsymbol{x}, t)$, respectively.

In the following examples, we exploit the spatial periodicity of the flow field. To that end, we adopt the approach by Encinar & Jiménez (2019) and use a spatial Fourier basis to project the fields. This procedure is usually known as spectral linear stochastic estimation (SLSE). Equation (B10) becomes

$$\hat{L}(k_x, k_z) = \frac{\langle \hat{\Psi}(k_x, k_z, t)\, \hat{\Phi}^\dagger(k_x, k_z, t)\rangle_t}{\langle \hat{\Phi}(k_x, k_z, t)\, \hat{\Phi}^*(k_x, k_z, t)\rangle_t}, \tag{B11}$$

where $\widehat{(\cdot)}$ denotes the Fourier transform, $(\cdot)^\dagger$ is the complex conjugate, and k_x, k_z are the wavenumbers in the x, z directions, respectively. It can be shown (see Tinney *et al.* 2006; Encinar & Jiménez 2019) that the optimal estimator is

$$\Phi_I(\boldsymbol{x}_i, t) = \sum_j L(x_i - x_j, z_i - z_j)\, \Psi(x_j, z_j, t). \tag{B12}$$

B.3. *Linear mapping*

As a first validation case, we consider a linear mapping function

$$\text{source} \quad \Phi(\boldsymbol{x}, t) = \Phi_I(\boldsymbol{x}, t) + \Phi_R(\boldsymbol{x}, t), \tag{B13a}$$

$$\text{target} \quad \Psi_+(\boldsymbol{x}, t) = \Phi_I(\boldsymbol{x}, t). \tag{B13b}$$

The exact informative and residual fields are normalised such that their standard deviations are $\langle \Phi_I \Phi_I \rangle = 1$ and $\langle \Phi_R \Phi_R \rangle = 1$, respectively. The instantaneous reconstructed fields are displayed in figure 18. To ease the comparison, the time instant is the same as in figure 17.

We can observe that aIND accurately reconstructs the informative and residual fields. The SLSE is also able to reconstruct the mapping, which is expected since the mapping is linear. On the contrary, EPOD fails to obtain the correct informative/residual field despite the linear character of the decomposition. Instead, it tends to reconstruct the original field Φ.

B.4. *Nonlinear mapping*

As a second validation case, we consider the nonlinear mapping function

$$\text{source} \quad \Phi(\boldsymbol{x}, t) = \Phi_I(\boldsymbol{x}, t) + \Phi_R(\boldsymbol{x}, t), \tag{B14a}$$

$$\text{target} \quad \Psi_+(\boldsymbol{x}, t) = \Phi_I^2(\boldsymbol{x}, t) - 0.2\, \Phi_I(\boldsymbol{x}, t). \tag{B14b}$$

The exact informative and residual fields are normalised such that their standard deviations are $\langle \Phi_I \Phi_I \rangle = 1$ and $\langle \Phi_R \Phi_R \rangle = 0.2$, respectively. The instantaneous reconstructed fields are displayed in figure 19 at the same time instant as in figure 17.

In this case, SLSE fails to correctly split the flow into the informative and residual fields. The same applies to EPOD: although the reconstruction of the informative resembles the original (due to higher correlation between the original and the informative terms from (B14a)), the error $\Phi_I - \Phi_I^{EPOD}$ is significant everywhere. A similar error is observed for the residual field, which is not correctly identified by EPOD. The aIND, similar to the

Figure 18. Instantaneous reconstructed fields of the informative and residual components for (B13) with (*a,b,c*) aIND, (*d,e,f*) SLSE, and (*g,h,i*) EPOD: (*a,d,g*) the reconstructed informative field, (*b,e,h*) the reconstructed residual field, and (*c,f,i*) the error between the exact and reconstructed informative fields. The colourmap for the errors range between $\pm 0.5\langle \Phi_i \Phi_I \rangle$.

previous example, accurately reconstructs the informative and residual fields. The small discrepancies in $\Phi_I - \Phi_I^{IND}$ occur at the locations where $\Phi_I \approx 0$, and stem from the approach followed to compute Φ_I^{IND}. Note that aIND accurately reconstructs the analytical mapping as shown in figure 20.

Appendix C. Analytical solution for Gaussian distributions

For the special case in which all the components in $\boldsymbol{\Phi}_I$, $\boldsymbol{\Psi}_+$ are jointly normal distributed variables, we can write their mutual information as (Cover & Thomas 2006)

$$I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi}_I) = \frac{1}{2} \log \left(\frac{|\Sigma(\boldsymbol{\Phi}_I)| \, |\Sigma(\boldsymbol{\Psi}_+)|}{|\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_I)|} \right). \tag{C1}$$

In (C1), $|\cdot|$ denotes the matrix determinant, $\Sigma(\boldsymbol{\Phi}_I)$ is the covariance matrix of $\boldsymbol{\Phi}_I$ (and similarly for $\boldsymbol{\Psi}_+$), a square matrix whose i, j entry is defined as

$$\Sigma(\boldsymbol{\Phi}_I)[i, j] = \langle \Phi_{I,i}(t) \, \Phi_{I,j}(t) \rangle_t, \tag{C2}$$
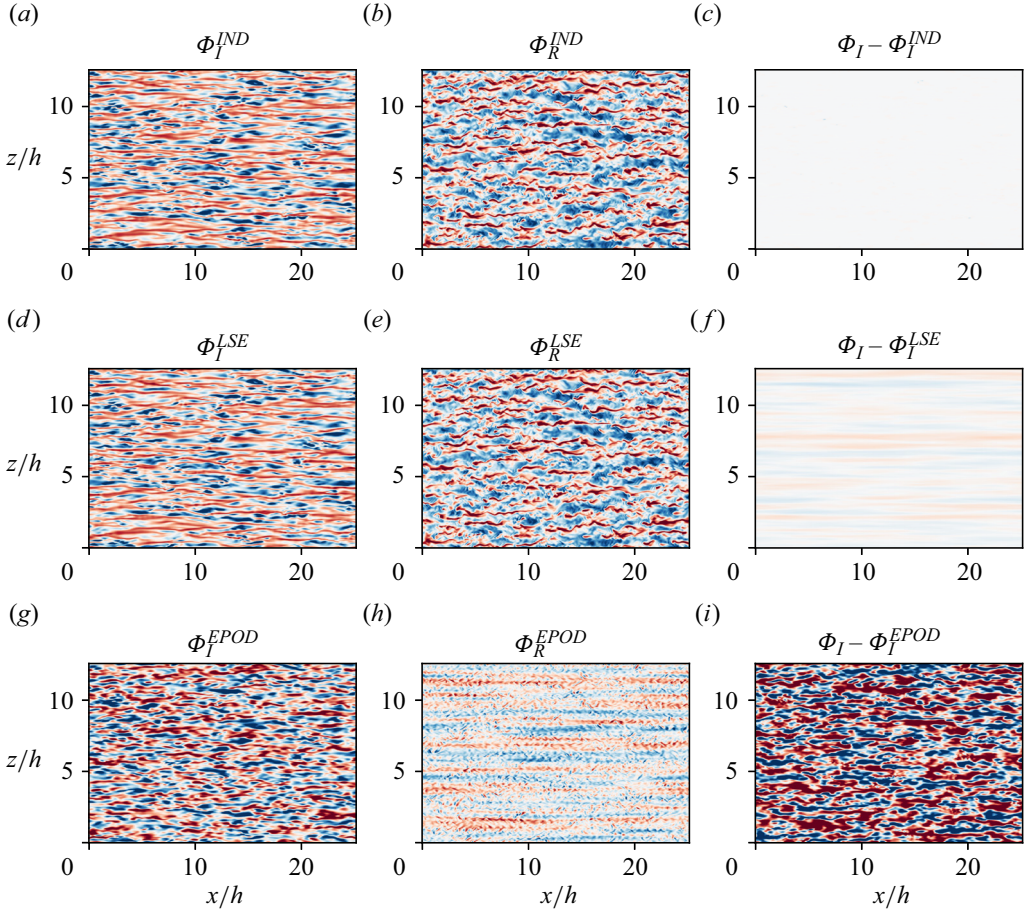
1000 A95-29
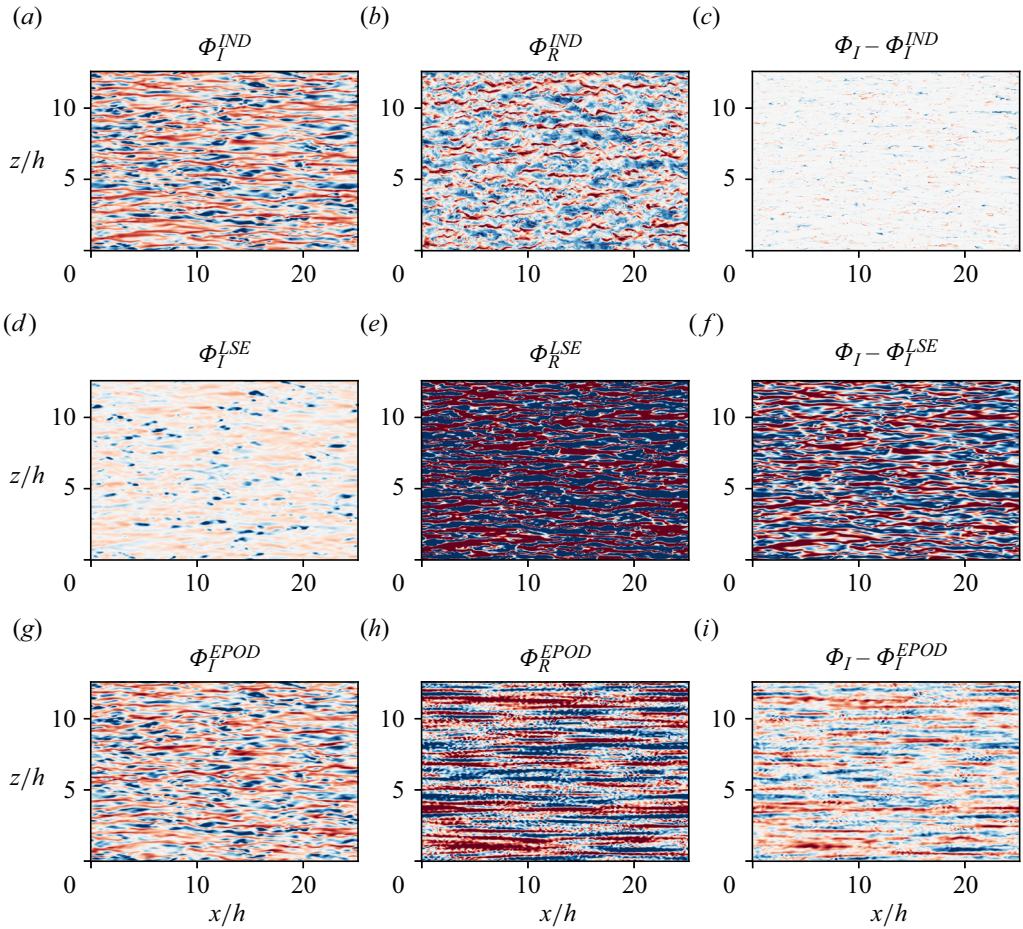
Figure 19. Instantaneous reconstructed fields of the informative and residual components for (B 14) with (*a,b,c*) aIND, (*d,e,f*) SLSE, and (*g,h,i*) EPOD: (*a,d,g*) the reconstructed informative field, (*b,e,h*) the reconstructed residual field, and (*c,f,i*) the error between the exact and the reconstructed informative field. The colourmap for the errors range between $\pm 0.4 \langle \Phi_i \Phi_I \rangle$.

Figure 20. Mapping of (*a*) the linear example from (B 13) and (*b*) the nonlinear example from (B 14). The dashed lines correspond to the analytic solution, and the solid lines correspond to the solution computed with aIND.

where $\Phi_{I,i}$ is the ith element of $\boldsymbol{\Phi}_I$. The covariance matrix $\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_I)$ can be written in block matrix form as

$$\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_I) = \begin{bmatrix} \Sigma(\boldsymbol{\Psi}_+) & \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{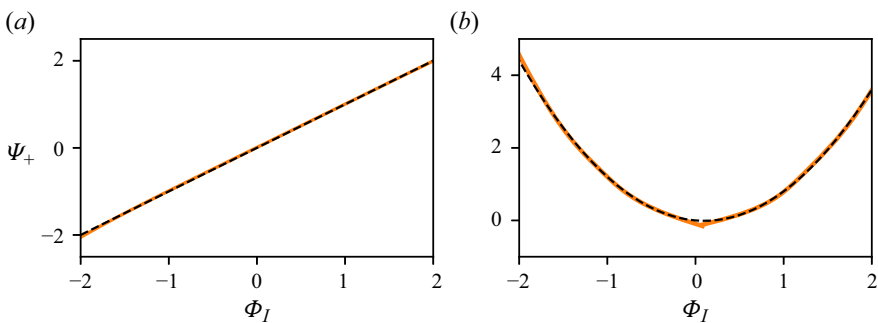\Phi}_I) \\ \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)^{\mathrm{T}} & \Sigma(\boldsymbol{\Phi}_I) \end{bmatrix}, \tag{C3}$$

where $\Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)$ is the cross-covariance matrix

$$\Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)[i,j] = \langle \Psi_{+,i}(t)\, \Phi_{I,j}(t) \rangle_t. \tag{C4}$$

The mutual information in (C1) is maximised when $|\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_I)| = 0$, provided that $|\Sigma(\boldsymbol{\Phi}_I)| \neq 0$. Using the block determinant identity (Johnson & Horn 1985; Barnett, Barrett & Seth 2009) gives

$$|\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_I)| = |\Sigma(\boldsymbol{\Psi}_+)|\,|\Sigma(\boldsymbol{\Phi}_I) - \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)\, \Sigma(\boldsymbol{\Psi}_+)^{-1}\, \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)^{\mathrm{T}}| \tag{C5a}$$

$$= |\Sigma(\boldsymbol{\Phi}_I+)|\,|\Sigma(\boldsymbol{\Psi}_+) - \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)^{\mathrm{T}}\, \Sigma(\boldsymbol{\Phi}_I)^{-1}\, \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_I)|. \tag{C5b}$$

The second term in (C5a) (resp. (C5b)) is the residual of a linear regression of $\boldsymbol{\Psi}_+$ on $\boldsymbol{\Phi}_I$ (resp. $\boldsymbol{\Phi}_I$ on $\boldsymbol{\Psi}_+$) (Barnett *et al.* 2009). Therefore, the mutual information in (C1) is maximised when $\boldsymbol{\Phi}_I$ is a linear function of $\boldsymbol{\Psi}_+$, or *vice versa*. However, only when $\boldsymbol{\Psi}_+$ is a function of $\boldsymbol{\Phi}_I$, $H(\boldsymbol{\Psi}_+|\boldsymbol{\Phi}_I) = 0$, as required by (2.5).

We assume that the number N_Φ of elements in $\boldsymbol{\Phi}$ is larger than the number N_Ψ of elements of $\boldsymbol{\Psi}_+$, so that if we find

$$\boldsymbol{\Phi}_I = \boldsymbol{L}\boldsymbol{\Psi}_+, \tag{C6}$$

then we can find the inverse mapping

$$\boldsymbol{\Psi}_+ = \boldsymbol{L}^{-1}\boldsymbol{\Phi}_I. \tag{C7}$$

The mutual information in (2.7) can be expanded as

$$I(\boldsymbol{\Psi}_+; \boldsymbol{\Phi}_R) = \frac{1}{2}\log\left(\frac{|\Sigma(\boldsymbol{\Phi}_R)|\,|\Sigma(\boldsymbol{\Psi}_+)|}{|\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_R)|}\right), \tag{C8}$$

which will be equal to zero for $|\Sigma(\boldsymbol{\Psi}_+ \oplus \boldsymbol{\Phi}_R)| = |\Sigma(\boldsymbol{\Psi}_+)|\,|\Sigma(\boldsymbol{\Phi}_R)|$. From the block determinant identity, this requires

$$|\Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_R)\, \Sigma(\boldsymbol{\Psi}_+)^{-1}\, \Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_R)^{\mathrm{T}}| = 0. \tag{C9}$$

In a general scenario, this requires

$$\Sigma(\boldsymbol{\Psi}_+, \boldsymbol{\Phi}_R)[i,j] = 0, \tag{C10}$$

namely

$$\langle \Psi_{+,i}(\Phi_j - \boldsymbol{L}[j,m]\,\Psi_{+,m}) \rangle_t \equiv \langle \Psi_{+,i}\Phi_j \rangle_t - \langle \Psi_{+,i}\,\boldsymbol{L}[j,m]\,\Psi_{+,m} \rangle_t = 0, \tag{C11}$$

where $\Psi_{+,k}$ is the kth element of $\boldsymbol{\Psi}_+$ and repeated indices imply summation. The solution to (C11) is given by Adrian & Moin (1988), and it correspond to the LSE:

$$\boldsymbol{L}[j,m] = \frac{\langle \Psi_{+,i}\Phi_j \rangle_t}{\langle \Psi_{+,j}\Psi_{+,m} \rangle_t}. \tag{C12}$$

Therefore, for the special case in which all variables involved are jointly normal distributed variables, the solution to IND is LSE. From the previous results, it is

straightforward to prove that the solution to aIND when Φ, Ψ_+ are jointly distributed is given by

$$\Phi_I(t) = \frac{\langle \Psi_+ \Phi \rangle_t}{\langle \Psi_+ \Psi_+ \rangle_t} \Psi_+(t). \tag{C13}$$

We conclude by emphasising that the similarity between IND and higher-order versions of LSE does not extend to the most likely case where all the variables are not jointly normal distributed. In this scenario, higher-order versions of LSE attempt to obtain a better reconstruction of Φ using Ψ_+, which will not fulfil the condition $H(\Psi_+|\Phi_I) = 0$, as discussed in the last paragraph of § A.1.

Appendix D. Computation of Δx^{max} for the turbulent channel flow

The aIND requires the value of $\Delta x_\square^{max} = (\Delta x_\square^{max}, \Delta z_\square^{max})$ for each informative component $\square = u$, v and w. To that end, we calculate their relative energies as functions of Δx, Δz and the wall-normal distance:

$$E_I^u(\Delta x, \Delta z, y) = \frac{\|u_I^2\|}{\|u^2\|}, \quad E_I^v(\Delta x, \Delta z, y) = \frac{\|v_I^2\|}{\|v^2\|}, \quad E_I^w(\Delta x, \Delta z, y) = \frac{\|w_I^2\|}{\|w^2\|}. \tag{D1a–c}$$

The parametric sweep is performed using data for a channel flow at $Re = 180$ in a computational domain of size $\pi h \times 2h \times \pi/2h$ in the streamwise, wall-normal and spanwise directions, respectively.

Figure 21 displays E_I^u, E_I^v and E_I^w as functions of Δx and Δz. Note that due to the symmetry of the flow, $E_I^u(\Delta x, \Delta z, y) = E_I^u(\Delta x, -\Delta z, y)$ (similarly for E_I^v and E_I^w). For E_I^u and E_I^v, the maximum is always located at $\Delta z = 0$, which is the plane displayed in figures 21(a,b). For the spanwise component, the maximum value of E_I^w is offset in the spanwise direction, and its location varies with y. Figure 21(c) displays the horizontal section that contains its global maximum, which is located at $y^* \approx 6$. This offset is caused by the fact that w motions travel in the spanwise direction until they reach the wall and affect the wall shear stress.

Close to the wall, we find high values of E_I^u, with peak value approximately 60 % at $y^* \approx 8$, and $\Delta x_u^{max}(y) \approx -h$, following an almost linear relationship with y. Farther from the wall ($y > 0.2h$), Δx_u^{max} becomes more or less constant, although it should be noted that in this region, the values of E_I^u for a fixed y are low and relatively constant. This may induce some numerical uncertainty in the particular value of Δx_u^{max}, but the overall results are not affected. In contrast, high values of E_I^v are located in a compact region further away from the wall ($y^* \approx 19$), and they tend to zero at the wall. The values $\Delta x_v^{max}(y)$ lie close to $-1.2h$ in this region, following a negative linear relationship with y. As before, $\Delta x_v^{max}(y)$ remains relative constant in low E_I^v regions. Finally, although not shown, $\Delta x_w^{max}(y)$ and $\Delta z_w^{max}(y)$ lie in the intervals $[-h, -0.7h]$ and $\pm[0.1h, 0.2h]$, respectively, approaching zero at the wall. Nevertheless, E_w^I becomes negligible for $y > 0.2h$.

We close this appendix by noting that although not explored in the present study, Δx^{max} computed with aIND might correspond to potential locations for sensor placement, since it maximises the mutual information with the target variable (Lozano-Durán & Arranz 2022).
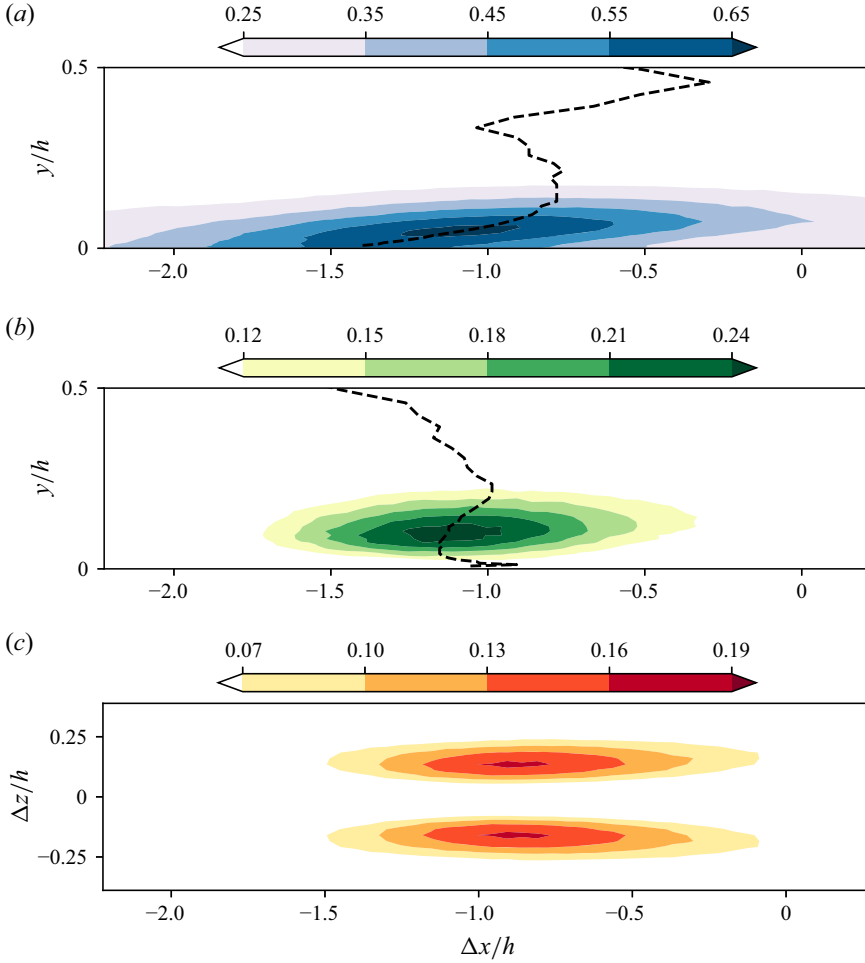
Figure 21. Informative regions quantifying the relative energy contained in the informative components u_I, v_I and w_I for $\Delta T^* = 25$: (a) E_I^u and (b) E_I^v in the $\Delta z = 0$ plane; (c) E_I^w in the plane $y^* \approx 6$. In (a,b), the black dashed line corresponds to $\Delta x^{max}(y)$.

Appendix E. Validity of aIND of u with respect to τ_x

Figure 22 displays the mutual information between $u_R(x_0, y_0, z_0)$ for $y_0^* \approx 10$ and $\tau_{x,+}(x_0 - \Delta x_u^{max} - \delta x, z_0 - \Delta z_u^{max} - \delta z)$ as a function of $\delta x = [\delta x, \delta z]$, denoted as $I(u_R; \tau_{x,+})(\delta x)$. The mutual information is normalised by the total Shannon information of the wall shear stress, $H(\tau_x)$, such that $I(u_R; \tau_{x,+})(\delta x)/H(\tau_x) = 0$ means that u_R contains no information about the wall shear stress at δx, and $I(u_R; \tau_{x,+})(\delta x)/H(\tau_x) = 1$ implies that u_R contains all the information about $\tau_{x,+}(\delta x)$. Note that aIND seeks to minimise $I(u_R; \tau_{x,+})(\mathbf{0})$. The results show that value of the $I(u_R; \tau_{x,+})(\delta x)/H(\tau_x)$ remains always low, reaching a maximum approximately 0.06 at $\delta x \approx -1.2h$ along the streamwise direction. Hence we can conclude that the residual term contains a negligible amount of information about the wall shear stress at any point in the wall, and aIND is a valid approximation of IND. For the sake of completeness, we also display in figure 22 the mutual information between u_I and the wall shear stress. Since $\tau_{x,+} = \mathcal{F}(u_I)$, the mutual information $I(u_I; \tau_{x,+})(\delta x)$ has to be equal to $H(\tau_x)$ at $\delta x = \mathbf{0}$, as corroborated
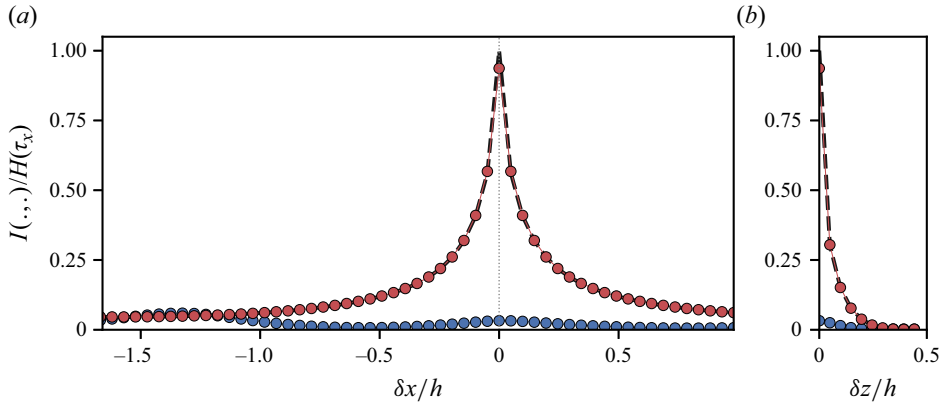
Figure 22. Mutual information between the streamwise wall shear stress and the residual field, $I(u_R; \tau_{x,+})(\delta x)$ (blue circles) and the informative field $I(u_I; \tau_{x,+})(\delta x)$ (red circles). The dashed line corresponds to $I(\tau_{x,+}; \tau_{x,+})(\delta x)$.

by the results. For larger distances, $I(u_I; \tau_{x,+})(\delta x)$ decays following the natural decay of $I(\tau_{x,+}; \tau_{x,+})(\delta x)$, with values below 0.1 after $|\delta x| \approx h$.

REFERENCES

ADRIAN, R.J. 2007 Hairpin vortex organization in wall turbulence. *Phys. Fluids* **19** (4), 041301.

ADRIAN, R.J., MEINHART, C.D. & TOMKINS, C.D. 2000 Vortex organization in the outer region of the turbulent boundary layer. *J. Fluid Mech.* **422**, 1–54.

ADRIAN, R.J. & MOIN, P. 1988 Stochastic estimation of organized turbulent structure: homogeneous shear flow. *J. Fluid Mech.* **190**, 531–559.

BAARS, W.J. & TINNEY, C.E. 2014 Proper orthogonal decomposition-based spectral higher-order stochastic estimation. *Phys. Fluids* **26** (5), 055112.

BAE, H.J. & LEE, M. 2021 Life cycle of streaks in the buffer layer of wall-bounded turbulence. *Phys. Rev. Fluids* **6** (6), 064603.

BARNETT, L., BARRETT, A.B. & SETH, A.K. 2009 Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* **103**, 238701.

BETCHOV, R. 1964 Measure of the intricacy of turbulence. *Phys. Fluids* **7** (8), 1160–1162.

BLACKWELDER, R.F. & KAPLAN, R.E. 1976 On the wall structure of the turbulent boundary layer. *J. Fluid Mech.* **76**, 89–112.

BORÉE, J. 2003 Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows. *Exp. Fluids* **35**, 188–192.

BROWN, G.L. & ROSHKO, A. 1974 On density effects and large structure in turbulent mixing layers. *J. Fluid Mech.* **64** (4), 775–816.

BRUNTON, S.L., NOACK, B.R. & KOUMOUTSAKOS, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52** (1), 477–508.

CERBUS, R.T. & GOLDBURG, W.I. 2013 Information content of turbulence. *Phys. Rev. E* **88**, 053012.

CHENG, C., LI, W., LOZANO-DURÁN, A. & LIU, H. 2019 Identity of attached eddies in turbulent channel flows with bidimensional empirical mode decomposition. *J. Fluid Mech.* **870**, 1037–1071.

CHOI, H., MOIN, P. & KIM, J. 1994 Active turbulence control for drag reduction in wall-bounded flows. *J. Fluid Mech.* **262**, 75–110.

CHUNG, Y.M. & TALHA, T. 2011 Effectiveness of active flow control for turbulent skin friction drag reduction. *Phys. Fluids* **23** (2), 025102.

CORRSIN, S. & KISTLER, A.L. 1954 The free-stream boundaries of turbulent flows. *NACA Tech. Note* TN-3133. National Advisory Committee for Aeronautics, Adv. Conf. Rep. 3123.

COVER, T.M. & THOMAS, J.A. 2006 *Elements of Information Theory*, 2nd edn. Wiley.

DEL ÁLAMO, J.C., JIMÉNEZ, J., ZANDONADE, P. & MOSER, R.D. 2006 Self-similar vortex clusters in the turbulent logarithmic region. *J. Fluid Mech.* **561**, 329–358.

DESHPANDE, R., MONTY, J.P. & MARUSIC, I. 2021 Active and inactive components of the streamwise velocity in wall-bounded turbulence. *J. Fluid Mech.* **914**, A5.

ENCINAR, M.P. & JIMÉNEZ, J. 2019 Logarithmic-layer turbulence: a view from the wall. *Phys. Rev. Fluids* **4**, 114603.

ERICHSON, N.B., MATHELIN, L., YAO, Z., BRUNTON, S.L., MAHONEY, M.W. & KUTZ, J.N. 2020 Shallow neural networks for fluid flow reconstruction with limited sensors. *Proc. R. Soc.* A **476** (2238), 20200097.

FARRELL, B.F. & IOANNOU, P.J. 2012 Dynamics of streamwise rolls and streaks in turbulent wall-bounded shear flow. *J. Fluid Mech.* **708**, 149–196.

GHAEMI, S. & SCARANO, F. 2013 Turbulent structure of high-amplitude pressure peaks within the turbulent boundary layer. *J. Fluid Mech.* **735**, 381–426.

GRANERO-BELINCHON, C. 2018 Multiscale information transfer in turbulence. Thesis, Université de Lyon.

GROENENDIJK, R., KARAOGLU, S., GEVERS, T. & MENSINK, T. 2021 Multi-loss weighting with coefficient of variations. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 1469–1478. IEEE.

GUASTONI, L., GÜEMES, A., IANIRO, A., DISCETTI, S., SCHLATTER, P., AZIZPOUR, H. & VINUESA, R. 2021 Convolutional-network models to predict wall-bounded turbulence from wall quantities. *J. Fluid Mech.* **928**, A27.

GUERRERO, B., LAMBERT, M.F. & CHIN, R.C. 2020 Extreme wall shear stress events in turbulent pipe flows: spatial characteristics of coherent motions. *J. Fluid Mech.* **904**, A18.

HAMMOND, E.P., BEWLEY, T.R. & MOIN, P. 1998 Observed mechanisms for turbulence attenuation and enhancement in opposition-controlled wall-bounded flows. *Phys. Fluids* **10** (9), 2421–2423.

HORNIK, K., STINCHCOMBE, M. & WHITE, H. 1989 Multilayer feedforward networks are universal approximators. *Neural Netw.* **2** (5), 359–366.

HUANG, C.-W., KRUEGER, D., LACOSTE, A. & COURVILLE, A. 2018 Neural autoregressive flows. In *Proceedings of the 35th International Conference on Machine Learning, Proceedings of Machine Learning Research*, vol. 80, pp. 2078–2087.

HUANG, N.E., SHEN, Z., LONG, S.R., WU, M.C., SHIH, H.H., ZHENG, Q., YEN, N.-C., TUNG, C.C. & LIU, H.H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc.* A **454** (1971), 903–995.

HWANG, J. & SUNG, H.J. 2018 Wall-attached structures of velocity fluctuations in a turbulent boundary layer. *J. Fluid Mech.* **856**, 958–983.

IWAMOTO, K., SUZUKI, Y. & KASAGI, N. 2002 Reynolds number effect on wall turbulence: toward effective feedback control. *Intl J. Heat Fluid Flow* **23** (5), 678–689.

JIMÉNEZ, J. 2018 Coherent structures in wall-bounded turbulence. *J. Fluid Mech.* **842**, P1.

JIMÉNEZ, J. & HOYAS, S. 2008 Turbulent fluctuations above the buffer layer of wall-bounded flows. *J. Fluid Mech.* **611**, 215–236.

JOHANSSON, A.V., HER, J.-Y. & HARITONIDIS, J.H. 1987 On the generation of high-amplitude wall-pressure peaks in turbulent boundary layers and spots. *J. Fluid Mech.* **175**, 119–142.

JOHNSON, C.R. & HORN, R.A. 1985 *Matrix Analysis*. Cambridge University Press.

JOVANOVIĆ, M.R. & BAMIEH, B. 2005 Componentwise energy amplification in channel flows. *J. Fluid Mech.* **534**, 145–183.

KAISER, A. & SCHREIBER, T. 2002 Information transfer in continuous processes. *Physica* D **166** (1), 43–62.

KIM, J. 1985 Turbulence structures associated with the bursting event. *Phys. Fluids* **28**, 52–58.

KIM, J., MOIN, P. & MOSER, R. 1987 Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.* **177**, 133–166.

KINGMA, D.P. & BA, J. 2017 Adam: A method for stochastic optimization. arXiv:1412.6980

KLINE, S.J., REYNOLDS, W.C., SCHRAUB, F.A. & RUNSTADLER, P.W. 1967 The structure of turbulent boundary layers. *J. Fluid Mech.* **30** (4), 741–773.

KUTZ, J.N., FU, X. & BRUNTON, S.L. 2016 Multiresolution dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **15** (2), 713–735.

LE CLAINCHE, S. & VEGA, J.M. 2017 Higher order dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **16** (2), 882–925.

LEE, T.-W. 2021 Scaling of the maximum-entropy turbulence energy spectra. *Eur. J. Mech.* B *Fluids* **87**, 128–134.

LIANG, X.S. & LOZANO-DURÁN, A. 2016 A preliminary study of the causal structure in fully developed near-wall turbulence. *CTR – Proceedings of the 2016 Summer Program*, pp. 233–242.

LONG, J., SHELHAMER, E. & DARRELL, T. 2015 Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE.

LOZANO-DURÁN, A. & ARRANZ, G. 2022 Information-theoretic formulation of dynamical systems: causality, modeling, and control. *Phys. Rev. Res.* **4**, 023195.

LOZANO-DURÁN, A., BAE, H.J. & ENCINAR, M.P. 2019 Causality of energy-containing eddies in wall turbulence. *J. Fluid Mech.* **882**, A2.

LOZANO-DURÁN, A., FLORES, O. & JIMÉNEZ, J. 2012 The three-dimensional structure of momentum transfer in turbulent channels. *J. Fluid Mech.* **694**, 100–130.

LOZANO-DURÁN, A., GIOMETTO, M.G., PARK, G.I. & MOIN, P. 2020 Non-equilibrium three-dimensional boundary layers at moderate Reynolds numbers. *J. Fluid Mech.* **883**, A20.

LOZANO-DURÁN, A. & JIMÉNEZ, J. 2014 Time-resolved evolution of coherent structures in turbulent channels: characterization of eddies and cascades. *J. Fluid Mech.* **759**, 432–471.

LUHAR, M., SHARMA, A.S. & MCKEON, B.J. 2014 On the structure and origin of pressure fluctuations in wall turbulence: predictions based on the resolvent analysis. *J. Fluid Mech.* **751**, 38–70.

LUMLEY, J.L. 1967 The structure of inhomogeneous turbulent flows. *Atmospheric Turbulence and Radio Wave Propagation* (ed. A.M. Yaglom & V.I. Tartarsky), pp. 166–178. Nauka.

MARTÍNEZ-SÁNCHEZ, A., LÓPEZ, E., LE CLAINCHE, S., LOZANO-DURÁN, A., SRIVASTAVA, A. & VINUESA, R. 2023 Causality analysis of large-scale structures in the flow around a wall-mounted square cylinder. *J. Fluid Mech.* **967**, A1.

MATERASSI, M., CONSOLINI, G., SMITH, N. & DE MARCO, R. 2014 Information theory analysis of cascading process in a synthetic model of fluid turbulence. *Entropy* **16** (3), 1272–1286.

MCKEON, B.J. 2017 The engine behind (wall) turbulence: perspectives on scale interactions. *J. Fluid Mech.* **817**, P1.

MCKEON, B.J. & SHARMA, A.S. 2010 A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.* **658**, 336–382.

MEINHART, C.D. & ADRIAN, R.J. 1995 On the existence of uniform momentum zones in a turbulent boundary layer. *Phys. Fluids* **7** (4), 694–696.

MEZIĆ, I. 2013 Analysis of fluid flows via spectral properties of the Koopman operator. *Annu. Rev. Fluid Mech.* **45** (1), 357–378.

MOIN, P. & MOSER, R.D. 1989 Characteristic-eddy decomposition of turbulence in a channel. *J. Fluid Mech.* **200**, 471–509.

MOISY, F. & JIMÉNEZ, J. 2004 Geometry and clustering of intense structures in isotropic turbulence. *J. Fluid Mech.* **513**, 111–133.

MURRAY, N. & UKEILEY, L.S. 2003 Estimation of the flow field from surface pressure measurements in an open cavity. *AIAA J.* **41** (5), 969–972.

NAGUIB, A.M., WARK, C.E. & JUCKENHÖFEL, O. 2001 Stochastic estimation and flow sources associated with surface pressure events in a turbulent boundary layer. *Phys. Fluids* **13** (9), 2611–2626.

PANTON, R.L. 2001 Overview of the self-sustaining mechanisms of wall turbulence. *Prog. Aerosp. Sci.* **37** (4), 341–383.

REYNOLDS, O. 1895 IV. On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Phil. Trans. R. Soc.* A **186**, 123–164.

ROBINSON, S.K. 1991 Coherent motions in the turbulent boundary layer. *Annu. Rev. Fluid Mech.* **23** (1), 601–639.

ROWLEY, C.W., MEZIĆ, I., BAGHERI, S., SCHLATTER, P. & HENNINGSON, D.S. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**, 115–127.

SCHEWE, G. 1983 On the structure and resolution of wall-pressure fluctuations associated with turbulent boundary-layer flow. *J. Fluid Mech.* **134**, 311–328.

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

SCHMID, P.J. 2022 Dynamic mode decomposition and its variants. *Annu. Rev. Fluid Mech.* **54** (1), 225–254.

SCHMID, P.J., LI, L., JUNIPER, M.P. & PUST, O. 2011 Applications of the dynamic mode decomposition. *Theor. Comput. Fluid Dyn.* **25**, 249–259.

SCHMIDT, O.T. 2020 Bispectral mode decomposition of nonlinear flows. *Nonlinear Dyn.* **102**, 2479–2501.

SCHMIDT, O.T. & SCHMID, P.J. 2019 A conditional space–time POD formalism for intermittent and rare events: example of acoustic bursts in turbulent jets. *J. Fluid Mech.* **867**, R2.

SHANNON, C.E. 1948 A mathematical theory or communication. *Bell Syst. Tech. J.* **27** (3), 379–423.

SHAVIT, M. & FALKOVICH, G. 2020 Singular measures and information capacity of turbulent cascades. *Phys. Rev. Lett.* **125**, 104501.

SILLERO, J.A., JIMÉNEZ, J. & MOSER, R.D. 2014 Two-point statistics for turbulent boundary layers and channels at Reynolds numbers up to $\delta^+ \approx 2000$. *Phys. Fluids* **26** (10), 105109.

DE SILVA, C.M., HUTCHINS, N. & MARUSIC, I. 2016 Uniform momentum zones in turbulent boundary layers. *J. Fluid Mech.* **786**, 309–331.

SIROVICH, L. 1987 Turbulence and the dynamics of coherent structures. Part I. Coherent structures. *Q. Appl. Math.* **45**, 561–571.

SMITS, A.J., MCKEON, B.J. & MARUSIC, I. 2011 High-Reynolds number wall turbulence. *Annu. Rev. Fluid Mech.* **43** (1), 353–375.

TAIRA, K., BRUNTON, S.L., DAWSON, S.T.M., ROWLEY, C.W., COLONIUS, T., MCKEON, B.J., SCHMIDT, O.T., GORDEYEV, S., THEOFILIS, V. & UKEILEY, L.S. 2017 Modal analysis of fluid flows: an overview. *AIAA J.* **55** (12), 4013–4041.

TAIRA, K., HEMATI, M.S., BRUNTON, S.L., SUN, Y., DURAISAMY, K., BAGHERI, S., DAWSON, S.T.M. & YEH, C.-A. 2020 Modal analysis of fluid flows: applications and outlook. *AIAA J.* **58** (3), 998–1022.

TANOGAMI, T. & ARAKI, R. 2024 Information-thermodynamic bound on information flow in turbulent cascade. *Phys. Rev. Res.* **6**, 013090.

THEODORSEN, T. 1952 Mechanisms of turbulence. In *Proceedings of the 2nd Midwestern Conference on Fluid Mechanics*. Ohio State University.

TINNEY, C.E., COIFFET, F., DELVILLE, J., HALL, A.M., JORDAN, P. & GLAUSER, M.N. 2006 On spectral linear stochastic estimation. *Exp. Fluids* **41** (5), 763–775.

TOWNE, A., SCHMIDT, O.T. & COLONIUS, T. 2018 Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.* **847**, 821–867.

TOWNSEND, A.A. 1961 Equilibrium layers and wall turbulence. *J. Fluid Mech.* **11** (1), 97–120.

TREFETHEN, L.N., TREFETHEN, A.E., REDDY, S.C. & DRISCOLL, T.A. 1993 Hydrodynamic stability without eigenvalues. *Science* **261** (5121), 578–584.

WALLACE, J.M. 2016 Quadrant analysis in turbulence research: history and evolution. *Annu. Rev. Fluid Mech.* **48**, 131–158.

WALLACE, J.M., ECKELMAN, H. & BRODKEY, R.S. 1972 The wall region in turbulent shear flow. *J. Fluid Mech.* **54**, 39–48.

WANG, W., CHU, X., LOZANO-DURÁN, A., HELMIG, R. & WEIGAND, B. 2021 Information transfer between turbulent boundary layers and porous media. *J. Fluid Mech.* **920**, A21.

WILLIAMS, M.O., KEVREKIDIS, I.G. & ROWLEY, C.W. 2015 A data-driven approximation of the Koopman operator: extending dynamic mode decomposition. *J. Nonlinear Sci.* **25**, 1307–1346.

YUAN, Y. & LOZANO-DURÁN, A. 2024 Limits to extreme event forecasting in chaotic systems. *Physica* D **467**, 134246.

ZAKI, T.A. & WANG, M. 2021 From limited observations to the state of turbulence: fundamental difficulties of flow reconstruction. *Phys. Rev. Fluids* **6**, 100501.