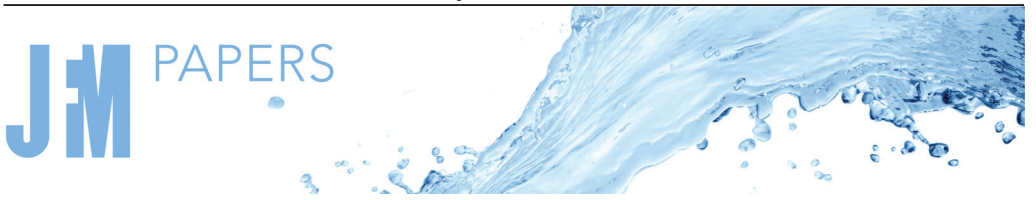
J. Fluid Mech. (2023), *vol*. 968, A10, doi:10.1017/jfm.2023.515



A Perron–Frobenius analysis of wall-bounded turbulence

Javier Jiménez†

School of Aeronautics, U. Politécnica Madrid, 28040 Madrid, Spain

(Received 21 January 2023; revised 30 April 2023; accepted 19 June 2023)

The Perron–Frobenius operator (PFO) is adapted from dynamical-system theory to the study of turbulent channel flow. It is shown that, as long as the analysis is restricted to the system attractor, the PFO can be used to differentiate causality and coherence from simple correlation without performing interventional experiments, and that the key difficulty remains the collection of enough data to populate the operator matrix. This is alleviated by limiting the analysis to two-dimensional projections of the phase space, and developing a series of indicators to choose the best parameter pairs from a large number of possibilities. The techniques thus developed are applied to the study of bursting in the inertial layer of the channel, with emphasis on the process by which bursts are reinitiated after they have decayed. Conditional averaging over phase-space trajectories suggested by the PFO shows, somewhat counter-intuitively, that a key ingredient for the burst recovery is the development of a low-shear region near the wall, overlaid by a lifted shear layer. This is confirmed by a computational experiment in which the control of the mean velocity profile by the turbulence fluctuations is artificially relaxed. The behaviour of the mean velocity profile is thus modified, but the association of low wall shear with the initiation of the bursts is maintained.

Key words: turbulent boundary layers, big data

1. Introduction

There is widespread agreement that physical phenomena have causes, but less consensus on what this may mean. Several questions come to mind. The first is whether the concept of cause has any meaning when the equations of motion are known, and whether, even if a definition could be agreed upon, it would be of any practical value. For example, Russell (1912) argued that, if the temporal evolution of a dynamical system is described by a set

† Email address for correspondence: javier.jimenezs@upm.es

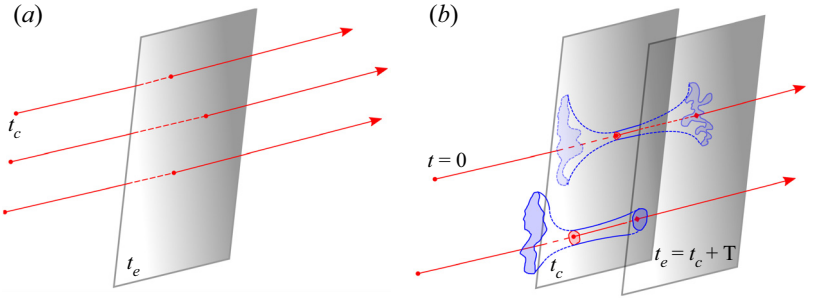


Figure 1. (*a*) The deterministic view of a dynamical system. The shaded plane represents the full phase space at one instant in time, and each trajectory is one possible evolution of the system. (*b*) The effect of dissipation and chaos. Phase points have to be substituted by neighbourhoods, and the flow becomes ill conditioned both forward and backwards in time.

of deterministic differential equations, causality is equivalent to knowledge of the initial conditions. This point of view can be traced to Newton and even to the classical world, and implies that the only causes of the state of the system at time t are the state of the system at any previous time. This is sketched in figure 1(*a*). Disregarding isolated singularities, any point $\boldsymbol{v}(t_e)$ in phase space is the 'effect' of all the points $\boldsymbol{v}(t_c < t_e)$ in a unique incoming trajectory. Conversely, $\boldsymbol{v}(t_e)$ is the 'cause' of all the points in that trajectory for which $t > t_e$.

However, Russell (1912) was probably thinking about reversible Hamiltonian mechanics and, although true in theory, his conclusions are not necessarily useful in more general cases. Many mechanical systems are dissipative, and identifying the (Russell 1912) cause of a particular state implies integrating ill-posed equations backwards in time. This certainly applies to Navier–Stokes turbulence, which is the system that mostly interests us here. Similarly, Russell (1912) knew little about deterministic chaos, but we now understand that most dynamical systems with many degrees of freedom are chaotic, and cannot in practice be uniquely integrated forward. The evolution of turbulence is closer to figure 1(*b*), in which $\boldsymbol{v}(t_e)$ has been substituted by a small neighbourhood, and the forward and backward trajectories become irregular or fractal cones formed by bundles of trajectories that contain the causes and effects of the points in the neighbourhood of $\boldsymbol{v}(t_e)$. Russell's question can be recast as whether, in such situations, something is retained of the deterministic picture in figure 1(*a*).

A related problem is whether something can be said about causality without performing interventional experiments. The usual answer is that it cannot, because the correlations that result from observations do not imply causation (Granger 1969; Pearl 2009). But the discussion in the previous paragraph suggests that this may not be the whole story, and that a sufficiently careful observation of the temporal evolution of a system may lead to the identification of the 'causal' trajectories that cross a neighbourhood of interest (Angrist, Imbens & Rubin 1996).

The coarse graining inherent in figure 1(*b*) suggests that the dynamical system can be simplified by partitioning the phase space into disjoint neighbourhoods of finite size, at least for a fixed temporal horizon. This is common practice in chaotic dynamical systems (Beck & Schlögl 1993) and, although the reasons given are often that it avoids singular measures in the statistics and that numerical experiments are anyway discrete, figure 1(*b*) suggests that there is a more fundamental justification. If chaos prevents us from predicting the behaviour of infinitesimally close neighbouring points, it makes little sense

to insist on treating them as if they were different, and we may as well consider finite-size neighbourhoods as our fundamental dynamical units. This has important consequences for the definition of the system. The main one is that the system propagator is substituted by the 'symbolic' dynamics of how often and in which order the system visits the different cells, and that the deterministic equations are substituted by the transition probabilities incorporated in the Perron–Frobenius (transfer) operator introduced in § 2.

Turbulence is well suited for these techniques, because it is a chaotic deterministic system with many degrees of freedom for which the (Navier–Stokes) equations are known. The difficulty is not how to integrate the equations, which are in principle within reach of a sufficiently powerful computer, but how to explain and predict turbulent flows in terms of simpler rules. Direct simulations are exact but expensive, and we would like to have reduced models that reproduce the flow, if not in full detail, at least well enough to provide general rules about its future behaviour and, ideally, about how that behaviour could be influenced. Shear-driven turbulence is particularly appropriate because it can be made statistically steady, as in pipes or channels, but also because it is believed to be partially controlled by linear processes (del Álamo & Jiménez 2006; McKeon & Sharma 2010; Jiménez 2013), and at least in part describable in terms of coherent structures that play the role of objects in a dynamical system (Adrian 2007; Jiménez 2018*a*).

Especially interesting is the regeneration cycle of wall-bounded turbulence, whose persistence has been explained by the interaction between the perturbations of the streamwise and cross-flow velocities (Jiménez 1994; Hamilton, Kim & Waleffe 1995; Waleffe 1997). There is fairly general consensus that the wall-normal velocity generates fluctuations of the streamwise velocity by deforming the mean shear, and that the shear interacts with the cross-flow fluctuations to amplify them (Orr 1907; Jiménez 2013, 2015). But this amplification is transient in most models (Butler & Farrell 1993; Farrell & Ioannou 1996; Schoppa & Hussain 2002; Jiménez 2013), and the details of how the cycle closes after the burst decays are unclear. The elucidation of this regeneration process is the underlying 'application' of our investigation, although much of the paper is dedicated to the development of the analytical procedure itself. Early examples of the use of transfer operators for burst identification in reduced-order models of the wall-turbulence cycle are Schmid, García-Gutiérrez & Jiménez (2018*a*) and Schmid *et al.* (2018*b*).

Note that most of the results of our analysis will not be causal in the sense of Granger (1969) or Pearl (2009), since they involve no intervention from the observer. But we are more interested in predictability and perhaps in coherence, and in the search for states of the system that best allow us to draw conclusions from partial flow information. The fundamental question of causality will be outsourced here to the equations of motion, and its direction to the direction of time. The main purpose of our analysis is to identify flow configurations in which the equations of motion give us the best possible information about the future of the system without necessarily solving them in detail, and which could perhaps lead to effective control strategies.

However, a simple interventional experiment will also be presented towards the end of the paper to help us confirm our conclusions, and to address certain limitations of the non-interventional analysis that will become apparent in the course of our discussion.

The organisation of the paper is as follows. Section 2 introduces the Perron–Frobenius operator, which is particularised to a small-box turbulent channel in § 3. Techniques for its use are developed in §§ 3.2 and 3.3, leading in § 4 to the study of conditional trajectories in phase space. Finally, the interventional experiment is described in § 5 and conclusions are offered in § 6.

2. The Perron–Frobenius operator

Assume a statistically stationary ergodic system

$$\boldsymbol{v}(t+T) = \boldsymbol{S}(T;t)\,\boldsymbol{v}(t), \tag{2.1}$$

for which temporal and ensemble averages can be interchanged. The probability density of the state variable, $\boldsymbol{v}$, over the cells of a partition $\{C_j | j = 1 \dots N\}$ of the phase space, can be approximated by the fractional distribution, $\boldsymbol{q} = \{q_j\}$, of the time spent by the system within each cell. After a sufficiently long time, or for a sufficiently large ensemble of experiments, these probabilities tend to an equilibrium distribution that we denote by $\boldsymbol{q}_\infty$. More locally, if we consider the probability distributions at two different times, $\boldsymbol{q}(t)$ and $\boldsymbol{q}(t+T)$, the two-dimensional Perron–Frobenius operator (PFO) $\hat{\boldsymbol{P}}^e$, relates the past to the future (Beck & Schlögl 1993),

$$\boldsymbol{q}(t+T) = \hat{\boldsymbol{P}}^e(T;t)\,\boldsymbol{q}(t). \tag{2.2}$$

Because probabilities represent the results of mutually independent tests, $\hat{\boldsymbol{P}}^e$ is linear and, for a finite partition, reduces to an $N \times N$ matrix, where N is the number of cells in the partition, which is potentially much larger than the number of degrees of freedom of the original dynamical system. We will assume $\hat{\boldsymbol{P}}^e$ to be independent of t.

When applied to a perfectly concentrated initial distribution, $\boldsymbol{q}^{(a)}(t) = \{\delta_{aj}\}$, where δ_{aj} is Kronecker's delta, the ath column of $\hat{\boldsymbol{P}}^e$ represents the probability that a system initially within the ath cell evolves into the different cells of the partition after the time interval T. Note that these concentrated initial probability distributions can be interpreted as non-interventional experiments, in which a statistical knowledge of the causal structure of the coarse-grained system can be gained by observing the system over a sufficiently long time (Angrist *et al.* 1996).

The PFO is equivalent to the Bayesian conditional probability matrix (Feller 1971), and can be estimated, after a sufficiently long observation, as (Ulam 1964)

$$Q_{ij}(t, t+T) \equiv Q_{ij}(T) = \mathrm{prob}_t(\boldsymbol{v}(t+T) \in C_i,\, \boldsymbol{v}(t) \in C_j), \tag{2.3}$$

normalised to unit column sums

$$P^e_{ij} = Q_{ij} \Big/ \sum_s Q_{sj}, \tag{2.4}$$

so that an input probability $\boldsymbol{q}(t)$ for which $\sum q_j = 1$ results in a similarly normalised output probability $\boldsymbol{q}(t+T)$. The matrix $\boldsymbol{P}^e$ is generally not symmetric, and there is a dual matrix

$$P^c_{ij} = Q_{ji} \Big/ \sum_s Q_{is}, \tag{2.5}$$

which generates $\boldsymbol{q}(t-T)$ given $\boldsymbol{q}(t)$, allowing us to estimate the statistical distribution of the causes of a given effect. Note that, even if

$$\boldsymbol{q}(t-T) = \hat{\boldsymbol{P}}^c(T)\,\boldsymbol{q}(t) \tag{2.6}$$

looks like the inverse of (2.2), $\hat{\boldsymbol{P}}^e$ is not the inverse of $\hat{\boldsymbol{P}}^c$, because the marginal probabilities $\boldsymbol{q}(t)$ and $\boldsymbol{q}(t+T)$ have different meanings in (2.2) and in (2.6). In the former, $\boldsymbol{q}(t)$ is observed, and $\boldsymbol{q}(t+T)$ is the conditional probability distribution at $t+T$ given

that observation, while their meaning in (2.6) is reversed. One of the effects of the coarse-grained partition is to destroy any reversibility that might have been present in the original dynamical system.

Another consequence of discrete partitions is to suppress the semigroup character of the dynamical system, by which $\boldsymbol{S}(T_1 + T_2) = \boldsymbol{S}(T_1) \circ \boldsymbol{S}(T_2)$. Indeed, even if the original dynamical system is Markovian in the sense that its future depends only on its present state (i.e. on its 'initial conditions'), the discretised system is generally not Markovian. The cells of almost any partition of a high-dimensional phase space are projections of infinite 'cylinders' whose base is the cell, and whose 'axis' extends along all the neglected system dimensions. Two trajectories that intersect a cell at a given time may actually intersect its cylinder at very different places along the axis, and the only way to distinguish different trajectories is often to consider the sequence of cells visited over their entire past. Even this may not be enough, and very little is known about partitions that preserve Markovianity in high-dimensional systems (Beck & Schlögl 1993, § 3.6). The transfer operator bypasses this limitation by acting on the transition probabilities, and is again Markovian in the sense that $\boldsymbol{q}(t + T)$ formally only depends on $\boldsymbol{q}(t)$ (Feller 1971, § X), but we regain Markovianity at the expense of losing determinacy, and we will see in § 3.1 that the semigroup property, $\boldsymbol{P}^{\varepsilon}(nT) = \boldsymbol{P}^{\varepsilon}(T)^n$, is very quickly lost for the approximate transfer operator of turbulent channels.

There are several reasons why $\boldsymbol{P}^{\varepsilon}$ is not a perfect estimator of the true operator $\hat{\boldsymbol{P}}^{\varepsilon}$, but the most important one has to do with the existence of an attractor. Dissipative systems, such as turbulence, typically evolve towards a lower-dimensional attracting subset of the full phase space, and the observations used in (2.4) only reflect the statistics of this subset. As such, $\boldsymbol{P}^{\varepsilon}$ is a restriction of $\hat{\boldsymbol{P}}^{\varepsilon}$ to the system attractor, and contains little or no information about how the system reacts outside it. It is thus useful in modelling the physics, where the interest is on how the system evolves in time, but it may need additional information in control applications, where we may wish to act in ways outside the attractor.

There are two ways in which the PFO can be used to analyse a complex dynamical system. The first one is to treat it as a matrix whose properties reflect the behaviour of the attractor as a whole. 'Stochastic' matrices like $\boldsymbol{P}^{c}$ or $\boldsymbol{P}^{\varepsilon}$, with non-negative elements and unit column sums, have useful properties that have been extensively studied, especially if care is exercised in dealing with the zero entries that represent cells that are never visited by the system (Lancaster 1969). Their best known property is that they possess a unit leading eigenvalue with a real eigenvector with non-negative entries, which can be interpreted as a probability distribution over the partition. For $\boldsymbol{P}^{\varepsilon}$, this eigenvector satisfies, $\boldsymbol{q}_1(t + T) = \boldsymbol{P}^{\varepsilon}(T)\boldsymbol{q}_1(t) = \boldsymbol{q}_1(t)$, and defines a probability density that remains invariant as the system evolves, and which is therefore identical to the natural invariant density, $\boldsymbol{q}_\infty$, mentioned earlier in this section. The subdominant eigenvalues control the approach to $\boldsymbol{q}_\infty$ when the initial distribution is different from the natural one, as well as whether the attractor can be partitioned into approximately disjoint subsets (Froyland 2005).

As already mentioned, these are examples of global properties that apply to the full attractor. The same is true of other approximation strategies, such as proper orthogonal decomposition (POD, Berkooz, Holmes & Lumley 1993) or dynamic-mode decomposition (Schmid 2010), which use ergodicity to minimise global errors of reduced models but are of limited use for our purpose. The root of the problem is that neither is the PFO a true matrix, nor is the probability distribution $\boldsymbol{q}$ a true vector. Neither of them forms a linear space (e.g. $2\boldsymbol{q}$ or $-\boldsymbol{q}$ are not probability distributions), and concepts required for most global optimisation strategies, such as norms or inner products, take a different meaning. For example, the min–max property of eigenvectors (Courant & Hilbert 1953) that is

behind some of the global methods mentioned above, is unlikely to survive in the presence of constraints such as $q_j \geq 0$. We are more interested in local analyses that use the PFO as a joint probability distribution to compute state-dependent conditional averages, and to study the expected short-term behaviour of the system in the neighbourhood of a particular cell. Our goal is to find whether some cells are more likely than others to form the basis for better predictions, and are thus 'more causal'.

The main problem in constructing the PFO is the choice of a partition in which the two-dimensional histogram (2.3) can be populated with a given set of data. Consider a system that has been projected onto a subspace of dimension D, each of whose variables is discretised into n_b bins. The D-dimensional bins in the histogram can be indexed as a vector of length $N = n_b^D$, so that the $N \times N$ matrix $\boldsymbol{Q}$ has

$$n_Q = N^2 = n_b^{2D} \tag{2.7}$$

elements. Interpreting (2.3) as a probability density requires that the number of available 'training' data be sufficiently larger than (2.7). For example, using $O(10)$ data per bin results in expected errors in Q_{ij} of the order of $1/\sqrt{10} \approx 30\,\%$. If we assume a moderate $n_b \approx 10$, this implies a minimum of $n_D \approx 10^{2D+1}$ data, and limits in practice the dimension to $D \lesssim 2$, requiring $n_D \lesssim 10^5$ data. Increasing it to $D = 3$ would increase the required number of data by a factor of 100, as well as the cost of creating and analysing them. We will see below that this limit can be relaxed a little because D is the dimension of the attractor rather than of the phase space, but the attractor of a turbulent flow is unlikely to have $D < 3$, and the savings are at most a factor of $O(1)$. Our strategy will thus be to analyse two-dimensional projections of the phase space, using different combinations of physically motivated variables, and choose those pairs that produce interesting results.

Note that, although the discretised variables are treated as a single long vector, the probabilities in $\boldsymbol{Q}$ correspond to the simultaneous occurrence of D independent variables, allowing us to study effects depending on the coincidence of several 'causes' (Pearl 2009). Although we use $D = 2$ in this paper, this could in principle be extended to the increasingly unlikely coincidence of three or more independent causes, but, as mentioned above, this quickly runs into the limitation of the number of available data.

3. Application to minimal channels

Most of our analysis centres on a data set already used in Jiménez (2015). A pressure-driven spatially periodic turbulent channel flow is simulated between parallel plates separated by $2h$. The wall-parallel periods of the computational box are $L_x = \pi h/2$ and $L_z = \pi h/4$, and the nominal friction Reynolds number is $h^+ = hu_\tau/\nu = 950$, where x, y and z are the streamwise, wall-normal and spanwise coordinates, respectively, and the corresponding velocity components are u, v and w. Capital letters, as in $U(y)$, denote y-dependent ensemble averages, $\langle\,\rangle$, and lower-case ones are perturbations with respect to this average. Primes are reserved for root-mean-squared intensities, and the kinetic energy of the fluctuations is defined as $E = u^2 + v^2 + w^2$. The '+' superscript denotes 'wall' normalisation with the kinematic viscosity ν, and with the friction velocity $u_\tau = \sqrt{\nu \partial_y U}$. The code is standard fully dealiased Fourier–Chebyshev spectral, as in Kim, Moin & Moser (1987), and the mass flux is kept constant. Time is usually normalised with the eddy turnover time h/u_τ, and is then denoted by an asterisk, $t^* = u_\tau t/h$. More details can be found in Jiménez (2013).

To improve statistics, the simulation was extended in time to $t^* \approx 650$, and sampled at a time interval between frames, $\Delta t^* \approx 0.025$. Such simulations are minimal within a band of

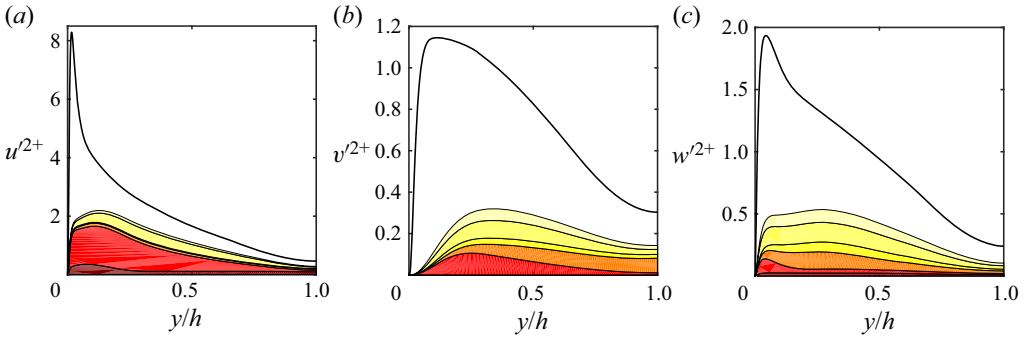


Figure 2. The coloured patches are profiles of the cumulative variance of the harmonics retained in this paper. From bottom to top, and dark to light: [00], [01] + [01*], [10], [20], [11] + [11*], [21] + [21*]. The solid black line is the total variance of the velocity component. (*a*) Streamwise velocity. (*b*) Wall-normal velocity. (*c*) Spanwise velocity.

wall distances $y/h \approx 0.2$–0.6 (Flores & Jiménez 2010), in the sense that a non-negligible fraction of the kinetic energy is contained in the first few largest wall-parallel Fourier modes. Closer to the wall, the flow contains a wider range of energy-containing scales, and cannot be considered minimal. Farther from it, the simulations cannot be directly compared with canonical turbulence, because some of the largest scales are missing. The range of wall distances mentioned above approximately includes a single largest structure that bursts irregularly. Since it was shown by Flores & Jiménez (2010) that the typical interval between bursts is $t^* \approx 2$–3, the simulation analysed here contains several hundreds of bursts per wall, and approximately 100 samples per burst. Moreover, since the box is too small to allow healthy large scales in the central part of the channel, the two walls are treated as independent realisations (the cross-correlation coefficient is less than 0.05 for the variables discussed below). The total number of data snapshots is thus approximately 5×10^4.

If we define Fourier expansions of the three velocity components along x and z as

$$a(x, y, z) = \sum_{m,n} \tilde{a}_{mn}(y) \exp[\mathrm{i}(k_x x + k_z z)], \tag{3.1}$$

where a is the variable to be expanded, $k_x = 2\pi m/L_x$ and $k_z = 2\pi n/L_z$. Whenever there is no ambiguity in the variable being expanded, Fourier coefficients are designated as $[mn]$. As mentioned above, only the largest structures at a given distance from the wall can be expected to be describable by relatively few degrees of freedom whose dynamics can be easily studied, and our analysis only retains the first few modes, $m = 0, 1, 2$ and $n = -1, 0, 1$. Appendix A explains how modes with $n \neq 0$ are used as combinations of the $\pm n$ pair, resulting in two equivalent modes displaced spanwise by a quarter of a wavelength. Although spanwise homogeneity ensures that the interactions of these combinations with the $n = 0$ modes are statistically equivalent, they interact non-trivially among themselves, and both combinations are retained. They are designated, for example, as [21] and [21*]. Profiles of the cumulative variance of all the modes retained in the paper are given in figure 2. They show that their overall energy is a comparatively small but non-trivial fraction of the total, and we will see later that they follow fairly independent dynamics. In addition, the retained modes account for approximately 65 % of the tangential Reynolds stress, $-\langle uv \rangle$ (not shown). Note that, because of the small computational box, there is substantial energy in the [00] modes of u and w, whose only fluctuations are temporal. They can be considered as approximately modelling the spatial variation of the

mean velocity profile over wall patches of the order of the size of the computational box. An example of instantaneous flow field synthesised from the retained modes can be found in figure 17 in Appendix B.

This limited subset of data still contains a large number of degrees of freedom, because each Fourier component is a function of y with $O(100)$ grid points. Even if we will see later that the wall-normal resolution can be reduced to $O(10)$ points through judicious filtering, the raw degrees of freedom for each velocity component is $O(100)$ complex numbers, and we mostly restrict ourselves to analysing the behaviour of a few integrated 'summary variables' that represent global properties of the velocity within the chosen band of wall distances. In particular, if we are interested in the band $y \in (y_0, y_1)$, we follow Jiménez (2013, 2015) in using an integrated intensity,

$$I^2_{a,mn} = \frac{1}{y_1 - y_0} \int_{y_0}^{y_1} |\tilde{a}^+_{mn}|^2 \,\mathrm{d}y, \tag{3.2}$$

which stands for the velocity magnitude and, when $k_x \neq 0$, an average tilting angle

$$\psi_{a,mn} = -\arctan\left(\mathrm{Im} \frac{\displaystyle\int_{y_0}^{y_1} \tilde{a}^\dagger_{mn} \partial_y \tilde{a}_{mn} \,\mathrm{d}y}{\displaystyle k_x \int_{y_0}^{y_1} |\tilde{a}_{mn}|^2 \,\mathrm{d}y} \right), \tag{3.3}$$

where 'Im' is the imaginary part, and the dagger stands for complex conjugation. This angle varies from $-\pi/2$ to $\pi/2$, and describes the wall-normal structure of the phase of the Fourier mode.

Several other summary variables were considered, either based on physical arguments or on standard statistical methods (e.g. individual POD modes), but they did not add appreciably to the argument or to the conclusions. They are not discussed in the rest of the paper, except for the use of PODs as a filtering device to balance the wall-parallel and wall-normal resolutions of the retained flow fields, as explained in Appendix B.

Because the retained harmonics exclude the smallest scales, they can be trusted closer to the wall than the full flow, and all our results use an integration band $y^+ > 40$ and $y/h \leq 0.6$. Somewhat narrower or wider ranges were tested with little effect on the results.

Not all the summary variables defined in this way are mutually fully independent. Figure 3 presents their correlation coefficient

$$C_{ab} = \frac{\langle (a - \langle a \rangle)(b - \langle b \rangle) \rangle}{\langle (a - \langle a \rangle)^2 \rangle^{1/2} \, \langle (b - \langle b \rangle)^2 \rangle^{1/2}}. \tag{3.4}$$

Several things stand out. The u and v components form reasonably well-correlated pairs, particularly among similar summary variables and Fourier modes, but most quantities involving w are not well correlated with u and v, or among themselves. The correlation between the intensities of u and v are significant because, even if they involve integrated quantities rather than the variables themselves, they reflect the generation of the tangential Reynolds stress, $-uv$. The higher modes, [11] and [21], tend to be better correlated among different variables than the lower ones, [10] and [20]. In particular, three of the highest correlations in figure 3 are $C(I_{v01}, I_{w01*}) \approx 0.70$, $C(I_{u11}, I_{w11*}) \approx 0.74$ and $C(I_{v11}, I_{w11*}) \approx 0.59$, notwithstanding the generally poor correlation between the summaries of w and those of other velocity components. Interestingly, these correlations come in $[m1, m1*]$ pairs, representing flow structures offset from each other by a quarter of a spanwise wavelength. They correspond to the inclined 'rollers' that have often been described in wall-bounded flows.

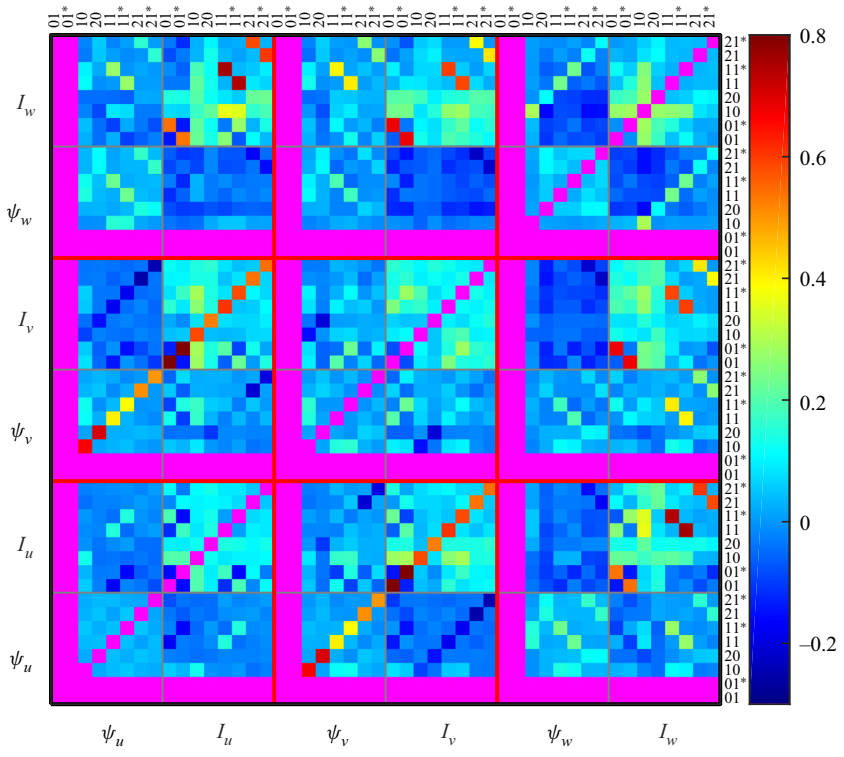


Figure 3. Correlation coefficient among the different summary quantities. Large squares outlined in red correspond to the three velocity components. Smaller squares outlined in grey are summary variables, and the smallest cells within each grey square are Fourier modes, in the order [01], [01*], [10], [20], [11], [11*], [21], [21*], from bottom to top and from left to right. The main diagonal has been blocked for clarity, as well as the inclinations for modes with $k_x = 0$, which are undefined.

Somewhat surprisingly, angles and intensities are generally uncorrelated, including the (I_{v10}, ψ_{v10}) pair that was shown by Jiménez (2013, 2015) and Encinar & Jiménez (2020) to be particularly useful, because its joint probability distribution is traversed by the flow in a physically interpretable way. This shows that correlations and coherence are different concepts. As a simple example, the temporal evolutions of $\sin(t)$ and $\cos(t)$ are orthogonal and uncorrelated, but they form a coherent pair in the sense that they transverse a one-dimensional circular sub-manifold of their phase space.

It could be tempting to use as summary variables the eigenvectors of the dominant eigenvalues of the matrix in figure 3, since they represent combinations of variables that optimally explain the variance of the data (Berkooz *et al.* 1993), but they turn out to be especially bad at describing the dynamics. This can best be understood by looking at the joint probability density of (I_{v10}, ψ_{v10}) in figure 4(*a*). It is clear that I_{v10} does not explain ψ_{v10}, nor *vice versa*, which is precisely why the pair can be used to define two-dimensional causal combinations.

Nevertheless, we have already mentioned that correlation does not imply causation, and figure 3 can at most be taken as indicative of which variables could be considered as causally related. We have just discussed reasons why a causal variable pair may not be correlated. We will see in § 3.2 how the opposite situation allows us to draw conclusions about the dynamics of specific structures.

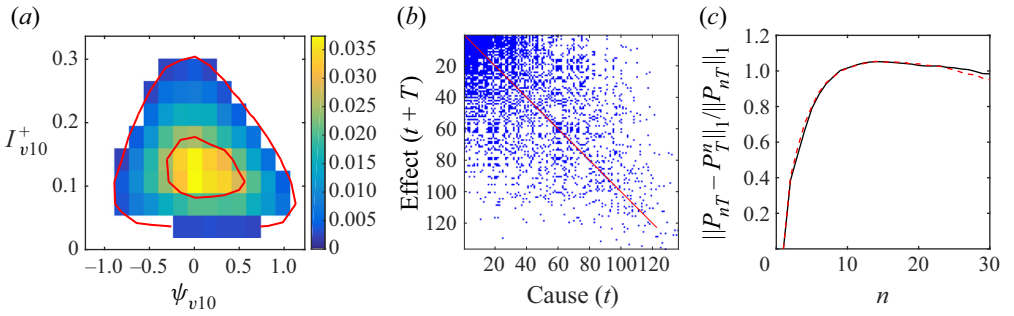


Figure 4. (*a*) Two-dimensional joint probability distribution, $\boldsymbol{q}_\infty$, for the integrated inclination and amplitude of the [10] mode of the wall-normal velocity component (i.e. $k_x = 2\pi/L_x$, $k_z = 0$), averaged over $y^+ > 40$ and $y/h < 0.6$. The inclination is partitioned in 15 equal bins, and the amplitude in 13 bins. The red contours contain 30 % and 95 % of the probability mass, and only cells within the outer contour are plotted. (*b*) Binary map of the non-zero elements of the joint probability $\boldsymbol{Q}(T^* = 0.076)$ in (2.3), obtained by collating the variables in (*a*) into a single vector. Rows and columns with zero sum have been eliminated, and the rest are arranged in order of decreasing column sum. The diagonal is highlighted in red. (*c*) The L_1-norm Markov test for the Perron–Frobenius matrices. ——, $\boldsymbol{P}^e$; – – –, $\boldsymbol{P}^c$.

3.1. *The transfer operator of the minimal channel*

We saw at the end of § 2 that the main problem in constructing the PFO for a given system is the choice of variables to express the underlying partition. Figure 4 presents results for the minimal channel just described, using as variables the inclination angle of the wall-normal velocity, ψ_{v10}, and its root-mean-squared amplitude, I_{v10}, whose joint probability distribution is shown in figure 4(*a*). Most of the distribution is contained within the inner probability contour, but the outer fringe is interesting because Jiménez (2015) showed that its upper edge can be modelled as a linearised burst in which the mean shear amplifies the velocity perturbations by tilting them forward (Orr 1907; Jiménez 2013).

The construction of the PFO starts by organising the 15×13 partition of the parameter space of figure 4(*a*) into a single vector of length 195, and constructing the two-time joint distribution, $\boldsymbol{Q}(t, t+T)$, from all the snapshots in the data sequence. The interval used in figure 4, $T^* = 0.076$, is chosen from the experience in Jiménez (2015) and Encinar & Jiménez (2020), and is the time taken by the system to traverse an increment $\Delta\psi \approx 0.3$ along the upper edge of figure 4(*a*). It is also the time over which Jiménez (2015) shows that the flow can be linearly predicted in that region of the phase plane.

Columns and rows of $\boldsymbol{Q}$ that contain only zeros are discarded, and those with non-zero elements are normalised using (2.4). The reduced $\boldsymbol{Q}$ is shown in figure 4(*b*), where cells have been sorted in order of decreasing column sum for graphical purposes. The restriction of the full $\hat{\boldsymbol{P}}^e$ to the on-attractor estimate $\boldsymbol{P}^e$ is done when discarding zeros at this step.

We can now quantify the effect of the limited number of data discussed at the end of § 2. The 5×10^4 snapshots used here have to populate the 138^2 matrix $\boldsymbol{P}^e$ in figure 4(*b*), giving an average of 2.5 phase points per matrix element. In practice, they range from $O(100)$ data for the better populated matrix elements to zero for elements outside the attractor. The limitation is not as strict for the smoother distribution in figure 4(*a*), which is essentially an eigenvector of the PFO (see § 2). Its dimension is just 138 cells, and each cell represents $O(300–1000)$ data points. Tests with partitions of the order of 10×10 cells did not qualitatively change the results described below, but attempts to use much finer partitions ran into problems at the interesting edge of the distribution.

As a consequence, we restrict ourselves to two-dimensional projections of the phase space, and explore in § 3.2 which pairs of summary variables give more interesting results. This is supplemented in § 3.3 by conditionally averaging other variables over these projection planes, recovering part of the three-dimensional dynamics.

In addition, to ensure that the noise in our results is a consequence of the discrete partition rather than of an insufficient number of data, the distribution in figure 4(*a*), and similar later ones, are only drawn within the probability isocontour containing 95 % of the total probability mass. Each cell along this contour contains $O(100)$ data snapshots. The global averages in § 3.2 are also computed within this high-probability region, and the analysis was repeated with half the number of data, with similar conclusions.

Figure 4(*c*) tests the non-Markovian behaviour of $\boldsymbol{P}^c$ and $\boldsymbol{P}^e$ discussed in § 2, which would imply that $\boldsymbol{P}(nT) = \boldsymbol{P}(T)^n$ for all n and T. The figure shows the relative Frobenius norm of the difference between $\boldsymbol{P}(nT)$ and $\boldsymbol{P}(T)^n$ for the shortest available time interval, $T^* = 0.025$, which should be the most deterministic case. The relative difference between uncorrelated stochastic matrices depends on the ratio between the standard deviation and the mean of individual matrix columns, but is approximately unity for cases such as those in figure 4. It is clear that the two matrices being tested become essentially uncorrelated after $O(10)$ time steps, corresponding to one quarter of a turnover time. Assuming Markovianity therefore modifies the dynamics of the system, and we use $\boldsymbol{P}(nT)$ as our basic operator from now on.

Figure 5 shows how the PFO can be used to extract the probability distributions of the causes and effects of a given observation. Figure 5(*a*) assumes that we know that the system is within the cell marked with a solid circle at $t = 0$. The conditional probability distribution at $t = T$ is given by the corresponding column of the transfer operator, $\boldsymbol{P}^e$, and is displayed in the figure in dashed blue contours. Conversely, the conditional probability distribution of causes at $t = -T$ is the corresponding column of the backwards operator $\boldsymbol{P}^c$. It is displayed in solid black lines, and the difference between the two distributions illustrates the temporal evolution of the system in the clockwise direction of the figure, as in Jiménez (2015).

The segregation into forward and backward distributions does not hold for all cells. Figure 5(*b*) applies the procedure to a cell in the high-probability core of the invariant density distribution. Its forward and backward distributions are marked as in figure 5(*a*), but they overlap each other and are difficult to tell apart.

Figure 5(*c*) is a representation of this mean displacement for all the cells in the distribution. The arrows join the centre of each reference cell to the mean position of its effects after a given time interval. Figure 5(*d*) does the same for the causes, and both figures show a mean clockwise displacement of the system along the upper edge of the distribution (see figure 3(*b*) in Jiménez 2015, for comparison). In addition to this circular displacement, the arrows spiral towards the centre of the distribution in figure 5(*c*), and outwards in figure 5(*d*). This tendency increases for longer time intervals, and is due to the non-Markovian component of the probability evolution.

Any random displacement from the periphery tends to move towards the most probable locations in the central part of the distribution, and random displacements ending in the periphery are most likely to come from the core. This is best seen in figure 5(*e*,*f*), which is computed in the same way as figure 5(*c*,*d*) after randomly shuffling the time stamps of the flow snapshots. In fact, since temporal relations are destroyed by the shuffling, effects and causes become randomly chosen states of the system and their conditional average coincides with the overall mean of the invariant distribution. These randomised figures are independent of the time interval.

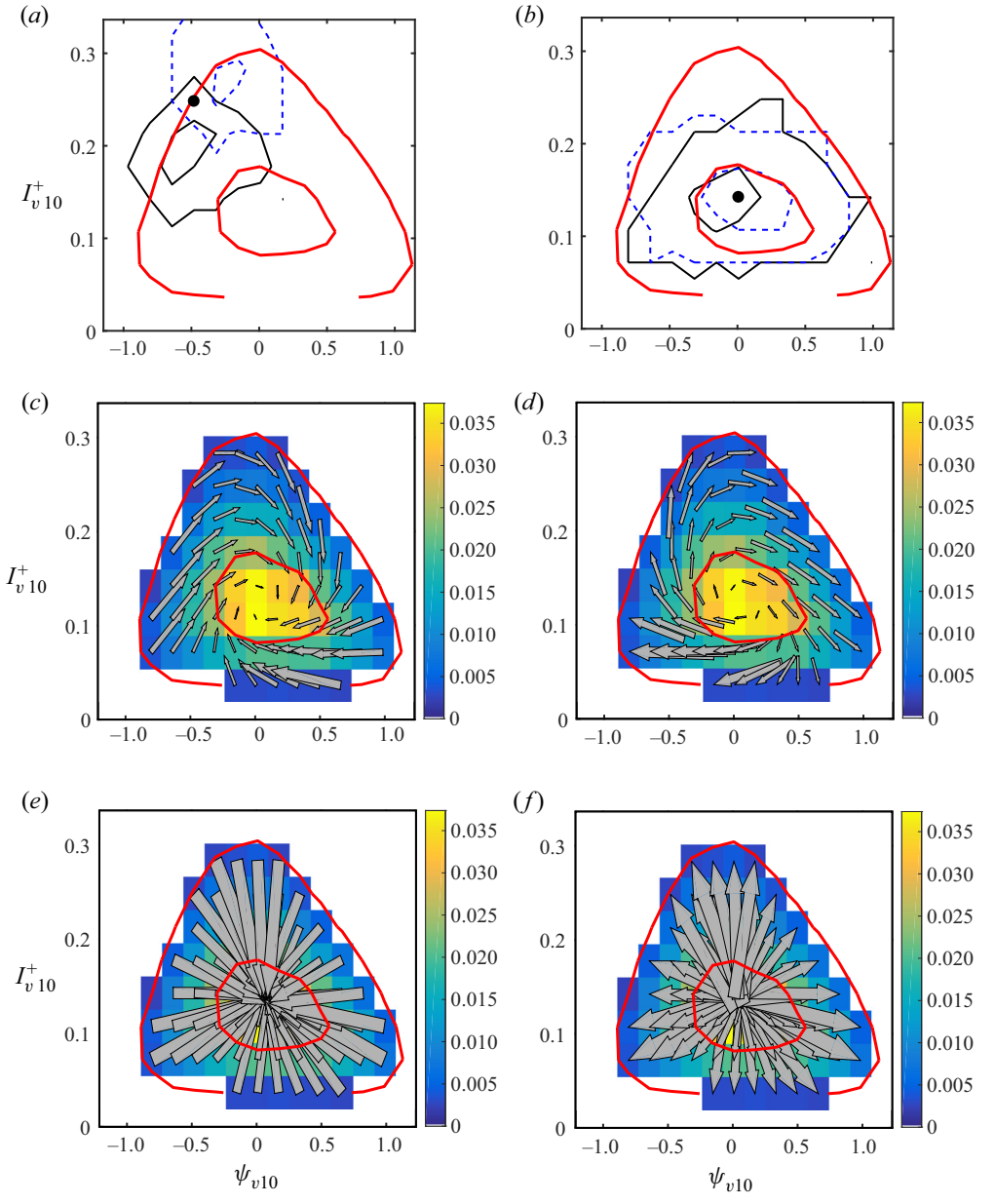


Figure 5. (*a*) For the variables in figure 4(*a*), and an interval $T^* = 0.076$, the solid black contours are the probability distribution of possible T-precursors to an observation of the cell marked with a solid circle, and the dashed blue contours are the distribution of possible effects after T. Contours contain 30 % and 95 % of the probability mass. (*b*) As in (*a*), for an observation in the core of the invariant density distribution. (*c*) Mean system displacement in the parameter plane. The coloured background is the invariant density, and the arrows join the cell taken as cause with the mean system location after time T. (*d*) As in (*c*), but the arrows join the mean location of the systems that will pass through the cell taken as reference, after time T. The red contours contain 30 % and 95 % of the invariant density. (*e*,*f*) As in (*c*,*d*), but using randomised time stamps for the data.

3.2. *Quality indicators*

We have mentioned several times that the main problem in constructing the PFO is collecting enough data to populate the two-dimensional histogram, $\boldsymbol{Q}$, making it unpractical to consider distributions over more than two independent variables. We have also mentioned that our strategy is to test many possible variable pairs in the hope of identifying couples whose statistical behaviour is optimal, but the 36 variables used in figure 3 can be paired into 630 possible ways, and automating the search requires indicators that are simpler to implement than the visual inspection of the two-dimensional plots in figure 5. Four such indicators are discussed in this section.

The statistical uncertainty of the displacement vectors is addressed in figure 6(*a*), which displays the ratio between the standard deviation of the conditionally averaged displacement of the system over a given time and its mean. To compensate for the different magnitudes of the two variables in the figure, which generally have incompatible units, each of them is normalised with its global standard deviation before computing the conditional statistics. The result is a measure of the error bars associated with each of the arrows in figure 5(*c*).

Having a small relative standard deviation does not guarantee that a quantity is physically relevant. Inspection of figure 5(*c–f*) reveals that deterministic and random evolutions behave differently with respect to the asymmetry between causes and effects. The displacement vectors of the causes and effects rotate in the same direction in figure 5(*c*,*d*), because both represent the deterministic evolution of the system. But the randomised vectors in figure 5(*e*,*f*) point in opposite directions, because they move from the conditioning cell towards the densest part of the distribution, independently of the direction of time. As a consequence, we can define a 'determinacy' index for an observation cell $\boldsymbol{v}_0$ as the normalised inner product

$$C_{ce}(\boldsymbol{v}_0, T) = \frac{(\boldsymbol{v}^e - \boldsymbol{v}_0) \cdot (\boldsymbol{v}_0 - \boldsymbol{v}^c)}{\|\boldsymbol{v}^e - \boldsymbol{v}_0\| \|\boldsymbol{v}_0 - \boldsymbol{v}^c\|}, \tag{3.5}$$

where $\boldsymbol{v}^e - \boldsymbol{v}_0$ and $\boldsymbol{v}_0 - \boldsymbol{v}^c$ are, respectively, the conditionally averaged displacement of effects and causes over the time interval $\pm T$. As in figure 6(*a*), variables are normalised with their standard deviation before computing (3.5). This index is an indication of how deterministic is the evolution of the system in the neighbourhood of $\boldsymbol{v}_0$, and of how much information is gained by the observation of the variable pair. When the system is completely deterministic in the subspace being considered, $C_{ce} \approx 1$, and when it is essentially random, $C_{ce} \approx -1$. Figure 6(*b*) displays C_{ce} for the data in figure 5, and shows that the evolution of this particular Fourier mode is deterministic almost everywhere in this parameter plane.

Along the upper edge of the distribution, this agrees with the physically based conclusions of Jiménez (2015), but not along its lower edge, where both Jiménez (2015) and Encinar & Jiménez (2020) conclude that the average displacement is opposite to the predictions of the model that explains the upper edge, and that the uncertainty of the displacements is too large for the means to be trusted. The high uncertainty in this region is clear in figure 6(*a*), but figure 6(*b*) suggests that this part of the distribution is also deterministic. Part of the reason is the longer time interval used in figure 6(*a*) compared with 6(*b*). The apparent randomness of the evolution increases for longer intervals, as the non-Markovian behaviour takes over. The determinacy index is almost unity in figure 6(*b*), where the displacements are of the order of one distribution cell, but decreases to $C_{ce} \approx 0.8$ when the figure is drawn for the more physically relevant time interval used in figure 6(*a*), and decreases further to $C_{ce} \approx 0.5$ for the even longer interval

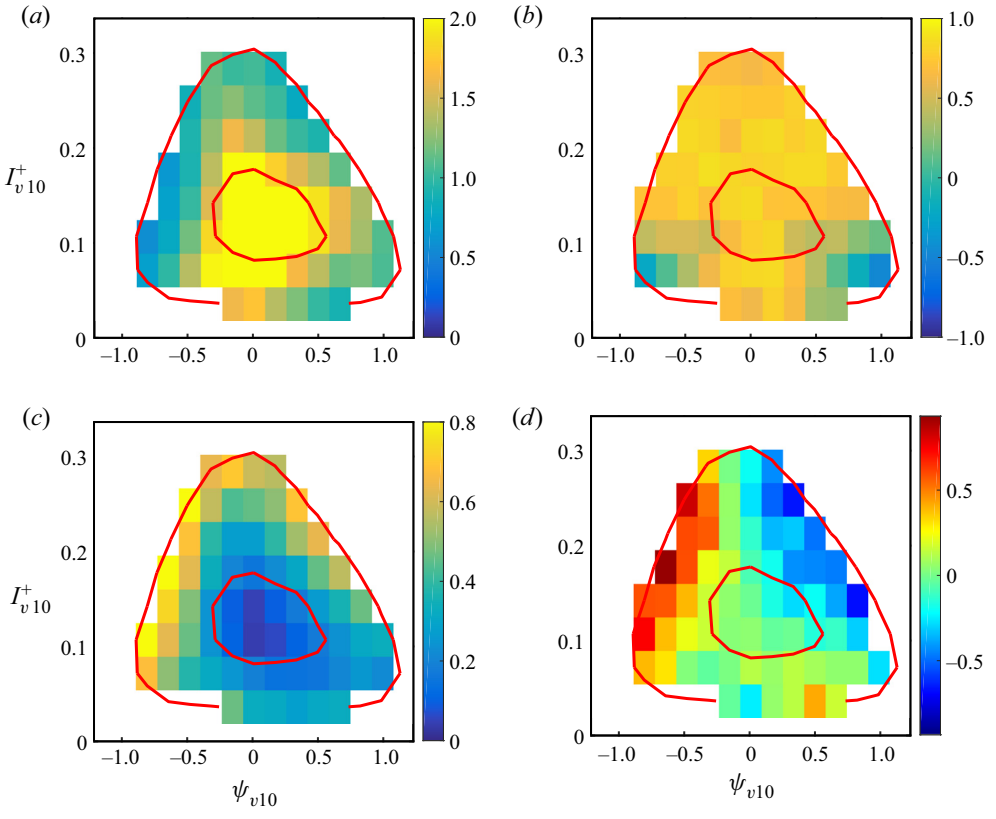


Figure 6. As in figure 4. (*a*) Ratio between the averaged displacement of the effects and their standard deviation. The aspect ratio of the geometry normalises each variable with its overall standard deviation to compensate for the different units. (*b*) Determinacy index (3.5) between the average displacement of causes and effects. Drawn for $T^* = 0.025$. (*c*) Hellinger segregation index (3.6) between the forward and backwards conditional distributions, as function of the observation cell. (*d*) Kullback–Leibler information gain (3.8) from the distributions of the effects and the causes, measured in bits. Warm colours represent creation of information, and cooler ones represent information loss. All panels refer to the [10] mode of the wall-normal velocity, and use only cells within the 95 % probability contour of $\boldsymbol{q}_\infty$. In all panels, except (*b*), $T^* = 0.076$.

used in Jiménez (2015). Encinar & Jiménez (2020), who use a different method from the one above, and a different set of data, compute a figure of merit equivalent to the relative dispersion in figure 6(*a*). Normalising their time offset with the average distance, $\bar{y}$, from the wall of their filtered fields (Flores & Jiménez 2010; Jiménez 2015), it varies between $u_\tau T/\bar{y} = 0.048$ and 0.19. The resulting standard deviations are negligible for the shortest of those intervals, but large enough to reverse some of the displacements for the largest one. When these values are applied to the present case, assuming $\bar{y} \approx 0.3$ for our integration band, the time interval in figure 6(*b*) is $u_\tau T/\bar{y} = 0.087$, and that in figure 6(*a*) is $u_\tau T/\bar{y} = 0.26$, explaining the apparent discrepancy between figures 6(*a*) and 6(*b*).

Figure 6(*c*) quantifies the temporal segregation between the conditional probability distributions of causes and effects in figure 5(*a*,*b*). The distance between two normalised probability distributions $\boldsymbol{q}^{(1)}$ and $\boldsymbol{q}^{(2)}$ can be characterised by the Hellinger norm

(Nikulin 2001), defined as

$$H^2(\boldsymbol{q}^{(1)}, \boldsymbol{q}^{(2)}) = \frac{1}{2}\sum_j \left(\sqrt{q_j^{(1)}} - \sqrt{q_j^{(2)}}\right)^2, \tag{3.6}$$

which vanishes for $\boldsymbol{q}^{(1)} = \boldsymbol{q}^{(2)}$, and reaches its maximum, $H = 1$, for disjoint distributions. In the case of figures 5(*a*,*b*) and 6(*c*), the distance between the conditional distributions of causes and effects varies from $H \approx 0.9$ at the edge of the density distribution, where they are clearly different, to $H \approx 0.1$ at the centre, where past and future are almost indistinguishable.

The information provided by the indices (3.5) and (3.6) is related but not identical. While a high value of (3.6) implies that causes and effects are different, a high value of (3.5) also shows that the directions of the mean drift associated with each of them are similar, and that the flow of probability can be described as a smooth vector field.

When figure 6(*a–c*) is considered, together the panels suggest that the top-right and top-left edges of the probability distribution are populated by systems which evolve in a fairly deterministic manner, while the lower edge of the distribution, and especially its central core, are more random.

Finally, figure 6(*d*) addresses the question of whether this evolution has any effect in the probability distribution of the variables used in this section; in essence, whether the effects conditioned to a given cell are more or less organised than its causes. The Kullback–Leibler information of a distribution $\boldsymbol{q}^{(1)}$, relative to a reference distribution $\boldsymbol{q}^{(2)}$, is defined as

$$K(\boldsymbol{q}^{(1)}, \boldsymbol{q}^{(2)}) = \sum_j q_j^{(1)} \log_2(q_j^{(1)}/q_j^{(2)}), \tag{3.7}$$

which is measured in bits, is always non-negative and only vanishes when $\boldsymbol{q}^{(1)} = \boldsymbol{q}^{(2)}$. Intuitively, it describes how much more organised is $\boldsymbol{q}^{(1)}$ compared with $\boldsymbol{q}^{(2)}$. Note that (3.7) is only finite if the support of $\boldsymbol{q}^{(1)}$ is within the support of $\boldsymbol{q}^{(2)}$, so that K can be understood as a measure of how much information is gained by restricting $\boldsymbol{q}^{(2)}$ to one of its subsets. Here, we will always use as reference the invariant distribution $\boldsymbol{q}_\infty$, so that K is guaranteed to exist both for the distribution $\boldsymbol{q}^c$ of the conditional causes and for the distribution $\boldsymbol{q}^e$ of the effects. This choice also implies that a distribution with $K = 0$ is statistically indistinguishable from $\boldsymbol{q}_\infty$, and represents an unconstrained set of phase points. The assumption that the system is restricted to a single cell at $t = 0$ almost guarantees that information is lost when this concentrated distribution is allowed to spread in the past or in the future, but the information contained in the two distributions cannot be compared directly, because they do not generically share a common support. Figure 6(*d*) displays the difference

$$K^{ce} = K(\boldsymbol{q}^e, \boldsymbol{q}_\infty) - K(\boldsymbol{q}^c, \boldsymbol{q}_\infty), \tag{3.8}$$

between the information of conditional effects and of conditional causes with respect to the reference. It is positive along the left (growth) edge of the distribution, and negative along the right (decay) edge, suggesting that coherence is first created and later destroyed as the system drifts clockwise. Because the system is stationary, the two effects cancel, and the mean generation of information vanishes.

The real power of the four indicators just described is in the large-scale screening of less obvious variable pairs. Figure 7 displays the value of the determinacy (3.5) and segregation (3.6) indices for all the combinations of modal inclination and intensity of the three

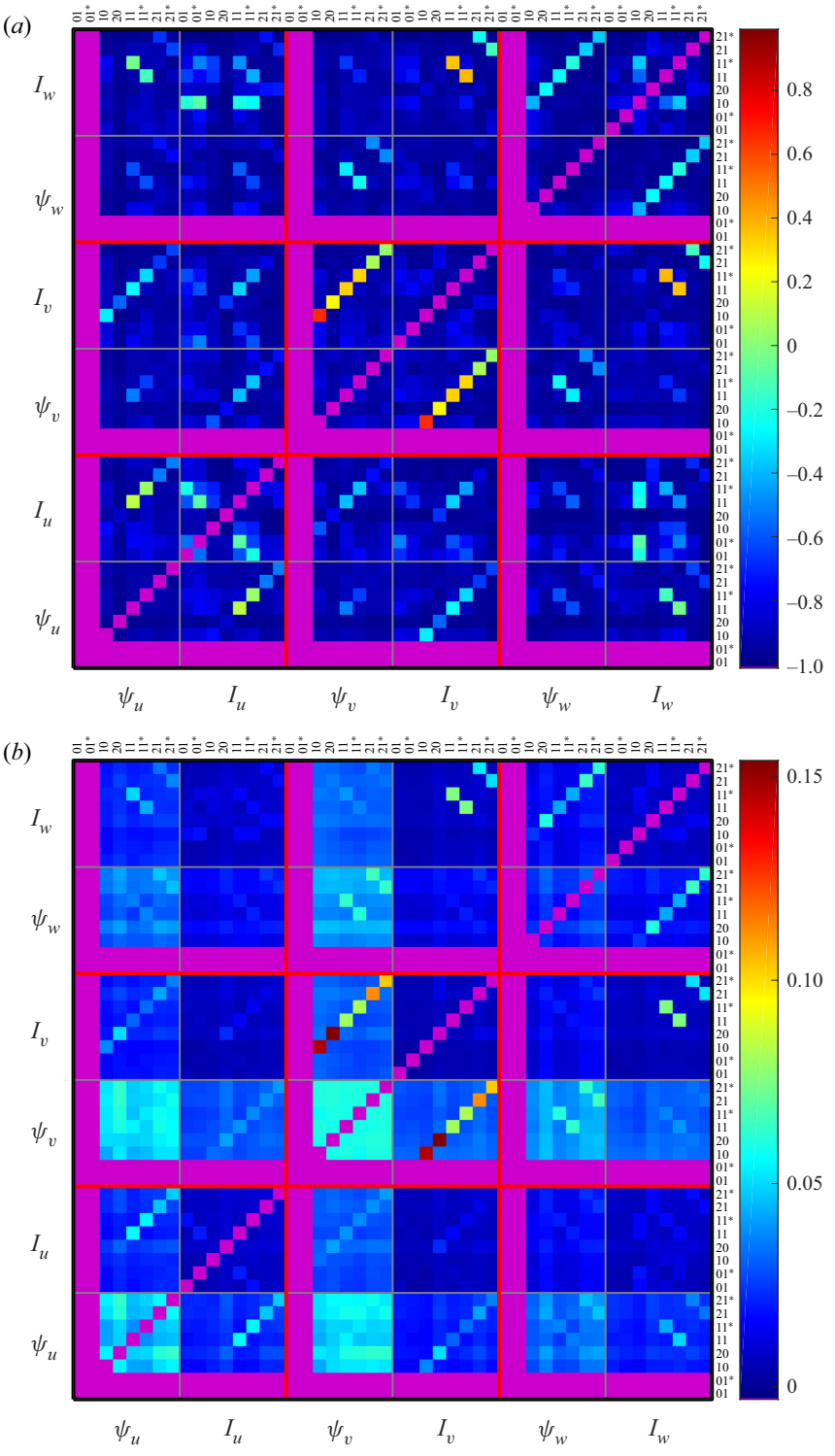


Figure 7. As in figure 3. (*a*) Determinacy index (3.5), averaged over the invariant distribution for different combinations of modal inclination and intensity. (*b*) Segregation index (3.6). The main diagonal and the inclinations of the $k_x = 0$ modes are blocked in magenta in both cases.

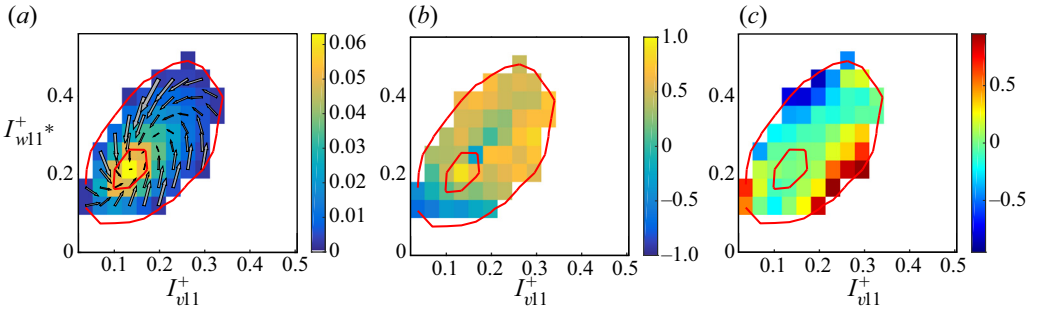


Figure 8. As in figures 5 and 6, for the roller variables, (I_{v11}, I_{w11*}). (*a*) Quiver plot of the effects, as in figure 5(*c*), $T^* = 0.076$. (*b*) Determinacy index (3.5), as in figure 6(*b*), $T^* = 0.025$. (*c*) Kullback–Leibler information gain (3.8), as in figure 6(*d*), $T^* = 0.076$.

velocity components, averaged over the corresponding distributions. Not surprisingly, the most deterministic combinations are those involving the inclination and intensity of the same Fourier mode, but it is interesting that v_{10}, which was the mode selected on physical grounds in Jiménez (2015), is also the most deterministic by a substantial margin. Slightly less deterministic is v_{20}, which is a harmonic of v_{10} for which the same theory applies. The meandering modes, v_{11}, v_{21}, u_{11}, u_{21}, are less organised but not fully incoherent. The spanwise velocity is not well described by an inclination and an intensity, but an interesting pair is (I_{v11}, I_{w11*}) or (I_{v11*}, I_{w11}), which was already identified as an inclined roller from the correlations in figure 3. Its appearance in the two indicators in figure 7 shows that this roller has its own causal dynamics, and that the same is true for the second harmonic, (I_{v21}, I_{w21*}). Interestingly, the streamwise-uniform roller (I_{v01}, I_{w01*}) does not appear in figure 7, even if it is one of strongest pairings in the correlations in figure 3, and one of the largest contributors to the fluctuation energies in figure 2. Such two-dimensional streamwise structures do not interact with the shear and, even if they grow to be strong, have little own dynamics.

Figure 8 displays the drift diagram and two quality indicators for the (I_{v11}, I_{w11*}) roller. The two variables are relatively well correlated, probably by continuity, but the distribution rotates counter-clockwise. This implies that the roller is first created as an ejection (or sweep) of the wall-normal velocity, and later spreads spanwise. Figure 8(*c*) shows that, as in figure 6(*d*), coherence is created when the structure strengthens, and destroyed when it weakens.

3.3. *Conditional averages*

The analysis in the previous sections reveals which variable pairs evolve in a deterministic way, but says little about the associated flow fields. A first step in that direction is figure 9, which displays the averages of other flow variables conditioned to the basic (ψ_{v10}, I_{v10}) pair. Most conditional averages provide relatively little information. They are either distributed almost uniformly over the invariant probability distribution, or track the evolution of the base variables, saying little more than that strong structures are strong in most variables. A few are more interesting.

Figure 9(*a*) shows that the streamwise velocity becomes more non-uniform in the streamwise direction (u_{10}) as the wall-normal velocity bursts. It follows from the choice of conditioning variables that this is also true for the wall-normal velocity, but is not true for the spanwise component (not shown). The coupling of u_{10} and v_{10} is required by continuity,

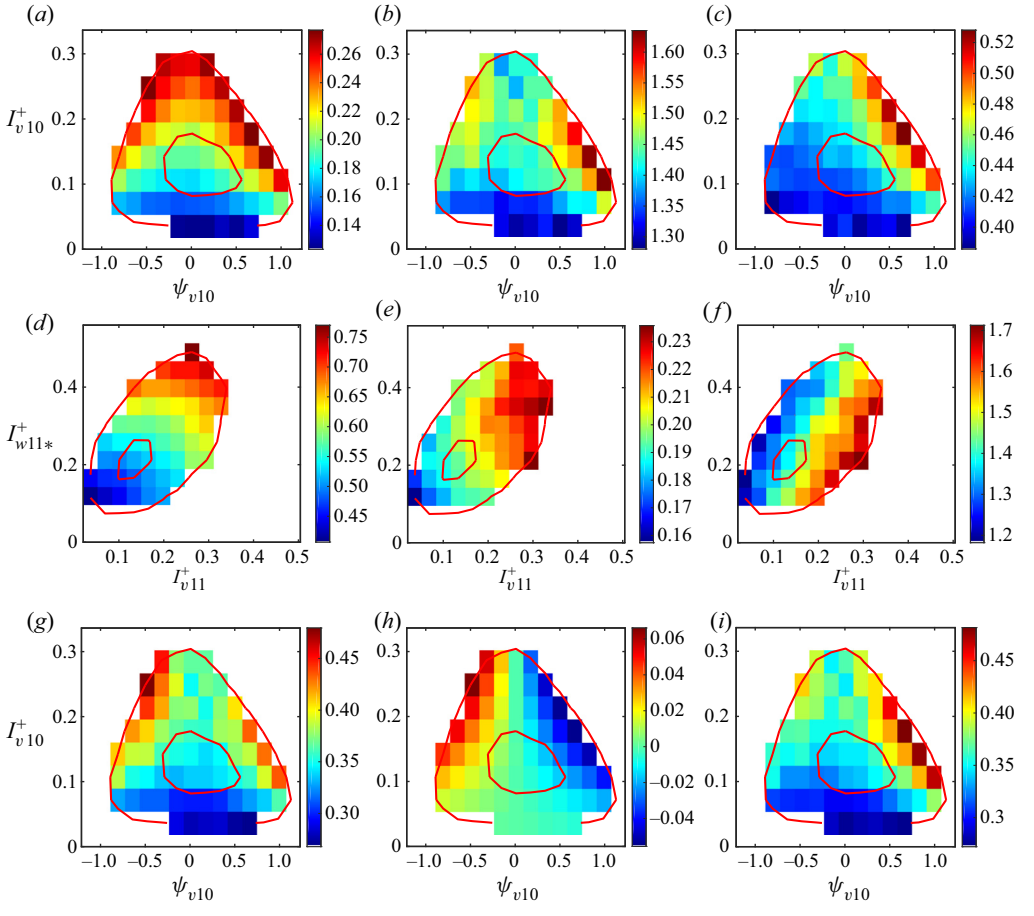


Figure 9. (*a–c*) Average intensity of different velocity modes, conditioned to the variable pair (ψ_{v10}, I_{v10}): (*a*) u_{10}, (*b*) all modes of u with $k_z \neq 0$, (*c*) $\sqrt{v_{11}^2 + w_{11}^2}$. (*d–f*) Conditioned to (I_{v11}, I_{w11*}): (*d*) u_{11}, (*e*) u_{10}, (*f*) all modes of u with $k_z \neq 0$. (*g–i*) Conditional tangential Reynolds stress, θ: (*g*) all the retained harmonics, (*h*) θ_{10}, (*i*) all harmonics with $k_z \neq 0$.

and the two variables are part of the two-dimensional burst described in Orr (1907) and Jiménez (2013). However, self-sustaining turbulence requires three-dimensionality, and this is provided by the $k_z \neq 0$ modes in figure 9(*b*). This figure displays the sum of all the modes of u that form a possibly non-uniform streak. They grow along the decay leg of the burst, and are therefore probably consequences, rather than precursors of the burst. It can be shown that most of the growth in figure 9(*b*) is due to the [01] mode, which measures the intensity of a streamwise-uniform streak, and is also responsible for most of the asymmetry of the figure along the ψ axis. The higher modes, [11] and [21], only grow weakly, and do so along the ascending leg of the burst. On the other hand, the cross-flow (v, w) roller grows along the descending branch, as shown in figure 9(*c*) for the [11] harmonic. This is also true for the [01] streamwise roller, which has similar intensity, but much less for the [21] mode, which is weaker and less coherent. In all these cases, the two signs of k_z are grouped in the figure.

When the intensity of the x-dependent modes is substituted by a non-uniformity index

$$\Omega_{a,mn} = I_{a,mn}/I_{a,0n}, \tag{3.9}$$

figure 9 changes somewhat. The growth of the roller in figure 9(*c*) is transferred to the bottom of the distribution. Its non-uniformity begins with w_{11} in the descending leg of the burst, and moves to u_{11} and v_{11} along its left-going bottom edge. This wavy quasi-streamwise roller is very weak in that region, but also very non-uniform, and it is tempting to conclude that its waviness is part of what eventually triggers a subsequent burst.

Figure 9(*d–f*) shows conditional averages in the plane of the (I_{v11}, I_{w11*}) meandering roller. Figure 9(*d*) is equivalent to figure 9(*a*), and confirms that the roller is a three-dimensional structure in which the streamwise velocity increases as v_{11} and w_{11*} do. The effect is much stronger for u_{11} than for u_{11*}, showing that the active streamwise velocity is collocated with the wall-normal component, v_{11}, rather than with w_{11*}. The three-dimensionality extends to other harmonics, and figure 9(*e*) shows that one effect is to enhance the streamwise non-uniformity of the spanwise-uniform component u_{10}. This figure is dual to figure 9(*a*), and suggests that the intensification of the roller takes place along the upper edge of the (ψ_{v10}, I_{v10}) burst, although where exactly cannot be decided from this representation. For example, the conditional I_{v10}, which would be a direct indicator of the position along the (ψ_{v10}, I_{v10}) burst, does not produce a clear signal in the (I_{v11}, I_{w11*}) plane. Figure 9(*f*) shows the conditional $k_z \neq 0$ component of u, which is the same quantity in figure 9(*b*), and gives more information. Figure 9(*b*) shows that the streak of u grows along the descending leg of the clockwise evolution of (ψ_{v10}, I_{v10}), while figure 9(*f*) shows that it grows along the ascending leg of the counter-clockwise evolution of (I_{v11}, I_{w11*}). Both legs therefore presumably correspond to the same stage of the flow, in agreement with the location in figure 9(*c*) of the high roller intensity.

Also interesting is figure 9(*g–i*), which shows the conditional tangential Reynolds stress, $\theta_{mn} = -\mathrm{Re}(\tilde{u}_{mn}\tilde{v}^{\dagger}_{mn})$, integrated as in (3.2). Although the quantity in the equations of motion is $\partial_y\theta$, rather than θ itself, a positive stress tends to make the mean velocity profile more turbulent, steeper near the wall, and negative ones tend to lower the wall shear. Figure 9(*g*) displays the conditional Reynolds stress due to all the retained harmonics, which we saw in § 3 to account for approximately two thirds of the total flow stress. It is positive everywhere, and stronger in the upper edge of the distribution, where other flow features are also strong. It is also asymmetric in the (ψ_{v10}, I_{v10}) plane, stronger during the growth of the burst than along its decay. The reason is shown in figure 9(*h*), which displays the stress due to the Orr (1907) harmonic, θ_{10}. It is almost antisymmetric in ψ_{v10}, positive during the growth of the burst and negative during its decay. This is what makes the burst transient, since its decay undoes the effect of the growth. The [10] mode is the only one that generates counter-gradient stresses. Figure 9(*i*) displays the tangential stresses from all the other modes. They are positive everywhere, and the net effect of the burst, although transient in itself, is to generate a three-dimensional structure along its decay leg.

Although not shown in the figure, it is interesting that, when θ_{10} is conditioned to the (I_{v11}, I_{w11*}) plane, it is also negative in the uppermost tip of the roller distribution, confirming our previous conclusion that strong rollers correspond to the burst decay.

In a loose sense, both the down-going right-hand edge of the probability distribution in figure 9(*a–c*) and the up-going right-hand edge of figure 9(*d–f*), portray the evolution of a non-uniform u-streak into a (v, w) roller. In the top row of figure 9 we see the decay of the streak, and in the middle one we see the growth of the roller. The comparison of the conditioned variables in both sets of figures suggests that the coherent right-hand edge of the evolution of the roller corresponds to the decay of the burst, so that the approach to the lower-right vertex of the triangle in figure 9(*a–c*) corresponds to the top of the distribution in figure 9(*d–f*). In this interpretation, the low-intensity evolution of the burst along the

bottom edge of the distribution in figure 9(*a–c*) should correspond to the decay of the roller along the left edge of the distribution in figure 9(*d–f*). This phase of the burst will be examined in more detail in the next section.

4. Conditional trajectories

It may be useful at this point to recall that the maps in figures such as 9 are not the PFO, but its leading eigenvector rearranged as a two-dimensional matrix for human convenience. The operator itself is the larger matrix in figure 4(*b*), or a stack of such matrices at different time intervals. The maps in figure 9 are probability distributions of the state of the flow in phase space, and the delta-function distribution used to generate figure 5(*a*,*b*) represents a measurement that particularises that probability to a combination of variables describing a given flow configuration. For the rest of the paper we will study these collapsed probabilities, and the statistical properties of the trajectories connecting two or more such states. The PFO then becomes mostly a guide to which cells can be expected to provide the most interesting information, and to how the trajectories connecting them should be chosen.

For example, consider the statistical characterisation of trajectories that satisfy conditions at two different times, such as in the recurrence test in which approximate periodic behaviour is identified by monitoring when a trajectory approaches itself after a given delay (Kawahara, Uhlmann & van Veen 2012). The statistical equivalent is whether the conditional probability distribution of the effects of a given cell includes the conditioning cell after some delay. In practice, this reduces to identifying among the diagonal of the matrix $\boldsymbol{P}^e(T)$ those cells that result in maximum probability of recurrence for each time delay. An example is figure 10(*a*), which plots the maximum of the diagonal of the PFO for the (ψ_{v10}, I_{v10}) plane. It is unity at $T = 0$, when trajectories are still at their initial position, and quickly decays to approximately 0.05, which is the probability that a random trajectory intersects some cell in the core of the invariant distribution $\boldsymbol{q}_\infty$. However, the curve peaks again at $T^* = 0.66$, and, interestingly, at twice that delay, $T^* = 1.35$. Moreover, since the PFO contains information about all the cells in the distribution, it allows us to recover which conditioning cell is responsible for the probability maximum, and therefore which cell has the highest probability of recurring. This is marked as 'B' (for burst) in figure 10(*b*), and turns out to be an extreme high-amplitude event beyond the 95 % threshold used up to now as the practical edge of our distribution. In the trajectories discussed in this section, closed symbols mark the position at $t = 0$, and the open triangles along some trajectories are equispaced by $\Delta t^* = 0.025$.

Two other cells are labelled in figure 10(*b*), marking the right- ('R') and left-hand ('L') corners of the triangular distribution. We will maintain this nomenclature for the rest of the section, with some adjustments in the location of the cells. Trajectories spanning the up-going (L $\rightarrow$ B), down-going (B $\rightarrow$ R) and bottom (R $\rightarrow$ L) legs of the periphery of the triangle will be denoted as growth, decay and recovery trajectories, respectively. The growth and decay legs form the burst (Jiménez 2013). The upper half of these two legs is deterministic, and can be predicted linearly (Jiménez 2015), but linearised bursts do not recur. There is no obvious theory for the bottom recovery leg, which is required if bursting is to explain self-sustaining turbulence. Most of this section is dedicated to analysing the recovery process.

Out of our 5×10^4 snapshots, only eight trajectories cross the extreme B in figure 10. Most of them do not recur, and the line of open black triangles in figure 10(*b*) traces the average conditional trajectory during the recurrence period. As is true for most trajectories, it approaches the high-probability core of the distribution. However, individual trajectories

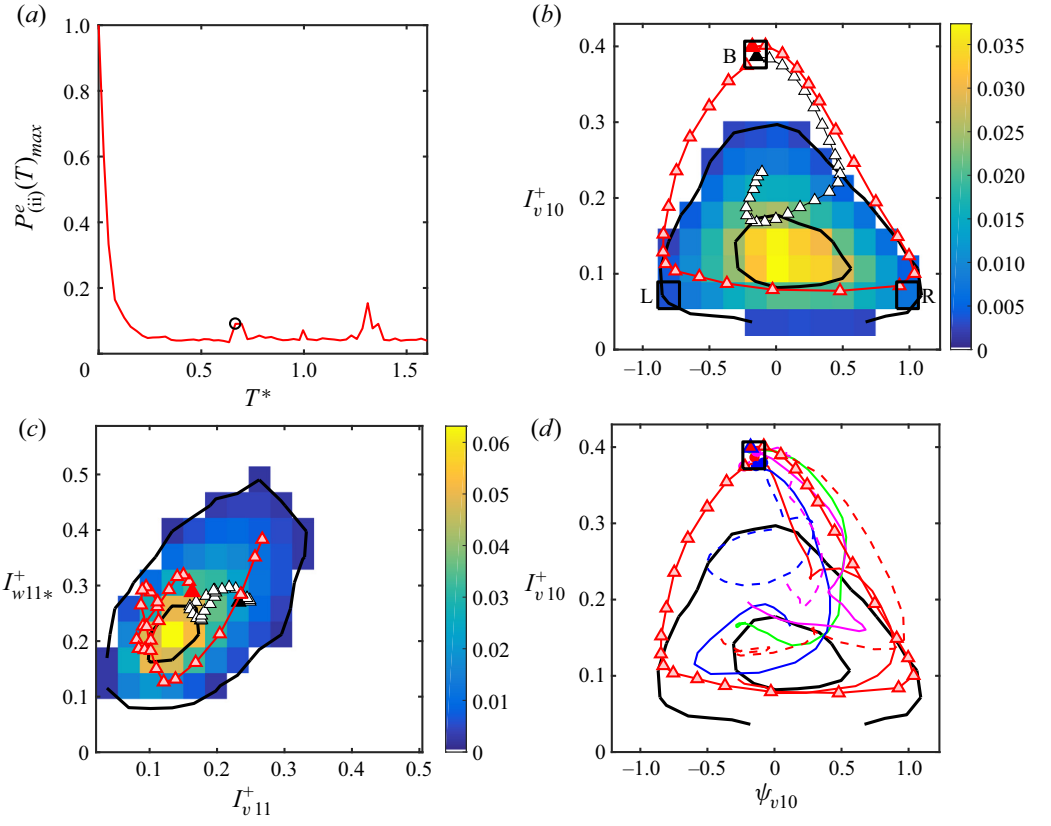


Figure 10. (*a*) Recurrence test for trajectories in the (ψ_{v10}, I_{v10}) plane. See text for explanation. (*b*) Black symbols are the mean trajectory passing through the most probable recurrent cell, for the period marked by the dark circle in (*a*). The red symbols are the only recurrent orbit crossing that cell, which is thus responsible for the local maximum in (*a*). (*c*) As in (*b*), in the (I_{v11}, I_{w11*}) plane of the meandering roller. (*d*) As in (*b*) but all the non-recurrent trajectories are plotted as simple lines, each one starting with a solid circle.

can be tested, and the line of red triangles in figure 10(*b*) shows the trajectory responsible for the peak in figure 10(*a*). Centring ourselves on this orbit, its growth, decay and recovery legs last approximately $T^* = 0.25$, 0.25 and 0.16, respectively, for a total recurrence time $T^* \approx 0.66$, as in figure 10(*a*). The total length of its two bursting legs, $T^* \approx 0.5$, also agrees with the width of the bursting correlations for this flow in Jiménez (2015).

Although the recurrent trajectory very approximately closes on itself in the plane of figure 10(*b*), its recurrence is weaker when more variables are included. For example, figure 10(*c*) plots the same trajectories in the plane of the (I_{v11}, I_{w11*}) roller. As before, the recurring trajectory is displayed as red triangles. It loiters for a while near the high-probability core of the (I_{v11}, I_{w11*}) distribution, and joins the counter-clockwise coherent circulation during the growth period of the (ψ_{v10}, I_{v10}) burst. The black triangles of the mean trajectory are again very different from the recurrent one, and never join the coherent circulation.

However, the divergence between the recurrent trajectory and the other trajectories that cross B is mostly a long-time phenomenon. This can be seen in figure 10(*d*), which is equivalent to figure 10(*b*) but plots all the individual trajectories going through B. The recurrent orbit is still plotted with open red triangles, while other trajectories are plotted

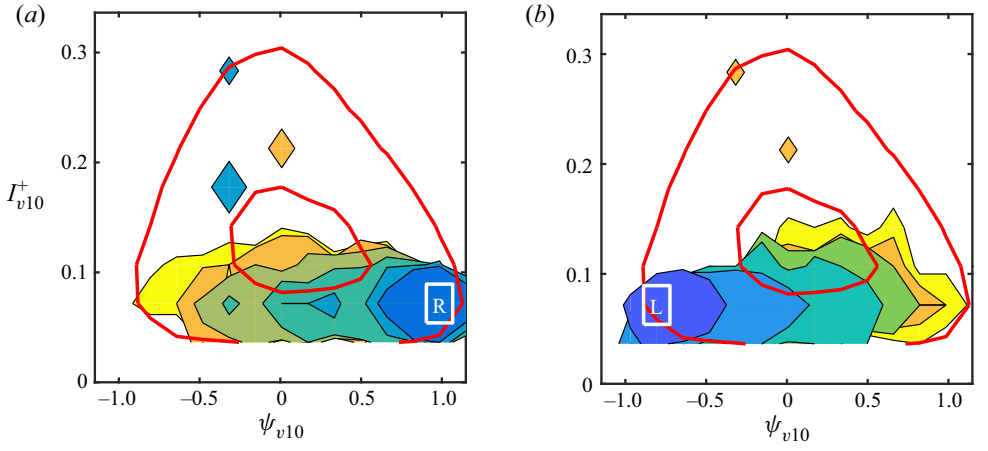


Figure 11. Probability distributions of the causes and effects for several delay intervals, as in figure 5(*a*,*b*). (*a*) Effects conditioned to the cell marked as R in the lower right-hand corner of the invariant density distribution. From blue to yellow, $T^* = 0.025(0.025)0.15$. (*b*) As in (*a*), for the causes leading to the cell marked as L in the lower left-hand corner.

as lines of different colours without symbols, in no particular order. All the trajectories in the figure start from the same cell and behave similarly for a while. It is only after they have decayed to approximately half their initial I_{v10} that they deviate towards the high-probability core. A similar plot in the (I_{v11}, I_{w11*}) plane, not shown, shows the same trends, although slightly more complicated because trajectories do not start in the same cell any more. The recurrent orbit is special in that it starts with a relatively weak and decaying roller, which only strengthens towards the end of the orbit. This is probably the reason why its (ψ_{v10}, I_{v10}) projection is able to proceed undisturbed for a relatively long time.

That the approximately recurrent orbit includes an infrequent extreme event suggests that it may not be very relevant to the flow statistics, or even to its evolution. Its interest resides in that it includes a recovery leg that traverses the lower edge of the parameter space in the 'wrong' direction, showing that coherent recovery connections exist, and suggesting how to study them. The question is whether other trajectories exist that cross the lower edge of the distribution in the same direction, even if they do not follow the recurrent orbit around the full (ψ_{v10}, I_{v10}) plane.

Figures 11 and 12 address this question. Figure 11(*a*,*b*) is equivalent to the cause and effect distributions in figure 5(*a*), but applied to the R and L corner cells for several values of the delay interval. Even if we saw in figure 6(*a*) that the lower edge of the distribution is a region of large scatter for the system drift, figure 11(*a*) shows that the flow propagates steadily from R towards L as the delay increases, and that it only spills slowly into the high-probability core of the distribution. The same is true for the causes of cell L in figure 11(*b*). When similar plots are drawn for a cell in the central part of the lower edge of the distribution, the drift is slower and the scatter higher (not shown), suggesting that the propagation in figure 11(*a*,*b*) can be interpreted as a direct connection between R and L that bypasses some of the intermediate states of the probability distribution.

This is further analysed in figure 12(*a–c*). Each of these panels represents a leg of the burst that traverse the periphery of the distribution in the (ψ_{v10}, I_{v10}) plane. For example, figure 12(*a*) represents the growing leg, from the cell marked as L to the one marked as B (note that the latter has been chosen substantially lower than in figure 10, to bring it within

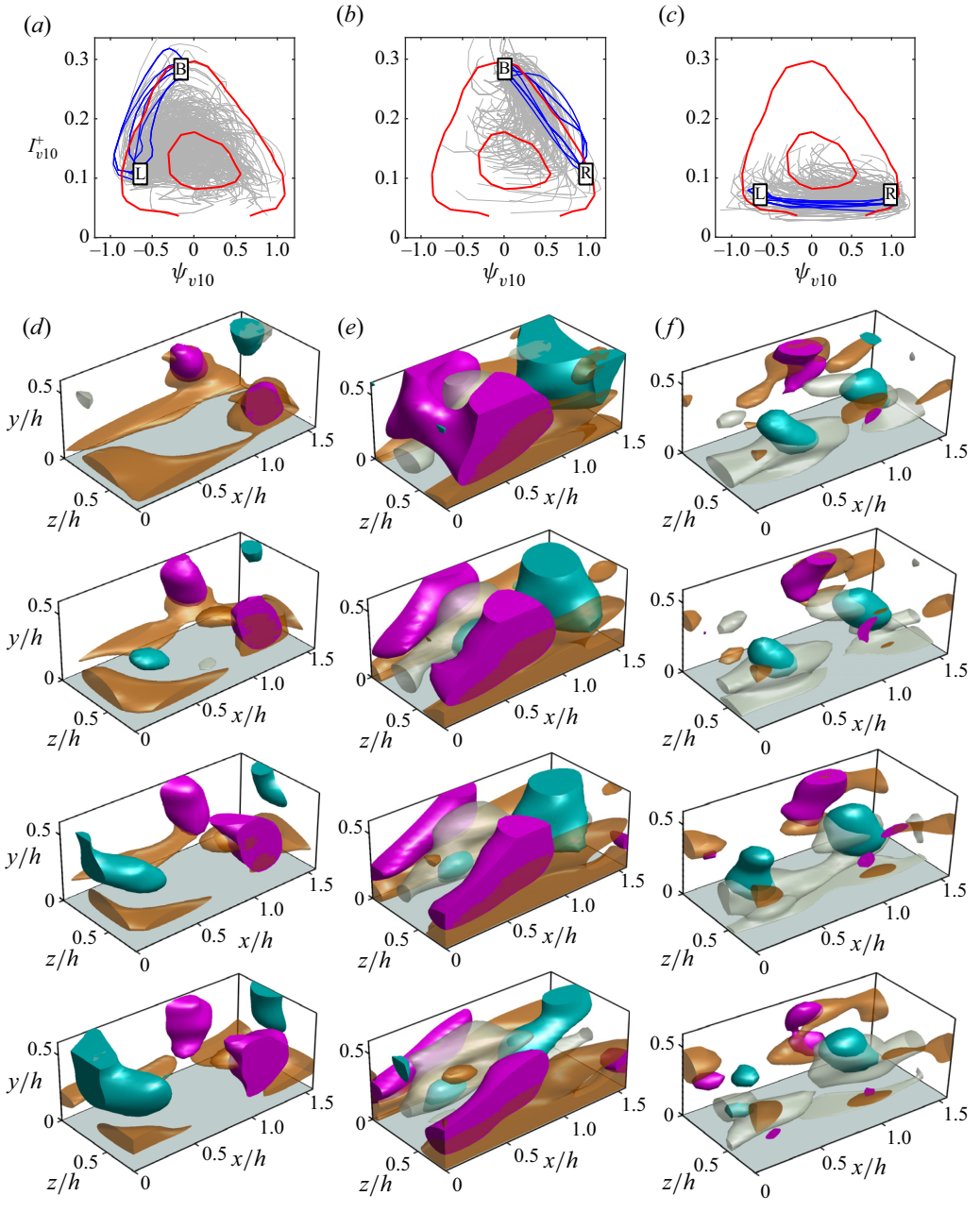


Figure 12. (*a–c*) Grey lines are the phase trajectories that cross a causal cell at $t = 0$. Blue ones are those that also cross the effect cell within a given range of time intervals. (*a*) Growth leg of the burst, from L → B in $T^* = 0.2$–0.3. (*b*) Decay from B → R in $T^* = 0.2$–0.3. (*c*) Recovery leg from R → L in $T^* = 0.15$–0.175. (*d–f*) Averaged evolution of the flow along the blue trajectories in (*a–c*), respectively, reconstructed from the retained harmonics. Flow is from bottom left to top right, and time increases from top to bottom. Translucent orange isosurface, $u^+ = 0.7$; translucent grey, $u^+ = -0.7$; cyan, $v^+ = 0.4$; purple, $v^+ = -0.4$. Panels (*d–e*) show $t^* = 0.075(0.05)0.225$. Panel (*f*) shows $t^* = 0(0.05)0.15$, and $v^+ = \pm 0.3$.

the 95 % probability contour). It takes the system $T^* \approx 0.25$ to move from one to the other, and the grey lines are the $O(500)$ trajectories that pass through L at $t = 0$. Of these, only the six trajectories drawn in blue also pass through B in the interval $T^* = 0.2$–0.3. Figure 12(*b*,*c*) is similarly drawn for the decay and recovery legs of the burst, respectively. The trajectories in these three legs are not continuations of each other. For example, there is a single trajectory linking cells L, B and R, in that order, three trajectories linking R, L and B and no trajectory linking B, R and L. The only trajectory approximating a full cycle in this plane is the one in figure 10.

This lack of continuity is a direct consequence of the non-Markovian behaviour discussed in § 2 and in figure 4(*c*). Cell R is not the initial condition for cell L, or for the recovery leg, but only the projection onto the (ψ_{v10}, I_{v10}) plane of many initial conditions, most of which will not fully follow that leg or continue into the growth one. In moving from probabilities based on one cell to two-point trajectories, we have in essence abandoned the two-dimensional PFO for a higher-dimensional version based on the state of the system at two points in time. In a different interpretation, we have moved from a Markovian probabilistic predictor to a more complex dynamics.

Figure 12(*d*–*f*) displays the mean evolution of the flow field, reconstructed from the retained harmonics along the blue trajectories in figure 12(*a*–*c*), respectively, with time increasing from top to bottom. Because the problem is homogeneous in x and z, the flow in the different trajectories has to be aligned to a common origin before averaging. This is done at $t = 0$ by translating each case so that the deepest low-speed perturbation of the streamwise velocity is located at the centre of the display box. To facilitate visual tracking, the frame of reference is then advected with a velocity $U^+_{ad} = 8$, which keeps the structures approximately stationary near the wall.

The forward tilting of the v-structures by the shear is clear in the growth and decay legs of the burst (figure 12*d*,*e*). The recovery leg in figure 12(*f*) is harder to interpret, in part because everything is much weaker than in the other two cases, but the clearest difference is that, while the growth and decay are dominated by high-streamwise-velocity structures near the wall, the high-speed regions of the recovery leg are predominantly farther into the flow. The near-wall layer only contains a weak discontinuous low-speed streak. Since the structures in figure 12(*d*–*f*) are defined with respect to a long-term-averaged velocity profile, this difference in organisation implies that the box-averaged instantaneous velocity profile differs among legs.

This is confirmed in figure 13(*a*,*b*), which shows that the burst is characterised by high velocities and steep profiles near the wall. Figure 13(*c*) shows that the recovery takes place in regions where the velocity is lower near the wall, and the shear is displaced away from it. The flow enters the recovery leg with a steep velocity gradient at $y/h \approx 0.2$, which decays as the recovery proceeds (red to blue). Figure 13(*d*–*f*) shows the evolution of the fluctuation energy at the same times. During the burst, in figure 13(*d*,*e*), the fluctuations stay relatively close to the wall ($y/h \approx 0.12$), and the energy peak extends all the way to it. During the recovery leg, in figure 13(*f*), a new peak grows at the location of the detached shear layer ($y/h \approx 0.22$), and forms a new secondary peak at the wall.

Note that the streamwise velocity perturbations in most panels of figure 12(*d*–*f*) are short and discontinuous, in agreement with the recent evidence in Jiménez (2022) that long streaks are not directly linked to the regeneration cycle of wall turbulence.

The conclusion that the recovery of the burst depends on a low-shear region near the wall is inconsistent with the intuitive notion that, since the shear is the ultimate source of turbulent energy, a higher shear should be a prerequisite for higher turbulent activity. Indeed, it has been known for some time that turbulent intensity and shear are correlated

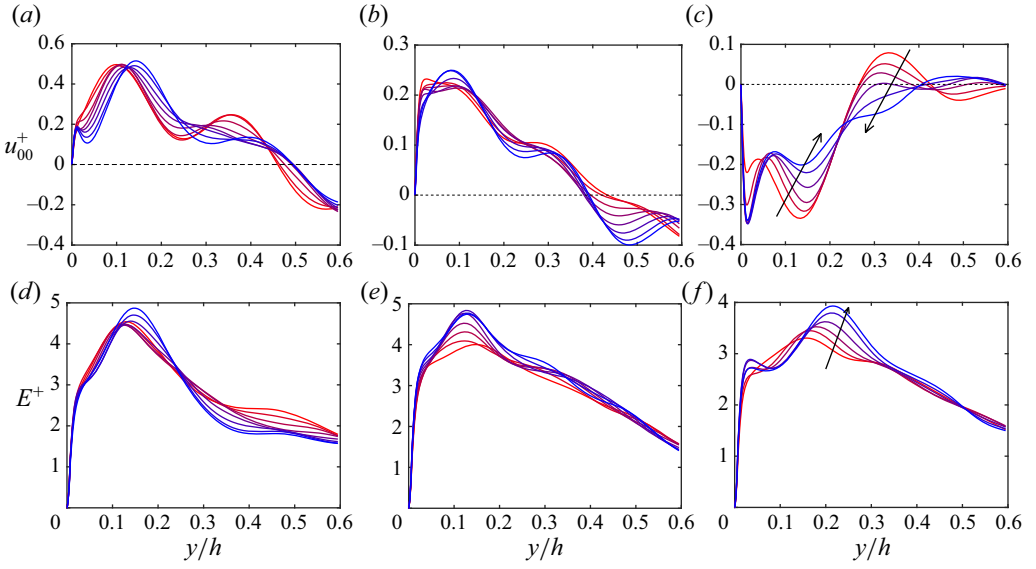


Figure 13. Box-averaged velocity profiles, for the trajectories marked in blue in figure 12. (*a–c*) Mean velocity, $\tilde{u}_{00}$. (*d–f*) Kinetic energy of the retained harmonics. (*a*,*d*) Growth leg. (*b*,*e*) Decay leg. (*c*,*f*) Recovery. Time increases from red to blue, separated by $t^* = 0.025$ among curves.

(Marusic, Mathis & Hutchins 2010; Jiménez 2012; Mathis *et al.* 2013), but the implied model is usually that the turbulence intensity evolves to be in equilibrium with the shear. Our discussion suggests that the causality is the other way around (figure 9), and that the shear is created by the Reynolds stresses of the fluctuations. In fact, since the mass flux in our channels is constant, a low shear near the wall implies a higher one further up. What our previous discussion suggests is that the decay of a burst induces a mild shear at the wall, which in turn steepens the velocity profile away from it. The steep off-wall profile is what triggers the new burst, which is the blue/magenta structure growing away from the wall in the downstream part of figure 12(*f*). Flores & Jiménez (2010) showed that stress waves travel to and from the wall in small-box simulations such as the present one, with a wall-normal velocity of the order of u_τ. It is difficult to decide from such kinematic observations which of the two directions is the primary causal one, but the discussion above suggests that at least the descending wave is causal, in agreement with previous reports that the structure of the logarithmic layer in wall-bounded flows is relatively independent from the details of the wall, which is therefore not the primary, or at least not the only, seat of causality (Townsend 1976; Mizuno & Jiménez 2013; Kwon & Jiménez 2021).

5. An interventional experiment

The analysis in the previous sections gives strong hints about which variables evolve coherently in wall-bounded flows, and about how this evolution can be interpreted in terms of causality within the attractor in phase space. However, the second question posed in the introduction, whether this information has any practical value, generally takes us outside the attractor, and can only be answered by more classical interventional experiments or by theoretical models. Strictly speaking, this is beyond the scope of the present paper, whose goal is to develop a methodology and to give examples of how it can be used to motivate

more classical work, but we present in this section an example of how that work might proceed, building on our discussion of figure 12.

The main result of that discussion is that a particular, weakly sheared, configuration of the mean velocity profile is a requirement for burst recovery. This is not a new idea: what is probably the oldest theory of how wall turbulence is controlled is based on the two-way interaction between shear and turbulence intensity (Malkus 1956). It is also known that forcing a locally steeper or shallower velocity profile leads to equilibrium intensities that correlate with the local shear (Tuerke & Jiménez 2013), and many low-order models of the turbulence cycle include a variable that stands for the mean velocity profile (Waleffe 1997).

However, the original linear version of the Malkus (1956) model was disproved by Reynolds & Tiederman (1967), who found no trace of the marginal instability of the velocity profile that that model assumes for turbulent channels, and the analysis in Waleffe (1997), although highly suggestive because the effect of a strong wall shear is to inhibit the instability of the streaks, only applies to permanent travelling waves in marginally turbulent low Reynolds number flows. Similarly, Tuerke & Jiménez (2013) refer to long-term flow averages, and it is unclear whether these are relevant to the short-term behaviour of an intermittent bursting cycle. The hypothesis to be tested is whether a feedback cycle involving the modification of the mean profile by the turbulent fluctuations, and the control of the latter by the former, can explain at least part of the mechanism that sets the frequency and amplitude of the bursts.

We do this by smoothing the evolution equation that links the fluctuations to the mean profile, defined as the deviation from the long-term mean velocity profile of the (x, z)-averaged streamwise velocity fluctuation, $\tilde{u}_{00}(y)$. In a regular channel it satisfies

$$\partial_t \tilde{u}_{00} = R - P + \nu \partial_{yy} \tilde{u}_{00}, \tag{5.1}$$

where $R = -\partial_y(\widetilde{uv})_{00}$ is the instantaneous mean Reynolds-stress gradient, and the pressure gradient P is determined by the ancillary flux-conservation condition

$$\int_0^{2h} \tilde{u}_{00}\,\mathrm{d}y = 0. \tag{5.2}$$

In our experiment, we substitute (5.1) by

$$\partial_t \tilde{u}_{00} = Q - P + \nu \partial_{yy} \tilde{u}_{00}, \tag{5.3}$$

where

$$\partial_t Q = (R - Q)/\tau, \tag{5.4}$$

which, after a transient in which the effect of the initial conditions decays exponentially, is solved by

$$Q(y, t) = \tau^{-1} \int_0^t \exp[(\xi - t)/\tau]\, R(y, \xi)\,\mathrm{d}\xi. \tag{5.5}$$

The modified right-hand side, Q, is therefore a smoothed version of R, with a smoothing time τ. The integral of R or Q across the channel can be considered as a body force that must be compensated by the pressure gradient, but it is easy to see from its expression that $\int R\,\mathrm{d}y = 0$ for impermeable walls, and that the same holds for Q after the initial transient.

Figure 14 shows some results from the experiment, using a smoothing time, $\tau^* = 2.27$, of the order of the expected bursting period (Flores & Jiménez 2010). Figure 14(*a*) shows the history of the friction Reynolds number for the natural and modified channels.

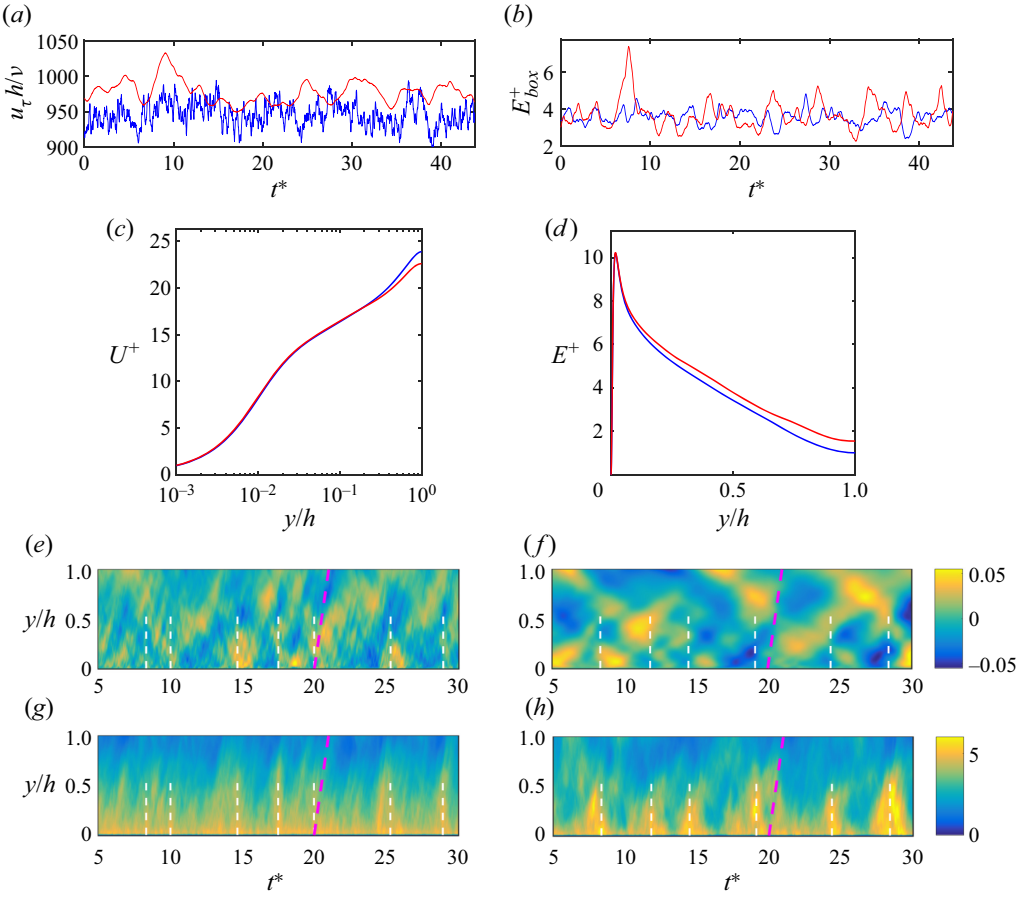


Figure 14. (*a*) Instantaneous friction velocity of the temporally smoothed and natural experiments, (5.1)–(5.4). (*b*) As in (*a*), for the box-averaged fluctuating kinetic energy, $E = u'^2 + v'^2 + w'^2$, measured with respect to its long-time average. (*c*) Mean velocity profiles. (*d*) Mean kinetic energy. In all cases: blue, natural channel; red, $\tau^* = 2.27$. (*e–h*) Temporal evolution of the mean profiles as functions of time. (*e,f*) Mean streamwise velocity fluctuation, with respect to its long-time average. (*g,h*) Kinetic energy. (*e,g*) Natural channel. (*f,h*) Case $\tau^* = 2.27$.

It only increases slightly for the smoothed case, from $\langle h \rangle^+ = 950$ to 975, although its temporal oscillations become much slower. More interesting is figure 14(*b*), which shows the oscillations of the box-averaged kinetic energy. They are somewhat deeper for the smoothed case, and substantially less frequent and more regular. An approximate count, using a method explained below, gives $\Delta t^* \approx 3.9$ for the mean distance between bursts in the natural case, and $\Delta t^* \approx 5.2$ in the smoothed one. Both are longer than in Flores & Jiménez (2010), who find $\Delta t^* \approx 2$ from the temporal spectrum of the integrated Reynolds stress, probably because their method is sensitive to weaker oscillations than the present one.

The effect on the integrated velocity profiles is slight. Figure 14(*c*) shows the mean velocity, and reveals that the main effect is to decrease the wake component above $y/h \approx 0.3$, but this is also the height at which this channel begins to be constrained by the numerical box, and where the profile in any case deviates from the natural one.

Figure 14(*d*) shows that the fluctuations of the kinetic energy are also slightly higher, as expected from figure 14(*b*).

Figure 14(*e–h*) shows the temporal evolution of the profiles, and give more information. Figure 14(*e*,*f*) shows the deviation, $\tilde{u}_{00}$, of the mean velocity profile with respect to its long-time average. Each vertical section of these images is an instantaneous box-averaged profile. Blue regions are slower that usual, and yellow ones are faster. They should be compared with the instantaneous profiles in figure 13(*a–c*), although the flows in figure 14 are not filtered or conditioned in any way. Figure 14(*e*,*g*) is the natural flow, and figure 14(*f*,*h*) is temporally smoothed. The white dashed vertical lines in figure 14(*e–h*) mark the time of the bursts of the kinetic energy, whose evolution is represented in figure 14(*g*,*h*). To detect them, the kinetic energy is integrated in $y/h \in (0,\, 0.4)$, and bursts are defined as intervals where the integrated energy rises above the level isolating the top 15 % of the time.

The most interesting differences are those between the mean velocities in figures 14(*e*) and 14(*f*). The evolution in the natural case in figure 14(*e*) is clearly more complex than the temporally smoothed case in figure 14(*f*), and there is no clear correlation between the mean velocity and the position of the bursts marked by the dashed white lines. The opposite is true for the smoothed case in figure 14(*f*), in which the mean profile rises and falls in a series of diagonal waves, and the lines marking the bursts correspond, even to the naked eye, to low-velocity intervals marked by bluish areas near the wall. The inclined red line in the four panels of the figure mark the friction velocity, $\mathrm{d}y/\mathrm{d}t = u_\tau$, which is known to be the vertical advection velocity of strong Reynolds-stress structures (Flores & Jiménez 2010; Lozano-Durán & Jiménez 2014). This agrees with the inclination of the rising and falling patterns in the energy evolution maps in figure 14(*g*,*h*). It also approximately describes the vertical advection velocity of the fine structure of the velocity profiles in figure 14(*e*), most probably because the mean profile is controlled by the Reynolds stress through (5.1). On the other hand, this influence is broken in figure 14(*f*), where the Reynolds stresses only acts indirectly on the mean profile because of the smoothing effect of (5.4), and the vertical advection is much slower. This strongly suggests that any correlation between figures 14(*f*) and 14(*h*) reflects a causal effect of the mean velocity on the burst, rather than the other way around.

This is directly tested in figure 15, which shows the conditionally averaged temporal evolution of different quantities obtained by centring them on the time of all the detected energy bursts. Figure 15(*a*) shows the conditional evolution of the mean profile, and clearly shows the low-velocity period preceding the burst, which is later substituted by a steeper wall profile created by the Reynolds stress of the burst, shown in figure 15(*b*). The conditional Reynolds stress is displayed in figure 15(*c*), which shows that it is a local effect due to the burst itself. An attempt to repeat this process for the natural flow in figure 15(*e*,*g*) fails, no doubt in part because the more complex structure of the flow field makes the identification of the bursts harder. In fact, even the conditioning of the burst on the burst position, as in figure 15(*b*), fails in that case. The experiment was repeated for smoothing times of the order of two or four bursting periods, with comparable results.

6. Conclusions

This paper can be divided in two parts. In the first one, up to § 3, we adapt the PFO of dynamical-system theory to the probabilistic description of the evolution over a phase-space partition of a turbulent channel flow. We show that the main difficulty for doing so is collecting enough data to populate the operator matrix, and we bypass it by restricting ourselves to two-dimensional projections of the phase space. This leads

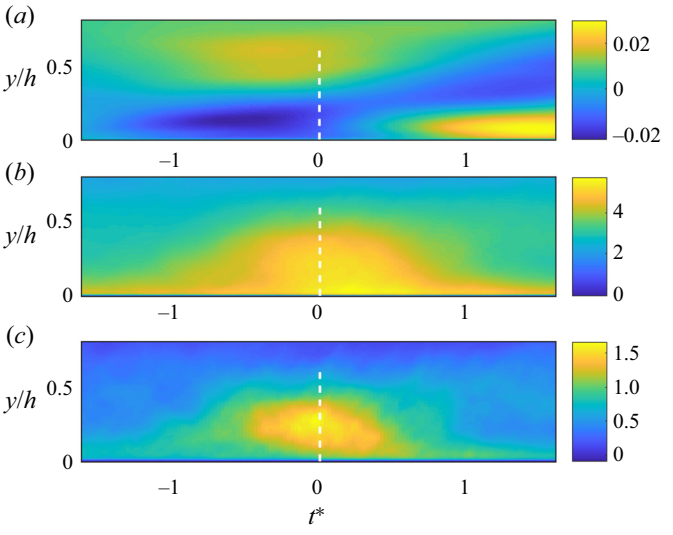


Figure 15. Time evolution of the (x, z)-averaged profiles conditioned to the bursts of the kinetic energy in the smoothed experiment. (*a*) Mean streamwise velocity fluctuation, as in figure 14(*f*). (*b*) Kinetic energy. This is the conditioning field. (*c*) Tangential Reynolds stress, $-(\widetilde{u}\widetilde{v})^{+}_{00}$.

to the question of how to choose the best pair of variables, and forces us to develop simple indicators of the quality of a particular representation. Several such indicators are developed, and shown to be interpretable in terms of causality and coherence within the attractor. It is argued that this last restriction allows us to draw conclusions about causality and information flux from flow histories, without the need for interventional experiments.

In particular, we show that we can use these indicators to distinguish between correlation and coherence, and to separate, for example, relatively weak structures that have their own dynamics, such as wavy rollers and streaks, from stronger and more correlated ones that have no dynamics of their own, such as streamwise-uniform streaks and rollers.

We show in § 3.2 how the indicators allow us to differentiate less promising variable pairs from those more likely to be useful in developing coherent physical models. Out of 630 possibilities, two promising pairs are found for the case analysed here. The first one is the intensity and inclination of the wall-normal velocity, which was already used by Jiménez (2015) to represent an approximately linear Orr (1907) burst, and the second is a more novel inclined wavy vortex.

The rest of the paper applies the techniques derived in the first sections to analyse the Orr (1907) burst, with emphasis on the poorly understood recovery process by which bursts are re-initiated after they decay. As was the case with previous attempts to use massive searches to choose among different analysis possibilities (e.g. Jiménez 2018*b*, 2020), the present one mostly suggests mechanisms that have to be confirmed by more classical means, mainly because of the original limitation to on-attractor dynamics. In this case, the PFO guides us in the choice of phase-space trajectories that connect interesting flow configurations within a known range of time intervals, including trajectories describing the recovery process, in what we note is equivalent to extending the PFO to higher dimensions. At least in our relatively small computational box, conditional averaging over these connections shows, somewhat counter-intuitively, that the key ingredient for regeneration is the development of a low-shear region near the wall. New bursts are seeded from a detached shear layer overlying it. Their Reynolds stress returns the shear to the wall,

and no new burst is possible until the decay of the old one again detaches the shear. It is not known whether this process generalises to larger boxes containing more than one burst.

To extend our conclusions outside the attractor, we finally perform a simple computational experiment in which the control of the mean shear by the burst is relaxed. The behaviour of the mean profile is thus modified, but the association of low wall shear with the initiation of the bursts is shown to be maintained.

Funding. This work was supported by the European Research Council under the Caust grant ERC-AdG-101018287.

Declaration of interests. The author reports no conflict of interest.

Author ORCIDs.
Javier Jiménez https://orcid.org/0000-0003-0755-843X.

Appendix A. Modes with non-zero spanwise wavenumber

The modes with non-zero spanwise wavenumber require special treatment. For $k_z = 0$, the variable a is expressed as

$$a(x) = \sum_{\alpha} \tilde{a}_\alpha \mathrm{e}^{\mathrm{i}\alpha x} + \text{c.c.}, \tag{A1}$$

where 'c.c.' stands for the complex conjugate. A translation $x \to x + \Delta x$ appears as a phase in $\tilde{a}_\alpha \to \tilde{a}_\alpha \exp(\mathrm{i}\alpha\Delta x)$, and the variables used in the main text are chosen to be invariant to such translations. For modes depending on z, the expansion becomes,

$$a(x,z) = \sum_{\alpha,\beta} \tilde{a}_{\alpha,+\beta} \exp(\mathrm{i}(\alpha x + \beta z)) + \tilde{a}_{\alpha,-\beta} \exp(\mathrm{i}(\alpha x - \beta z)) + \text{c.c.}. \tag{A2}$$

Each (α, β) mode has associated two independent coefficients, which transform differently under spanwise translations. The physically relevant expansion is

$$a(x,z) = \sqrt{2} \sum_{\alpha,\beta} \left[\tilde{a}^{+}_{\alpha,\beta} \cos(\beta z) + \mathrm{i}\tilde{a}^{-}_{\alpha,\beta} \sin(\beta z) \right] \mathrm{e}^{\mathrm{i}\alpha x} + \text{c.c.}, \tag{A3}$$

where $\tilde{a}^{\pm}_{\alpha,\beta} = \tilde{a}_{\alpha,+\beta} \pm \tilde{a}_{\alpha,-\beta}$. We can choose to observe the flow at any convenient spanwise location. Keeping the cosine term in (A3) corresponds to observing at $z = 0$, while keeping the sine corresponds to $\beta z = \pi/2$, but neither $\tilde{a}_{\alpha,+\beta}$ nor $\tilde{a}_{\alpha,-\beta}$ are physical observations. The new coefficients, $\hat{a}^{\pm}_{\alpha,\beta}$, are also independent Fourier coefficients, and are used throughout the manuscript.

Appendix B. Proper orthogonal filtering

The POD was introduced to fluid mechanics by Berkooz *et al.* (1993) as a variant of the older principal component analysis approximation of Pearson (1901). It seeks to represent a vector field $\boldsymbol{u}$ as an expansion of POD modes $\boldsymbol{\phi}_{(k)}$

$$\boldsymbol{u}_{(n)} = \sum_{k=1}^{n} \hat{u}_k \boldsymbol{\phi}_{(k)}, \tag{B1}$$

in such a way that the two-point covariance of $\boldsymbol{u}_{(n)}$ approximates as well as possible the covariance of the true vector field, even when the order of the expansion is much less

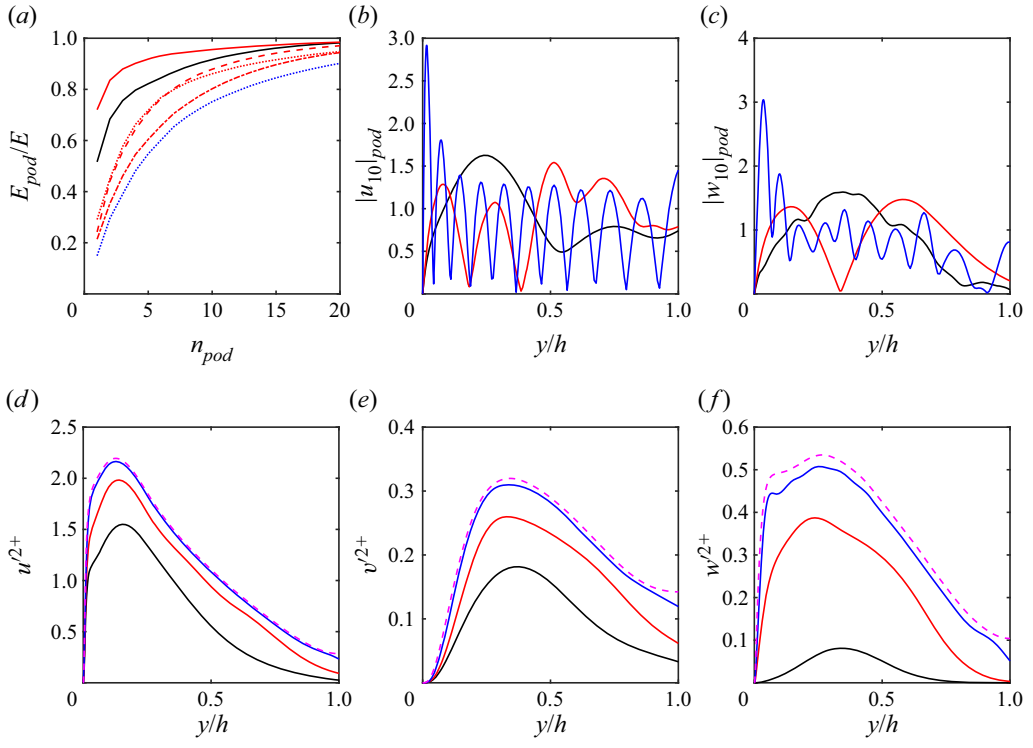


Figure 16. (*a*) Energy fraction of the different harmonics contained in first n_{pod} modes. Solid black, [00] mode; solid red, [10]; dashed red, [10]; chaindotted red, [20]; dotted red, [11]; dotted blue, [21]. (*b*,*c*) Wall-normal intensity profiles of typical POD modes of the [10] harmonic. Black, $n_{pod} = 1$; red, $n_{pod} = 5$; blue, $n_{pod} = 20$. (*b*) u_{10}. Panel (*c*) shows w_{01}. (*d–f*) Intensity profiles of all the retained harmonics, approximated with different number of POD modes. Colours as in (*b*). The dashed magenta line, is the true intensity profile; (*d*) u, (*e*) v, (*f*) w.

than the number of degrees of freedom in $\boldsymbol{u}$. As already mentioned in §§ 2 and 3, POD modes are not good choices to reduce a dynamical system to a few variables, because they minimise the integral of the error over the attractor as a whole, instead of, as has been our interest in this paper, differentiating among different phase-space neighbourhoods. However, they can approximate the flow field using a somewhat reduced number of spatial modes.

It is clear that our choice to represent the flow using nine Fourier modes in the (x, z) plane is a drastic reduction of numerical resolution in that plane, but there is no simple equivalent way of reducing the resolution along the non-homogeneous direction, y. The result is a strongly anisotropic flow field with spurious thin wall-parallel layers which are not justified by the wall-parallel resolution. POD modes provide a useful basis to balance the resolution in a statistically significant way. In fact, Fourier modes are POD modes along homogeneous directions, such as x and z, and we saw in figure 2 that the few retained modes account for a significant fraction of the kinetic energy (most of the mathematical properties mentioned in this appendix are drawn from Berkooz *et al.* (1993), although this will not be mentioned from now on). The POD analysis can be extended to y by expanding individual Fourier modes, and figure 16(*a*) shows the fraction of the total energy of each Fourier mode contained in its first n POD modes. The first few modes do not represent the energy well, but approximately 20 modes account for most of it. Because the POD modes in (B1) are eigenvectors of the covariance matrix, successive eigenvectors contain

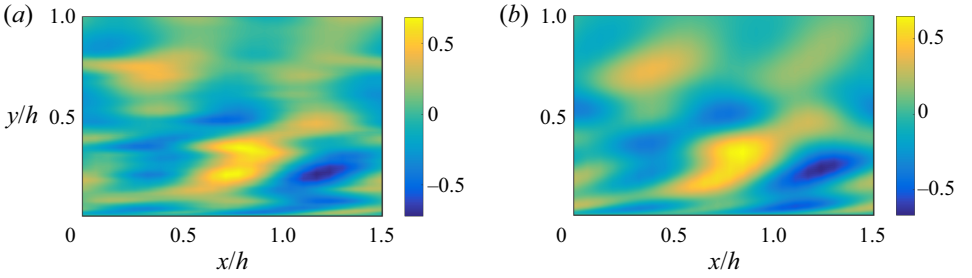


Figure 17. Example of the effect of POD filtering on a streamwise section of the w field. (a) Reconstructed with the nine retained Fourier modes and full resolution in y. (b) Using 20 PODs in y.

an increasing number of 'waves', as shown in the examples in figure 16(b,c). Broadly speaking, the 20 modes mentioned above are equivalent to a wall-normal resolution of ten points, which is a reasonable compromise for the number of wall-parallel Fourier modes. Figure 16(d–f) shows that this resolution is enough to account for most of the energy profile of the retained Fourier approximation.

Several precautions are necessary for a consistent approximation. In the first place, expansion (B1) optimises the approximation of the flow field in terms of the L_2 norm defined over the collocation nodes of the numerical grid, while it makes more physical sense to use an integral norm $\int |\boldsymbol{u}|^2 \,\mathrm{d}y$. This requires scaling the flow field by the square root of the grid spacing, and undoing the scaling upon reconstruction. Similarly, although figure 16 displays modes as associated with individual velocity components, it is important to compute the PODs over the three velocities at the same time, catenated as a single compound vector. This minimises the total kinetic energy of the approximation error, and ensures, among other things, that the reconstructed field satisfies continuity.

An example of the effect of filtering the w field with the first 20 PODs is shown in figure 17. Most of the results in this paper are obtained from flow fields filtered with 20 POD modes. There is very little difference between these results and those from unfiltered variables, except for the somewhat cleaner flow fields such as the one in figure 17(b).

References

Adrian, R.J. 2007 Hairpin vortex organization in wall turbulence. *Phys. Fluids* **19**, 041301.

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

Angrist, J.D., Imbens, G.W. & Rubin, D.B. 1996 Identification of causal effects using instrumental variables. *J. Am. Stat. Assoc.* **91**, 444–455.

Beck, C. & Schlögl, F. 1993 *Thermodynamics of Chaotic Systems*. Cambridge University Press.

Berkooz, G., Holmes, P. & Lumley, J.L. 1993 The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.* **25**, 539–575.

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

Courant, R. & Hilbert, D. 1953 *Methods of Mathematical Physics I*. Wiley.

Encinar, M.P. & Jiménez, J. 2020 Momentum transfer by linearised eddies in turbulent channel flows. *J. Fluid Mech.* **895**, A23.

Farrell, B.F. & Ioannou, P.J. 1996 Generalized stability theory. Part II: nonautonomous operators. *J. Atmos. Sci* **53**, 2041–2053.

Feller, W. 1971 *An Introduction to Probability Theory and its Applications*, 2nd edn, vol. 2. Wiley.

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

Froyland, G. 2005 Statistically optimal almost-invariant sets. *Physica* D **200**, 205–219.

GRANGER, C.W.J. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438.
HAMILTON, J.M., KIM, J. & WALEFFE, F. 1995 Regeneration mechanisms of near-wall turbulence structures. *J. Fluid Mech.* **287**, 317–348.
JIMÉNEZ, J. 1994 On the structure and control of near wall turbulence. *Phys. Fluids* **6**, 944–953.
JIMÉNEZ, J. 2012 Cascades in wall-bounded turbulence. *Annu. Rev. Fluid Mech.* **44**, 27–45.
JIMÉNEZ, J. 2013 How linear is wall-bounded turbulence? *Phys. Fluids* **25**, 110814.
JIMÉNEZ, J. 2015 Direct detection of linearized bursts in turbulence. *Phys. Fluids* **27**, 065102.
JIMÉNEZ, J. 2018*a* Coherent structures in wall-bounded turbulence. *J. Fluid Mech.* **842**, P1.
JIMÉNEZ, J. 2018*b* Machine-aided turbulence theory. *J. Fluid Mech.* **854**, R1.
JIMÉNEZ, J. 2020 Monte Carlo science. *J. Turbul.* **21**, 544–566.
JIMÉNEZ, J. 2022 The streaks of wall-bounded turbulence need not be long. *J. Fluid Mech.* **945**, R3.
KAWAHARA, G., UHLMANN, M. & VAN VEEN, L. 2012 The significance of simple invariant solutions in turbulent flows. *Annu. Rev. Fluid Mech.* **44**, 203–225.
KIM, J., MOIN, P. & MOSER, R.D. 1987 Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.* **177**, 133–166.
KWON, Y. & JIMÉNEZ, J. 2021 An isolated logarithmic layer. *J. Fluid Mech.* **916**, A35.
LANCASTER, P. 1969 *Theory of Matrices*. Academic.
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.
MALKUS, W.V.R. 1956 Outline of a theory of turbulent shear flow. *J. Fluid Mech.* **1**, 521–539.
MARUSIC, I., MATHIS, R. & HUTCHINS, N. 2010 Predictive model for wall-bounded turbulent flow. *Science* **329**, 193–196.
MATHIS, R., MARUSIC, I., CHERNYSHENKO, S.I. & HUTCHINS, N. 2013 Estimating wall-shear-stress fluctuations given an outer region input. *J. Fluid Mech.* **715**, 163–180.
MCKEON, B.J. & SHARMA, A.S. 2010 A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.* **658**, 336–382.
MIZUNO, Y. & JIMÉNEZ, J. 2013 Wall turbulence without walls. *J. Fluid Mech.* **723**, 429–455.
NIKULIN, M.S. 2001 Hellinger distance. In *Encyclopaedia of Mathematics*. EMS Press.
ORR, W.M. 1907 The stability or instability of the steady motions of a perfect liquid, and of a viscous liquid. Part I: A perfect liquid. *Proc. R. Irish Acad.* **27**, 9–68.
PEARL, J. 2009 Causal inference in statistics: an overview. *Stat. Surveys* **3**, 95–146.
PEARSON, K. 1901 On lines and planes of closest fit to systems of points in space. *Phil. Mag.* **6**, 559–572.
REYNOLDS, W.C. & TIEDERMAN, W.G. 1967 Stability of turbulent channel flow, with application to Malkus' theory. *J. Fluid Mech.* **27**, 253–272.
RUSSELL, B. 1912 On the notion of cause. *Proc. Aristotelian Soc.* **13**, 1–26.
SCHMID, P.J. 2010 Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28.
SCHMID, P.J., GARCÍA-GUTIÉRREZ, A. & JIMÉNEZ, J. 2018*a* Description and detection of burst events in turbulent flows. *J. Phys.: Conf. Ser.* **1001**, 012015.
SCHMID, P.J., SCHMIDT, O.T., TOWNE, A. & HACK, M.J.P. 2018*b* Analysis and prediction of rare events in turbulent flows. In *CTR Summer Program Proceedings*, pp. 139–148. Center for Turbulence Research.
SCHOPPA, W. & HUSSAIN, F. 2002 Coherent structure generation in near-wall turbulence. *J. Fluid Mech.* **453**, 57–108.
TOWNSEND, A.A. 1976 *The Structure of Turbulent Shear Flow*. Cambridge University Press.
TUERKE, F. & JIMÉNEZ, J. 2013 Simulations of turbulent channels with prescribed velocity profiles. *J. Fluid Mech.* **723**, 587–603.
ULAM, S.M. 1964 *Problems in Modern Mathematics*. Interscience.
WALEFFE, F. 1997 On a self-sustaining process in shear flows. *Phys. Fluids* **9**, 883–900.