Description and detection of burst events in turbulent flows

P J Schmid\(^1\), A García-Gutierrez\(^2\) and J Jiménez\(^2\)
\(^1\)Department of Mathematics, Imperial College London, London SW7 2AZ, UK
\(^2\)School of Aeronautics, U. Politénica Madrid, 28040 Madrid, Spain
E-mail: pjschmid@imperial.ac.uk

Abstract. A mathematical and computational framework is developed for the detection and identification of coherent structures in turbulent wall-bounded shear flows. In a first step, this data-based technique will use an embedding methodology to formulate the fluid motion as a phase-space trajectory, from which state-transition probabilities can be computed. Within this formalism, a second step then applies repeated clustering and graph-community techniques to determine a hierarchy of coherent structures ranked by their persistencies. This latter information will be used to detect highly transitory states that act as precursors to violent and intermittent events in turbulent fluid motion (e.g., bursts). Used as an analysis tool, this technique allows the objective identification of intermittent (but important) events in turbulent fluid motion; however, it also lays the foundation for advanced control strategies for their manipulation. The techniques are applied to low-dimensional model equations for turbulent transport, such as the self-sustaining process (SSP), for varying levels of complexity.

1. Introduction and motivation
Turbulent fluid motion is characterized by a wide range of motion, involving broad spatial as well as temporal scales (e.g. [1]): while large, coherent structures are known to play a critical and persistent role in the sustainment of turbulence, intermittent events linked to their breakdown and subsequent reconstitution are also observed. This co-existence of different dynamic events is perhaps best encapsulated by minimal flow-units [2] or self-sustaining model processes [3, 4, 5]. In the former case, spatial scales are progressively removed until turbulent fluid motion can no longer be sustained, thus restricting the essential dynamics to its minimal ingredients. In the latter case, hydrodynamic stability results are harnessed together with a feedback loop that describes the re-triggering (and eventual maintenance) of the instabilities. In either case, the final result shows the interplay of slowly moving coherent structures that intermittently – and often violently – collapse via a fast nonlinear process, before rising again and starting another cycle [6, 7].

Describing this combined, multiscale process is challenging and shall be attempted in this study. To this end, it seems advantageous to describe the dynamic evolution of the flow fields as a path in phase space. In this description, the coherent part of the quasi-cyclic process can be described by motion in certain regions of phase space, while intermittent events are represented by fast and far excursions from these regions. Characteristic turbulence features and quantities, such as production, dissipation and intermittency, can be associated with these events, and the frequency of their occurrence in simulations and experiments can be gauged. Identifying these
building blocks of persistent and intermittent fluid structures is thus instrumental in a complete description and categorization of turbulent fluid motion; at the same time, the same structures and their catchment basins in phase space can be used to design efficient control strategies to avoid or mitigate violent events.

The distinction between persistence and intermittency is made within a probabilistic framework, together with the above phase-space description. A phase-space trajectory that remains (with high probability) in the vicinity of a given region describes a coherent structure. On the other hand, a trajectory that shows a high probability of diverging from its current neighbourhood in phase space represents a transitory state. The transition probabilities of phase space trajectories are thus closely related to the persistence of the structures they represent.

The same probabilities, together with the geometric properties of the phase-space trajectories, can be used to combine structures that behave similarly. To this end, we interpret the fluid motion as a weighted graph in phase space – connecting one region in phase space (a node in the graph) with a neighbouring region (a connected node in the graph), with a given transition probability between them (the weight of the edge between the two nodes). This point of view allows the application of advanced clustering algorithms to reorganize the graph (and thus the fluid motion) into communities (distinct and well-defined structures). In the end, we aim at the identification of the building blocks of turbulent flows and a minimal description of motion involving these blocks.

2. Mathematical background
The above formulation of fluid motion, given in the form of snapshot sequences, proceeds in three steps. First, the data need to be embedded in phase space, commonly via projection onto an appropriate lower-dimensional basis. Second, in this phase space the transition probabilities are computed by following the phase-space trajectory. In a third step, the raw transition probability matrix is deflated (lumped) by detecting and combining states of similar dynamic behavior. This last step is performed repeatedly, until a desired level of coarse-graining is reached.

2.1. Embedding in phase space
Numerical simulations or experimental measurements of turbulent flows require the resolution of a wide range of scales; this yields data sequences of many degrees of freedom. However, not all degrees of freedom are independent and dynamically relevant. This fact justifies the projection of the full motion onto an appropriate basis and the approximation of the fluid motion by tracking the coefficients of the basis vectors. Various choices for the basis vectors exist and have been explored over past decades. Proper orthogonal decomposition (POD) [8, 9, 10] is a popular option; POD-modes preserve orthogonality of the spatial structure and can be shown to optimally preserve the energy in the retained modes. They are computed directly from the data matrix, i.e., the matrix whose columns represent the snapshots of the flow field, by a singular value decomposition or, alternatively, by forming correlation matrices and computing their eigenvalues. While POD-modes focus on spatial orthogonality, they yield multi-frequential time traces. If single-frequency structures are desired (for example, to describe the shedding structures behind a bluff body in terms of single Strouhal numbers), a dynamic mode decomposition (DMD) can be taken [11, 12, 13]. In this case, spatial orthogonality is sacrificed to gain a single-frequency decomposition of the data. Other alternatives are resolvent modes [14], that is, the principal singular vectors of a harmonically driven linear system.

In either case, we can describe a temporally evolving system with many spatial degrees of freedom by a set of temporally evolving coefficients. These coefficients form a phase space, and the original time-evolving flow field can be represented by a trajectory in this phase space. By truncating the number of basis vectors, we have control over the dimensionality of this phase space. As a representation of a general phase space embedding, we decompose the data matrix
D into a product of three matrices: the spatial basis $U$, the amplitude matrix $\Sigma$, and the time-dynamics $V$.

$$D = U \Sigma V^H = U W^H$$

(1)

We assume that the matrix $\Sigma$ is diagonal, as it allows the one-to-one association of a spatial structure (column in $U$) with a specific dynamic signature (column in $V$). Without any additional constraints on $U$ and $V$, the above decomposition is not unique. Nonetheless, we will not further specify constraints and accept the decomposition in its generality. The columns of $W$ describe the time traces in phase space and will form the foundation for further analysis.

2.2. Transition probabilities

The phase-space trajectory is then used to determine state transition probabilities. To this end, we seek a gridding of the high-dimensional phase space which ascribes a finite hyper-cube volume in phase space to each instance in the data sequence. A naïve gridding approach would be prohibitively expensive, as the phase space would grow in dimension and the number of hyper-cube volumes would increase exponentially. A different approach has to be adopted.

Rather than gridding the entire phase space, we define hyper-cubes around time instances of the phase-space trajectory and introduce new hyper-cubes as the trajectory leaves any of the current hyper-cubes. In this process, we also keep a linked list that stores information about the temporal evolution along the trajectory as to the propagation direction from hyper-cube to hyper-cube. In this manner, the curse of dimensionality associated with the naïve approach can be avoided, and the final set of hyper-cubes is always less than the number of snapshots in our data matrix. More specifically, the tessellation of the phase-space trajectory is accomplished by a subdivision algorithm that, starting with a coarse assignment to large hyper-cubes, iteratively refines the hyper-cube volumes (by factors of two in subsequent coordinate directions) until a sufficient degree of fidelity of the phase-space motion is assured. As a result, we have access to information about the location and the motion of the trajectory in phase space, discretized by our set of hyper-cubes.

With this information, we next compute a transition probability matrix. Each row of the matrix represents a box containing a trajectory at a given time, while each column represents a box into which the trajectory moves within a single time-step. The entry in the matrix at a given row $i$ and column $j$ then measures the probability of having moved from box $j$ to box $i$ over one time step. This probability can be approximated by measuring the number of instances where motion between $i$ and $j$ is detected (within one time step) normalized by the total number of instances in $i$. Mathematically, we can express this as

$$P_{ij} = \frac{m(B_i \cap F^{-1}(B_j))}{m(B_i)}$$

(2)

where $m(B)$ denotes a suitable density measure in phase space (in our case, the number of instances contained in the box $B$), and $F^{-1}$ is the temporal backstep operator, i.e., we evaluate its argument at the previous time-step (see, e.g., [15]). The above formula states that we count the number of occurrences, where instances move from box $j$ to box $i$ over one time step and divide it by the number of total instances in box $i$. This assigns a probability measure to the event that instances along the trajectory that end up in box $i$ have come from box $j$ over one time step. This procedure provides an approximation of the (Markov) transition probability matrix between all boxes in our tessellation of phase space and is referred to as the Ulam–Galerkin method [16]. It is interesting to note that the transition probability matrix $P_{ij}$ is a finite-dimensional approximation of the Frobenius–Perron operator [17] (i.e., the adjoint of the Koopman operator).
This Markov transition matrix $P$ allows the description of phase-space dynamics in probabilistic terms. Starting in one particular hypercube $B_i$, we can compute the probabilities of remaining within the same hypercube or transitioning to a neighbouring hypercube. By repeated multiplication by $P$ we arrive at a probability distribution of possible outcomes. Furthermore, the eigenvectors of the $P$-matrix associated with near-unity eigenvalues constitute quasi-invariant flow states that represent persistent and highly observable coherent structures.

2.3. Community clustering and deflation

With the tessellation of the occupied phase space established and the transition probability matrix computed, we are now in a position to re-organize the various identified flow states to account for the temporal dynamics between them. This is best accomplished by interpreting the transition probability matrix as an adjacency matrix of a directed, weighted graph [18]: the flow states (high-dimensional boxes) are connected by non-zero transition probabilities. At this stage, the dynamics or persistence of flow states has not been taken into account. This latter property is brought out by regrouping similar flow states into a compound flow state; in the language of graphs, this regrouping is known as community clustering. It consists of a repartition of the network into groups of nodes such that intra-connectivity within the same group is maximized, while inter-connectivity from group to group is minimized [19, 20]. Using the transition probabilities above, the community clustering is equivalent to lumping dynamically similar structures and to generating a hierarchy of structures ranging from quasi-invariant structures (with low transition probabilities) to highly transitory structures (with high transition probabilities).

Computationally, we apply a simple yet efficient algorithm to optimize the community structure of our phase-space network, proposed in [21]. This algorithm targets a measure of inter-connectedness in networks based on the number and weight of links between communities. Mathematically, we define as

$$Q = \frac{1}{2m} \sum_{i,j} \left[ P_{ij} - \frac{k_i^{(in)} k_j^{(out)}}{2m} \right] \delta_{c_i,c_j}$$

the modularity of the graph. In the above expression, $P_{ij}$ stands for the transition probability between node $i$ and $j$ and $k_i^{(in, out)} = \sum_j P_{ij}$ is the sum of all probabilities associated with node $i$, going into or out of node $i$, respectively. The variable $c_i$ denotes the $i$-th community and $\delta$ is the standard Kronecker symbol. Finally, $m$ is defined as $m = \frac{1}{2} \sum_{ij} P_{ij}$, the sum of the weights of all links in an undirected network [22, 23].

The algorithm proceeds in two steps. First, a greedy optimization step maximizes the modularity $Q$ by placing the neighbors of node $i$ into neighbouring communities, if it increases $Q$. Once, no more improvement can be achieved by swapping nodes into adjacent communities, a second step commences which replaces the new communities by single nodes and updates the transition probabilities. After this step, we return to the maximization of $Q$ and continue to iterate until no further improvement can be made. The two-step process performs very efficiently, in particular, for sparsely connected graphs, as is the case for us.

Alternative clustering algorithms can be and have been used [15], such as the classical k-means method. In this case, however, the number of communities is prescribed and the resulting community association is rather uniform. These properties defy the purpose of an objective clustering where strongly non-uniform communities are expected.

We apply the community clustering algorithm to our transition probability matrix $P \in \mathbb{R}^{N \times N}$, which results into a regrouping of our $N$ initial boxes into $k$ communities, by maximizing the modularity of the underlying graph. The result of this process is a rectangular matrix
$Z_{(1)} \in \mathbb{R}^{N \times k}$ which uniquely assigns a given box $B_i$ to one of the identified communities; mathematically, we have
\[
Z_{(1),ij} = \begin{cases} 
1 & \text{if } B_i \in c_j, \\
0 & \text{otherwise.}
\end{cases}
\] (4)

The transformation
\[
Z_{(1)}^T P Z_{(1)} = P_{(1)}
\] (5)
then regroups the transition matrix $P$ into $P_{(1)} \in \mathbb{R}^{k\times k}$ which describes the transition probabilities between the newly found communities, rather than our original boxes. By repeating this clustering and lumping process we finally arrive at a transition probability matrix of desired size. We have
\[
Z_{(k)}^T \cdots Z_{(1)}^T P Z_{(1)} \cdots Z_{(k)} = P_{(k)}.
\] (6)

It is worth noting that the entries of the matrix $Z_{(i)}$ depend on the graph structure of the associated matrix $P_{(i-1)}$; the above formula thus expresses a complex and nonlinear transformation process.

Once we are satisfied with the number of identified communities, we can classify the identified communities with associated flow structures. By relating the communities back to the original snapshots, we can catalogue similar dynamical behaviour. In particular, communities that have a nearly unitary transition probability represent flow structures that occupy and remain (with high probability) in a specific region of phase space. Physically, these structures, while unsteady, show a great degree of persistence in the overall flow. In our case of turbulent flow, we expect the slowly meandering, streamwise elongated structures to fall into this category. On the other hand, communities that show a high transition probability to neighbouring communities signify transitory structures and can be thought of as precursors for a violent (and/or rare) event. These structures are far less observable, but highly controllable.

3. Low-dimensional models

While designed to apply to flow field sequences of fully turbulent flows, from numerical or experimental sources, we will demonstrate the key elements of the above methodology on a simplified yet relevant model of the underlying self-sustaining process that maintains turbulent fluid motion. We select the nine-dimensional system of nonlinear ordinary differential equations proposed in [24]. This model has been following earlier attempts at deriving a compact system by projecting the Navier–Stokes equations onto pertinent models that play a critical role in the sustainance of turbulence in generic wall-bounded flows. In our case, the flow is based on a uni-directional, parallel, sinusoidal shear flow, and the flow is considered in a doubly periodic computational domain with one inhomogeneous direction. Modelling the base flow profile, the streak and the downstream vortex, the spanwise modulations of the latter components, normal vortex modes and fully three-dimensional perturbations – all with appropriate spatial scales and unknown amplitudes – we arrive at an evolution problem for the nine coefficients $\{a_1, a_2, \ldots, a_9\}$ that capture the growth, instability, breakdown and regeneration of flow structures. The model [24] has strong similarity to an earlier eight-coefficient model [4]. Previous work has suggested that the self-sustaining process, described by the model equation, constitutes a universal characteristic of wall-bounded shear flow turbulence [3, 4, 25, 26]. The chosen model can be stated in the form of nine quadratically interacting ordinary differential equations according to
\[\dot{a}_1 + \zeta_1 a_1 = \zeta_1 + \xi_1 a_6 a_8 + \xi_1 a_2 a_3,\]  
(7a)

\[\dot{a}_2 + \zeta_2 a_2 = \xi_2 a_1 a_6 - \xi_3 a_5 a_7 - \xi_3 a_5 a_8 - \xi_3 a_1 a_3 - \xi_2 a_5 a_9,\]  
(7b)

\[\dot{a}_3 + \zeta_3 a_3 = \xi_3 a_1 a_7 + a_5 a_6,\]  
(7c)

\[\dot{a}_4 + \zeta_4 a_4 = -\xi_4 a_1 a_5 - \xi_4 a_2 a_6 - \xi_4 a_3 a_7 - \xi_4 a_3 a_8 - \xi_4 a_5 a_9,\]  
(7d)

\[\dot{a}_5 + \zeta_5 a_5 = \xi_5 a_1 a_4 + \xi_5 a_2 a_7 - \xi_5 a_2 a_8 + \xi_5 a_4 a_9 + \xi_5 a_3 a_6,\]  
(7e)

\[\dot{a}_6 + \zeta_6 a_6 = \xi_6 a_1 a_7 + \xi_6 a_1 a_8 + \xi_6 a_2 a_4 - \xi_6 a_3 a_5 + \xi_6 a_7 a_9 + \xi_6 a_8 a_9,\]  
(7f)

\[\dot{a}_7 + \zeta_7 a_7 = -\xi_7 a_1 a_6 - a_6 a_9 + \xi_7 a_2 a_5 + \xi_7 a_3 a_4,\]  
(7g)

\[\dot{a}_8 + \zeta_8 a_8 = \xi_8 a_1 a_5 + \xi_8 a_3 a_4,\]  
(7h)

\[\dot{a}_9 + \zeta_9 a_9 = \xi_9 a_2 a_3 - \xi_9 a_6 a_8,\]  
(7i)

where the coefficients in the above equations are given as

\[\zeta_1 = \frac{\beta^2}{Re}, \quad \zeta_2 = \frac{1}{Re} \left( \frac{4\beta^2}{3} + \gamma^2 \right), \quad \zeta_3 = \frac{\beta^2 + \gamma^2}{3Re}, \quad \zeta_4 = \frac{3\alpha^2 + 4\beta^2}{3Re},\]

\[\zeta_5 = \frac{\alpha^2 + \beta^2}{Re}, \quad \zeta_6 = \frac{3\alpha^2 + 4\beta^2 + 3\gamma^2}{3Re}, \quad \zeta_7 = \frac{\alpha^2 + \beta^2 + \gamma^2}{Re}, \quad \zeta_8 = 9\zeta_1 \]  
(8)

and

\[\xi_{11} = \xi_{62} = \xi_{66} = \xi_{92} = \sqrt{\frac{3}{2}} \frac{\beta\gamma}{\kappa_{\alpha\beta}}, \quad \xi_{12} = \xi_{24} = \xi_{25} = \xi_{91} = \sqrt{\frac{3}{2}} \frac{\beta\gamma}{\kappa_{\beta\gamma}},\]

\[\xi_{21} = \frac{5\sqrt{2}}{3\sqrt{3}} \frac{\gamma^2}{\kappa_{\alpha\gamma}}, \quad \xi_{22} = \frac{\gamma^2}{\sqrt{6\kappa_{\alpha\gamma}}}, \quad \xi_{23} = \xi_{53} = \frac{1}{\sqrt{6}} \kappa_2, \quad \xi_{31} = \xi_{55} = \frac{2}{\sqrt{6}} \kappa_1,\]

\[\xi_{32} = \frac{\beta^2 (3\alpha^2 + \gamma^2) - 3\gamma^2 (\alpha^2 + \gamma^2)}{\sqrt{6} \kappa_{\alpha\gamma} \kappa_{\beta\gamma} \kappa_{\alpha\beta}}, \quad \xi_{41} = \xi_{45} = \xi_{51} = \xi_{54} = \xi_{61} = \xi_{65} = \xi_{71} = \frac{\alpha}{\sqrt{6}},\]

\[\xi_{42} = \frac{10\alpha^2}{3\sqrt{6} \kappa_{\alpha\gamma}}, \quad \xi_{43} = \frac{\sqrt{3}}{2} \kappa_1, \quad \xi_{44} = \frac{\sqrt{3}}{2} \frac{\alpha^2 \beta^2}{\kappa_{\alpha\gamma} \kappa_{\beta\gamma} \kappa_{\alpha\beta}}, \quad \xi_{52} = \frac{\alpha^2}{\sqrt{6} \kappa_{\alpha\gamma}},\]

\[\xi_{63} = \frac{10}{3\sqrt{6}} \frac{\alpha^2 - \gamma^2}{\kappa_{\alpha\gamma}}, \quad \xi_{64} = \frac{2}{\sqrt{6}} \kappa_1, \quad \xi_{72} = \frac{\gamma^2 - \alpha^2}{\sqrt{6} \kappa_{\gamma}}, \quad \xi_{73} = \frac{1}{\sqrt{6}} \kappa_1,\]

\[\xi_{81} = \frac{2}{\sqrt{6}} \kappa_2, \quad \xi_{82} = \frac{\gamma^2 (3\alpha^2 - \beta^2 + 3\gamma^2)}{\sqrt{6} \kappa_{\alpha\gamma} \kappa_{\beta\gamma} \kappa_{\alpha\beta}} \]  
(9)

with

\[\kappa_{\alpha\gamma} = \sqrt{\alpha^2 + \gamma^2}, \quad \kappa_{\beta\gamma} = \sqrt{\beta^2 + \gamma^2}, \quad \kappa_{\alpha\beta} = \sqrt{\alpha^2 + \beta^2 + \gamma^2} \]  
(10)

and

\[\kappa_1 = \frac{\alpha \beta \gamma}{\kappa_{\alpha\gamma} \kappa_{\beta\gamma}}, \quad \kappa_2 = \frac{\alpha \beta \gamma}{\kappa_{\alpha\gamma} \kappa_{\alpha\beta}}.\]

(11)

The remaining, user-specified coefficients are the Reynolds number \( Re \), and the streamwise and spanwise wavenumbers \( \alpha \) and \( \gamma \), respectively. The wavenumber \( \beta \), which determines the length scale in the wall-normal direction, is taken as constant and set to \( \pi/2 \).

The coefficients \{\( a_1, a_3 \}\} collectively describe the mean flow and shear, the coefficients \{\( a_2, a_3, a_4 \}\} represent the rolls and streaks, and the coefficients \{\( a_5, a_6, a_7, a_8 \}\} characterize the burst event. For visualization purposes, we will thus combine the nine coefficients in this manner. We choose the dimension of our minimal unit as \( L_x = 4\pi \) and \( L_z = 2\pi \) from which we determine \( \alpha = 2\pi/L_x \) and \( \gamma = 2\pi/L_z \) and the remaining system coefficients. The Reynolds number (based on the half-channel height) is set to \( Re = 400 \).
4. Results
We integrate the governing equations (7) forward in time and, after an initial transient period, visualize the results in phase space. Due to the high dimensionality of the system, we choose to combine the nine coefficients into three groups characterizing the contributions from (i) the streamwise roll and streak ($\|(a_2, a_3)^T\|$), (ii) the mean shear ($\|1 - a_1\|$), and (iii) the three-dimensional instability structure responsible for the breakdown of the streak ($\|(a_4, a_5)^T\|$).

Figure 1 shows a typical phase-space trajectory, visualized by combining the nine coefficients in the manner described above. We observe a clumping of the trajectory in phase space that is interrupted by occasional and violent excursions from the cluster and an eventual and equally rapid return to the cluster. This phase-space behavior is further exemplified by displaying the burst component, consisting of the three-dimensional instability coefficients, as a function of time (see Figure 1(b)). We notice a signal that consists of spikes in amplitude that occur rapidly, intermittently and in varying intensity. This behaviour indicates the type of flow pattern described above, where a quasi-coherent motion is periodically but irregularly disrupted by violent events of a different nature. In our case, these events represent the breakdown of the streamwise roll and streak and its subsequent reconstitution.

The phase-space trajectory, represented by the original nine coefficients $a_1, ..., a_9$ over 2000 time units, is then tessellated such that each nine-dimensional hypercube does not contain more than 150 snapshots. From this tessellation the raw transition probability matrix is computed following (2). The fill pattern of this matrix is shown in Figure 2(a): it is clearly diagonally dominant, with few off-diagonal entries, reflecting the initial, sequential processing of the phase-space trajectory. No clustering or categorization of the fluid motion has yet been applied. Interpreting the transition probability matrix $P$ as the adjacency matrix of an associated directed graph, we can now apply the clustering algorithm and determine graph-communities that show an optimal level of graph modularity according to (3). The identified community allows us to regroup the snapshots and deflate the original matrix $P$ – describing the motion between...
more than 1200 hypercubes – into a transition probability matrix $P_{(1)}$ between nearly 500 communities. The fill pattern of this latter matrix is shown in Figure 2(b), where each symbol represents a nonzero entry in the matrix. The figure displays a substantially different pattern: the communities can be grouped into more persistent elements (with a strong concentration along the diagonal) and more transitory elements (where the diagonal is absent or sparsely populated). An approximate classification is indicated by dashed boxes in the figure. It appears that after only one application of the graph-community clustering algorithm, a separation of persevering, slow fluid motion and volatile events is already emerging. By repeated application of the clustering algorithm, we can further solidify this feature. Figure 2(c) shows the fill pattern of the transition matrix after four deflation applications; at this point, only 34 different communities are tracked. We notice that the gap along the diagonal consisting of communities $c_5, \ldots, c_{12}$, which was already notable in Figure 2(b), is now even more pronounced.

The layout of the fill pattern now allows the quantitative and objective description of
rare events. This is indicated in one example in Figure 2(c). The probability of resting in community $c_2$ is given by the diagonal entry in $P^{(4)}$; however, there is also a (smaller) probability of transitioning to community $c_{10}$. The probability of resting in community $c_{10}$ is, however, negligibly small (i.e., below the threshold of marking the entry in the matrix, taken as $10^{-4}$). We thus revert to community $c_0$ and then close the cycle by slowly moving again to community $c_2$. The slow-fast cycle is indicated by arrows in Figure 2(c). Instances of similar behaviour can be extracted and quantified in this manner from the transition probability matrix $P^{(4)}$. The flow states associated with communities $c_{0,...,4}$ can be labelled as quasi-stationary (with rather large residence times and low transition probabilities), while the flow states corresponding to communities $c_{5,...,12}$ represent highly transitory states with comparatively low residence time and high transition probabilities. In this manner, cyclical processes that consist of a combination of slow motion and intermittent, fast events can be detected objectively and described quantitatively.

5. Summary and conclusions

In this preliminary study, a computational framework for the description and analysis of rare and violent events embedded in a slowly varying fluid motion has been introduced. It consists of the analysis of an embedded phase-space trajectory using a probabilistic approach by estimating the transition probabilities between discretized regions of phase-space traversed by the flow. Employing techniques from graph clustering, the discretized regions are recombined and grouped into communities that describe persistent and highly observable fluid motion as well as transitory and highly controllable processes.

We applied this technique to a reduced-order minimal-channel model of the self-sustaining process (SSP) – a mechanism that is pivotal in sustaining turbulent fluid motion. The framework confirmed the existence of quasi-stationary and transitory states and extracted flow states of either kind objectively and without prior definitions.

In future work, data from direct numerical simulations will be used to detect burst events. Moreover, the above framework will be extended to devise optimized (sparse) control techniques to avoid or attenuate intermittently occurring fluid motion. The application to larger-dimensional systems should not pose significant difficulties, since the chosen data-structure for the phase-space trajectory avoids the curse of dimensionality of a more naïve approach and allows the efficient computation of the transition matrix even for higher-dimensional phase-space embeddings. Results of our analysis of turbulent data-bases will be reported elsewhere.

Acknowledgements

Support from the European Union under the ERC Advanced Grant “COTURB: Coherent Structures in Wall-bounded Turbulence” is gratefully acknowledged. We are grateful to S. Illingworth for his careful refereeing comment on the manuscript.

References

[1] Pope S 2000 Turbulent Flows (Cambridge University Press)
[2] Jimenez J and Moin P 1991 J. Fluid Mech. 225 213
[3] Hamilton J M, Kim J and Waleffe F 1995 J. Fluid Mech. 287 317
[4] Waleffe F 1995 Phys. Fluids 7 3060
[5] Cossu C and Hwang Y 2017 Phil. Trans. Roy. Soc. A 375 20160088
[6] Antonia R, Danh H and Prabhu A 1976 Phys. Fluids 19 1680
[7] Kim J 1985 Phys. Fluids 28 52
[8] Lumley J 1970 Stochastic Tools in Turbulence (Academic Press)
[9] Sirovich L 1987 Q. Appl. Math. 561
[10] Berkooz G, Holmes P and Lumley J 1993 Ann. Rev. Fluid Mech. 25 539
[11] Schmid P 2010 J. Fluid Mech. 656 5
[12] Rowley C, Mezic I, Bagheri S, Schlatter P and Henningson D 2009 J. Fluid Mech. 641 85
[13] Mezic I 2013 Ann. Rev. Fluid Mech. 45 357
[14] McKeon B and Sharma A 2010 J. Fluid Mech. 658 336
[15] Kaiser E, Noack B, Cordier L and Spohn A 2014 J. Fluid Mech. 754 365
[16] Ulam S 1960 A Collection of Mathematical Problems (Interscience Publisher NY)
[17] Lasota A and Mackey M 1994 Chaos, Fractals, and Noise: Stochastic aspects of dynamics (Springer Verlag)
[18] Chartrand G and Zhang P 2012 A First Course in Graph Theory (Dover Publishing)
[19] Girvan M and Newman M 2002 Proc. Nat. Acad. Sci. 99 7821
[20] Fortunato S and Barthelemy M 2007 Proc. Nat. Acad. Sci. 104 36
[21] Blondel V, Guillaume J L, Lambiotte R and Lefebvre E 2008 J. Stat. Mech. 10
[22] Dugué N and Perez A 2015 Directed Louvain: maximizing modularity in directed networks Ph.D. thesis
  Université d’Orléans
[23] Leicht E and Newman M 2008 Phys. Rev. Lett. 100 118703
[24] Moehlis J, Faisst H and Eckhardt B 2004 New J. Phys. 6
[25] Reddy S, Schmid P, Baggett J and Henningson D 1998 J. Fluid Mech. 365 269
[26] Jimenez J and Pinelli A 1999 J. Fluid Mech. 389 335