Data-driven classification of sheared stratified turbulence from experimental shadowgraphs

Adrien Lefauve *
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom

Miles M. P. Couchman
Department of Mathematics and Statistics, York University, Toronto, Ontario M3J 1P3, Canada

(Received 4 May 2023; accepted 26 January 2024; published 8 March 2024)

We propose a dimensionality reduction and unsupervised clustering method for the automatic classification and reduced-order modeling of density-stratified turbulence in laboratory experiments. We apply this method to 113 long shadowgraph movies collected in a "stratified inclined duct" experiment, where turbulence is generated by instabilities arising from a sheared buoyancy-driven counterflow at Reynolds numbers $\mathrm{Re} \approx 300$–5000, tilt angles $\theta = 1^\circ$–6°, and Prandtl number $\mathrm{Pr} \approx 700$. The method automatically detects edges representative of discrete density interfaces, extracts a low-dimensional vector of statistics representative of their morphology, projects these statistics onto a two-dimensional phase space of principal coordinates, and applies a clustering algorithm. Five clusters are detected and interpreted physically based on their typical interface morphology and an examination of representative frames, revealing distinct types of turbulence and mixing: laminarizing, braided, overturning, granular, and unstructured, as well as some intermediate types. The ratio of time spent in each cluster varies gradually across the (Re, θ) space. At intermediate values of $\mathrm{Re}\,\theta$, intermittent turbulence cycles between clusters in phase space and reveals at least two distinct routes to stratified turbulence. These insights demonstrate the potential of this method to reveal the underlying physics of complex turbulent systems from large experimental datasets.

DOI: 10.1103/PhysRevFluids.9.034603

I. INTRODUCTION

A. Context and motivation

Fluid flows, viewed under the dynamical systems lens, yield many examples of spatially and temporally coherent structures as the strength of nonlinearities within the flow (as quantified by the Reynolds number) increases. Examples include turbulent spots in plane Couette flow, puffs and slugs in pipe flow, bands in channel flow, vortex streets in the wake of bluff bodies, ring vortices in buoyant plumes, hairpin and horseshoe vortices in boundary layers, and billows and braids in mixing layers [1]. Flows characterized by more than one nondimensional parameter

*lefauve@damtp.cam.ac.uk

may exhibit richer dynamical behaviors, with a variety of distinct flow "states" or "regimes" emerging in different regions of their multidimensional parameter space. One salient example is the "surprisingly complex transition diagram" observed in Taylor-Couette flow, where a plethora of spatiotemporal dynamics are observed in the two-dimensional parameter space described by the Reynolds numbers of the inner and outer rotating cylinders [2]. Identifying such regimes and delineating their extent in parameter space is typically performed manually by a trained human eye based on an inspection of various properties of the observed coherent structures. In this article we demonstrate that the classification of regimes within complex flows, and the physical insight thus gained, can be enhanced by automated data-driven algorithms. We demonstrate this by revealing the rich turbulent states and transitions in an experiment, the "stratified inclined duct" (SID), which we argue represents a new paradigm for the study of turbulence, on par with the well-researched cylindrical pipe [3] or Taylor-Couette [4] experiments.

To understand the potential advantages of our data-driven approach, it is worth summarizing how such canonical flows are typically analyzed. Two broad questions are usually considered. First, how can we predict the emergence and evolution of coherent structures as nondimensional parameters are varied? This can be done from first principles (e.g., examining bifurcations of the governing equations and stability theory) or through consideration of more approximate and phenomenological models. Second, how do these coherent structures relate to flow phenomena of interest, e.g., wall drag or mixing? These two steps can thus be represented as:

$$\underset{\text{(equations)}}{\text{input parameters}} \xrightarrow{\text{step 1}} \text{regimes, coherent structures} \xrightarrow{\text{step 2}} \underset{\text{(solutions)}}{\text{useful output variables.}} \quad (1)$$

The intermediate focus on regimes and coherent structures is relevant because they generally play a leading-order role in governing the output. The coherent structure approach to modeling turbulence uses a "skeleton" of "simple invariant solutions" or "exact coherent states." Uncovering the role of such coherent structures allows us to build physical intuition in terms of mechanical processes and cause-and-effect relationships, thus bridging the gap between the governing equations and their solutions. Developing new data-driven techniques to discover distinct regimes and quantify their properties thus seems critical in deepening our understanding of turbulence.

B. Focus

We focus here on identifying distinct turbulent states in stratified shear flows, that is, turbulence energized by a mean shear between two counterflowing fluid layers having slightly different densities (satisfying the Boussinesq approximation). Coupling terms in the equations governing the evolution of the momentum and density fields mean that coherent velocity structures (e.g., shear and vortices), interact with coherent density structures (i.e., sharp density interfaces of enhanced gradient). This interaction is very complex and predictions from first principles are lacking. For instance, vortices may broaden density interfaces and/or actively sharpen them depending on the circumstances [5]. This has leading-order but poorly understood implications for the energy dissipation, the buoyancy flux across stable density interfaces and irreversible diapycnal mixing and its associated efficiency, which are variables of central importance in ocean and climate modeling [6]. In other words, progress is needed on steps 1 and 2 in expression (1).

C. Choice of problem

To improve our understanding of the regimes of sheared stratified turbulence under controlled laboratory conditions (focusing on step 1), we collected data from the SID experiment. Insightful experiments on the instability of stratified shear flow have a long history, dating back at least to the seminal papers of Reynolds in 1883 [7, Sec. 12] (who noted "it proved a very pretty experiment"), Taylor in 1927 [8], and Thorpe in 1971 [9]. The novelty of SID is that it sustains a highly dissipative two-layer exchange flow through a long tilted rectangular duct connecting two large reservoirs

of fluid at different densities. It allow us to explore regions of parameter space and record long time series of turbulent dynamics which were previously inaccessible to experiments and which remain prohibitively expensive to simulate numerically. Four regimes have been identified in SID in the two-dimensional space spanned by the Reynolds number and duct tilt angle: stable laminar flow, finite-amplitude interfacial ("Holmboe") waves, intermittent turbulence, and full turbulence. These regimes were first described in 1961 [10], before being independently rediscovered in 2014 [11]. Since then, measurements of the three-dimensional volumetric velocity and density fields [12] have allowed progress on steps 1 and 2. The regime diagrams were partially explained from first principles using energy budgets and dissipation arguments [13–15] as well as analytic theory [16], and the morphology and interaction of three-dimensional vortices with density interfaces were described and linked to a linear instability [17,18].

However, a limitation of SID research to date lies in the subjectivity with which the regimes are defined and the flows are classified. This causes at least three problems undermining both steps 1 and 2. First, this classification relies on choosing qualitative visual criteria with which to classify the flow, which are typically inconsistent between different individuals, especially in transitional regions where a flow exhibits elements of multiple regimes. Second, a classification into discrete regimes implicitly implies sharp transitions, whereas a trained eye recognizes that SID transitions smoothly between regimes (e.g., the turbulent periods and their intensity increase across the intermittent regime). Third, assigning a single regime label to the entire temporal evolution of an unsteady flow is reductive and brushes over its spatial and temporal complexity.

D. Approach and outline

This paper resolves the problems associated with the subjective nature of a human classification by applying an objective, automated, and physically interpretable classification frame by frame to a large experimental dataset of 113 shadowgraph movies. Recent work has demonstrated that density interfaces and their distribution and structure through a turbulent flow play a key role in the resulting mixing [5,19,20]. Our approach here will thus be to classify turbulence in terms of the nature of observed density interfaces using a robust processing pipeline starting from raw experimental measurements.

In Sec. II, we review the experimental dataset and its previous human classification. In Sec. III, we describe our new data-driven methodology for automatically discovering distinct turbulent clusters in a low-dimensional phase space. Our results are provided in Sec. IV. We first give a physical interpretation of the identified clusters in terms of their characteristic density interface morphology and coherent structures and highlight how they compare to the human-identified regimes. We then study the transitions of cluster prevalence in the space of input parameters, as well as different types of intermittent behaviors. Finally, we conclude in Sec. V and suggest avenues of future exploration.

II. EXPERIMENTAL DATASET AND HUMAN CLASSIFICATION

A. The SID setup

1. Principle and geometry

The SID setup is sketched in Fig. 1(a), consisting of a long rectangular duct connecting two large, closed reservoirs filled with salt (sodium chloride) solutions of different densities $\rho_0 \pm \Delta\rho/2$. The duct is $l = 2000$ mm long, $h = 50$ mm tall (streamwise aspect ratio $l/h = 40$), and $w = 100$ mm wide (spanwise aspect ratio $w/h = 2$), and each reservoir has a volume $V = 400$ L. As the gates isolating the duct from the reservoirs are opened, a two-layer exchange flow through the duct develops. The hydrostratic pressure differential in the reservoirs caused by the reduced gravity $g' = g\Delta\rho/\rho_0$ results in a (baroclinic) pressure gradient of opposite sign on either side of the neutral $\rho = \rho_0$ interface, driving the flow with a layer-averaged velocity $\pm u/4 = \sqrt{g'h}/2$ where u is the maximal peak-to-peak velocity scale. The flow can be further energized by inclining the entire apparatus by a small tilt angle $\theta > 0$. This accelerates the bottom layer of denser fluid downhill,

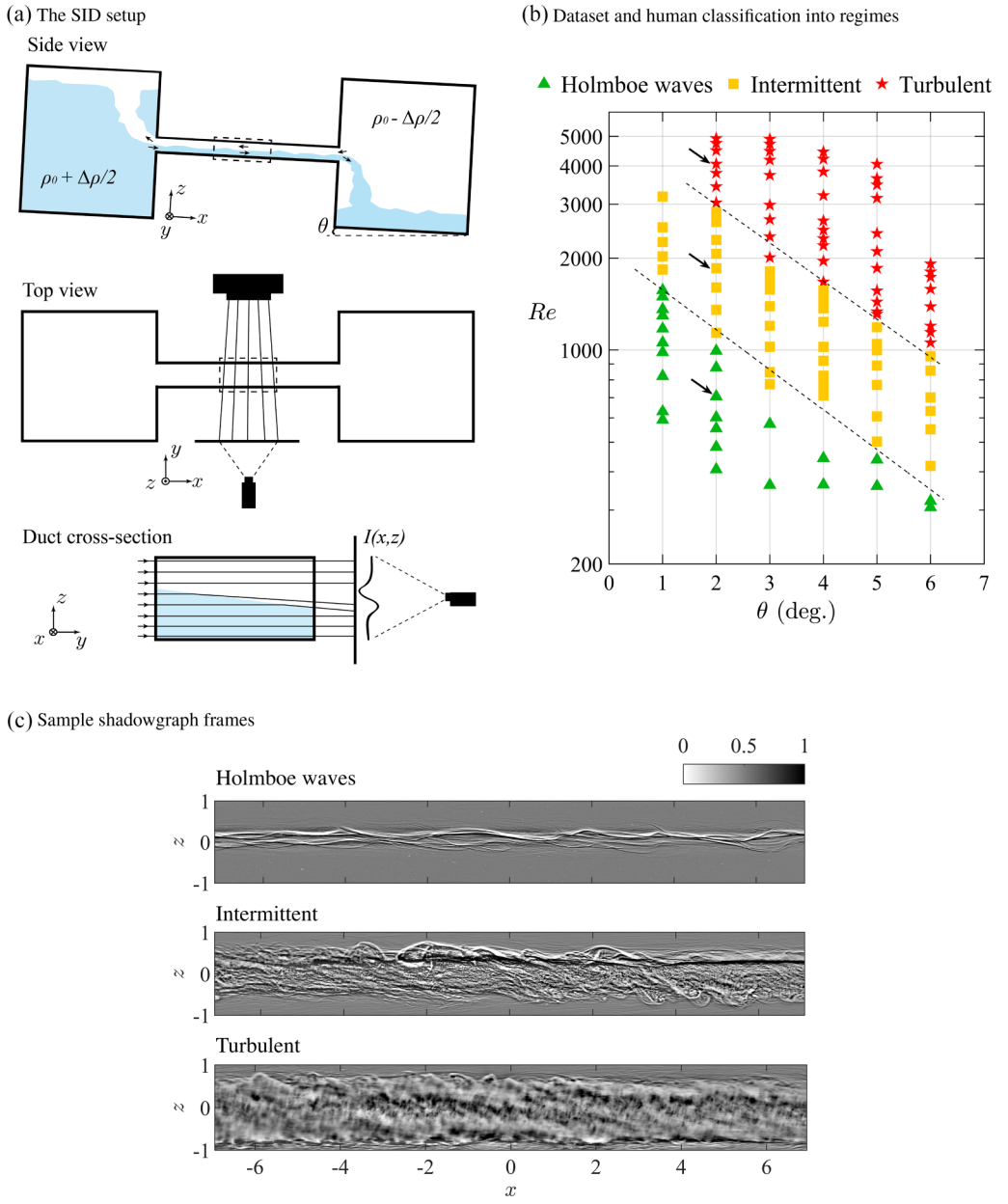

FIG. 1. Experimental data. (a) Sketch of the setup and shadowgraph measurements used to visualize the density field. (b) The 113 experiments in the space of input parameters (Re, θ) labeled with their human classified regime. The dashed lines roughly highlight regime transitions. Three arrows indicate the locations of the sample shadowgraph frames $\tilde{I}(x, z)$ presented in (c). The original images are $\approx$3400 $\times$ 450 pixels.

and the upper layer of buoyant fluid uphill, more than they would under the pressure gradient alone. We align the streamwise axis x along the duct, tilting the z axis by an angle $-\theta$ with respect to the true vertical (opposite to the direction of gravity). The apparatus used here differs from previous two generations (the first being used in Ref. [11] and the second being used in Refs. [12–18,21]) in that (i) the duct is located outside rather than inside the reservoirs resulting in cleaner visualizations

of the flow, (ii) the reservoirs are larger and closed by rigid lids (there are no free surfaces), (iii) the entire duct-reservoirs dumbbell assembly is tilted at once, and (iv) the duct connects to the reservoirs smoothly with trumpet-shaped ends (adding an extra length of 10%).

2. Dimensional analysis

We nondimensionalize lengths by $h/2$, velocities by $u/2 = \sqrt{g'h}$, and time by h/u and place the origin of the coordinate system at the center of the duct. The duct volume is thus $(x, y, z) \in [-40, 40] \times [-2, 2] \times [-1, 1]$. The three nondimensional dynamical parameters are (i) the tilt angle θ (variable from one experiment to the next), (ii) the Reynolds number $\mathrm{Re} = uh/(4\nu) = \sqrt{g'h}h/(2\nu)$ (variable through $\Delta\rho/\rho_0$ in the range $\mathrm{Re} = 300 - 5000$), and (iii) the Prandtl number $\mathrm{Pr} = \nu/\kappa = 700$ (fixed). In our experiments, we take the kinematic viscosity of water as $\nu = 1.05 \times 10^{-6}\ \mathrm{m^2\,s^{-1}}$ and the molecular diffusivity of salt as $\kappa = 1.5 \times 10^{-9}\ \mathrm{m^2\,s^{-1}}$.

3. Tilt and flow regimes

Previous SID experiments, simulations, and two-layer shallow water wave theory showed that the mean exchange flow rate through the duct was bounded by the nondimensional layer-averaged speed 0.5 as a consequence of "hydraulic control" [11,22]. This means that the additional power input caused by the tilt θ cannot be balanced by the viscous dissipation of a faster laminar flow and must instead be balanced by increasingly dissipative flow structures and interfacial mixing. This causes increasingly turbulent flow regimes with increased θ (energizing the flow) and with increased Re (reducing the weight of viscous dissipation), as demonstrated in Fig. 1(b). We return to a more detailed discussion of these regimes in Sec. II C.

4. Long time series

The out-of-equilibrium sheared stratified turbulence in SID persists until each reservoir has been filled with outflowing fluid to midlevel, which takes approximately $V/(h^2 w) = 1600$ advective time units (A.T.U.), i.e., a fixed nondimensional time set by the setup geometry. What makes SID particularly valuable is its ability to collect such long time series at high $\mathrm{Re} = O(10^3)$ and $\mathrm{Pr} = O(10^3)$, a region of parameter space exhibiting flow regimes and physics relevant to salinity- or turbidity-driven environmental and geophysical flows but currently prohibitive to direct computation.

5. Broader significance

From a dynamical systems point of view, SID is also valuable in that the stabilizing effects of density stratification give rise to a richer set of coherent structures, intermittency, and transitions at higher Re than in other canonical, unstratified flows (e.g., Ref. [23]). The discovery of new, high-Re building blocks for the skeleton of stratified turbulence is highly relevant to the broader efforts into modeling of multiphysics (e.g., rotating, multiphase, or magnetohydrodynamic) turbulence.

B. Shadowgraphs of the density field

1. Principle

The density field is visualized using a shadowgraph technique that involves shining light through the duct and measuring its refraction [Fig. 1(a)]. Approximately parallel light rays produced by a slide projector travel through the duct along the spanwise y direction and are projected onto a semitransparent screen. Any variations in the curvature (normal to the rays) of the local perturbation density field $\rho(x, y, z, t) - \rho_0$, and thus of the refractive index field $N(x, y, z, t)$, cause the rays to focus or defocus, varying the light intensity that reaches the screen. In the limit of weak variations,

the intensity of the image formed and recorded by video camera is (see, e.g., [24, Sec. 2.1])

$$I(x,z,t) = \beta\, I_0(x,z) \int_{-2}^{2} \Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \Big) \rho(x,y,z,t)\, dy. \tag{2}$$

Here β depends on $(\rho_0/N_0)\partial N/\partial\rho$ and the experimental geometry, and I_0 is the (approximately) uniform background intensity of the illumination. The shadowgraph signal I is thus particularly well suited to detecting density interfaces, the structures of interest here for distinguishing between different turbulent regimes. A sharp density gradient $\partial\rho/\partial z$ will result in an intensity I having a low (dark) and high (bright) peak on either side of it.

2. Acquisition

We carried out 113 individual experiments at six different tilt angles θ equally spaced between 1° and 6°. Each campaign was run at a fixed θ by initially filling the left reservoir with brine and the right reservoir with fresh water, resulting in a large $\Delta\rho/\rho_0$ and thus a large $\mathrm{Re} \approx$ 2000–5000. The duct was opened to start the exchange flow, and time was counted ($t = 0$) from the moment the gravity currents originating from either ends of the duct reached the center of the duct ($x = 0$). Shadowgraph movies were then recorded with a video camera tilted at the same angle θ as the setup to record natively in the (x, z) coordinate system, covering the full internal height of the duct (50 mm) and a width ≈325–425 mm (depending on the campaign), centered at $x = 0$. The frame rate was set between 5 and 100 fps depending on the speed of the flow, to achieve a typical frame spacing of 0.1 nondimensional A.T.U., and 400–600 A.T.U. were typically recorded (recalling that time is nondimensionalized in each experiment by $h/u = h/(2\sqrt{g'h})$). The experiment was then stopped by closing off the duct at both ends and mixing the fluid in both reservoirs, thus reducing $\Delta\rho/\rho_0$ and Re. The next experiment at a lower Re was then started, recorded, and so forth, until the lowest $\mathrm{Re} \approx$ 300–600. This allowed us to cover the (Re, θ) space of Fig. 1(b), consisting of 15 experiments at $\theta = 1^\circ$, 22 at $\theta = 2^\circ$, 19 at $\theta = 3^\circ$, 21 at $\theta = 4^\circ$, 20 at $\theta = 5^\circ$, and 16 at $\theta = 6^\circ$. The light intensity across all videos was normalized to yield $\tilde{I}(x,z,t)$, the initial transient ($t < 100$) discarded, and the temporal resolution subsampled by a factor of 10 to avoid excessive redundancy between subsequent frames, as further described in Appendix A 1. Three example frames are shown in Fig. 1(c). The final dataset consists of 113 movies totalling 50 155 frames, giving an average of 444 frames per experiment with a typical frame-to-frame spacing of 1 A.T.U. These data are freely available on the repository [25].

C. Human classification into flow regimes

Qualitative flow visualizations, including shadowgraph movies, have previously been used to classify SID measurements into four flow regimes: laminar (L), Holmboe waves (H), intermittent (I), and turbulent (T), as we have done for our current dataset in Fig. 1(b). Such a manual classification was first introduced by Macagno and Rouse [10] (hereafter MR61), and subsequently rediscovered (without knowledge of MR61) by Meyer and Linden [11] (hereafter ML14), using almost identical descriptions. We quote and compare their descriptions of the four qualitatively different flow regimes below:

(L) "uniform laminar motion with straight streamlines" (MR61) and "an undisturbed density interface separating the two layers" (ML14);

(H) "laminar motion with regular waves" (MR61) and "the flow is wave-dominated and exhibits Holmboe modes on the interface, with characteristic cusplike wave breaking" (ML14);

(I) "incipient turbulence, with waves which break and start to show irregularity and randomness" (MR61) and "intermittent state, which exhibits a rich range of spatiotemporal behavior and an interfacial region that contains features of Kelvin-Helmholtz-like structures and of the other two lower-dissipation states: thin interfaces and Holmboe-like structures" (ML14);

(T) "pronounced turbulence and active mixing across the interface" (MR61) and "turbulent high-dissipation interfacial region typically containing Kelvin-Helmholtz-like structures sheared in the direction of the mean shear and connecting both layers" (ML14).

Figure 1(c) shows examples of a shadowgraph frame in the H, I, and T regimes. Note that the dataset in this paper does not contain flows in the L regime (found at lower values of Re, θ than considered in Fig. 1(b) because their flat, sharp interface and steadiness render them uninteresting to our analysis. The main difference between the I and T regimes is that the latter never relaminarizes.

As explained in the Introduction, this classification of an entire movie into a single regime causes three problems: (i) arbitrariness and inconsistency, (ii) implicit assumption of sharp transitions, and (iii) neglect of spatial and temporal complexity. For example, the temporal variations between various quasilaminar wave structures and more turbulent structures is essential to the fascinating intermittent regime, which sometimes exhibits quasiperiodic laminar-turbulent cycles with a wide range of flow structures and for which a probabilistic description appears needed. Three-dimensional velocity and density experimental data have also revealed different "flavors" of turbulence, even at comparable values of turbulent kinetic energy dissipation proportional to $\mathrm{Re}\,\theta$ [13,15], with low-θ flows having more extreme enstrophy but less overturning events than high-θ flows [21]. Such valuable insight cannot easily be drawn by eye from numerous shadowgraph movies. This motivates the need for an automated classification based on physically interpretable coherent structures, which we formalize and apply next.

III. DIMENSIONALITY REDUCTION AND CLASSIFICATION

A. Overview of the method

This section introduces the automated pipeline that takes an entire dataset of shadowgraph images, collected across the two-dimensional parameter space spanned by Reynolds number Re and tilt angle θ, and determines a natural grouping of these data into distinct clusters in another low-dimensional space, where the number of clusters and their properties are initially unknown. Figure 2 summarizes our approach, and each step is detailed in the following sections.

In summary, each shadowgraph image is first transformed into a collection of binary edges delineating the locations of sharp density gradients (see Sec. III B). A variety of geometrical properties of each interface are then computed and their statistics are used to form a low-dimensional vector of characteristics (Sec. III C). A principal component analysis is then performed on the entire set of morphology vectors, demonstrating that the dominant trends in the morphology statistics data may be captured in a two-dimensional subspace (Sec. III D). The two-dimensional vectors representing all shadowgraph frames are then automatically classified (Sec. III E), revealing five clusters with distinct properties. We interpret these clusters in Sec. IV by analyzing the inverse mappings ($\mathcal{C} \to \mathcal{D}, \mathcal{E}, \mathcal{S}, \mathcal{I}$), comparing the clusters to human-identified regimes ($\mathcal{C} \to \mathcal{H}$), and studying the temporal trajectories within or between clusters. The associated codes and data can be downloaded from Ref. [26].

We note that the traditional human classification approach (sketched in gray, left-column Fig. 2) mapped a collection of frames (a movie) directly to the one-dimensional space of regimes $\mathcal{H}$, containing three possible values: Holmboe wave (H), intermittent turbulence (I), and sustained turbulence (T). By contrast, our automated classification approach provides a series of objective, repeatable mappings which remain easily interpretable. This interpretability distinguishes our approach from other data-driven approaches relying on neural networks (e.g., autoencoders).

B. Edge detection ($\mathcal{S} \to \mathcal{E}$)

1. Canny edge detection

A Canny edge-detection algorithm [27], implemented using Matlab's function `edge`, was applied to each shadowgraph frame (containing $\approx$1.5 million pixels) in order to transform the grayscale

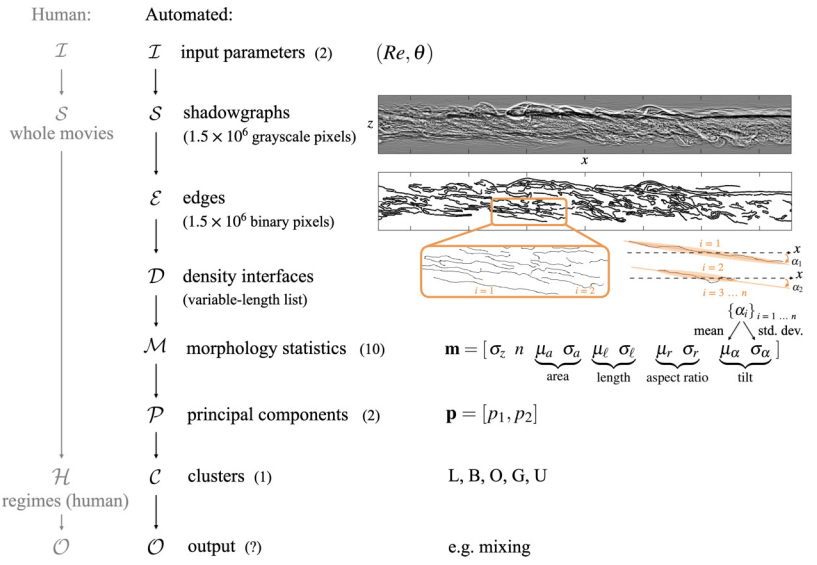

FIG. 2. Dimensionality reduction pipeline used to assign a cluster label to each shadowgraph frame. The spaces and successive mappings of the new automated classification approach (second column from left, in black) are compared to the traditional human approach (left, in gray). The bracketed numbers indicate the dimension of each space, highlighting a dramatic reduction in dimensionality.

shadowgraph image of the density field to a binary image delineating the positions of sharp density gradients. The Canny algorithm works by first lightly smoothing the grayscale image (normalized to have values between 0 and 1) with a Gaussian filter (here with an isotropic standard deviation of 5 pixels) and then numerically computing the gradient of the filtered image. In order to pick out edges, two thresholds are considered. If a pixel gradient is higher than the upper threshold (set here at 0.5), then the pixel is accepted as an edge. If a pixel gradient is below the lower threshold (set here at 0.05), then it is rejected. If a pixel gradient is between the two thresholds, then it is accepted only if it is connected to a pixel that is above the upper threshold (an edge). This double thresholding improves the detection of true weak edges (avoiding true negatives) and robustness to noise (avoiding false positives). Identical thresholds were used for all frames, and the detected contours were only weakly dependent on the thresholds used. Only the strongest gradients are detected, having excellent signal-to-noise ratio, ensuring the robustness of this method. The optimal thresholds are relatively easy to set: If the threshold is set too low, then the entire frame becomes covered in spurious, noisy edges, whereas if the threshold is set too high, then no edges are detected. Within the optimal range, changing the thresholds by a factor of 2 yielded less than a 1% average change in the edge statistics [Eq. (3)] subsequently computed and used for our analysis.

2. Results

Figure 3 shows examples of the binary edge images $e(x, z) \in \mathcal{E}$ corresponding to the shadowgraphs $\tilde{I}(x, z) \in \mathcal{S}$ of Fig. 1(c). Typically pixel thin, the edges are rendered here with much thicker black contours for better visualization. The Holmboe (H) flow (top) exhibits a relatively sharp but undulating density interface with cusped waves, and typically multiple edges, or filaments, stacked on top of one another. The thicker layer of intermediate density of the intermittent (I) flow (middle) results in more numerous edges, some of which are shorter along x and more tilted with respect to the x axis. The even thicker intermediate layer in the turbulent (T) flow is so turbulent that few edges are detected within it (at low $|z|$), as the three-dimensional nature of the turbulence blurs the resulting shadowgraph image, but some edges are detected at the upper and lower edges of the frame

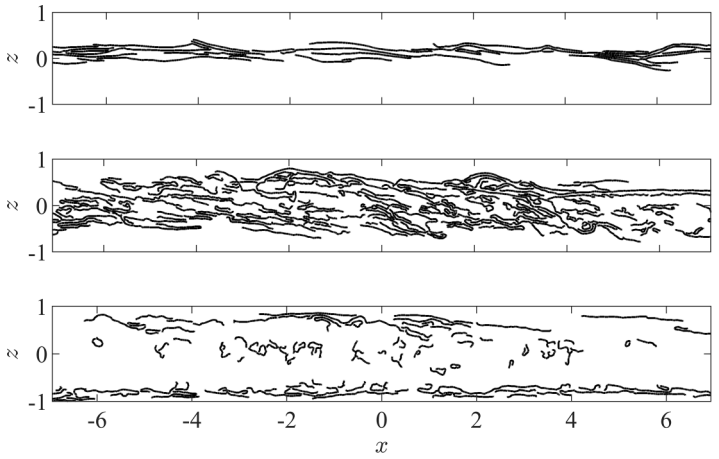

FIG. 3. Sample binary edge frames ($\mathcal{E}$ space), delineating the locations of sharp density interfaces (black), mapped from the shadowgraphs ($\mathcal{S}$ space) of Fig. 1(c) in the three human-classified regimes: Holmboe (H, top), intermittent (I, middle), and turbulent (T, bottom). Edges are here rendered much thicker than detected for better visualization.

($|z| \approx 1$). The edges within the intermediate layer tend to be relatively short and tilted away from the horizontal, revealing instability and overturning motions, whereas the edges on either side of it are longer and flatter, revealing more quiescent stable interfaces. This is consistent with gradient Richardson number profiles $Ri_g(z)$ [13, Fig. 4] (quantifying the competition between destabilizing shear and stabilizing stratification); low $Ri_g \approx 0.15$ were found throughout the turbulent layer while higher $Ri_g > 1$ were found at the interfaces on either side.

C. Discrete density interfaces and morphology statistics ($\mathcal{E} \to \mathcal{D} \to \mathcal{M}$)

1. Connected components

Having generated a binary image of sharp density interfaces, we now wish to quantify the morphological properties of each connected edge. This is achieved by first applying a connected component algorithm to the binary image, using Matlab's function `bwconncomp`. A connected component is defined here as a set of pixels that are connected on any of their four sides or four corners (often referred to as a two-dimensional pixel connectivity of eight). In other words, two adjoining pixels are considered part of the same density interface if they are connected along the horizontal, vertical, or diagonal direction. This algorithm returns n discrete density interfaces for a given frame, characterized by a list of pixels belonging to each.

2. Edge properties

We then compute the morphology of each density interface within a given frame using Matlab's function `regionprops` with the following four arguments: "Area," returning the actual number of pixels in the interface; "MajorAxisLength" and "MinorAxisLength," returning the length (in pixels) of the major and minor axes, respectively, of the ellipse that has the same normalized second central moments as the interface; and "Orientation," returning the angle between the x axis and the major axis of this ellipse (with positive angles denoting anticlockwise rotations). The areas and lengths are then converted from pixels to nondimensional units (like the axes of Fig. 3) so that all frames can be compared consistently with meaningful physical values. This yields, for each frame, a list of n nondimensional density interface areas $\{a_i\}_{i=1,\ldots,n}$, lengths $\{\ell_i\}$ (from the major axis), aspect ratio $\{r_i\}$ (ratio of major to minor axes), and tilts $\{\alpha_i\}$ (0 is aligned with $\hat{x}$, and $\pm 90^\circ$

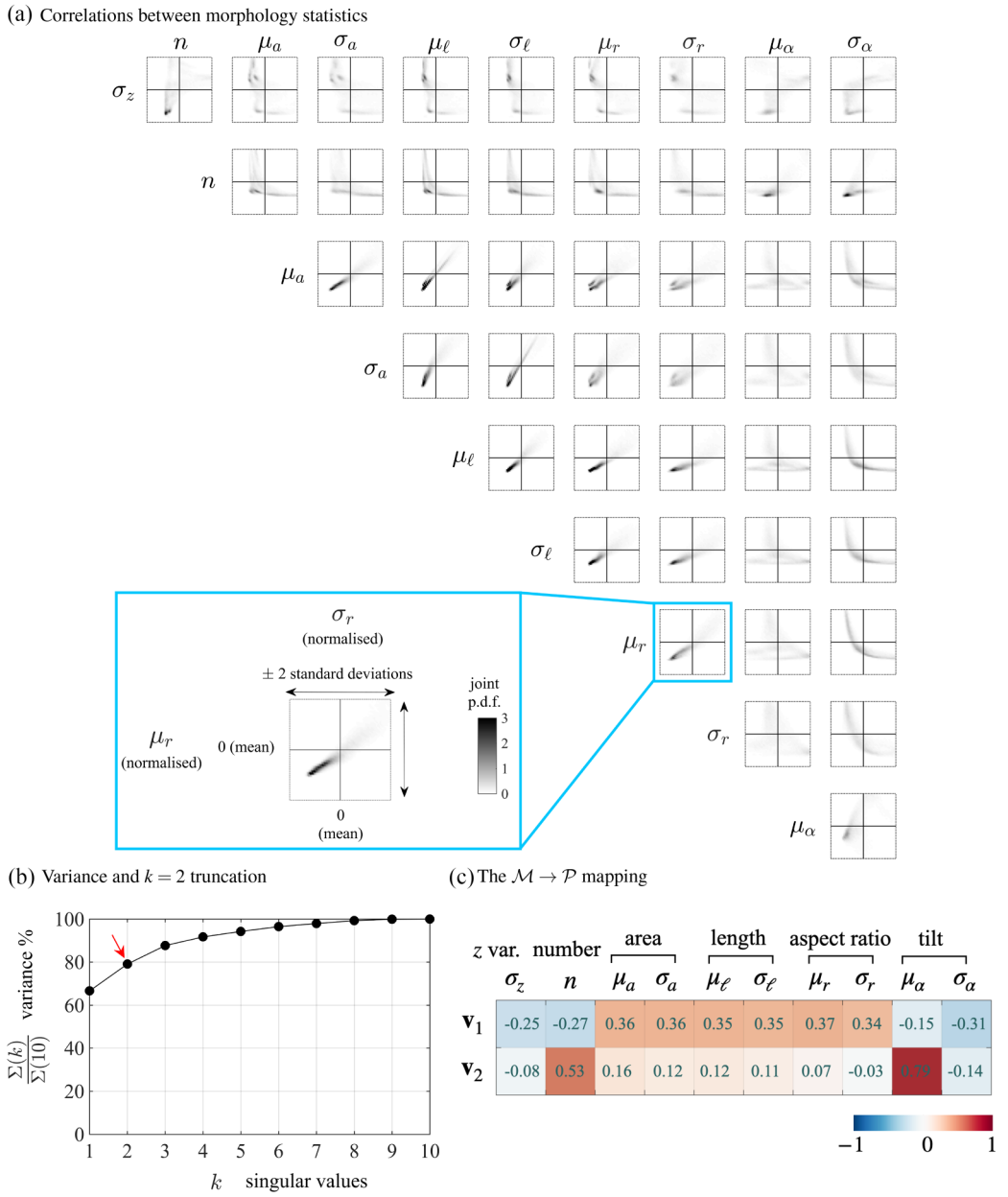

FIG. 4. (a) Joint probability density functions (p.d.fs) of the 10 statistics characterizing the morphology of density interfaces [see Eq. (3)]. The example inset (in blue) shows correlations between μ_r (the mean aspect ratio of density interfaces) and σ_r (the variability in their aspect ratio). (b) Cumulative variance captured by the principal components (or singular values); the red arrow highlights that $k = 2$ captures 79% of the total. (c) Coordinates of the first two principal unit vectors $\mathbf{v}$ in terms of the original basis of morphology statistics.

are aligned with $\pm\hat{z}$ respectively). These lists can be seen as belonging to the space $\mathcal{D}$, describing the morphological properties of the edges within each frame, which can be conveniently visualized by plotting histograms of a, ℓ, r, α (which we will show in Sec. IV B).

3. Morphology vector

Finally, in order to distil the properties of the multiple edges within a frame into a single vector, we compute the mean (μ) and standard deviation (σ) of the distributions $\{a_i\}$, $\{\ell_i\}$, $\{r_i\}$, $\{\alpha_i\}$, characterizing the center of mass and moment of inertia of their histograms. We verified that higher-order moments (e.g., the skewness and kurtosis) are not needed as they add very little additional signal. To this eight-dimensional vector, we added two further components: the number of interfaces n, and the moment of inertia of the edges with respect to the vertical coordinate $\sigma_z = \sqrt{\langle \langle e \rangle_x (z - \langle e \rangle_{x,z})^2 \rangle_z}$, where larger values indicate a greater spread of density interfaces around the mean position. This yielded the following 10-dimensional vector of morphology statistics for each frame:

$$\mathbf{m} = [\,\sigma_z \;\; n \;\; \underbrace{\mu_a \;\; \sigma_a}_{\text{area}} \;\; \underbrace{\mu_\ell \;\; \sigma_\ell}_{\text{length}} \;\; \underbrace{\mu_r \;\; \sigma_r}_{\text{aspect ratio}} \;\; \underbrace{\mu_\alpha \;\; \sigma_\alpha}_{\text{tilt}}\,] \;\in\; \mathcal{M}. \tag{3}$$

The 50 155 row vectors generated from all frames are then arranged into a single tall, skinny matrix denoted $\mathbf{M} \in \mathbb{R}^{50155\times 10}$.

D. Principal component analysis ($\mathcal{M} \to \mathcal{P}$)

1. Motivation

Before clustering, it is worth exploiting potential correlations between the morphological features within $\mathcal{M}$ (see definition (3)) by performing a principal component analysis (PCA) [28]. Our goal is to further reduce the dimensionality of the data to assist with the interpretation of the clustering results and to improve the effectiveness of the clustering algorithm due to the "curse of dimensionality," i.e., the dramatic increase in the volume of the clustering space with increasing dimensions.

2. Correlations and rank-two approximation

We start by normalizing each column of $\mathbf{M}$ to have zero mean and unit standard deviation and obtain $\tilde{\mathbf{M}}$. Figure 4(a) shows the correlation between all $(10 \times 9)/2 = 45$ pairs of variables, highlighting the structure of the covariance matrix $\tilde{\mathbf{M}}^T\tilde{\mathbf{M}}$. We find many direct and inverse correlations, motivating the need to express these data in a basis of linearly uncorrelated variables, called principal components. We perform the singular value decomposition (SVD) of $\tilde{\mathbf{M}}$ (for more details, see Appendix A 2) and plot in Fig. 4(b) the cumulative variance $\Sigma(k) = \sum_{l=1}^{k} \sigma_l^2$ captured by the first k singular values. We find that the first $k = 2$ principal components capture 79% of the total variance, and given the convenience of visualizing a two-dimensional phase space, we choose to truncate the SVD at $k = 2$ to obtain a new rank-two data matrix $\mathbf{P} \in \mathbb{R}^{50155\times 2}$ containing two-dimensional vectors $\mathbf{p} \in \mathcal{P}$.

3. Principal component space

Figure 4(c) plots the coordinates of the two new orthogonal basis vectors $\mathbf{v}_1, \mathbf{v}_2$ spanning $\mathcal{P}$ (pointing in the directions of maximal variance) in the previous basis spanning $\mathcal{M}$. They explicitly show the linear combination of morphology statistics making up each principal component, thereby defining the $\mathcal{M} \to \mathcal{P}$ map. Simply put, principal vector $\mathbf{v}_1$ weighs almost equally the area a, length ℓ, and aspect ratio s with positive values ($\approx$0.35, in orange) and the remaining statistics with negative values (≈ -0.3, in light blue), except the mean tilt μ_α, which is weaker. Principal vector $\mathbf{v}_2$, on the other hand, weighs much more heavily this mean tilt and the number of interfaces n (in red). Physically, the first principal vector $\mathbf{v}_1$ may thus be thought of encoding how "laminar" the flow is, with a large component P_1 indicating long, horizontal, undisturbed filaments characteristic of laminar shear layers, which would get broken up in more turbulent flows leading to smaller values of P_1. The second principal vector $\mathbf{v}_2$ may be thought of encoding a measure of instability in the flow, strongly encoding the tilt angle and number of interfaces, with a large component P_2 indicating

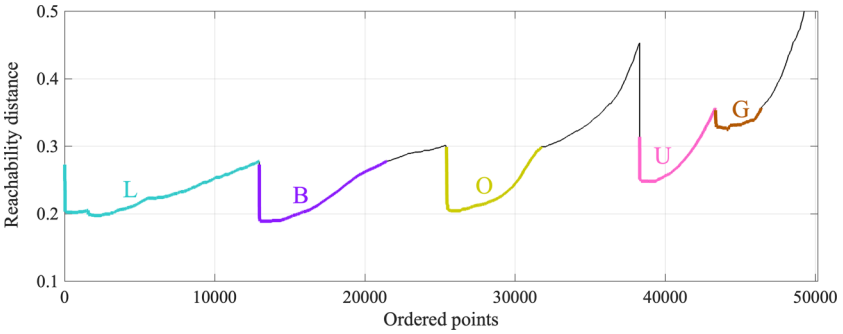

FIG. 5. The five distinct clusters (colored, labeled L-G) detected by the OPTICS algorithm, corresponding to local minima in the reachability distance (defined in Appendix A 3). Unclustered data are colored black. The distance monotonically increases above the y-axis limit of 0.5 for the last $\approx$1000 unclustered points.

unstable motions and broken-up interfaces (high n). We also note that given the high degree of correlation among area, length, and aspect ratio (encoding the level of laminarity), deleting some of these metrics from the morphology vectors (3) would still result in similar clustering results. By contrast, deleting tilt angle α would yield different results given that it does not contain any redundant partner features. More advanced algorithms, such as sparsity-promoting PCA [29], could also have been used at this stage, although the relatively simple PCA used here already yields sufficient compression and physical insight.

E. Clustering ($\mathcal{P} \to \mathcal{C}$)

1. Algorithm

Having reduced the data within each frame of the density field to a point in the two-dimensional space $\mathcal{P}$, we now perform a clustering analysis on the set of all frames. We choose to use the density-based clustering algorithm "Ordering Points To Identify the Clustering Structure" (OPTICS) [30], recently applied to the discovery of distinct regions of turbulent mixing within an ocean microstructure dataset [31]. OPTICS has three main advantages over other algorithms: it determines a natural hierarchical clustering of the data automatically, meaning that the user does not have to specify the number of clusters to be detected in advance, it identifies clusters of arbitrary shape and density, and it is robust to noise and outliers [32]. OPTICS uses a metric called the "reachability distance" d_{reach} to compute the pairwise distances between the rows of $\mathbf{P}$ (for more details, see Appendix A 3). Small values of d_{reach} indicate a high local density of frames in $\mathcal{P}$.

2. Results

The output reachability plot is shown in Fig. 5. It is generated sequentially from left to right: starting from an arbitrary row in $\mathbf{P}$, at each step OPTICS moves to the next closest point based on the reachability distance, and plots d_{reach} at that step. OPTICS thus steps toward a region with a high local density, as reflected by a steady decrease in d_{reach}. Having visited each point in this dense region, it then automatically moves to the next closest points in sparser intervening regions, reflected by a larger d_{reach}, before entering a new locally dense region (if one exists), etc. Different arbitrary starting points may cause the order of valleys (clusters) in the reachability plot to change along the x axis but does not affect the actual clustering results.

The five valleys in Fig. 5 reveal five clusters, with lower minima indicating denser clusters. Using this reachability plot, we manually split and label the five clusters L, B, O, G, U, excluding some local maxima between clusters which we consider unclustered. Overall, 72% of all points belong to clusters (26% in L, 17 % in B, 13% in O, 6% in G, and 10% in U) and 28% are unclustered. The reachability plot also highlights a hierarchical clustering: on a broader level, the dataset is split

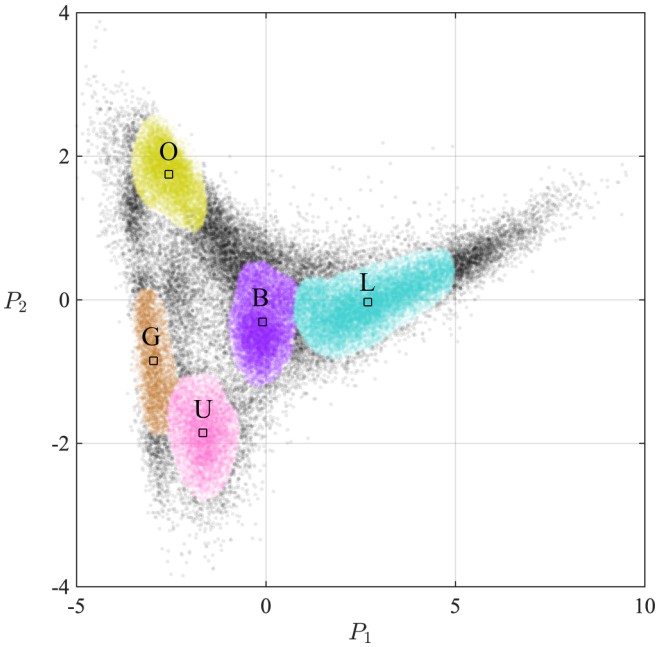

FIG. 6. Clusters (as identified and colored in Fig. 5) plotted in the space of the two dominant principal components of the density interface morphology statistics. The relation between the two principal component coordinates and the 10 morphology statistics is illustrated in Fig. 4(c).

into two larger basins (one containing L, B, O and the other containing U, G), suggesting broader similarities between the density structures within these basins, which then distinguish themselves on a finer level. The physical interpretation of these clusters is given in the next section, including a direct comparison of our clustering results to the previous human classification of this data (see Sec. IV D).

IV. RESULTS AND PHYSICAL INTERPRETATION

Having identified five clusters, we now interpret their properties by first analyzing their distribution in the clustering space, and then working back to the original shadowgraphs. We then use this insight to build a picture of the dominant dynamics across the parameter space of SID.

A. In terms of principal components ($\mathcal{C} \rightarrow \mathcal{P}$)

Figure 6 shows the location of the five clusters L, B, O, U, G in the subspace of the first two principal components. Each translucent circle represents one of the 50 155 shadowgraph frames of the density field as is colored according to its assigned cluster based on the reachability plot in Fig. 5. Darker shading indicates a greater local density of points. Black circles denote the 28% unclustered frames which did not form regions of sufficient local density to be considered part of a cluster.

The data organize into a rough (nonconvex) triangle. As the data were originally normalized using a z score (see Sec. III D), the origin corresponds to the mean of the data, and P_1 and P_2 measure the number of standard deviations away from the mean in the directions of the respective PCA vectors. The vast majority of points belong to the rectangle $[-5, 10] \times [-4, 4]$, and all five clusters belong to the smaller rectangle $[-4, 5] \times [-3, 3]$. The five black square symbols denote the centroid of each cluster, which we interpret next by projecting back to the morphology space $\mathcal{M}$.

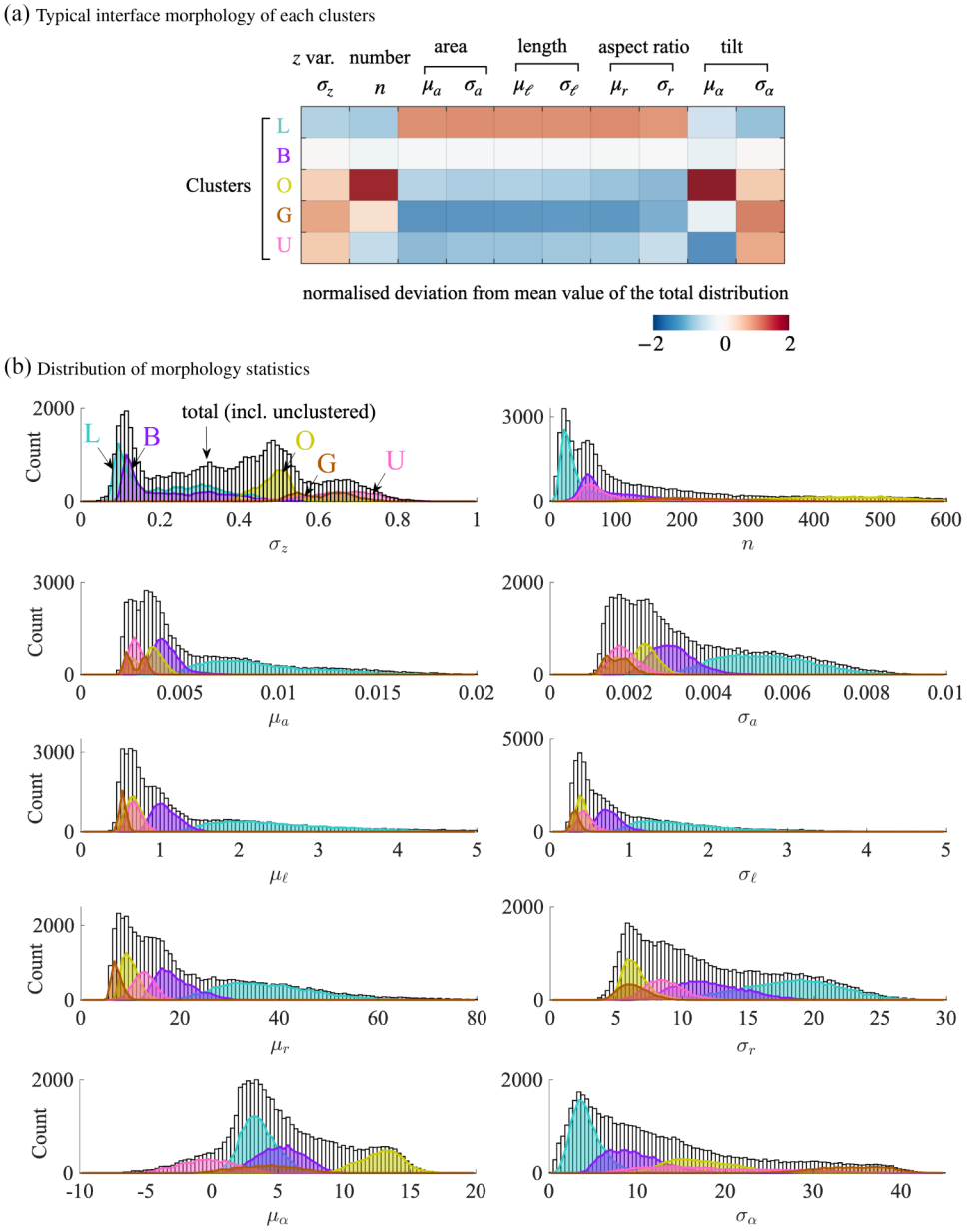

FIG. 7. (a) Density morphology properties [vector **m**, see Eq. (3)] of cluster centroids (black squares, Fig. 6), expressed as normalized deviations from the total distributions. (b) Histograms of the distributions and the contribution of each cluster in physical, nondimensional space $\mathcal{M}$.

B. In terms of density interfaces morphology ($\mathcal{C} \to \mathcal{M}$)

Figure 7(a) illustrates the values of the original 10 *normalized* morphology statistics corresponding to the cluster centroids plotted in Fig. 6. For each cluster, white denotes a property equal to the average of the total distribution of 50 155 points, whereas blue and red colors denote properties below and above the average, respectively. Figure 7(b) shows histograms of the underlying distribution of *physical (non-normalized)* morphology statistics $\mathbf{m} \in \mathcal{M}$ [see Eq. (3)]. White bars show

the total distribution and colors show the contribution of each cluster. From a combined analysis of Figs. 7(a) and 7(b) we deduce the following typical descriptions of each cluster. These descriptions will be illustrated with representative shadowgraph images from each cluster in Fig. 8 (discussed in Sec. IV C).

(L:) The density interfaces are typically concentrated in a single set (i.e., a unimodal distribution of edges having low σ_z) are scarce (low $n \approx 10$–30) but have a relatively large area (high μ_a) and are long (high $\mu_\ell \approx 1.5$–3), slender (high $\mu_r \approx 30$–45), and flat (low tilt $\mu_\alpha \approx 0°$–$6°$). They show significant spatial variability (in any instantaneous frame) in their area, length, aspect ratio (high $\sigma_a, \sigma_\ell, \sigma_r$) but not in their tilt (low σ_α). These properties suggest relatively stable and laminar-like flow snapshots, as will be illustrated by a typical image in Fig. 8.

(B:) The density interfaces have the most average properties among all clusters: average vertical spread σ_z, number ($n \approx 30$–150), length, aspect ratio, and tilt, both in mean values and spatial variability (e.g., mean $\mu_\alpha \approx 2°$–$10°$ and standard deviation $\sigma_\alpha \approx 5°$–$15°$). These properties suggest less stable, intermediate snapshots.

(O:) The density interfaces are fairly spread out in z ($\sigma_z \approx 0.4$–0.6), and are by far the most numerous ($n \approx 300$–600) and the most tilted ($\mu_\alpha \approx 9°$–$16°$, $\sigma_\alpha \approx 10°$–$25°$), but they are short ($\mu_\ell \approx 0.5$–$1$) and thick ($\mu_r \approx 5$–$15$). These properties suggest very unstable snapshots which feature a large number of distinct density interfaces.

(G:) The density interfaces are greatly spread out in z ($\sigma_z \approx 0.5$–0.8) and have strong variability in tilt ($\sigma_\alpha \approx 30°$–$40°$), but they are the smallest, shortest ($\mu_\ell \approx 0.5$), and thickest ($\mu_r \approx 5$–10). They are about average in number ($n \approx 100$–300) and in mean tilt ($\mu_\alpha \approx 0°$–$10°$). These properties suggest turbulent mixing across a thicker layer than in cluster O but with fewer and more stable detectable interfaces, perhaps due to weaker density gradients.

(U:) The density interfaces resemble those in cluster G but are much less numerous ($n \approx 40$–120) and flatter, with a mean tilt that is (uniquely) frequently negative ($\mu_\alpha \approx -5°$ to $5°$) and very variable ($\sigma_\alpha \approx 7°$–$30°$). These properties suggest fewer distinct density interfaces, perhaps comprising primarily the strongest (and therefore most stable) density gradients on either side of the mixing layer, together with a few weaker and more three-dimensional, small-scale gradients within it. Once integrated across the entire spanwise direction in Eq. (2), they significantly blur the mixing layer, resulting in few detected edges.

The level of detail in which each cluster may be interpreted in terms of the morphological properties of its density interfaces is significantly more informative than the prior qualitative descriptions of distinct flow regimes chosen by the human eye. These quantitative features appear to suggest increasing levels of turbulence from cluster L to U, perhaps most clearly shown by the histogram of σ_z showing the increasing spread of density interfaces along z from L to U. In the next section we confirm and illustrate these findings using the original shadowgraph and edge images.

C. In terms of edges and shadowgraphs ($\mathcal{C} \to \mathcal{E} \to \mathcal{S}$)

Figure 8 shows examples of shadowgraphs $\in \mathcal{S}$ and edges $\in \mathcal{E}$, consisting of the frames nearest to each of the five cluster centroids (labelled L to U in Fig. 6) and of five additional unclustered frames chosen to span the sparser, intervening unclustered regions (LB, BO, OG, BG, and LU).

Frame L illustrates the morphological description of the L cluster given in the previous section. However, it does not intuitively qualify as the most laminar flow. Its parameters ($\theta = 3°$ and $\mathrm{Re} = 1809$) place it instead on the upper end of the intermittent regime [see Fig. 1(b)]. This frame thus captures a relaminarization phase, as the shadowgraph still exhibits small-scale structure from a past turbulent phase, which is (only partially) mirrored in the edge image by the stacking of multiple flat and stable density interfaces. This frame represents what we may call *laminarizing turbulence*. As a reference point, a stable laminar flow with a density profile having a single flat, sharp (pixel-thin) interface centered at $z = 0$ would give $\mathbf{m} \approx [0,\ 1,\ 0.0056,\ 0,\ 16,\ 0,\ 3500,\ 0,\ 0,\ 0] \Rightarrow \mathbf{p} \approx [1287,\ 248]$, i.e., very far off the top right vertex of the triangle in the (P_1, P_2) plot.

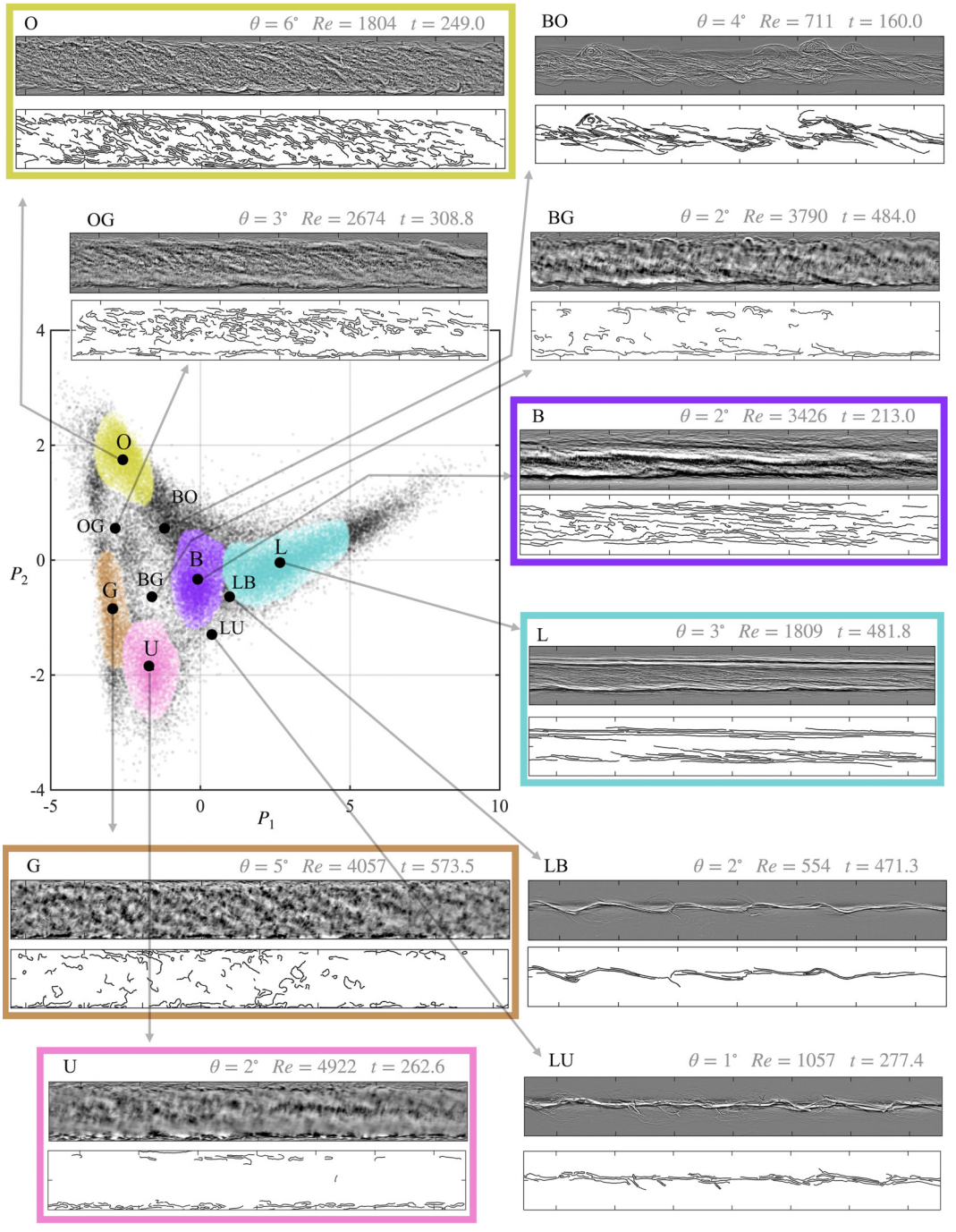

FIG. 8. Shadowgraph and corresponding edge images for the centroid of each cluster (colored) and additional examples of frames between clusters. The input parameters (Re, θ) and time t are indicated in the top-right corner of each frame. Frames are scaled to have equal height, resulting in a range of lengths.

Frames LB and LU intuitively qualify as the least turbulent—despite their proximity to more turbulent clusters B and U—and illustrate the well-known Holmboe wave regime. The Holmboe instability is caused by the interaction of two vorticity waves on either side of a broad shear layer with an internal gravity wave on the sharper density interface at the center of the shear layer [33]. This linear instability saturates nonlinearly at finite amplitude, giving a pair of counter-propagating modes with a distinctive cusped shape, which persists for arbitrarily long run times. At these amplitudes only minimal interfacial mixing takes place, and hence these frames represent the "lower end" of what we may call *Holmboe wave turbulence*.

Frame B illustrates what we may call *braided turbulence* owing to its pair of stacked, flat, central "braids" with a strong density curvature evidenced by the strong contrast in shadowgraph intensity (black and white shades). These two strong interfaces define an essentially three-layer density profile. Weaker contrasts (shades of gray) reveal a number of weaker interfaces, which do not extend all the way to the duct walls.

Frame O illustrates *overturning turbulence* owing to numerous short, unstable interfaces spanning most of the duct height. Most "wisps" correspond to weak density curvature and contrasts in the shadowgraphs, which sets it apart from braided turbulence. Frame BO illustrates a transient roll up and mixing of two main braidlike interfaces (as in B) creating more tilted wisps (as in O).

Frames G and U illustrate what we may call *granular turbulence* and *unstructured turbulence*, respectively. Both have strong three-dimensional density curvature spanning the duct height. We recall from the reachability plot (Fig. 5) and the (P_1, P_2) plot (Fig. 6) that clusters G and U are adjacent and might in fact be nearly considered to be one larger single cluster, as these frames confirm. However, frame G has slightly stronger contrast and hence a larger number of detected edges, especially at mid-depth, some of which are nearly circular and give an overall granular appearance. The interfacial turbulence in frame U has much less structure, presumably because of higher three-dimensionality and dynamically active lengthscales. The edges in U tend to be flatter than in G and located almost exclusively near the top and bottom walls, where the thick intermediate layer of mixed fluid meets boundary layers of relatively laminar and unmixed fluid. The eye appears to detect in U a pair of two turbulent braids in lighter shades of gray sandwiching a slightly darker region, but this large scale pattern is too weak to be detected by the edges algorithm.

Frame OG illustrates an intermediate stage between overturning and fully granular turbulence, featuring a mix of both types. We note that the set of input parameters (Re, θ) would not allow us to classify, *a priori*, the turbulence in OG as intermediate between O and G. Similarly, frame BG illustrates how braids (B) grow thicker and more blurred, as an intermediate stage before granularity (G). A frame BU halfway between clusters B and U would likely look similar.

The data-driven discovery of these different types of turbulence constitute the first key finding of this paper. Future improvements may consider classifying, within each frame, various "dynamically distinct regions" in the spirit of Ref. [34] (who used the density gradient field). Unsupervised neural networks have also been used to detect the laminar-turbulent boundary in transitional boundary layers [35] (using the velocity field), while others subdivided the domain and applied clustering to identify "the regions containing streaks, turbulent spots [...] and developed turbulence" [36].

D. In terms of human classified regimes ($\mathcal{C} \rightarrow \mathcal{H}$)

Figure 9 shows the overlap in the two-dimensional principal component space $\mathcal{P}$ between the human-classified regimes H, I, T $\in \mathcal{H}$ (see Sec. II C) and the clusters L-U $\in \mathcal{C}$ obtained using our data-driven technique. All frames belonging to movies labeled "Holmboe" are colored in green (left panel) and similarly for movies labeled "Intermittent" (middle panel, in yellow), and "Turbulent" (right panel, in red). The insets in each panel show the fraction of frames belonging to each cluster.

The Holmboe (H) regime spans only clusters L and B, with a few unclustered frames located either around the top right vertex of the triangle towards stable laminar flow or outside the lower

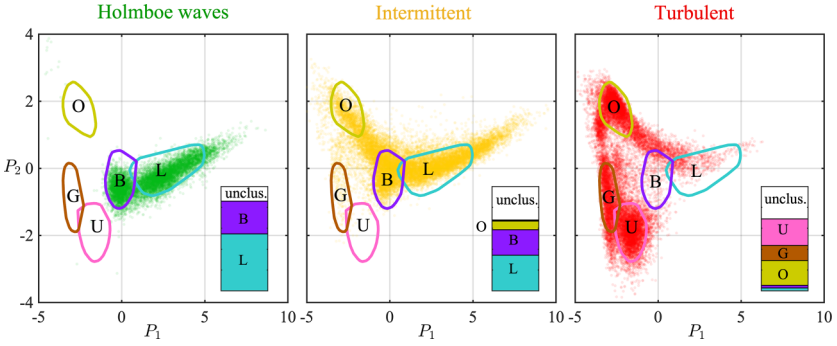

FIG. 9. Comparison between the location of clusters in principal component space (approximated by their convex hull perimeter) and the three human-classified regimes (semitransparent colored circles): Holmboe wave (left), Intermittent (center) and Turbulent (right). Insets denote the fraction of frames in each cluster.

boundary of L and B (see frames LB and LE in Fig. 8). This means that waves and turbulence that are distinctively of Holmboe type display a range of density interfaces akin either to a slightly perturbed laminar flow, laminarizing turbulence (in L), or braided turbulence (in B).

The intermittent (I) regime spans clusters L and B in approximately equal measure, although slightly more towards their upper boundary (contrary to the H regime), as well as, to a smaller extent, cluster O. It also contains a much larger proportion of unclustered frames, which are located around the top right vertex and between B and O. This means that flows classified as intermittently turbulent display a smooth continuum of most of the features we would expect: stable laminar flow, as well as overturning, braided, and laminarizing turbulence.

The turbulent (T) regime spans clusters O, G, U in approximately equal measure, with the fourth quarter of frames being unclustered. This means that flows classified as fully turbulent feature more overturning than any other flows and that they are unique in displaying granular and unstructured turbulence. Moreover, the T regime has the widest distribution of frames across $\mathcal{P}$, covering all clusters to some extent as well as most of the unclustered regions. This suggests that the variability (in time and across $\mathcal{I}$) of density interfaces is richer in the T than in the I regime.

E. In terms of input parameters ($\mathcal{C} \to \mathcal{I}$)

We now study the distribution of time spent in each cluster in Fig. 10, shown by a horizontal colored bar for each of the 113 experiments at the point (Re, θ) at which they were run. Complementary p.d.f.s corresponding to the entire data binned in Re and θ are shown in the top and right columns of the plot, respectively.

Moving from left to right (increasing θ) at intermediate Re $\approx$ 1500–3000, we find almost exclusively braided turbulence (B) at $\theta = 1^\circ$, which gradually gives way to more (and eventually exclusively) overturning turbulence (O) at $\theta = 6^\circ$. Moving from the bottom to top (increasing Re) at intermediate $\theta \approx 2^\circ$–5°, we find a gradual decrease in laminarizing turbulence (L), followed first by an increase in braided turbulence (B) at Re $\approx$ 500 and then in overturning turbulence (O) at Re $\approx$ 1000 (as well as intermediate, unclassified types). We then find a decrease in unclassified turbulence and an increase in fully three-dimensional turbulence, first in cluster G at Re $\approx$ 2500, and eventually in cluster U at Re $\gtrsim$ 3500. The binned p.d.f.s highlight that this change of type of turbulence across $\mathcal{I}$ occurs through gradual shifts in the time spent in the respective clusters, as Re and θ are varied.

This automated, quantitative characterization of the parameter space of high-Re stratified turbulence constitutes the second key finding of this paper. Future improvements should note that our results have limited precision since the experimental dataset does not uniformly sample the entire (Re, θ) plane with fine resolution. Further, although our time series are relatively long (on average

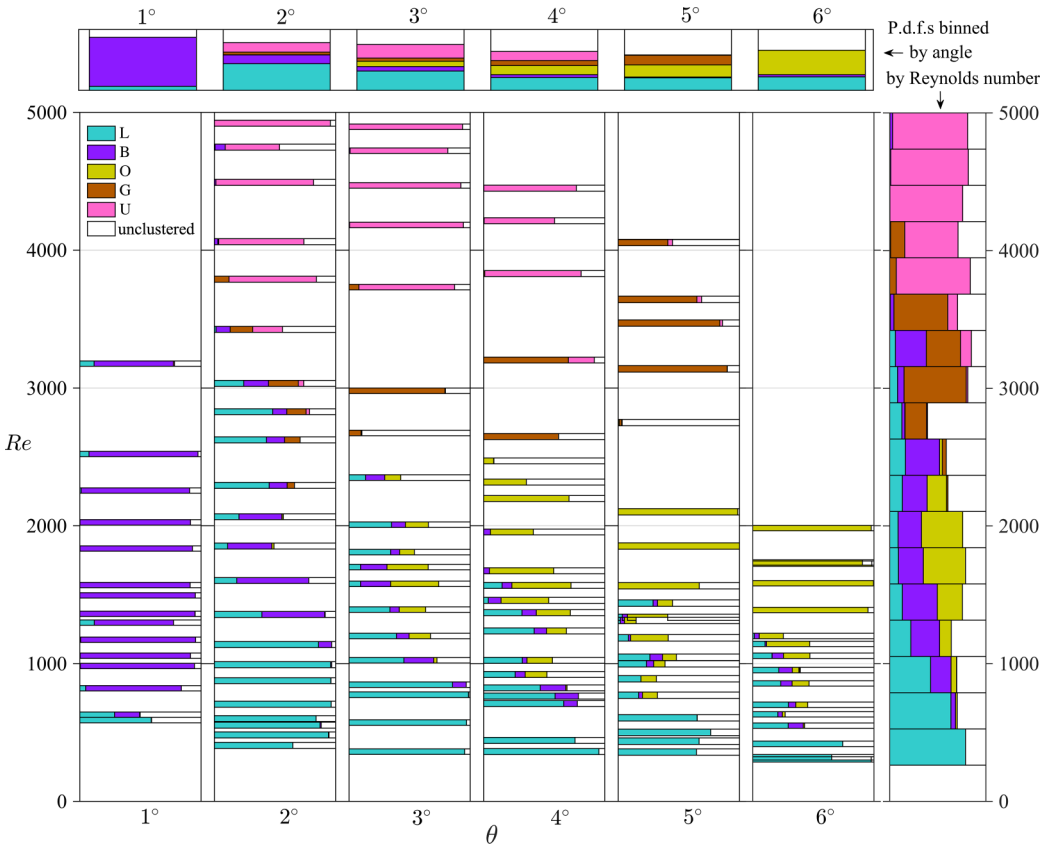

FIG. 10. Ratio of frames (time spent) in each cluster for each of the 113 experiments organized in parameter space (Re, θ). Some overlap occurs between bars at neighboring Re values. The top row and right column show overall p.d.f.s. The full time series of 21 representative experiments are shown in Fig. 11.

444 A.T.U.), even longer time series would help improve convergence towards the full underlying dynamical picture.

F. Temporal intermittency and transitions

Figure 11 examines the temporal dynamics of 21 experiments selected across the input space $\mathcal{I}$ (see Fig. 10), with increasing Re from bottom to top, and increasing θ from left to right. Each panel shows the frame-by-frame trajectories in phase space $\mathcal{P} = (P_1, P_2)$ using translucent black symbols and each vertical bar shows the corresponding time series for which cluster each frame belongs to (for clarity, unclustered frames are colored based on the cluster having the nearest centroid). Each row shows a set of three experiments with approximately matched $\mathrm{Re}\,\theta$, decreasing from $\approx 1 \times 10^4$ (top row) to $\approx 2 \times 10^3$ (bottom row). Previous SID theory, experiments [13,15] and direct numerical simulations [37] showed that the product $\mathrm{Re}\,\theta$ is proportional to the dynamic range of stratified turbulence, quantified by, the buoyancy Reynolds number $\mathrm{Re}_b = (L_O/L_K)^{4/3}$. This parameter Re_b measures the separation between the Ozmidov $L_O = (\epsilon/N^3)^{1/2}$ and Kolmogorov $L_K = (\nu^3/\epsilon)^{1/4}$ turbulent lengthscales, where ϵ is the averaged turbulent kinetic energy dissipation and N is the averaged buoyancy frequency [15, Sec. 5.1].

Tracking the most dissipative turbulence along the top row of Fig. 11 (highest $\mathrm{Re}\,\theta$), we find different types of turbulence, each of which is tightly grouped in phase space. At the lower angle

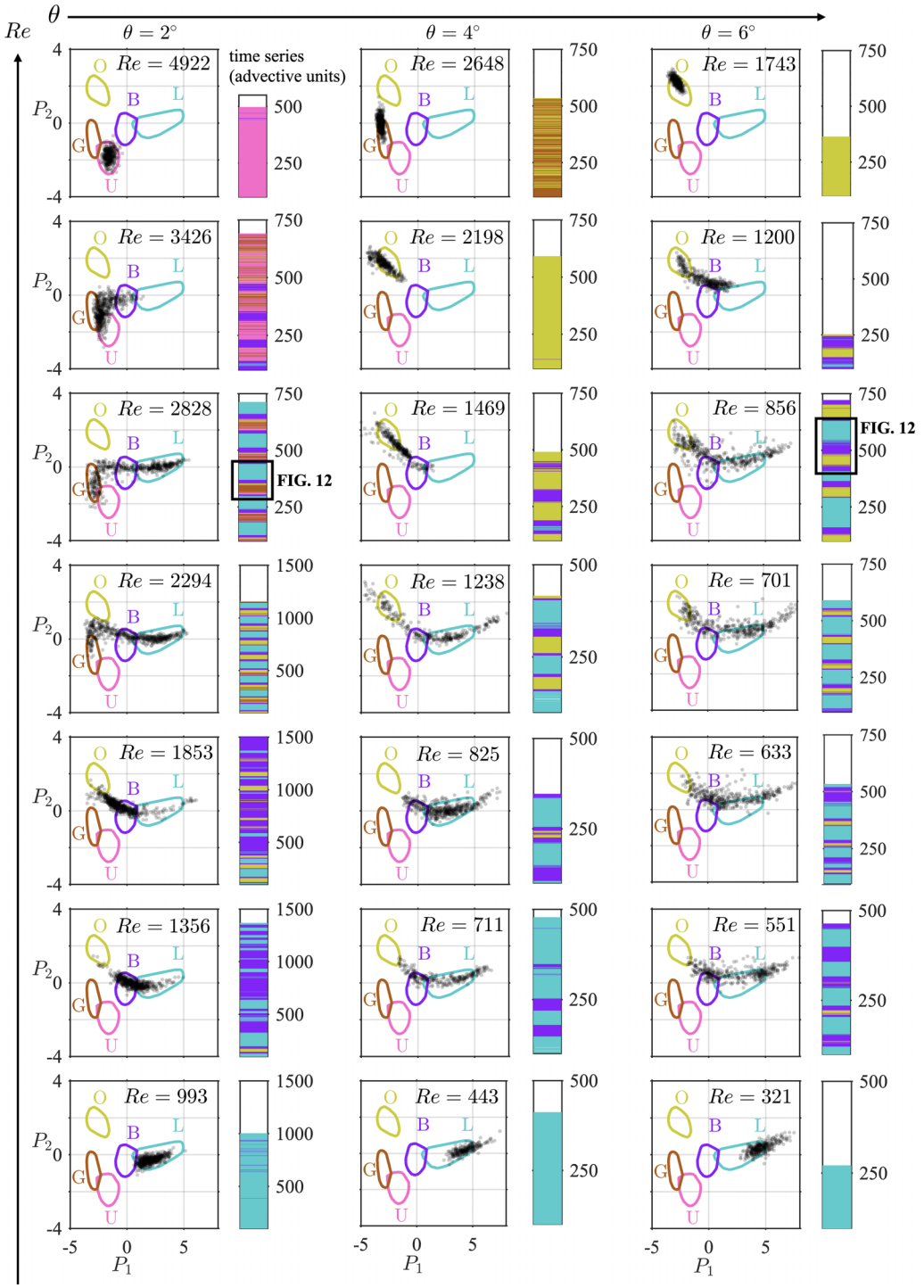

FIG. 11. Temporal dynamics in experiments at increasing Re (from bottom to top) and three tilt angles: $\theta = 2^\circ$ (left column), $\theta = 4^\circ$ (middle), and $\theta = 6^\circ$ (right). Individual frames are marked by transluscent black symbols and cluster boundaries (convex hulls) are denoted in color. Time series indicating which cluster the frames belong to, for $t \geqslant 100$, are shown by a vertical bar on the right. Unclustered frames are colored based on the nearest cluster. Note the involuntary variations in recording time between experiments.

$\theta = 2^\circ$, turbulence is exclusively unstructured, staying within or very near cluster U. At the intermediate $\theta = 4^\circ$, it shifts to being granular (primarily in G) with brief excursions to the sparser (unclustered) space towards O. At the higher $\theta = 6^\circ$, it is exclusively of overturning type (in or near O). At slightly lower values of $\mathrm{Re}\,\theta$ (second row), the trajectories are less tightly grouped in phase space and intermittency appears. At $\theta = 2^\circ$ and $\mathrm{Re} = 3426$, turbulence now cycles between U, G and B. At $\theta = 4^\circ$ and $\mathrm{Re} = 2198$, turbulence shifted to O, but without clear intermittency. At $\theta = 6^\circ$ and $\mathrm{Re} = 1200$, turbulence is intermittent among O, B, and L (despite the short time series for this dataset).

At slightly lower values of Re again (third row), the trajectories are even more spread out and intermittency is now generic at all angles. At $\theta = 2^\circ$ and $\mathrm{Re} = 2828$, the gray data cloud now covers G, B, and L, with a distinctive "upward bend" between G and B (avoiding the O cluster). The time-series cycles quasiperiodically between them with period $T \approx 120$–140 advective units. Importantly, the transitions from L to G always pass through B, just like the transitions from G to L turbulence; moreover, both transitions follow identical trajectories in this two-dimensional projection of phase space. At $\theta = 4^\circ$ and $\theta = 6^\circ$, the gray data cloud assumes a fundamentally different shape, among the O, B, and L clusters, with a distinctive "downward bend." The time series are again quasiperiodic, with period $T \approx 120$–220. At $\theta = 6^\circ$ and $\mathrm{Re} = 856$, the relaminarizations are more complete, and the transitions to and from O pass through B with identical trajectories, just like at $\theta = 2^\circ$. Moving down in Re again (fourth to seventh rows), we find that these two types of intermittency persist for a wide range of Re but that the most turbulent phases (in G at $\theta = 2^\circ$ and in O at $\theta = 4^\circ, 6^\circ$) gradually become shorter and are replaced by longer phases in B (fourth to sixth row) and eventually in L (seventh row), which here corresponds in fact to Holmboe waves.

The two distinct quasiperiodic dynamics at intermediate Re (G-B-L vs O-B-L) reveal two fundamentally different routes to turbulence in SID and constitute the third key finding of this paper. SID intermittency appears to organizes around at least two inherently different "slow manifolds" in different regions of (Re, θ). Future work is still needed to characterize these slow manifolds and their corresponding orthogonal fast manifolds, in order to understand the specific flow structures responsible for the laminar-turbulent transition in different regions of (Re, θ).

Figure 12 makes a first step in this direction by contrasting these two distinct transitional phases by zooming in on two quintessential time series of Fig. 11 (black boxes, first and third columns of third row) and showing six representative shadowgraph frames spanning a single cycle. Just before the start of a cycle (frames numbered "1," Fig. 12), signs of a growing instability are weak enough that the density interfaces are still classified as laminarizing turbulence (L). At the start of a cycle (frames 2), the instabilities have grown in amplitude and are now classified as braided turbulence (B). However, we find significant differences between the left and right frames, representing distinct routes to turbulence. The left frame shows turbulence arising first from localized small-scale structures on the upper density interface, while the right frame shows turbulence arising from a more extensive roll-up. These two transitions add complexity to our understanding of stratified turbulence and to the classical paradigm of a Kelvin-Helmholtz breaking billow, featuring a "hot" growing phase, a "Goldilocks" energetic phase, and a "cold" fossilization decaying phase [38,39]. After the active turbulence phase (frames 3), the stabilizing phase significantly differs between the two experiments, from the first stage in B (frames 4) to the last stage in L (frames 5 and 6). Turbulence and mixing subside, leaving a three-layer stratification with a partially mixed intermediate layer having a complex, temporally evolving structure. These particular time series (top bars) also highlight the general feature that the excursions in B during the relaminarization phase G/O $\rightarrow$ B $\rightarrow$ L are consistently longer than during the unstable transitional phase L $\rightarrow$ B $\rightarrow$ G/O.

The above finding that apparently distinct transitions pass through the same cluster (either B or L, see Fig. 12) suggests that consideration of a higher-dimensional phase space [greater than the two-dimensional space $\mathcal{P} = (P_1, P_2)$ considered here] may yield a deeper understanding of the underlying higher-dimensional intermittent dynamics. Future work considering a higher-dimensional space of principal components $\mathcal{P} = (P_1, P_2, P_3, \ldots)$ may thus be able to resolve the bursting and relaxation dynamics in the fast manifold [40, Sec. 6.7.2] that may be orthogonal to the first two

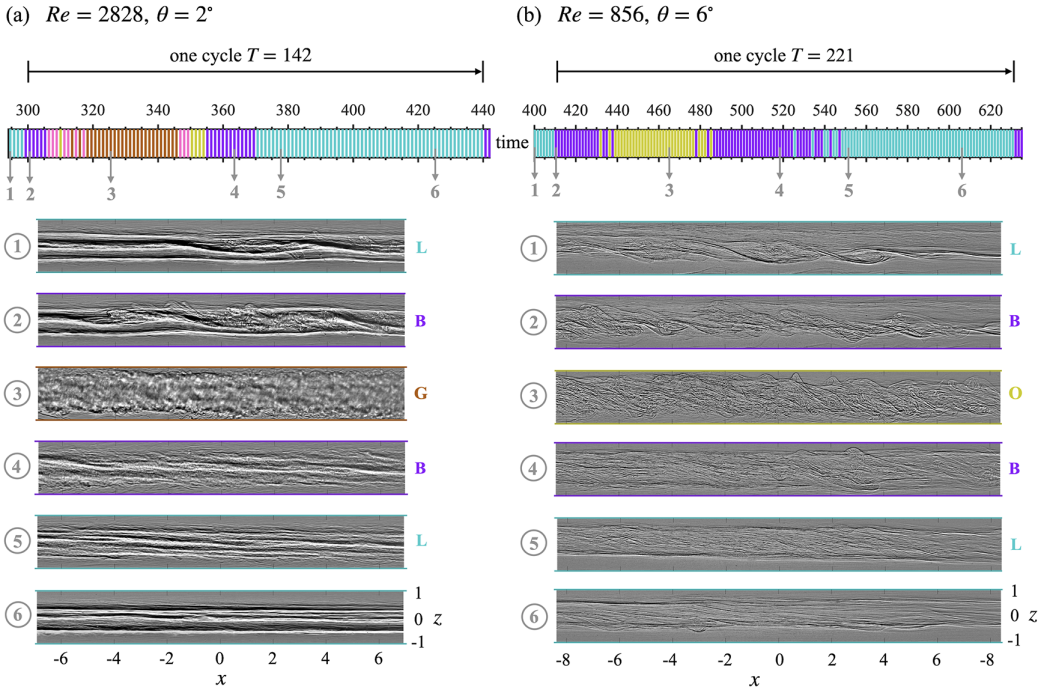

FIG. 12. Detailed transitional states present in the two distinct intermittent cycles identified in Fig. 11 (left and right column, third row). We contrast the G-B-L dynamics (left column) to the O-B-L dynamics (right) during a single cycle. Time series are provided in the top bar of each column (colored by nearest cluster) while frames 1–6 illustrate shadowgraphs at six representative times.

principal vectors. Candidates for this analysis are cluster-based Markov chain models [36,41], network models [42], and the Perron-Frobenius (transfer) operator approximated with Ulam's method [43]. The latter has the potential to identify the density interface structures "that have a higher probability of detaching from the slow manifold and use these structures as precursors (or predictors) of impending violent events" [40].

V. CONCLUSIONS

We have developed a data-driven method that automatically classifies shadowgraph measurements of turbulence and applied it to data collected from the SID experiment, in which sheared stratified turbulence is sustained at high Reynolds and Prandtl numbers. A set of 50 155 individual frames sampled from 113 shadowgraph movies were considered, spanning the relevant Reynolds number (Re) − tilt angle (θ) parameter space. These movies provide, over hundreds of advective time units, a spanwise-integrated view of the curvature of the density field (caused by a spatially varying salinity), allowing us to identify various turbulent states based on a clustering of the morphology of density interfaces embedded within the flow. The three key results in Sec. IV can be summarized as follows.

A. Instantaneous classification of turbulence

A physical interpretation of the identified clusters in Fig. 8 revealed five distinct types of stratified turbulence and mixing: laminarizing (L), braided (B), overturning (O), granular (G), and unstructured (U), as well as intermediate types. The strength of our automated classification approach is that it is objective, quantitative, sensitive to fine details of the flow, and can be readily

applied to a vast number of instantaneous frames. Its results differ from and complement the prior human classification of movies into the Holmboe wave, intermittently turbulent and turbulent flow regimes. Our clustering reveals that flows belonging to the same regime generally have a different temporal "mix" of types of turbulence as the parameters (Re, θ) are varied within the regime. Our data-driven approach can also be easily generalized and adapted to classify stratified turbulence in other insightful experiments, such as the stratified Taylor-Couette flow [44–50].

B. Dynamical map of parameter space

The fractions of time spent in each cluster in Fig. 10 followed gradual variations across the input space (Re, θ). Simply speaking, by increasing Re, laminarizing turbulence gradually gives way to more braided turbulence, then overturning turbulence, and eventually granular and unstructured turbulence. By increasing θ, braided, near-horizontal turbulence gradually gives way to overturning turbulence. This finding confirms the hypothesis in Ref. [21, Sec. 6.4] (who used a dataset of simultaneous, three-dimensional velocity and density measurements in 16 experiments) that high-Re/low-θ turbulence has more extreme enstrophy events and that low-Re/high-θ turbulence has more density overturnings. Recalling that the product $\mathrm{Re}\,\theta$ controls the rate of dissipation of turbulent kinetic energy, and that high-Re flows have a wider spectral inertial subrange, it follows that high-Re flows dissipate comparatively more at small scales (increasing extreme enstrophy) while high-θ flows dissipate comparatively more at large scales (increasing overturning). These differences in overturns statistics likely affect the energetics and efficiency of mixing [38,51].

C. Distinct temporal routes to turbulence

The phase-space trajectories of Fig. 11 revealed different dynamical behaviors with decreasing energy dissipation and dynamic range $\mathrm{Re}\,\theta$. The most dissipative turbulence remains localized in cluster U, G, or O (at low, medium, and high θ, respectively) for hundreds of advective time units. By contrast, in less dissipative turbulence, temporal intermittency gradually appears as the trajectories cycle between clusters G-B-L (at low θ) and O-B-L (at high θ), generically with remarkable quasiperiodicity. Shadowgraph snapshots at selected times during the cycles in Fig. 12 illustrated these two different transition pathways to and from stratified turbulence, i.e., laminarizing turbulence "curving" in phase space either towards granular or overturning turbulence and back. We hypothesized the existence of a low-dimensional slow manifold composed of L and O/G and of a faster, higher-dimensional manifold currently projected onto B. This paves the way for a future more accurate identification of the structures responsible for the quasiperiodic cycles of instability, turbulence, and relaminarization.

D. Outlook: Reduced-order modeling

Returning to the approach introduced in expression (1) of the Introduction, we conclude that this paper has advanced the characterization of coherent structures and their link to input parameters (step 1). The way in which the different density interface morphologies identified here, the time spent in and between each cluster, and the intermittent flow history affect the useful output variables such as mixing (step 2) remains a fundamental challenge in the community [5,52]. This challenge (reduced-order modeling) remains to be tackled with full velocity and density datasets, now available in SID experiments [13] and direct numerical simulations [37]. With such data, other computer-vision techniques would be worth exploring, such as autoencoders, which have the advantage of being nonlinear, in particular when disentanglement in the latent space is imposed. Further machine learning techniques could also learn the mapping between cluster dynamics and measures of mixing to yield predictions in unseen datasets. For example, Ref. [53] used a deep convolutional neural network to learn the relation between a low-dimensional representation of turbulence (vertical profiles of buoyancy frequency and turbulent kinetic energy dissipation) and mixing efficiency, outperforming standard parametrizations. As a further example, Ref. [54] used clustering and

sparse approximation methods to discover directly from data the dominant balance relations in equation space, which could be applied to stratified turbulence to uncover the various physical mixing processes active locally within a spatially extended domain and across parameter space.

All data are freely available. The original shadowgraph and edge datasets can be downloaded at Ref. [25], and the reduced datasets of interface morphology, principal components and clusters can be downloaded at Ref. [26].

ACKNOWLEDGMENTS

We thank X. Jiang, G. Kong, and the technicians of the G. K. Batchelor Laboratory for their help with the experiments. We are also grateful to P. F. Linden and S. B. Dalziel for their support and to C. P. Caulfield for insightful discussions on the implications of this work. The experimental facility was funded by the ERC grant "Stratified Turbulence And Mixing Processes" (STAMP, No 742480). A.L. acknowledges funding from a Leverhulme Early Career Fellowship and a NERC Independent Research Fellowship (NE/W008971/1).

APPENDIX: METHODS (COMPLEMENTING SEC. III)

1. Postprocessing of shadowgraph data

The following four steps ensured that all 113 raw movies (taking a total of ≈2 TB storage) could be used efficiently for automated classification. First, each movie was cropped vertically precisely at $z = \pm 1$ to only keep the internal wall-to-wall flow, resulting in a typical frame resolution of 3400×450 pixels (≈1.5 MPixels with 8-bit depth, i.e., 1.5 MB each). Second, the temporal mean signal (background pattern) was removed $I'(x, z, t) = I(x, z, t) - \langle I \rangle_t(x, z)$, and each frame was rescaled by calculating the 5th percentile I'_5 and 95th percentile I'_{95} of the distribution of I' and setting those of the rescaled frame as 0 and 1, respectively, by $\tilde{I} = I'/(I'_{95} - I'_5)$. All pixels having $\tilde{I} < 0$ (5% of the data) were set to 0, and all pixels having $\tilde{I} > 1$ (5% of the data) were set to 1. This ensured that all frames had comparable brightness and dynamic range, and that density interfaces could be later extracted with identical edge detection parameters. Third, we discarded the first 100 A.T.U. of each movie (i.e., keeping only data for $t \geqslant 100$) to conservatively remove any transients associated with initial gravity currents (estimated to reach the ends of the duct $x = \pm 40$ at an estimated $t \approx 80$ due to their speed ≈0.5). Fourth, we subsampled the time series by a factor of 10, to avoid excessively redundant temporal data. This yielded an average temporal resolution across all movies of $0.10 \times 10 = 1.0$ A.T.U. (standard deviation 0.4). The final dataset of 50 155 frames takes 60 GB storage if stored in the minimal 8-bit precision.

2. Principal components analysis

The data matrix $\mathbf{M}$ is first normalized to transform the values of the 10 morphology statistics into standard scores (or "z scores"), where 0 corresponds to the mean value across all frames, and $\pm\lambda$ to λ standard deviations above and below the mean, thereby ensuring that all characteristics are weighted equally. The SVD of the normalized data matrix is then $\tilde{\mathbf{M}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$. Here $\mathbf{\Sigma} = \{\sigma_m\}_{m=1,\ldots,10}$ is the diagonal matrix of positive singular values arranged in decreasing order, and $\mathbf{U} = \{\mathbf{u}_m\}_{m=1,\ldots,10} \in \mathbb{R}^{50155\times 10}$ and $\mathbf{V} = \{\mathbf{v}_m\}_{m=1,\ldots,10} \in \mathbb{R}^{10\times 10}$ are the real orthogonal matrices containing the left and right singular vectors, respectively. The cumulative variance in Fig. 4(b) shows that $\Sigma(2)/\Sigma(10) = 0.79$, noting that the total variance $\Sigma(10)$ is the sum of the 10 eigenvalues of the covariance matrix $\tilde{\mathbf{M}}^T\tilde{\mathbf{M}}$. This justifies the truncated approximation $\tilde{\mathbf{M}} \approx \tilde{\mathbf{M}}_t = \mathbf{U}_t\mathbf{\Sigma}_t\mathbf{V}_t^T$ where we only retain the first two columns of $\mathbf{U}_t$, $\mathbf{V}_t$ and the first 2×2 block of $\mathbf{\Sigma}_t$. In other words, $\tilde{\mathbf{M}}_t = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T$, which approximates each frame (row) $f = 1, \ldots, 50\,155$ as the sum $\approx \sigma_1 u_{f1}\mathbf{v}_1^T + \sigma_2 u_{f2}\mathbf{v}_2^T$ using the top two right transposed singular vectors $\mathbf{v}_1^T, \mathbf{v}_2^T \in \mathbb{R}^{1\times 10}$. Rewriting this as $p_1\mathbf{v}_1^T + p_2\mathbf{v}_2^T$ highlights that we effectively project each rank-two-approximated frame from

the 10-dimensional space $\mathcal{M}$ to the two-dimensional space $\mathcal{P}$ spanned by the two normal unit vectors $\mathbf{v}_1^T$, $\mathbf{v}_2^T$, and obtain the vector of coordinates $\mathbf{p} = [p_1, p_2] = [\sigma_1 u_{f1}, \sigma_2 u_{f2}]$. In matrix form, this mapping thus transformed $\tilde{\mathbf{M}} \in \mathbb{R}^{50155\times 10}$ to the new data matrix $\mathbf{P} = \tilde{\mathbf{M}}_t \mathbf{V}_t = \mathbf{U}_t \mathbf{\Sigma}_t \in \mathbb{R}^{50155\times 2}$ containing the 50 155 two-dimensional row vectors $\mathbf{p}$.

3. OPTICS algorithm

OPTICS computes the pairwise distances between all points based on a metric called the reachability distance:

$$d_{\text{reach}}(\mathbf{p}_i, \mathbf{p}_j) = \max(\, d_{\text{euclidean}}(\mathbf{p}_i, \mathbf{p}_j),\ d_{\text{core}}(\mathbf{p}_i),\ d_{\text{core}}(\mathbf{p}_j)), \tag{A1}$$

where $d_{\text{euclidean}}$ denotes the standard Euclidean distance, and d_{core} denotes the core distance, i.e., the radius of the hypersphere around each point (in our case a two-dimensional vector $\mathbf{p}$) that encloses exactly minPts neighbors. OPTICS starts from an arbitrary initial point (random initialization), and generates the reachability plot sequentially, moving to the next point in the dataset with the closest d_{reach} and plotting it at each step. Therefore, the shape of the reachability plot, and thus the detected clusters, do not depend on the choice of initial point and the algorithm only needs to be run once. The key property of the reachability distance is that it penalizes points in sparser regions having larger d_{core} by increasing their perceived distance from points in denser regions. A small d_{core} indicates that the point sits in a dense regions of the space $\mathcal{P} = (P_1, P_2)$. OPTICS thus requires one user-specified parameter "minPts," which can be interpreted as the minimum number of points required to form a cluster. Extreme values of 50 155 (our total number of points) or 1 would yield a single cluster or a cluster per point, respectively. To find a useful intermediate value, we progressively decreased minPts, leading to distinct valleys (meaningful clusters) in the reachability plot (as shown in Fig. 5), which are robust for $600 \lesssim \text{minPts} \lesssim 1200$. Decreasing minPts below 600 leads to a much greater number of clusters clearly dominated by noise. The results in this paper were computed using minPts $= 800$.

[1] M. Van Dyke, *An Album of Fluid Motion* (Parabolic Press, Standford, CA, 1982).

[2] C. Andereck, S. Liu, and H. Swinney, Flow regimes in a circular Couette system with independently rotating cylinders, J. Fluid Mech. **164**, 155 (1986).

[3] D. Barkley, Theoretical perspective on the route to turbulence in a pipe, J. Fluid Mech. **803**, P1 (2016).

[4] D. Feldmann, D. Borrero-Echeverry, M. J. J. Burin, K. Avila, and M. Avila, Routes to turbulence in Taylor-Couette flow, Philos. Trans. R. Soc. A **381**, 20220114 (2023).

[5] C. P. Caulfield, Layering, instabilities, and mixing in turbulent stratified flows, Annu. Rev. Fluid Mech. **53**, 113 (2021).

[6] M. C. Gregg, E. A. D'Asaro, J. J. Riley, and E. Kunze, Mixing efficiency in the ocean, Annu. Rev. Mar. Sci. **10**, 443 (2018).

[7] O. Reynolds, An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels, Phil. Trans. R. Soc. **174**, 935 (1883).

[8] G. I. Taylor, An experiment on the stability of superposed streams of fluid, Math. Proc. Cambr. Phil. Soc. **23**, 730 (1927).

[9] S. A. Thorpe, Experiments on the instability of stratified shear flows: Miscible fluids, J. Fluid Mech. **46**, 299 (1971).

[10] E. O. Macagno and H. Rouse, Interfacial mixing in stratified flow, J. Eng. Mech. Div. Proc. Am. Soc. Civ. Eng. **87**, 55 (1961).

[11] C. R. Meyer and P. F. Linden, Stratified shear flow: Experiments in an inclined duct, J. Fluid Mech. **753**, 242 (2014).

[12] J. L. Partridge, A. Lefauve, and S. B. Dalziel, A versatile scanning method for volumetric measurements of velocity and density fields, Meas. Sci. Technol. **30**, 055203 (2019).
[13] A. Lefauve, J. L. Partridge, and P. F. Linden, Regime transitions and energetics of sustained stratified shear flows, J. Fluid Mech. **875**, 657 (2019).
[14] A. Lefauve and P. F. Linden, Buoyancy-driven exchange flows in inclined ducts, J. Fluid Mech. **893**, A2 (2020).
[15] A. Lefauve and P. F. Linden, Experimental properties of continuously forced, shear-driven, stratified turbulence. Part 2. Energetics, anisotropy, parameterisation, J. Fluid Mech. **937**, A35 (2022).
[16] M. Duran-Matute, S. J. Kaptein, and H. J. H. Clercx, Regime transitions in stratified shear flows: The link between horizontal and inclined ducts, J. Fluid Mech. **956**, A4 (2023).
[17] A. Lefauve, J. L. Partridge, Q. Zhou, C. P. Caulfield, S. B. Dalziel, and P. F. Linden, The structure and origin of confined Holmboe waves, J. Fluid Mech. **848**, 508 (2018).
[18] X. Jiang, A. Lefauve, S. B. Dalziel, and P. F. Linden, The evolution of coherent vortical structures in increasingly turbulent stratified shear layers, J. Fluid Mech. **947**, A30 (2022).
[19] M. M. P. Couchman, S. M. de Bruyn Kops, and C. P. Caulfield, Mixing across stable density interfaces in forced stratified turbulence, J. Fluid Mech. **961**, A20 (2023).
[20] J. J. Riley, M. M. P. Couchman, and S. M. de Bruyn Kops, The effect of Prandtl number on decaying stratified turbulence, J. Turbulence **24**, 330 (2023).
[21] A. Lefauve and P. F. Linden, Experimental properties of continuously forced, shear-driven, stratified turbulence. Part 1. Mean flows, self-organisation, turbulent fractions, J. Fluid Mech. **937**, A34 (2022).
[22] A. Atoufi, L. Zhu, A. Lefauve, J. R. Taylor, G. A. Lawrence, S. B. Dalziel, R. R. Kerswell, and P. F. Linden, Stratified inclined duct: Two-layers hydraulics and instabilities, J. Fluid Mech. **977**, A25 (2023).
[23] E. Deusebio, C. P. Caulfield, and J. R. Taylor, The intermittency boundary in stratified plane couette flow, J. Fluid Mech. **781**, 298 (2015).
[24] A. Lefauve, Waves and turbulence in sustained stratified shear flows, Ph.D. thesis, University of Cambridge, 2018.
[25] X. Jiang, G. Kong, and A. Lefauve, Shadowgraph visualisations of salt-stratified turbulence obtained in a stratified inclined duct (SID) laboratory experiment (2023), doi:10.17863/CAM.104471.
[26] A. Lefauve and M. M. P. Couchman, Research data supporting "Data-driven classification of sheared stratified turbulence from experimental shadowgraphs" (2024), doi:10.17863/CAM.104427.
[27] J. Canny, A computational approach to ddge detection, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-8**, 679 (1986).
[28] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, 1st ed. (Cambridge University Press, Cambridge, UK, 2019).
[29] H. Shen and J. Z. Huang, Sparse principal component analysis via regularized low rank matrix approximation, J. Multivariate Anal. **99**, 1015 (2008).
[30] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, Ordering points to identify the clustering structure, ACM SIGMOD Rec. **28**, 49 (1999).
[31] M. M. P. Couchman, B. Wynne-Cattanach, M. H. Alford, C. P. Caulfield, R. R. Kerswell, J. A. MacKinnon, and G. Voet, Data-driven identification of turbulent oceanic mixing from observational microstructure Data, Geophys. Res. Lett. **48**, e2021GL094978 (2021).
[32] J. Han, M. Kamber, and J. Pei, Cluster analysis: Basic concepts and methods, in *Data Mining: Concepts and Techniques* (Elsevier, Amsterdam, 2012).
[33] J. R. Carpenter, E. W. Tedford, E. Heifetz, and G. A. Lawrence, Instability in stratified shear flow: Review of a physical interpretation based on interacting waves, App. Mech. Rev. **64**, 060801 (2011).
[34] G. D. Portwood, S. M. de Bruyn Kops, J. R. Taylor, H. Salehipour, and C. P. Caulfield, Robust identification of dynamically distinct regions in stratified turbulence, J. Fluid Mech. **807**, R2 (2016).
[35] G. Narasimhan, C. Meneveau, and T. A. Zaki, Large eddy simulation of transitional channel flow using a machine learning classifier to distinguish laminar and turbulent regions, Phys. Rev. Fluids **6**, 074608 (2021).
[36] F. Foroozan, V. Guerrero, A. Ianiro, and S. Discetti, Unsupervised modelling of a transitional boundary layer, J. Fluid Mech. **929**, A3 (2021).

[37] L. Zhu, A. Atoufi, A. Lefauve, J. R. Taylor, G. A. Lawrence, S. B. Dalziel, R. R. Kerswell, and P. F. Linden, Stratified inclined duct: direct numerical simulations, J. Fluid Mech. **969**, A20 (2023).
[38] A. Mashayek, C. P. Caulfield, and M. H. Alford, Goldilocks mixing in oceanic shear-induced turbulent overturns, J. Fluid Mech. **928**, A1 (2021).
[39] K. M. Smith, C. P. Caulfield, and J. R. Taylor, Turbulence in forced stratified shear flows, J. Fluid Mech. **910**, A42 (2021).
[40] P. J. Schmid, Data-driven and operator-based tools for the analysis of turbulent flows, in *Advanced Approaches in Turbulence* (Elsevier, Amsterdam, 2021), pp. 243–305.
[41] E. Kaiser, B. R. Noack, L. Cordier, A. Spohn, M. Segond, M. Abel, G. Daviller, J. Östh, S. Krajnović, and R. K. Niven, Cluster-based reduced-order modelling of a mixing layer, J. Fluid Mech. **754**, 365 (2014).
[42] H. Li, D. Fernex, R. Semaan, J. Tan, M. Morzyński, and B. R. Noack, Cluster-based network model, J. Fluid Mech. **906**, A21 (2021).
[43] O. Junge and P. Kolta, Discretization of the Frobenius-Perron operator using a sparse Haar tensor basis: the sparse Ulam method, SIAM J. Numer. Anal. **47**, 3464 (2009).
[44] F. Caton, B. Janiaud, and E. J. Hopfinger, Primary and secondary Hopf bifurcations in stratified Taylor-Couette flow, Phys. Rev. Lett. **82**, 4647 (1999).
[45] R. L. F. Oglethorpe, C. P. Caulfield, and A. W. Woods, Spontaneous layering in stratified turbulent Taylor-Couette flow, J. Fluid Mech. **721**, R3 (2013).
[46] C. Leclercq, J. L. Partridge, P. Augier, S. B. Dalziel, and R. R. Kerswell, Using stratification to mitigate end effects in quasi-Keplerian Taylor-Couette flow, J. Fluid Mech. **791**, 608 (2016).
[47] R. Ibanez, H. L. Swinney, and B. Rodenborn, Observations of the stratorotational instability in rotating concentric cylinders, Phys. Rev. Fluids **1**, 053601 (2016).
[48] J. Park, P. Billant, J.-J. Baik, and J. M. Seo, Competition between the centrifugal and strato-rotational instabilities in the stratified Taylor-Couette flow, J. Fluid Mech. **840**, 5 (2018).
[49] D. Petrolo and S. Longo, Buoyancy transfer in a two-layer system in steady state: Experiments in a Taylor-Couette cell, J. Fluid Mech. **896**, A27 (2020).
[50] G. Meletti, S. Abide, S. Viazzo, A. Krebs, and U. Harlander, Experiments and long-term high-performance computations on amplitude modulations of strato-rotational flows, Geophys. Astrophys. Fluid Dyn. **115**, 297 (2021).
[51] A. Mashayek, B. B. Cael, L. Cimoli, M. H. Alford, and C. P. Caulfield, A physical–statistical recipe for representation of small-scale oceanic turbulent mixing in climate models, Flow **2**, E24 (2022).
[52] C.-c. P. Caulfield, Open questions in turbulent stratified mixing: Do we even know what we do not know? Phys. Rev. Fluids **5**, 110518 (2020).
[53] H. Salehipour and W. R. Peltier, Deep learning of mixing by two "atoms" of stratified turbulence, J. Fluid Mech. **861**, R4 (2019).
[54] J. L. Callaham, J. V. Koch, B. W. Brunton, J. N. Kutz, and S. L. Brunton, Learning dominant physical processes with data-driven balance models, Nat. Commun. **12**, 1016 (2021).