PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Extreme events in transitional turbulence

Sébastien Gomé[1], Laurette S. Tuckerman[1] and Dwight Barkley[2]

[1]Laboratoire de Physique et Mécanique des Milieux Hétérogènes, CNRS, ESPCI Paris, PSL Research University, Sorbonne Université, Université de Paris, Paris 75005, France

[2]Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

SG, 0000-0002-4423-6213; LST, 0000-0001-5893-9238; DB, 0000-0003-4317-3705

Research

Check for updates

Cite this article: Gomé S, Tuckerman LS, Barkley D. 2022 Extreme events in transitional turbulence. *Phil. Trans. R. Soc. A* **380**: 20210036. https://doi.org/10.1098/rsta.2021.0036

Received: 22 August 2021

Accepted: 4 January 2022

One contribution of 14 to a theme issue 'Mathematical problems in physical fluid dynamics (part 2)'.

Subject Areas:

fluid mechanics, mathematical modelling, statistical physics

Keywords:

rare events, transitional turbulence, extreme values, Large Deviation theory

Author for correspondence:

Dwight Barkley

e-mail: d.barkley@warwick.ac.uk

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5884439.

Transitional localized turbulence in shear flows is known to either decay to an absorbing laminar state or to proliferate via splitting. The average passage times from one state to the other depend super-exponentially on the Reynolds number and lead to a crossing Reynolds number above which proliferation is more likely than decay. In this paper, we apply a rare-event algorithm, Adaptative Multilevel Splitting, to the deterministic Navier–Stokes equations to study transition paths and estimate large passage times in channel flow more efficiently than direct simulations. We establish a connection with extreme value distributions and show that transition between states is mediated by a regime that is self-similar with the Reynolds number. The super-exponential variation of the passage times is linked to the Reynolds number dependence of the parameters of the extreme value distribution. Finally, motivated by instantons from Large Deviation theory, we show that decay or splitting events approach a most-probable pathway.

This article is part of the theme issue 'Mathematical problems in physical fluid dynamics (part 2)'.

1. Introduction

The route to turbulence in many wall-bounded shear flows is a spatio-temporal process that results from the interplay between the tendency for turbulence to decay or for it to proliferate. Individual decay and proliferation events occur extremely rarely near the critical Reynolds number for the onset of sustained turbulence, and

THE ROYAL SOCIETY PUBLISHING

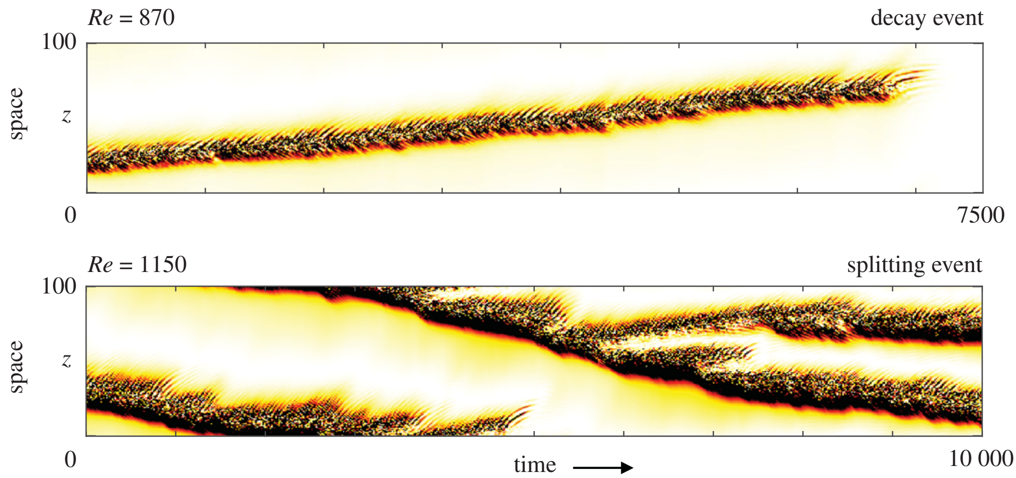

Figure 1. Evolution of turbulence in channel flow at two different Reynolds numbers. Turbulence is seen as black and is localized to only a portion of space. White corresponds to laminar (or nearly laminar) flow. The motion of the turbulent patch is seen in a frame of reference moving with the mean flow in the channel and the system is periodic in spatial coordinate z. At $Re = 870$ the localized band of turbulence maintains an approximately constant width and intensity for a considerable time and then abruptly transitions to laminar flow in a decay event. At $Re = 1150$ the localized turbulent band is wider and noticeably asymmetric. In this case, the band splits into two bands. In the vicinity of $Re = 1000$, both of these key events become extremely rare and the mean exit time from the one-band state becomes very large. Results are obtained by a numerical simulation in an oblique domain represented in figure 3. (Online version in colour.)

this makes measuring, let alone understanding the onset of turbulence in these flows both fascinating and challenging. In this paper, we investigate these rare events.

Figure 1 illustrates individual decay and proliferation (splitting) events of interest. These have been obtained from numerical simulations of pressure-driven flow in a channel. The spatio-temporal diagrams of figure 1 display the evolution of such localized turbulent bands at two Reynolds numbers. Simulations begin after some initial equilibration time. It can be seen that the one-band state is metastable—it persists for significant time before transitioning to another state, either laminar flow, as in the upper panel, or a two-band state, as in the lower one. The corresponding phase-space picture for the governing Navier–Stokes equations is sketched in figure 2. Trajectories spend a significant time in a region of phase space associated with a single turbulent band, $\mathcal{A}$, before exiting the region and going to laminar flow or to the two-band state. Repeated simulations starting from one-band states (in the region $\mathcal{A}$) show that the exit times are distributed exponentially, so that decay and splitting events are effectively governed by a memoryless, Poisson process (see [1–6] and references therein).

A typical study consists of the following. For each value of the Reynolds number, Re, a large number of events is generated, from which the mean lifetime is determined by averaging the lifetimes observed in the sample events. This is the Monte Carlo (MC) approach. The process is repeated for a range of Re to obtain the mean lifetimes to decay $\tau_d(Re)$ and to split $\tau_s(Re)$. These lifetimes are observed to depend super-exponentially on Reynolds number as sketched in figure 2*b*, and are approximated by a double exponential form: $\tau_d(Re) \sim \exp(\exp(a_d Re + b_d))$ and similarly for $\tau_s(Re)$. (Figure 7 discussed below contains actual measured mean lifetimes for channel flow.) The timescales cross at a critical value Re_c. Below Re_c decay events occur more frequently, while above Re_c splitting events occur more frequently. The crossover between these cases is a key mechanism in the onset of sustained turbulence in wall-bounded shear flow. This crossing point is not, however, the focus of the present study.

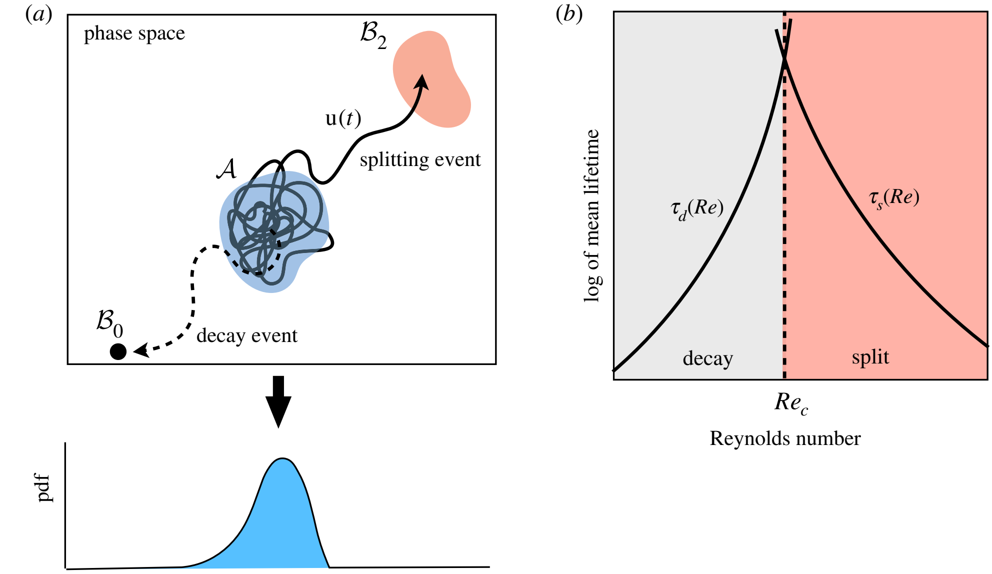

Figure 2. (*a*) Illustration of the phase space of the Navier–Stokes equations. Time evolving flow fields u(t) are seen as trajectories. The one-band state corresponds to a region $\mathcal{A}$ in the phase space in which trajectories u(t) spend considerable time before exiting and transitioning either to laminar flow $\mathcal{B}_0$ or to the two-band state $\mathcal{B}_2$. The fluctuations of observables, such as the turbulence fraction, are described by extreme value distributions. (*b*) Schematic showing the dependence of mean lifetimes on Reynolds number, *Re*. Lifetimes vary super-exponentially with *Re*, with τ_d increasing and τ_s decreasing with *Re*. The timescales cross at a critical value Re_c. Below Re_c, decay occurs more frequently while above Re_c, splitting occurs more frequently. (Online version in colour.)

The present study focuses instead on two key issues associated with the rare events themselves. The first is the efficient numerical computation of mean lifetimes. In shear flows, τ_d and τ_s become extremely large near Re_c, making brute force MC estimation of mean times exceedingly expensive. Hence we turn to a more sophisticated class of algorithms that sample rare events by advancing ensembles of trajectories, removing (pruning) unfavourable and duplicating (cloning) favourable ones. In particular, we will employ the Adaptative Multilevel Splitting (AMS) algorithm proposed by Cérou & Guyader [7–9]. (This nomenclature of 'splitting' in the algorithm is unrelated to the splitting of turbulent bands.) This algorithm impressively paved the way for quantitative study of low-dimensional stochastic systems, as pioneered by Rolland & Simonnet [10], Rolland *et al.* [11] or Lestang *et al.* [12]. It was recently applied to large-dimensional fluid-dynamical systems such as atmospheric dynamics [13,14] and bluff-body flow [15]. Rolland [16] extended the application of this rare-event technique to transitional turbulence, first for transition in a stochastic reduced-order model [17] of pipe flow, and then for the collapse of homogeneous turbulence in plane Couette flow [18].

The second main focus of our study is the origin of the super-exponential dependence of mean lifetimes on Reynolds number, and in particular the connection to extreme values of fluctuations within the one-band state. Goldenfeld *et al.* [19] proposed a mechanism to account for the super-exponential dependence of decay lifetimes of Reynolds number. The essential insight is that the decay process is governed by extreme values and that a linear variation of Reynolds number translates via extreme value distributions to a super-exponential variation in lifetimes. This mechanism was investigated and refined by Nemoto & Alexakis [20,21] in a numerical study of decay events in pipe flow. We will follow a similar analysis applied to both decay and splitting events in channel flow. Finally, the possible connection to the large deviation framework is considered through the computation of most-probable pathways and mean reactive times for rare events.

2. Methods

We will now describe two very different types of methods: first, those we use for solving the Navier–Stokes equations governing channel flow, and, second, our implementation of the AMS algorithm for capturing rare events.

(a) Integration of Navier–Stokes equations in a transitional flow unit

The turbulent bands that are the subject of our study are illustrated in figure 3. We impose a mean velocity U on the flow between the two parallel rigid plates. Lengths are non-dimensionalized by the half-gap h between the plates, velocities by $3U_{\mathrm{bulk}}/2$ (which is the centreline velocity of the parabolic laminar flow with mean velocity U_{bulk}), and time by the ratio between them. The Reynolds number is defined to be $Re = 3U_{\mathrm{bulk}}h/(2\nu)$. The non-dimensionalized equations that we simulate are the incompressible Navier–Stokes equations

$$\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U}\cdot\nabla)\mathbf{U} = -\nabla p + \frac{1}{Re}\nabla^2\mathbf{U} \tag{2.1a}$$

and

$$\nabla\cdot\mathbf{U} = 0 \tag{2.1b}$$

Since the bands are found to be oriented obliquely with respect to the streamwise direction, we use a periodic numerical domain which is tilted with respect to the streamwise direction of the flow, shown as the black rectangle in figure 3. This is common in studying turbulent bands [22,23] and more specifically those in transitional plane channel flow [6,24,25]. The x-direction is chosen to be aligned with a typical turbulent band and the z-coordinate to be orthogonal to the band. The relationship between streamwise–spanwise coordinates (x', z') and tilted band-oriented (x, z) coordinates is

$$\mathbf{e}_{x'} = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_z \tag{2.2a}$$

and

$$\mathbf{e}_{z'} = -\sin\theta\,\mathbf{e}_x + \cos\theta\,\mathbf{e}_z. \tag{2.2b}$$

The usual wall-normal coordinate is denoted by y. The field visualized in figure 3 comes from an additional simulation we carried out in a domain of size $(L_{x'}, L_y, L_{z'}) = (200, 2, 120)$ aligned with the streamwise–spanwise coordinates.

Equations (2.1) are completed by rigid boundary conditions in y, periodic boundary conditions in x and z, and imposed flux $2/3$ in the streamwise direction x' and 0 in the spanwise direction z':

$$\mathbf{U}(x+L_x, y, z) = \mathbf{U}(x, y, z+L_z) = \mathbf{U}(x, y, z) \quad \mathbf{U}(x, \pm 1, z) = 0 \tag{2.3a}$$

and

$$\frac{1}{2}\int_{-1}^{+1} \mathrm{d}y\,\mathbf{U}(x, y, z) = \frac{2}{3}\mathbf{e}_{x'} = \frac{2}{3}(\cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_z). \tag{2.3b}$$

To integrate (2.1) with boundary conditions (2.3), we use the parallelized pseudospectral C++ code ChannelFlow [26], which employs a Fourier–Chebychev spatial discretization. The velocity field can be decomposed into the stationary laminar parabolic base flow $\mathbf{U}_{\mathrm{base}} = (1-y^2)\mathbf{e}_{x'}$ and the deviation $\mathbf{u} \equiv \mathbf{U} - \mathbf{U}_{\mathrm{base}}$ which satisfies the same equations and boundary conditions as $\mathbf{U}$ but with zero flux instead of (2.3*b*). A Green's function method is used to impose the flux in each direction. More specifically, for each periodic direction, one computes and uses the pressure gradient such that the resulting flow field will have the desired bulk velocity, e.g. [27,28]. Throughout our study, we present the deviation $\mathbf{u} = (u, v, w)$ so as to highlight the difference with the dominant laminar flow $\mathbf{U}_{\mathrm{base}}$ and the motion of flow features with respect to the bulk velocity.

The angle in this study is fixed at $\theta = 24^\circ$, as has been used extensively in the past [6,22,24]. The orientation of the domain imposes a fixed angle on turbulent bands, and choosing a short length for the x-direction of the domain suppresses any large-scale variation along the bands. Thus, these simulations effectively capture the dynamics of infinitely long bands that only interact along their

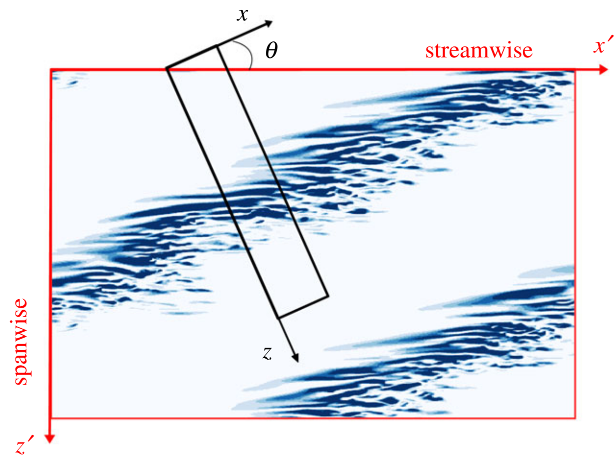

Figure 3. Visualization of a turbulent band in a domain periodic in the streamwise and spanwise directions (red bounding box) at $Re = 1000$. Colours show transverse energy $\frac{1}{2}(v^2 + w'^2)$ in the plane $y = 0.8$, from our numerical simulation in a box of size $L_{x'} = 200$, $L_{z'} = 120$. Illustration of the associated tilted computational domain (black) at angle $\theta = 24°$. (Online version in colour.)

perpendicular direction, preventing complex 2D interactions that are possible for finite-length bands [29,30]. In this way, localized bands in the tilted channel geometry are similar to localized puffs in pipe flow.

Our domain Ω has dimensions $(L_x, L_y, L_z) = (6.6, 2, 100)$ and a numerical resolution of $(N_x, N_y, N_z) = (84, 64, 1250)$, exactly as in [6], thus allowing direct comparison with these prior results. The length $L_z = 100$ of our tilted domain corresponds to an inter-band distance above which a band is considered as isolated, while the domain width $L_x = 6.6$ is used because it corresponds to the natural spacing of streaks in channel flow in a 24° box [6,31]. For puffs in pipe flow, which are similar in many respects to the isolated bands considered here, Nemoto & Alexakis [21] conducted extensive computations showing that domain length had some effect on mean decay timescales, with $L = 50$ and $L = 100$ giving quantitatively different, but qualitatively similar results. Domain length is expected to have a quantitative effect on the splitting timescale; our domain length $L_z = 100$ has been selected as a compromise between accuracy and computational cost.

A semi-implicit time-stepping scheme is used to progress from $\mathbf{u}(t)$ to $\mathbf{u}(u + \mathrm{d}t)$, with time step $\mathrm{d}t = 1/32 = 0.03125$. Trajectories and associated quantities such as turbulence fraction are sampled at time intervals $\delta t = 32\mathrm{d}t = 1$. This sampling time is used throughout for collecting statistics and generating probability distributions. The computation of solutions of the Navier–Stokes equations discretized in space and time is called, as usual, direct numerical simulation or DNS.

(b) The adaptive multilevel splitting algorithm

Here, we present the essence of the AMS algorithm. We follow closely the method originally described in Cérou & Guyader [7], although here we consider a deterministic dynamical system, the Navier–Stokes equations (2.1), whereas Cérou and colleagues considered a stochastic process. The AMS algorithm has been applied recently to other deterministic fluid systems [12,15,18]. For the application of other rare-event algorithms to deterministic systems, see [32] and references therein.

(i) Setup

Let $\mathcal{A}$ and $\mathcal{B}$ be two states visited by trajectories of a dynamical system. More precisely, $\mathcal{A}$ and $\mathcal{B}$ are regions in phase space corresponding to particular flow states of interest. We commonly refer

Table 1. Definitions of designated levels of turbulent fraction or score function used throughout the paper.

symbol	definition
$h_{\mathcal{A}}$	hypersurface within $\mathcal{A}$, origin of trajectories, in practice one-band state
$h_{\mathcal{S}}$	hypersurface $\mathcal{S}$ close to and surrounding $\mathcal{A}$
$h_{\mathcal{B}}$	hypersurface within $\mathcal{B}$, destination of trajectories
$h_{\mathcal{B}_0}$	threshold for decay events in AMS
$h_{\mathcal{B}_2}$	threshold for splitting events in AMS
h_0	entrance of the *collapse zone* for decays for all *Re*
h_2	entrance of the *collapse zone* for splits for all *Re*
h_M	maximal value of F_t at fixed *Re*
h_{left}	left endpoint of fit between PDF of F_t and Fisher–Tippett distribution
h_{right}	right endpoint of fit between PDF of F_t and Fisher–Tippett distribution

Table 2. Values of designated levels of turbulent fraction or score function used throughout the paper.

Re	815	830	870	900	950	1000	1050	1100	1150	1200
$h_{\mathcal{A}}$	0.21	0.22	0.24	0.26	0.31	0.34	0.37	0.40	0.43	0.44
$h_{\mathcal{S}}$	0.17	0.18	0.21	0.23	0.27	0.375	0.41	0.44	0.46	0.47
$h_{\mathcal{B}_0}, h_{\mathcal{B}_2}$	0.0001	0.0001	0.0001	0.0001	0.0001	0.70	0.70	0.70	0.70	0.70
h_0, h_2	0.22	0.22	0.22	0.22	0.22	0.42	0.431	0.461	0.474	0.483
h_M	0.292	0.305	0.344	0.385	0.44	0.635	0.616	0.659	0.677	0.69
h_{left}	0.13	0.148	0.176	0.207	0.243	0.30	0.32	0.279	0.271	0.326
h_{right}	0.285	0.278	0.307	0.327	0.364	0.42	0.436	0.469	0.501	0.536

to $\mathcal{A}$ and $\mathcal{B}$ simply as states. The goal is to produce a large sample of the rare transitions from $\mathcal{A}$ to $\mathcal{B}$. In our case $\mathcal{A}$ will always be the one-band state, labelled as $\mathcal{A}$ in figure 2, while $\mathcal{B}$ will be either the laminar flow, labelled as $\mathcal{B}_0$, or else the two-band state, labelled as $\mathcal{B}_2$ in figure 2.

Perhaps the most crucial piece of the AMS algorithm is the specification of a score function, or reaction coordinate, ϕ, that quantifies transitions from $\mathcal{A}$ to $\mathcal{B}$. The score function $\phi(\mathbf{u})$ is a real-valued function of the flow field whose gradient is non-zero (at least everywhere of interest), and such that there exist real values $h_{\mathcal{A}}$ and $h_{\mathcal{B}}$, with $h_{\mathcal{A}} < h_{\mathcal{B}}$, such that $\phi(\mathbf{u}) < h_{\mathcal{A}}$ implies $\mathbf{u} \in \mathcal{A}$ while $\phi(\mathbf{u}) > h_{\mathcal{B}}$ implies $\mathbf{u} \in \mathcal{B}$. Note that for decay, the laminar state is a single point in phase space, so we will take $\mathcal{B}$ to be a set within its basin of attraction. Tables 1 and 2 list the various thresholds of the score function that we will use throughout the paper. The score function provides a smooth landscape for quantifying the progress of the transition between $\mathcal{A}$ and $\mathcal{B}$, as illustrated in figure 4*a*.

The algorithm also requires a value $h_{\mathcal{S}}$ and associated hypersurface $\mathcal{S}$, close to $\mathcal{A}$, given by

$$\mathcal{S} = \{\mathbf{u} | \phi(\mathbf{u}) = h_{\mathcal{S}}\}.$$

(ii) Initialization

The initialization step consists of generating a sample of N trajectories $\mathbf{u}_i(t)$, $i \in \{1, \ldots, N\}$, that start within $\mathcal{A}$, leave $\mathcal{A}$ at least as far as $\mathcal{S}$, and then either reach $\mathcal{B}$ or, more likely, return to $\mathcal{A}$ (figure 4*a*). In practice, the N initial conditions $\mathbf{u}_i(0)$ are obtained by taking N snapshots, equally

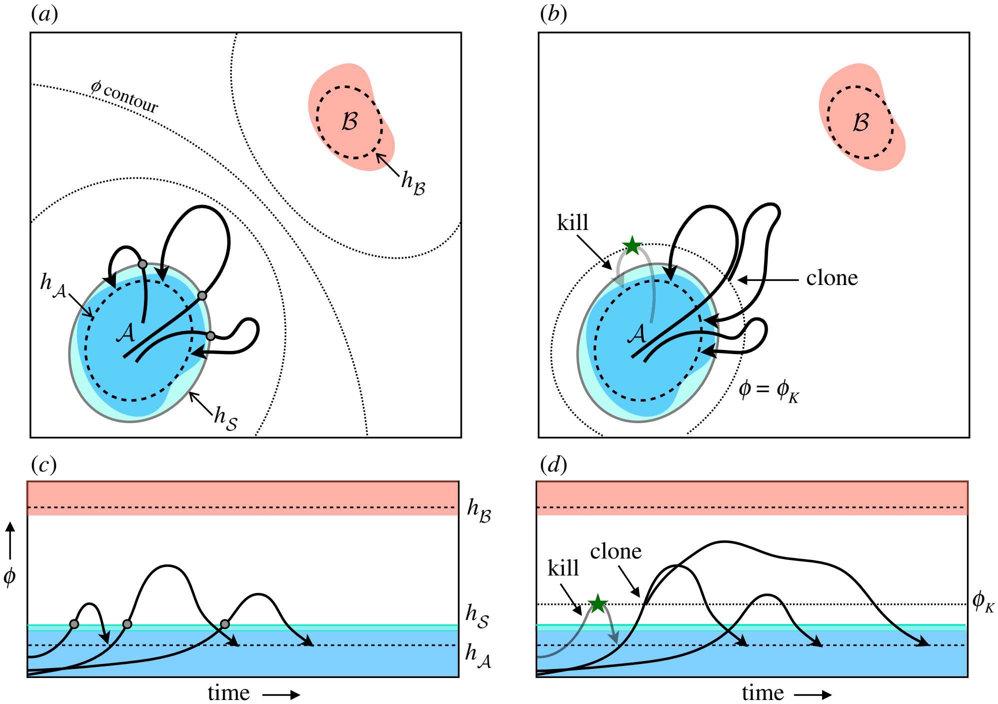

Figure 4. Schematic depiction of the AMS algorithm for a transition from $\mathcal{A}$ to $\mathcal{B}$. (*a*) The initialization of the algorithm. Contours are shown for the score function $\phi(\mathbf{u})$ and a hypersurface $\mathcal{S}$ surrounding $\mathcal{A}$. N trajectories are computed starting from random initial conditions in $\mathcal{A}$ that cross $\mathcal{S}$ and then either return to $\mathcal{A}$ or go to $\mathcal{B}$. (Here $N = 3$ and no initial trajectories reach $\mathcal{B}$.) (*b*) First iteration of the algorithm. The trajectory attaining the smallest maximum score function (here ϕ_K with $K = 1$) is killed, and a new trajectory is cloned from another randomly selected trajectory, resulting in an improved set of trajectories. The process is then iterated until a sufficient number of trajectories reach $\mathcal{B}$. Time series (*c*) and (*d*) correspond to the trajectories in (*a*) and (*b*). (Online version in colour.)

spaced in time, from a single trajectory that remains in $\mathcal{A}$ over a long time and thus samples the natural measure of states within $\mathcal{A}$.

The role of the hypersurface $\mathcal{S}$ is to ensure that after initialization, all trajectories in our sample have ventured from $\mathcal{A}$ at least as far as $\mathcal{S}$. Hence the maximum value of the score function obtained along each trajectory is at least $h_\mathcal{S}$. From the point of view of the score function, all trajectories in our initial sample have made some, possibly small, progress towards $\mathcal{B}$. Since $\mathcal{S}$ is chosen close to $\mathcal{A}$, the initialization step is not computationally demanding.

For the initialization and subsequent iterations, it is necessary to store the trajectories. In practice, we store full flow fields $\mathbf{u}_i(t_j)$ for each trajectory $i \in \{1, \ldots, N\}$ at sparsely spaced times $t_j = j\,\mathrm{d}T$, as a compromise between the large CPU times required for computing trajectories and the large memory needed to store them. The computations reported here all use a storage interval of $\mathrm{d}T = 320\,\mathrm{d}t = 10$, which is 10 times the sampling time δt used to collect statistics on trajectories.

(iii) Iteration

Iterative step m consists of discarding the K worst-performing trajectories and replacing them with trajectories obtained by cloning non-discarded trajectories. Specifically, we compute the maximal value $\phi_i^{(m)}$ of the score function along each trajectory and re-order the trajectories such that

$$\phi_1^{(m)} \le \phi_2^{(m)} \le \cdots \le \phi_K^{(m)} \le \cdots \le \phi_N^{(m)}.$$

We discard the K trajectories whose maximal values are lowest, in practice, a value $K^{(m)} \geq K$ because of possible equality of the maxima. Thus, in general we retain trajectories $\mathbf{u}_i$ such that $\phi_i^{(m)} > \phi_K^{(m)}$. We replace each discarded trajectory $\mathbf{u}_k(t)$ with a new trajectory constructed as follows:

(i) Choose at random (uniformly) one of the trajectories $\mathbf{u}_l(t)$ from the set of $N - K^{(m)}$ retained trajectories. Overwrite the trajectory $\mathbf{u}_k(t)$ with the part of the trajectory $\mathbf{u}_l(t)$ up to time t^{clone} at which the score function along $\mathbf{u}_l(t)$ first reaches $\phi_K^{(m)}$, i.e. $\phi(\mathbf{u}_l(t^{\text{clone}})) = \phi_K^{(m)}$(figure 4*b*). (Owing to the discrete sampling of stored trajectories, in practice, we copy trajectories until the score function first exceeds $\phi_K^{(m)}$.)

(ii) Modify $\mathbf{u}_l(t^{\text{clone}})$ with a low-amplitude multiplicative spectral perturbation as follows. Let

$$\boldsymbol{\eta}(x, y, z) = \sum_{m_x} \sum_{m_z} \sum_{m_y} \tilde{\boldsymbol{\eta}}_{m_x,m_y,m_z} s^{|m_x|+|m_y|+|m_z|} \, \mathrm{e}^{i(m_x k_x x + m_z k_z z)} T_{m_y}(y),$$

where each $\tilde{\boldsymbol{\eta}}_{m_x,m_y,m_z}$ is a vector whose components are uniform random complex numbers of modulus less than 1, s is a smoothing parameter such that $0 < s < 1$, and T_{m_y} is the Chebyshev polynomial of order m_y. Then the low-amplitude multiplicative perturbation at the cloning time is

$$\mathbf{u}_k(x, y, z, t^{\text{clone}}) = (\boldsymbol{I} + \epsilon \boldsymbol{\eta}(x, y, z))\mathbf{u}_l(x, y, z, t^{\text{clone}}) \tag{2.4}$$

where ϵ sets the size of the perturbation. The weak random perturbation is necessary to ensure that cloned trajectories do not exactly repeat the path of the trajectory from which they are cloned. Perturbations are always sufficiently weak that they leave the score function unchanged to at least four significant digits. Rolland [18] uses a similar approach in applying AMS to turbulence collapse in Couette flow. The remainder of the trajectory $\mathbf{u}_k(t)$ for $t > t^{\text{clone}}$ is obtained by simulating the new trajectory until it reaches $\mathcal{A}$ or $\mathcal{B}$ as before.

Once the $K^{(m)}$ discarded trajectories have been replaced (overwritten), we have a new set of N trajectories that are superior to the set at the start of the iteration, in the sense of being closer to reaching $\mathcal{B}$. Specifically, the maximum value of the score function for each of the new trajectories is now at least $\phi_K^{(m)}$. We increment m and repeat as necessary.

(iv) Stopping and post processing

Iterations end once the N samples have all reached $\mathcal{B}$. The final number of iterations is denoted by M. From the resulting trajectories and information gathered during the iteration process, we can construct estimators of relevant statistical quantities. Trajectories begin in $\mathcal{A}$, pass through $\mathcal{S}$ and terminate upon arrival at either $\mathcal{A}$ or $\mathcal{B}$. The estimator of the probability to go from $\mathcal{S}$ to $\mathcal{B}$ is given by [7]

$$\hat{p} = \prod_{m=1}^{M} \left(1 - \frac{K^{(m)}}{N}\right), \tag{2.5}$$

where $K^{(m)}$ is the number of trajectories eliminated at iteration m. The probability of going from $\mathcal{S}$ to $\mathcal{A}$ is $(1 - \hat{p})$ and that of going from $\mathcal{A}$ to $\mathcal{S}$ is 1.

The main quantity of interest is the mean first passage time τ from state $\mathcal{A}$ to state $\mathcal{B}$. For this, we will require the sample mean times available from the computations [8]. Let $T_{\mathcal{AS}} \equiv \inf\{t > 0, \mathbf{u}(t) \in \mathcal{S} | \mathbf{u}(0) \in \mathcal{A}\}$ and let $\overline{T}_{\mathcal{AS}}$ denote its sample mean obtained from trajectories whose initial conditions $\mathbf{u}(0)$ are selected from a long simulation lying within $\mathcal{A}$. Because $\mathcal{S}$ is close to $\mathcal{A}$, $\overline{T}_{\mathcal{AS}}$ is easily obtained from DNS (or from the initialization step of the AMS). Similarly, from the trajectories that cross $\mathcal{S}$ and return to $\mathcal{A}$ we can compute $\overline{T}_{\mathcal{SA}}$, the sample mean time to go from $\mathcal{S}$ to $\mathcal{A}$. Finally, from the N sample paths constructed as part of the AMS we can compute $\overline{T}_{\mathcal{SB}}$, the sample mean time to go from $\mathcal{S}$ to $\mathcal{B}$.

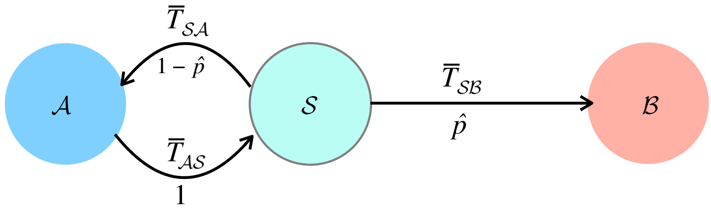

Figure 5. Schematic depiction of the data gathered via the AMS algorithm for a transition from $\mathcal{A}$ to $\mathcal{B}$ via $\mathcal{S}$. The probability $\hat{p}$ of transition from $\mathcal{S}$ to $\mathcal{B}$ is estimated, thus giving $1-\hat{p}$ as the probability for transition from $\mathcal{S}$ to $\mathcal{A}$. The sample mean times obtained for the two transitions are $\overline{T}_{\mathcal{SB}}$ and $\overline{T}_{\mathcal{SA}}$. From $\mathcal{A}$, all trajectories reach $\mathcal{S}$ (probability of one) and the sample mean time for this transition is $\overline{T}_{\mathcal{AS}}$. Trajectories begin at $\mathcal{A}$ and make some number of round trips between $\mathcal{S}$ and $\mathcal{A}$ before possibly reaching $\mathcal{B}$. (Online version in colour.)

From these quantities, the estimator for the mean first passage time τ is constructed as illustrated in figure 5. A trajectory going from $\mathcal{A}$ to $\mathcal{B}$ does so by going from $\mathcal{A}$ to $\mathcal{S}$ and back some number of times, n, before finally transitioning from $\mathcal{A}$ to $\mathcal{S}$ to $\mathcal{B}$. The probability of such a trajectory is $(1-\hat{p})^n\,\hat{p}$ and the mean time for all such trajectories is $(\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SA}})n+\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SB}}$. Summing over all possible n yields the estimator for τ:

$$\begin{aligned}\tau &= \sum_{n=0}^{\infty}(1-\hat{p})^n\hat{p}[(\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SA}})n+\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SB}}]\\ &= (\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SA}})\frac{1-\hat{p}}{\hat{p}}+(\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SB}}).\end{aligned} \tag{2.6}$$

We do not use separate notation for the true mean first passage time and this estimator of it. In describing the transition dynamics in terms of a Markov chain in figure 5, we rely on standard assumptions of the AMS algorithm, stated by Cérou *et al.* [8, p. 12].

The time $\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SA}}$ is the mean *non-reactive time*. This is the mean time for trajectories starting from within $\mathcal{A}$ to return to $\mathcal{A}$, conditioned on the fact that they reach $\mathcal{S}$. Similarly, $\overline{T}_{\mathcal{AS}}+\overline{T}_{\mathcal{SB}}$ is the mean *reactive time* for trajectories starting from within $\mathcal{A}$ to reach $\mathcal{B}$, conditioned on the fact that they do not return to $\mathcal{A}$. Neither the reactive time nor the non-reactive time is particularly large. What makes the mean first passage time large is that on average a trajectory will make many failed attempts to reach $\mathcal{B}$ so that the mean non-reactive time is multiplied by the large factor $(1-\hat{p})/\hat{p}$.

3. Computing mean passage times in channel flow

(a) Choice of the score function for band decay and splitting

The choice of the score function is critical for the AMS algorithm. In our case, we need functions that quantify the transition progress between the one-band state $\mathcal{A}$ and either the laminar state $\mathcal{B}_0$ (decay event) or the two-band state $\mathcal{B}_2$ (splitting event). We use slightly different score functions for decay and splitting.

We introduce the turbulent fraction, F_t, quantifying the proportion of the flow that is turbulent: $F_t=0$ for laminar flow, while $F_t=1$ for flow that is turbulent throughout the channel. For localized turbulent bands, the turbulent fraction is between zero and one. Specifically we define

$$e(z)\equiv\frac{1}{L_xL_y}\int_{-1}^{1}\int_0^{L_x}\frac{1}{2}(v^2+w'^2)\,\mathrm{d}x\,\mathrm{d}y \quad\text{and}\quad F_t\equiv\frac{1}{L_z}\int_0^{L_z}H(e(z)-e_{\text{thresh}})\,\mathrm{d}z, \tag{3.1}$$

where H is the Heaviside function. These quantities use the energy contained in the cross-channel and spanwise velocity components v and w', which is zero for laminar flow. Its cross-sectional

royalsocietypublishing.org/journal/rsta *Phil. Trans. R. Soc. A* **380**: 20210036

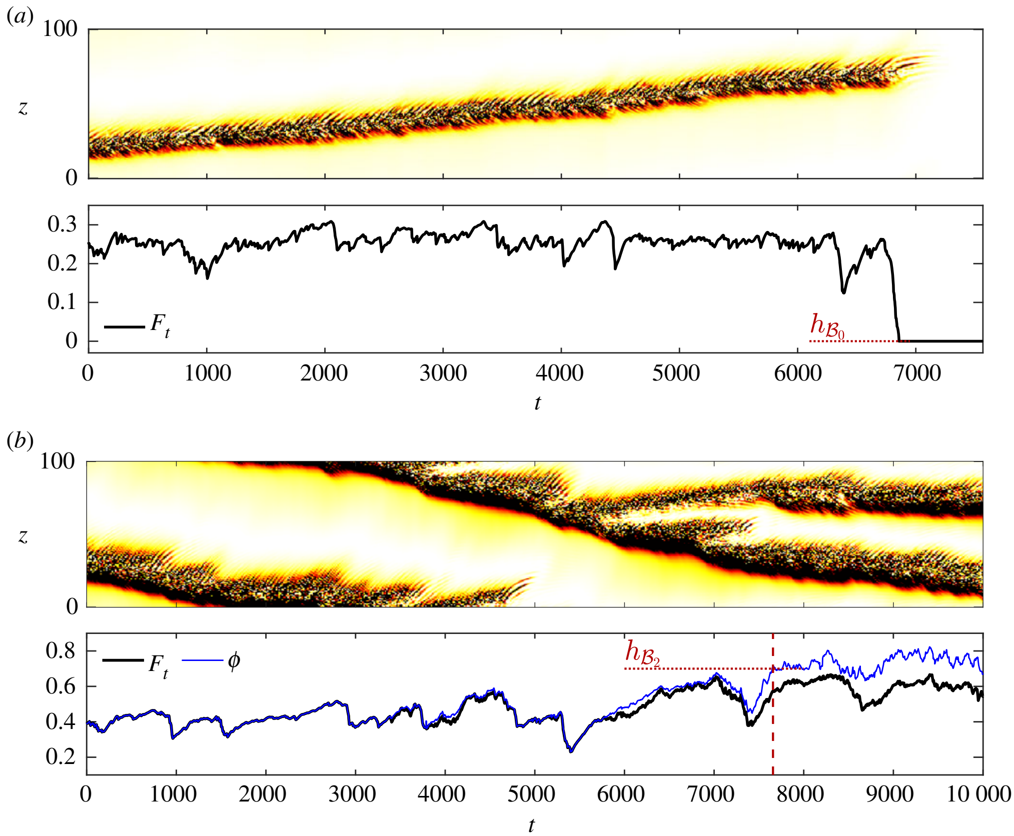

Figure 6. Evolution of the turbulent band during (*a*) a decay at $Re = 870$ and (*b*) a split at $Re = 1150$. Top: Spatio-temporal visualization. Colours show $(v^2 + w'^2)/2$ at $(x = 3.3, y = 0.8)$ (white: 0, black: 0.001). Bottom: Evolution of the turbulent fraction F_t (black curves) and of score function ϕ (thin blue curve) defined for splits in (3.2). (Online version in colour.)

integral $e(z)$ provides a good characterization of the turbulence as a function of z. We define the flow at location z to be turbulent if $e(z)$ exceeds the empirical threshold e_{thresh}, where $e_{\text{thresh}} = 1.1 \times 10^{-3}$. Figure 6*a* presents the typical life of a decaying band, repeated from figure 1, along with the corresponding time series of the turbulent fraction F_t. Local minima of F_t occur at local contractions of the band, which are themselves small detours towards the laminar state. Then F_t drops sharply to zero when the band transitions to the laminar state. In practice, we take $\phi = F_t$ and replace $<$ with $>$ (and max with min) as necessary in the algorithm. We define the system to be in $\mathcal{B}_0$ if $\phi < h_{\mathcal{B}_0} = 0.0001$ independently of Re, since all trajectories attaining this value of F_t are in the basin of attraction of the laminar state. The value $h_{\mathcal{A}}$ is taken as the most-probable value of the score function from a long simulation of the one-band state. As a result, $h_{\mathcal{A}}$ depends on Reynolds number. The level $h_{\mathcal{S}}$ is chosen to be approximately $0.8\,h_{\mathcal{A}}$. (See also tables 1 and 2 for definitions and values of all of these levels.)

We now consider the transition from one to two bands. Unlike for band decay, we have found that the turbulent fraction is not an adequate score function for band splitting. Figure 6*b* illustrates the difficulty. We see that before attaining the two-band state, multiple attempts to split occur. These deviations from the one-band state are characterized by widening of the initial band, possibly leading to the opening of a laminar gap between two turbulent regions. The resulting downstream turbulent patch then either decays, leading to a one-band state, or gains in intensity, ultimately leading to a steady second turbulent band whose shape and energy level are comparable to those of the initial band. The problem with using F_t as a score function is that while it captures the widening of the single band, it does not select for the intensification of downstream patches that results in a persistent secondary band. In figure 6*b*, the branching which will eventually lead to a new band occurs at $t \approx 5400$, but it is only at $t \approx 7660$ that this band becomes wider and more intense, acquiring some permanence and stability. It is this latter event that we will define as the split.

We have constructed an empirical but successful score function ϕ that encompasses the entire process of band stretching, captured by F_t, as well as separation into multiple bands and subsequent intensification of downstream bands. As can be seen by comparing the blue and black curves in figure 6*b*, ϕ does not differ greatly from F_t, but the difference is crucial for the performance of the AMS algorithm. The score function is given as follows. Consider the flow to consist of n_b turbulent bands, i.e. n_b distinct regions in which $e(z) > e_{\text{thresh}}$, as defined in (3.1). We associate with each turbulent band its width W_i in z, the laminar gap length L_i upstream until the next turbulent band, and finally its average energy E_i. We consider the mother band to be the band whose upstream laminar gap is maximal. Its properties are labelled (W_1, L_1, E_1), and the other bands i are ordered by downstream distance from the mother band. We then define the following empirical score function for splits:

$$\phi = F_t + \sum_{i=1}^{n_b} \frac{l_i}{L_z} \left(\frac{E_i}{E_{\max}}\right)^{\alpha} = \frac{1}{L_z} \sum_{i=1}^{n_b} \left[W_i + l_i \left(\frac{E_i}{E_{\max}}\right)^{\alpha} \right]. \tag{3.2}$$

Here, $E_{\max} \equiv \max_{1 \le i \le n_b} E_i$ and $l_i \equiv \sum_{j=2}^{i} L_j$ is the total laminar gap between band i and the mother band, which can describe continuously the collapse or splits of multiple child bands. The exponent α is chosen empirically to balance energy localization and turbulence spreading. In practice, we use $\alpha = 0.5$, in order to counteract the decrease in turbulent fraction usually observed after a split, as shown on figure 6*b* at $t = 7\,500$. In this way, we have enhanced the turbulent fraction by adding a function of band intensity E_i and of the total laminar distance l_i to the mother band. In this study, the level $h_{\mathcal{B}_2} = 0.7$ is found to capture a successful split: the presence of a lasting secondary band whose profile and intensity are similar to those of the initial band. We take $h_{\mathcal{S}} \simeq 1.2 h_{\mathcal{A}}$, with $h_{\mathcal{A}}$ the most probable value of (3.2) in the one-band state.

We have introduced a number of numerical parameters that could affect the performance and the accuracy of the computations. Of these, the selection of $h_{\mathcal{B}_2}$ and ϵ require the most care. Referring to figure 6*b* one sees that the threshold $h_{\mathcal{B}_2}$ must correctly capture the completion of a splitting event. As with the difficulty in defining a good score function for splitting, this is a reflection of our lack of good understanding of the splitting process. As can be seen in figure 6*a*, this issue does not arise for decay since the score function of the laminar state is known to be zero. Concerning the perturbation size ϵ used in the cloning, equation (2.4), one would ideally choose this to be small and independent of *Re*. In practice, we have found it necessary to vary ϵ with *Re*, and as discussed at the end of §3b, the current algorithm applied to decay events sometimes requires ϵ to be larger than desired. (See the electronic supplemental material for further discussion of the perturbation size ϵ and also the sample size N.)

(b) Simulating rare events with AMS

We have used the AMS algorithm to compute the mean decay and splitting times of an isolated turbulent band in a channel. These mean times are plotted as a function of Reynolds number in figure 7, where we also include previous results obtained via standard MC simulations [6]. The AMS results overlap with the MC data, but also substantially extend the range of accessible time scales. Both the AMS and MC results use the same tilted computational domain, the same spatial resolution, and the same underlying time-stepping code, as described in §2(a). This permits direct comparison of the two methods.

Figure 7 confirms the super-exponential dependence of the time scales found for decay and splitting events in wall-bounded shear flows [3–6]. From fits with $\tau_d = \exp(\exp(a_d\, Re + b_d))$ and $\tau_s = \exp(\exp(a_s\, Re + b_s))$ in the decay and split regimes, we find $Re_c \simeq 980$ as an improved

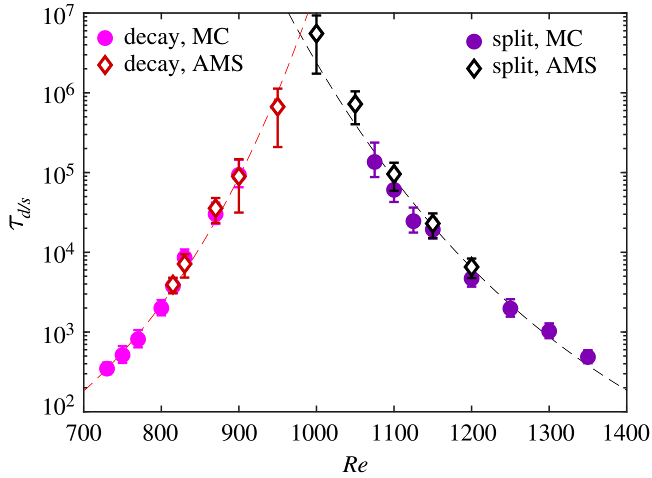

Figure 7. Mean decay times (red, magenta) and splitting times (black, purple) of turbulent bands as a function of Reynolds number, estimated with the Monte Carlo method (MC, circles) or with the AMS (diamonds). Error bars give confidence intervals for MC and are computed from multiple realizations of the algorithm for AMS. Dashed lines are best fits to double exponential form using the combined AMS and MC data: $\tau_d \simeq \exp[\exp(3.9 \times 10^{-3}\,Re - 1.09)]$; $\tau_s \simeq \exp[\exp(-2.6 \times 10^{-3}\,Re + 5.27)]$. (Online version in colour.)

estimate of the crossing Reynolds number for this flow configuration. (Previous fits to the MC data gave a crossing Reynolds number of 965.)

We recall a few details from the MC computations in [6]. The initial fields for the simulations are taken from snapshots of long-lasting bands simulated at $Re \in [900 - 1050]$. The Reynolds number is then changed to the desired value. Decay and splitting times from the start of the simulation are recorded. From these, the mean times and associated error bars are obtained [6]. The MC estimate of the transition probability $\hat{p}_{\mathrm{MC}}$ is computed from the multiple simulations by counting the number of decays or splits relative to the number of passages through $\mathcal{S}$. Typically $N = 40$ decay and splitting events are obtained at each Reynolds number. Fewer than $N = 40$ events were obtained by MC at the largest values of τ. With such techniques, only time scales $\tau < 10^5$ are currently accessible in practice.

The AMS initial fields are created from long-lasting bands, as in the MC method, except that each initial field is simulated for an additional relaxation time of 500 before commencing the AMS algorithm. The number of trajectories we seek to discard at each AMS iteration is $K = 1$. At each value of Re, the AMS algorithm is run N_{AMS} times, with each realization computing a sample of N trajectories. Each realization gives a value of τ calculated using (2.6), where $\overline{T}_{\mathcal{AS}} + \overline{T}_{\mathcal{SA}}$ is computed by DNS as part of the initialization step, $\overline{T}_{\mathcal{AS}} + \overline{T}_{\mathcal{SB}}$ is obtained from the AMS trajectories, and $\hat{p}$ is obtained via (2.5). Then the final estimate of τ is obtained by averaging over the N_{AMS} independent realizations.

Table 3 compares estimates of the transition probability $\hat{p}$ from the MC and AMS strategies. Both methods yield comparable estimates when MC results can be obtained. We emphasize that lifetimes τ change by orders of magnitude over the range of Re of interest, so we do not seek more than about one digit of accuracy in their values. The overall gain in computational speed achieved by the AMS over MC is measured by the total CPU time. One component of this cost is the CPU time per trajectory, for which the AMS shows a typical improvement of order $\mathcal{O}(10)$ and even $\mathcal{O}(100)$ for the low-transition-probability cases we considered; see $Re = 1000$ in table 4. For higher-transition-probability cases, AMS does not outperform MC because AMS requires N_{AMS} realizations to compensate for the variability in individual realizations. For low-transition-probability cases such as $Re = 1000$, only AMS is capable of inducing the very rare trajectories

Table 3. Results of Monte Carlo (MC) and Adaptative Multilevel Splitting (AMS). N is the number of samples for MC or for a single realization of AMS. For AMS, N_{AMS} is the number of realizations of the algorithm and ϵ is the perturbation amplitude used in cloning. The estimated transition probability and mean first passage time obtained by MC and AMS are $\hat{p}_{MC}$, τ_{MC} and $\hat{p}$, τ, respectively.

	Monte Carlo (MC)			Adaptive Multilevel Splitting (AMS)			
Re	N	$\hat{p}_{MC}$	τ_{MC}	ϵ	$N_{AMS} \times N$	$\hat{p}$	τ
870	40	0.081	3.0×10^4	5×10^{-4}	9×50	0.081	3.6×10^4
900	40	0.013	9.3×10^4	1×10^{-3}	7×50	0.015	8.9×10^4
1000	—	—	—	1×10^{-3}	3×50	0.00029	5.5×10^6
1150	40	0.047	2.1×10^4	1×10^{-5}	9×50	0.046	2.2×10^4

Table 4. Performance of Monte Carlo (MC) and Adaptative Multilevel Splitting (AMS). N is the number of samples for MC or for a single realization of AMS. For AMS, N_{AMS} is the number of realizations of the algorithm and ϵ is the perturbation amplitude used in cloning. The estimated CPU time per successful trajectory is given, as well as the total CPU time (both in processor hours on an HPE SGI 8600 computer).

	Monte Carlo (MC)			Adaptive Multilevel Splitting (AMS)			
Re	N	CPU_{traj}	CPU_{tot}	ϵ	$N_{AMS} \times N$	CPU_{traj}	CPU_{tot}
870	40	2500	1×10^5	5×10^{-4}	9×50	360	1.6×10^5
900	40	7500	3×10^5	1×10^{-3}	7×50	330	1.2×10^5
1000[a]	40	4×10^5	2×10^7	1×10^{-3}	3×50	1000	1.5×10^5
1150	40	5000	2×10^5	1×10^{-5}	9×50	500	2.2×10^5

[a]For $Re = 1000$, no estimator of the time scale could be achieved by Monte Carlo, so the CPU times are extrapolated from the AMS estimator τ.

which are out of reach for the MC method (see electronic supplemental material for further comparisons).

The results from AMS show larger variability than those from MC, especially for decay cases, as seen by the error bars on figure 7. It is known that the standard deviation of the estimated probability for AMS will decrease as $1/\sqrt{N}$ (at least in ideal cases) [10,33]. For our high-dimensional system, N is restricted by computational costs. Using N larger than 100 is not practical and we typically use $N = 50$. We observe that the large variability between different realizations of the AMS algorithm is associated with variability in the initialization, especially the extent to which the initial trajectories are a representative sample.

The amplitude ϵ of the perturbation that we use in cloning trajectories is chosen to promote separation of the trajectories. The only issue occurs for rare decay ($Re \in [900 - 950]$) where the amplitude must be increased ($\epsilon > 10^{-2}$ at $Re = 950$). In these cases, cloned trajectories resulting from small-amplitude perturbations separate from one another only after having reached their minimum F_t value. Hence they do not acquire an improved score function, causing the algorithm to stagnate. The reason for this is that the duration of the approach to the minimum of F_t is shorter than the Lyapunov time of the system. This limitation of our current procedure has been observed in other studies [15,18] and has been addressed in [18] by anticipating branching. This technique clones trajectories prior to where one would in the standard algorithm, thus promoting the separation of trajectories near the minimum of F_t.

4. Extreme value description of decay and splitting trajectories

The super-exponential dependence of lifetime of turbulence on Reynolds numbers seen in figure 7 is ubiquitous for decay and splitting events in wall-bounded shear flows, e.g. [3–6,34]. Goldenfeld *et al.* [19] have formulated a hypothesis explaining decays through extreme value theory. The main idea is to associate the decay of a turbulent patch to the statistical distribution of the largest fluctuation over some space–time interval. If the maximum amplitude of fluctuations becomes lower than some threshold, then the multiple fluctuations comprising a turbulent zone will all laminarize. This connects laminarization to the distribution of extrema of a set of random variables. Just as the Central Limit Theorem states that under very general conditions the limit of the sum of independent and identically distributed random variables is a Gaussian, the Fisher–Tippett–Gnedenko theorem [35] states that the extrema of a set of n independent and identically distributed variables should follow a Fisher–Tippett distribution. Goldenfeld *et al.* assumed that the decay threshold depends on Re and approximated that dependence locally as linear. This linear dependence translates into a super-exponential dependence of the lifetimes on Re via properties of the Fisher–Tippett distribution.

In a study of the decay of turbulent puffs in pipe flow, Nemoto & Alexakis [21] found that the maximal vorticity over the domain followed a Fréchet distribution, a member of the Fisher–Tippett family. Moreover, they found that the parameters of this distribution depend linearly on Re over a range of 75 in Re near the critical value Re_c. Similar to the Goldenfeld *et al.* argument, this linear dependence on parameters translates to a super-exponential dependence of the lifetimes on Re. Thus, Nemoto & Alexakis were able to directly relate extreme values to the super-exponential evolution with Re of the puff decay times in pipe flow. Other quantities related to the distance to the laminar attractor have been shown to follow the extreme value law [36,37], particularly when a maximal or minimal value is extracted from a divided time series [38].

Here, we explore these ideas for both the decay and splitting of turbulent bands in channel flow over a substantial range of Re. To do so, we must link the rare events (decay or split) with some observable that follows an extreme distribution. Rather than speculate on which variable or combination of variables are mechanistically responsible for driving decay and splitting events, we choose to focus on F_t for both transitions. Our reasoning is that turbulence fraction is a useful observable of general interest that is easily obtainable in computations and experiments. As we show below, the turbulent fluctuations and reaction pathways project onto F_t and allow us to analyse the connection between fluctuations and the rare events. As a practical matter, it is helpful to study distributions of a quantity that is (or is closely related to) the score function used to obtain rare events.

(a) Probability densities of turbulent fraction

We begin by showing in figure 8 the probability density function (PDF) of the turbulent fraction $p(F_t)$ for a variety of Reynolds numbers. These PDFs have been constructed from MC simulations that start, after initial equilibration time, from the one-band state $\mathcal{A}$ and terminate at the end of a decay or split. The distributions have a clear asymmetry about their maxima and they have broad tails that depend on Re: the low-F_t tails are present at lower Re while high-F_t tails are present at higher Re. To our knowledge, this is the first report of $p(F_t)$ in any transitional shear flow.

We find that the central portions of these PDFs are closely approximated by Fisher–Tippett distributions. The cumulative distribution function (CDF) of the Fisher–Tippett (also called Generalized Extreme Value) distribution that we will use is

$$\mathbb{P}(X \le h) = P_{FT}(h) \equiv 1 - \mathrm{e}^{-(1+\xi(\mu-h)/\sigma)^{-1/\xi}}, \tag{4.1}$$

where the location μ, scale σ and shape ξ are fitting parameters. Equation (4.1) is the CDF for minima of a set of random variables, and it is this form that fits our data. We fit $p(F_t)$ with the Fisher–Tippett density $p_{FT}(h) = \mathrm{d}P_{FT}/\mathrm{d}h$ shown as dashed curves on figure 8. (The resemblance of the abbreviation FT for Fisher–Tippett and the notation F_t for turbulent fraction is coincidental.)

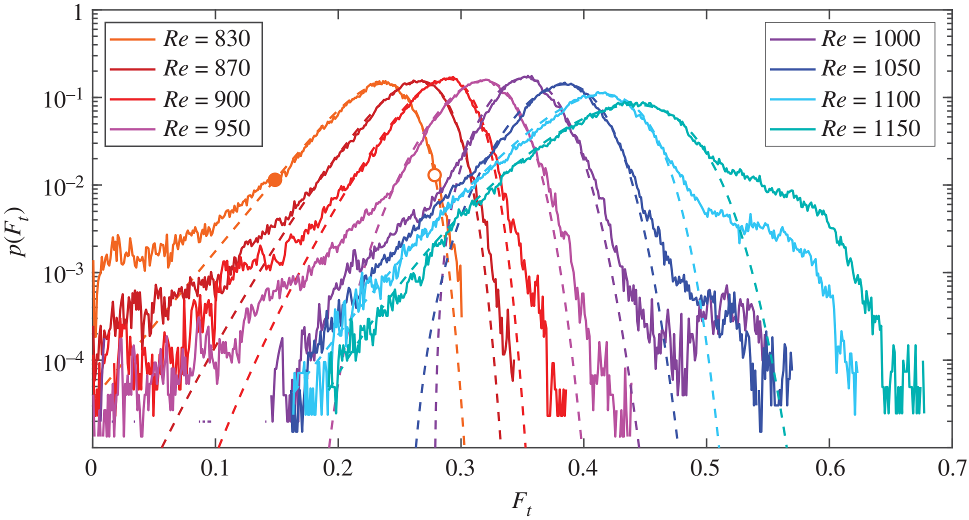

Figure 8. Probability density function of the turbulent fraction around the one-banded state $\mathcal{A}$. Dashed lines correspond to fits with a Fisher–Tippett probability density, the derivative of (4.1). Fits are carried out over intervals $[h_{\text{left}}, h_{\text{right}}]$, shown for the case $Re = 830$ by coloured and white points. Values of h_{left} and h_{right} are given in table 2. (Online version in colour.)

Figure 8 shows that the central region near the maximum of each PDF fits well with the Fisher–Tippett distribution inside a range spanning from h_{left} to h_{right}. As an example, these lower and upper bounds of the fit are indicated by coloured and white circles for $Re = 830$. The quality of the fit is particularly good for $Re < 1000$ but shows some noticeable deviations at $Re = 1000$ and $Re = 1050$. The fitting parameter values as a function of the Reynolds number are plotted in figure 10*a*, which will be discussed below.

The turbulence fraction F_t defined in equation (3.1) is not a maximum of a set of independent quantities (although it includes a Heaviside function which, like the maximum, is a non-analytic operation). Hence, it is not obvious that F_t should be governed by an extreme value distribution. Even in the case of vorticity maxima, Nemoto & Alexakis noted that it is not possible to fully justify Fisher–Tippett distributions since vorticity is correlated in space and time and hence the maxima are not independent. At present we do not have a formal justification for the fits used in figure 8 other than that the distributions are clearly non-Gaussian and are fit reasonably well with the Fisher–Tippett form. We hypothesize that the strong spatio-temporal correlations within the localized turbulent bands play a significant role in the statistics, but we leave this for further investigation. The only way the fits will enter into the analysis that follows is via their parameterization. In this regard the fits give us a useful representation of the PDFs in terms of three parameters depending on *Re*. It is nevertheless possible that the distributions are of some other type.

The Nemoto & Alexakis approach requires many numerical simulations of rare events in order to obtain the tails of probability distributions. Here, the AMS approach is particularly useful as it produces large samples of the rare event trajectories that reach destination $\mathcal{B}$. From the AMS data one can reconstruct the CDF of any observable X depending on a field $\mathbf{u}$ as follows. Each point on a trajectory $\mathbf{u}(t)$ is known to be on a segment from $\mathcal{A}$ to $\mathcal{S}$, from $\mathcal{S}$ to $\mathcal{A}$ or from $\mathcal{S}$ to $\mathcal{B}$. (See figure 5.) Hence the CDF can be decomposed into a weighted sum of independent CDFs conditioned on the location of $\mathbf{u}$:

$$\mathbb{P}(X \le h) = \frac{\tau_{\mathcal{AS}}}{\tau}\,\mathbb{P}(X \le h|\mathcal{C}_{\mathcal{AS}}) + \frac{\tau_{\mathcal{SA}}}{\tau}\,\mathbb{P}(X \le h|\mathcal{C}_{\mathcal{SA}}) + \frac{\tau_{\mathcal{SB}}}{\tau}\,\mathbb{P}(X \le h|\mathcal{C}_{\mathcal{SB}}), \tag{4.2}$$

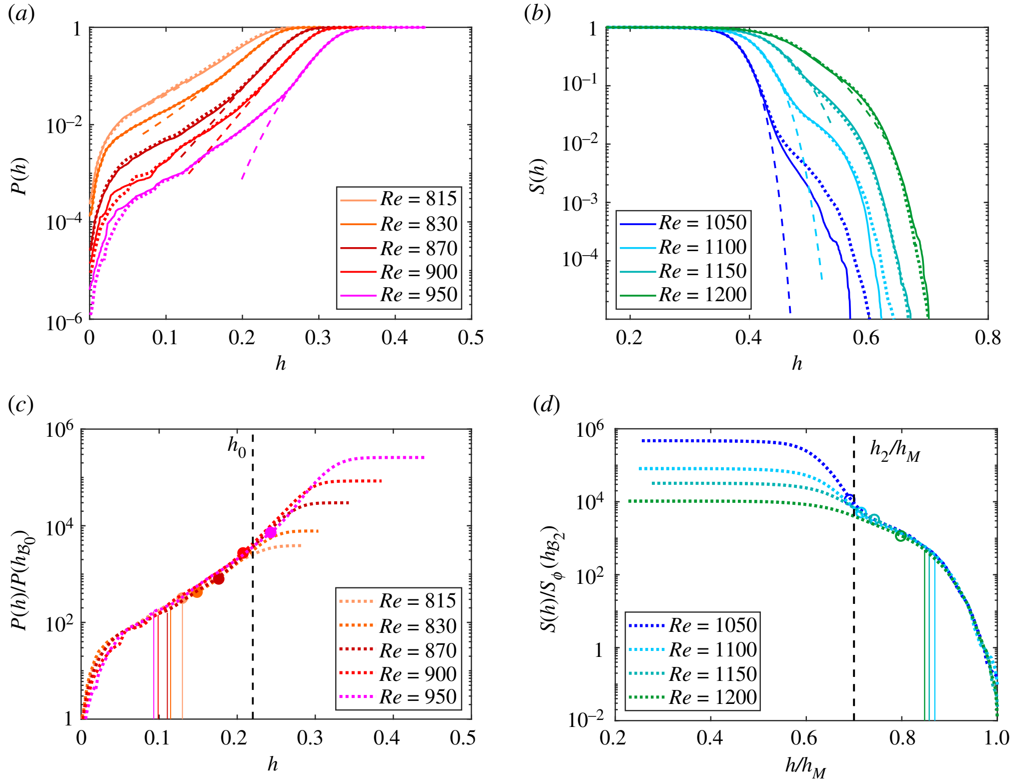

Figure 9. (*a*) Cumulative distribution function $P(h) = \mathbb{P}(F_t \leq h)$ for band decay and (*b*) survival function $S(h) \equiv 1 - P(h) \equiv \mathbb{P}(F_t \geq h)$ for band splitting at values of *Re* indicated in the legend. Continuous lines are obtained from Monte Carlo and dotted lines are from the AMS algorithm. Dashed lines correspond to fits to a Fisher–Tippett distribution (4.1). (*c*,*d*) Distributions from the AMS algorithm rescaled by $P(h_{\mathcal{B}_0})$ and $S_\phi(h_{\mathcal{B}_2}) \equiv 1 - P_\phi(h_{\mathcal{B}_2})$. In the splitting case (*d*), the range in F_t is rescaled by $h_M(Re) = \max(F_t)$. Coloured points in (*c*) show h_{left}, the lower bounds of the fit to the PDF with a Fisher–Tippett density function (figure 8). Similarly the open points in (*d*) show the upper bounds h_{right}. Vertical lines show the break-even points defined in the text. (Online version in colour.)

where $\mathcal{C}_{\mathcal{AS}}$ (resp. $\mathcal{C}_{\mathcal{SA}}$ and $\mathcal{C}_{\mathcal{SB}}$) is the conditional event that a field $\mathbf{u}$ lies on a trajectory that goes from $\mathcal{A}$ to $\mathcal{S}$ (resp. from $\mathcal{S}$ to $\mathcal{A}$ or to $\mathcal{B}$). The weights are the relative time spent in each segment, where

$$\begin{aligned}\tau &= \tau_{\mathcal{AS}} + \tau_{\mathcal{SA}} + \tau_{\mathcal{SB}}\\ &= \frac{1}{\hat{p}}\,\overline{T}_{\mathcal{AS}} + \frac{1-\hat{p}}{\hat{p}}\,\overline{T}_{\mathcal{SA}} + \overline{T}_{\mathcal{SB}}.\end{aligned}$$

We refer the reader back to equation (2.6) for the formula for τ in terms of $\overline{T}_{\mathcal{AS}}$, etc. The individual CDFs in (4.2) are obtained in the standard way by rank ordering the sample data and performing a cumulative summation followed by normalization.

Figure 9*a* shows the CDF $P(h) = \mathbb{P}(F_t \leq h)$ for the low-*Re* decay cases and figure 9*b* shows its complement, the survival function $S(h) \equiv 1 - P(h) \equiv \mathbb{P}(F_t \geq h)$, for the high-*Re* splitting cases. Results from the MC simulations are shown as continuous curves, while those from AMS have been included as dotted curves. It can be seen that the distribution functions constructed from AMS improve the quality of the tails from MC, particularly in the range $900 \leq Re \leq 1100$ where MC systematically underestimates the tails associated with rare transitions. (We note, however, that even with the improvements from the AMS, there remain some sampling effects in the weak tails.) Dashed curves show the Fisher–Tippett CDFs obtained by fitting the PDFs of F_t shown in figure 8.

(b) Timescales from extreme value distributions

We can now apply the Nemoto & Alexakis approach [21] to our decay and splitting data. The essential idea is to scale the CDFs and obtain forms that separate into approximately *Re*-independent portions and *Re*-dependent portions that can be fit to Fisher–Tippett distributions. From this it is possible to express the mean timescales for decay and splitting directly in terms of the parameters of the Fisher–Tippett distributions.

We will first describe the decay case and afterwards summarize the differences for the splitting case. Recall that in the decay case the score function for AMS is just the turbulence fraction and the boundary of the laminar state is $h_{\mathcal{B}_0} = 0.0001$, meaning that trajectories $\mathbf{u}(t)$ that reach the threshold $F_t(\mathbf{u}) = \phi(\mathbf{u}) = h_{\mathcal{B}_0}$ from above are considered to have undergone transition to the laminar state. As shown in figure 9*c*, by rescaling CDFs by their value at the threshold $P(h_{\mathcal{B}_0})$, the low-probability tails for different *Re* nearly collapse to a common curve. More specifically, we observe that below a value h_0, indicated on the plot, the ratio $P(h)/P(h_{\mathcal{B}_0})$ depends only weakly on *Re*. (Moreover, some of this dependence is likely due to sampling errors of the low-probability tails.) Flow fields $\mathbf{u}$ such that $F_t(\mathbf{u}) \in [h_{\mathcal{B}_0}, h_0]$, called the *collapse zone* in the following, are in an intermediate state that can either recover (missed decay) or die (successful decay). This process is not a strong function of *Re*. Above h_0, the rescaled CDFs depend strongly on *Re*, varying by over an order of magnitude over the *Re* range shown. Significantly, however, for almost all *Re* this portion of the CDFs lies within the region that is well fit by the Fisher–Tippett distribution. Concretely, the coloured points in figure 9*c* indicate the left-most values of h for each *Re* for which the Fisher–Tippett fits are good and in almost all cases, these points are below h_0, with the point for $Re = 950$ slightly above h_0.

Following Nemoto and Alexakis, we can connect the CDFs to decay lifetime τ_d. The algebraic statement is

$$\tau_d = \frac{\delta t}{P(h_{\mathcal{B}_0})} = \delta t \frac{P(h_0)}{P(h_{\mathcal{B}_0})} \frac{1}{P(h_0)} \simeq \underbrace{\delta t \frac{P(h_0)}{P(h_{\mathcal{B}_0})}}_{\Pi_d} \underbrace{\frac{1}{1 - \mathrm{e}^{-(1+\xi(\mu - h_0)/\sigma)^{-1/\xi}}}}_{f_d(Re)}, \tag{4.3}$$

which we will explain in steps.

The first equality can be understood as follows [21]. Consider estimating τ_d by MC simulation with N_{decay} independent realizations of decay events. Then $\tau_d = T_{\mathrm{total}}/N_{\mathrm{decay}}$, where T_{total} is the total combined time to decay for all realizations. Further letting $T_{\mathrm{total}} = \delta t\, N_{\mathrm{total}}$, where N_{total} is the total number of sample points on all trajectories and δt is the sample time, we have $\tau_d = \delta t\, N_{\mathrm{total}}/N_{\mathrm{decay}}$. Finally, from N_{decay} simulations that terminate at $h_{\mathcal{B}_0}$, we have $P(h_{\mathcal{B}_0}) = N_{\mathrm{decay}}/N_{\mathrm{total}}$, since there are N_{decay} out of N_{total} sample points with $F_t \leq h_{\mathcal{B}_0}$. In practice, we construct $P(h_{\mathcal{B}_0})$ from AMS simulations via (4.2) with a sampling time $\delta t = 1$.

The remainder of (4.3) consists of multiplying and dividing by $P(h_0)$ and then applying the previous observations about figure 9*c* to decompose (4.3) into a factor Π_d, that depends only weakly on *Re*, and $1/P(h_0)$, that depends strongly on *Re*. Furthermore, we approximate $P(h_0)$ by the Fisher–Tippett distribution evaluated at h_0. The *Re*-dependence of $f_d \simeq 1/P(h_0)$ is contained in the *Re*-dependence of the parameters μ, σ and ξ. We return to this after discussing the splitting case.

In almost all respects the splitting analysis is the same as that of the decay case. The only important differences comes from the fact that the score function ϕ for splitting (3.2) is not the turbulence fraction F_t. However, ϕ and F_t are closely related, both in terms of expression (3.2) and in terms of the values they take during band splitting in figure 6*b*. A split is deemed to have occurred when $\phi(\mathbf{u}(t))$ reaches $h_{\mathcal{B}_2}$ from below. Hence, analogously with (4.3), the time scale for splits is related to the survival function of ϕ evaluated at $h_{\mathcal{B}_2}$:

$$\tau_s = \frac{\delta t}{\mathbb{P}(\phi > h_{\mathcal{B}_2})} = \frac{\delta t}{S_\phi(h_{\mathcal{B}_2})}, \tag{4.4}$$

where S_ϕ is the survival function for ϕ. While one could analyse distributions of the score function ϕ, the turbulence fraction is ubiquitous in this field and the distributions in figures 8 and 9*b* are of general interest. Hence it is preferable to work with these distributions, even though it will be necessary to rescale the CDF in figure 9*b* using $S_\phi(h_{\mathcal{B}_2})$. This is not as awkward as it may seem since $S_\phi(h_{\mathcal{B}_2}) = N_{\text{split}}/N_{\text{total}}$, by the same argument as above for decay. Hence, while we write the normalization in terms of S_ϕ, it is not necessary to have access to this CDF to know the normalization, which is determined simply from the number of sample points and the number of splitting cases. To collapse the CDFs we must also rescale the horizontal axis of figure 9*b*. We rescale by h_M, the maximum value of F_t observed at each Re. This was unnecessary in the decay case because the minimum value of F_t is achieved at the Re-independent termination value $h_{\mathcal{B}_0}$.

Figure 9*d* shows the rescaled CDFs for band splitting. We observe that the low-probability tails for different Re collapse well to a common curve $h \geq h_2$, while for $h < h_2$ the rescaled CDFs depend strongly on Re. Also shown as points in figure 9*d* are the upper limits for which the curves are well approximated by Fisher–Tippett distributions. These points are above, or nearly above h_2 in all cases. Hence, we can again exploit this to approximate the splitting time scale in terms of parameters of the Fisher–Tippett distributions. Starting from (4.4) the algebra is

$$\tau_s = \frac{\delta t}{S_\phi(h_{\mathcal{B}_2})} = \delta t\, \frac{S(h_2)}{S_\phi(h_{\mathcal{B}_2})} \frac{1}{S(h_2)} \simeq \underbrace{\delta t\, \frac{S(h_2)}{S_\phi(h_{\mathcal{B}_2})}}_{\Pi_s} \underbrace{\mathrm{e}^{(1+\xi(\mu-h_2)/\sigma)^{-1/\xi}}}_{f_s(Re)}. \tag{4.5}$$

We thus obtain an approximation for τ_s as a product of a factor Π_s, weakly dependent on Re, and a factor $f_s(Re)$, strongly dependent on Re via the parameters μ, σ, ξ, as well as h_2. Note that h_2/h_M is constant at the start of the collapse zone, but h_M depends on Re, and hence so does h_2. Values of h_2 and h_M, as well as h_0, are given in table 2.

Finally, the vertical lines in figure 9*c*,*d* indicate the break-even point for transition events to take place. These have been determined from DNS trajectories that originate in $\mathcal{A}$ as follows. For a given value of h, we compute the fraction of trajectories attaining $F_t = h$ that successfully transition to $\mathcal{B}_0$ or $\mathcal{B}_2$, without returning to $\mathcal{A}$. The value of h for which this fraction is $1/2$ is the break-even point. This is conceptually similar to finding where the *committor function* for a stochastic process [39] is equal to $1/2$, but here we condition on values of the turbulence fraction and not points in phase space. At $Re = 1050$ we have not obtained a sufficient number of DNS trajectories undergoing transition to $\mathcal{B}_2$ to estimate the break-even point, and hence this case is not included in figure 9*d*. We provide context for these break-even points in the next section.

(c) Super-exponential scaling

We now explore the connection between the observed super-exponential dependence of mean lifetimes on Re seen in figure 7 and the approximations to the mean lifetime given in (4.3) and (4.5). We have argued that the dominant dependence of mean lifetimes on Re is captured by the dependence of the functions f_d and f_s on Re. These functions depend on Re via the Fisher–Tippett parameters μ, σ and ξ of (4.1) which are shown in figure 10*a*. The location parameter μ varies linearly with Re, a feature which can already be seen in the Re-dependence of the maxima in figure 8. The Re-dependence of the scale σ and the shape ξ is less clear; their fluctuations may be due to their sensitivity to the fitting procedure. Since the quality of the fits in figure 8 is not improved by the inclusion of more simulation data, the fluctuations may indicate that $p(F_t)$ is not exactly of Fisher–Tippett form even near its maximum.

The parameter ξ plays an essential role in the family of Fisher–Tippett distributions, dividing them into three categories. Those with $\xi > 0$ are the Fréchet distributions (also known as type II

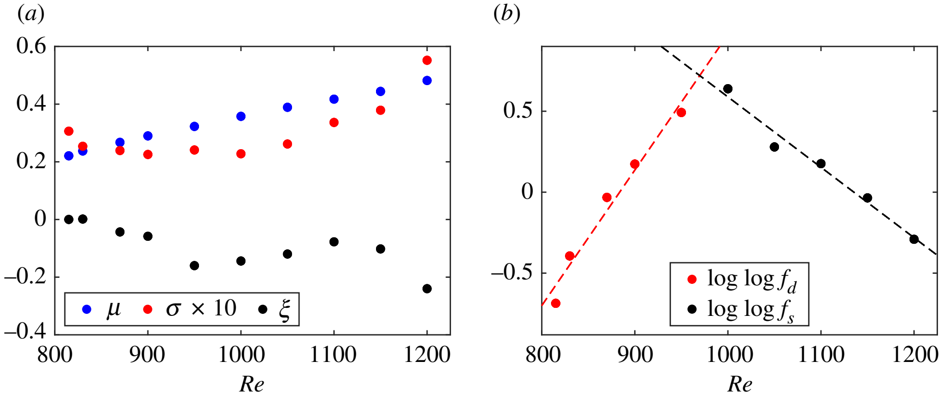

Figure 10. (*a*) Dependence of the three Fisher–Tippett parameters on *Re*. These have been obtained by fitting Fisher–Tippett distributions to the numerical PDFs $p(F_t)$ over ranges $h_{\rm left} \leq F_t \leq h_{\rm right}$ as seen in figure 8. (*b*) Dependence of log log f_d (4.3) and log log f_s (4.5) on *Re* using the parameters from (*a*). Dashed lines show linear fits. (Online version in colour.)

extreme value distributions), while $\xi < 0$ corresponds to Weibull (type III). Figure 10*a* shows that the central portions of most of the curves in figure 8 are best fit to Weibull distributions (ξ may be positive for $Re = 815$ and 830, but there is too much uncertainty in our fits to be sure). The limiting case $\xi = 0$ is the family of Gumbel distributions (type I), which will play a role in §5.

Figure 10*b* shows $\log\log f_d$ and $\log\log f_s$ from expressions (4.3) and (4.5) as a function of *Re* using the numerically obtained parameter values for each *Re*. Linear fits show that $\log\log f_d \simeq a_d\, Re + b_d$ and $\log\log f_s \simeq a_s\, Re + b_s$ over a range of nearly 200 in *Re* in each case. Hence both f_d and f_s depend super-exponentially on *Re* and are at least approximately of the form $[\exp(\exp(a\, Re + b))]$. Given the functional forms of f_d and f_s and the complicated dependence of the fitting parameters on *Re*, the double exponential dependence on *Re* is only an approximation. Nevertheless, we clearly observe a faster than exponential dependence on *Re* resulting from modest variation with *Re* of parameters of the Fisher–Tippett distribution characterizing the fluctuations in the one-band state.

The interpretation of these results comes from the mechanism proposed by Goldenfeld *et al.* [19] and subsequently refined by Nemoto & Alexakis [20,21]. We focus on the decay case, but similar statements apply to the splitting case. The picture is that the statistics of strong turbulent fluctuations are governed by extreme value distributions and this gives rise to the strong *Re* dependence of the probability $P(h_0)$ of states being in the collapse zone $h \leq h_0$; see figure 9*c*. Note that most trajectories that enter the collapse zone do not decay directly, but instead return to the one-band state $\mathcal{A}$. Only when trajectories achieve values of F_t below the break-even points (shown as vertical lines in the figure) are trajectories more likely to decay than to return to $\mathcal{A}$. The probability of decay becomes one at $h_{\mathcal{B}_0}$, since this defines the boundary we have chosen for the laminar state $\mathcal{B}_0$, and the rate of ultimate decay is given by $P(h_{\mathcal{B}_0})$ which is much less than $P(h_0)$. However, the ratio $P(h_0)/P(h_{\mathcal{B}_0})$ is nearly independent of Reynolds number. Hence up to a *Re*-independent multiplicative factor, the decay rate is determined from probability $P(h_0)$. The reason why the CDFs for different *Re* collapse over a range of turbulence fractions, and why this occurs for both decay and splitting processes, remains unexplained.

We end this section with a few observations and caveats. We observe that PDFs of F_t are well fit near their maxima by Weibull distributions, at least for most of the *Re* range investigated. This is distinctly different from the Fréchet distributions observed by Nemoto & Alexakis [21] for maximum vorticity in pipe flow. We note also that while F_t is a non-smooth function of the flow field, it is not given as an extremum over any feature of the field.

The purpose of decomposing the mean lifetimes (4.3), (4.5) and using the Fisher–Tippett parameter fits is not to obtain quantitatively accurate formulas for τ_d and τ_s, but to gain insight into the source of the super-exponential dependence on *Re*. In this regard we note that the biggest issue, both quantitative and conceptual, with this approach is that we rely on the existence of delimiters h_0 and h_2 that are simultaneously within the collapse zones and within the range in which the distributions are close to Fisher–Tippett form. As can be seen in figure 9*c*,*d*, this does not hold for $950 \lesssim Re \lesssim 1050$. This was also observed for puff decay in pipe flow: figure 10a in [21]. This does not invalidate the connection between extreme value statistics and the super-exponential scaling, but it does mean that there is a gap in using the Fisher–Tippett approximation at large time scales that at present we do not see how to close.

5. Transition pathways

Extreme value theory not only relates the super-exponential scaling of mean lifetimes to the distribution of fluctuations of the one-band state, it also provides a framework for understanding the rare pathways from one state to another. In a previous publication [6] we observed that the dynamics of band splitting were concentrated around a most-probable pathway in the phase space of large-scale Fourier coefficients. This motivates us to explore connections with *instantons* in the framework of Large Deviation theory for systems driven by weak random perturbations. See for example [40–42] and references therein. The concept is easily illustrated with the following stochastic differential equation

$$\dot{X} = -\nabla V(X) + \sqrt{\varepsilon}\eta, \tag{5.1}$$

where $X \in \mathbb{R}^d$, V is a potential, ε is a perturbation strength and η is Gaussian white noise. We assume that V has two local minima $\mathcal{A}$ and $\mathcal{B}$ separated by a saddle point and we consider transitions from $\mathcal{A}$ to $\mathcal{B}$. In the weak noise limit $\varepsilon \to 0$, transitions will be rare and the trajectories associated with these rare events will be concentrated around a most-probable path that connects states $\mathcal{A}$ and $\mathcal{B}$. This is the *instanton*. The dynamics along the instanton is such as to climb uphill from $\mathcal{A}$ to the saddle point under the influence of weak noise, and then to fall deterministically from the saddle to $\mathcal{B}$.

Examples of instantons in fluid systems are found for shocks in Burgers equations [41,43], and have been predicted and experimentally observed in rogue waves [44]. The concentration of transition paths around an instanton in a high-dimensional fully turbulent system was first observed by Bouchet *et al.* [13] in a 2D barotropic model of atmospheric dynamics. Schorlepp *et al.* [45] have used instanton calculus to investigate the most likely configurations to generate large vorticity or strain within turbulence in the 3D Navier–Stokes equations. This phenomenology can also apply to deterministic chaos, as in the solar system [46]. Rolland has discussed instantons specifically in relation to turbulent–laminar transition, both in a model system [16] and in plane Couette flow [18].

Rare transitions of the type considered here could exhibit instanton-type behaviour if turbulent fluctuations were to play the role of weak noise. A detailed investigation is outside the scope of this paper, but the current interest in the topic and the capacity of AMS to generate large numbers of rare transitions motivates us to briefly present transition paths for decays and splits. Examples of each are shown in figure 11. By binning samples from 200 transition paths we construct PDFs and then render isosurfaces of these PDFs to reveal the reactive tubes where paths concentrate. We include only reactive trajectories that leave $\mathcal{A}$ and terminate at the boundary of $\mathcal{B}_0$ or $\mathcal{B}_2$ without returning to $\mathcal{A}$.

The coordinates used for the PDF are chosen separately for decay and splitting. For decay, we show the decay of energy associated with the three velocity components of the flow, E_x, E_y, E_z

$$E_{(x,y,z)} \equiv \frac{1}{L_x L_y L_z} \int_{\Omega} \frac{1}{2}(u^2, v^2, w^2)\, \mathrm{d}\mathbf{x}.$$

Figure 11*a* shows that the reactive pathway from $\mathcal{A}$ to $\mathcal{B}_0$ is such that E_y decays most quickly, followed by E_z, followed by E_x so that the tube of reactive trajectories approaches $\mathcal{B}_0$ almost tangent to the E_x axis. This ordering of decay of energy components has been reported previously

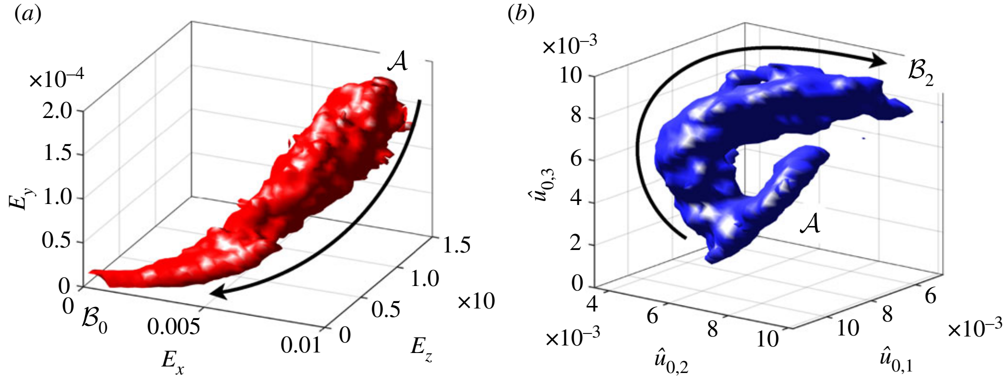

Figure 11. Joint probability density functions for reactive trajectories going from (*a*) $\mathcal{A}$ to $\mathcal{B}_0$ at $Re = 830$ (decay events) or (*b*) from $\mathcal{A}$ to $\mathcal{B}_2$ at $Re = 1150$ (splitting events). (*a*) Isosurface of $p(E_x, E_y, E_z)$ enclosing 90% of the total probability. (*b*) Isosurface of $p(\hat{u}_{0,1}, \hat{u}_{0,2}, \hat{u}_{0,3})$ enclosing 80% of the total probability. Two-hundred trajectories are computed in each case. (Online version in colour.)

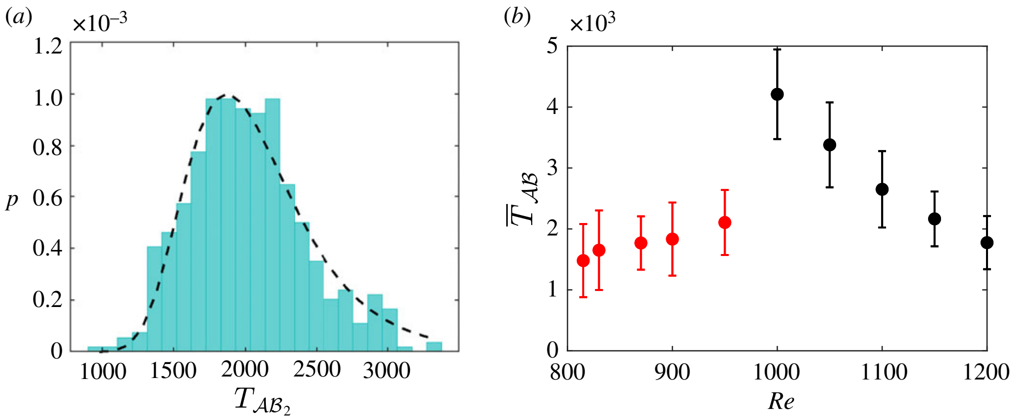

Figure 12. (*a*) Histogram of the reactive times $T_{\mathcal{AB}_2}$ at $Re = 1150$, estimated with the AMS on $N = 500$ trajectories. Dashed lines show a fit with a Gumbel distribution (5.2) with $\alpha = 1.9 \times 10^3$ and $\beta = 2.7 \times 10^{-3}$. (*b*) Mean reactive times $\overline{T}_{\mathcal{AB}_0}$ and $\overline{T}_{\mathcal{AB}_2}$, for different Re, estimated with the AMS. Error bars indicated one standard deviation. Reactive times are measured from a random point in $\mathcal{A}$ to the boundary of $\mathcal{B}_0$ or $\mathcal{B}_2$. (Online version in colour.)

[6,47]; here the 90% probability isosurface shows that almost every successful decay trajectory follows a similar path.

For splits, we use coordinates similar to those in [6], the first three z Fourier components $\hat{u}_{0,1}, \hat{u}_{0,2}, \hat{u}_{0,3}$ of u, averaged in x and y:

$$\hat{u}_{0,n} = \frac{1}{L_x L_z} \int \mathrm{d}y \left| \int \mathrm{d}x \int \mathrm{d}z \, u(x,y,z) \, \mathrm{e}^{-2\pi i n z/L_z} \right|.$$

Figure 11*b* shows that the reactive pathway from $\mathcal{A}$ to $\mathcal{B}_2$ for the case of splits consists of a highly curved tube. This shape arises from the non-monotonicity of the splitting trajectories in these coordinates, as seen in [6]. While a one-band state in $\mathcal{A}$ is characterized by high $\hat{u}_{0,1}$, the magnitude of $\hat{u}_{0,2}$ decreases at the beginning of a split before reaching its ultimate higher value in the two-band state in $\mathcal{B}_2$.

The transition pathways can also be described by the distribution of reactive times $T_{\mathcal{AB}}$. Reactive times have been characterized by Gumbel distributions

$$p_{\mathrm{Gum}}(T) = \beta\, \mathrm{e}^{-\beta(T-\alpha)} \exp\left(-\mathrm{e}^{-\beta(T-\alpha)}\right), \tag{5.2}$$

rigorously in simple stochastic ODEs in the weak noise limit [48], and observationally in one-dimensional stochastic PDEs [11,16] and in the decay of uniform turbulence in the Navier–Stokes equations [18]. We find that the distributions of reactive times $T_{\mathcal{AB}_0}$ for decays and $T_{\mathcal{AB}_2}$ for splits are consistent with Gumbel distributions for each *Re* and hence also with instanton-like behaviour. Figure 12*a* illustrates this for $Re = 1150$, but the relatively small number of computed reaction trajectories (around 500 for this *Re*) precludes drawing more definite conclusions. The mean duration of reactive trajectories and their standard deviation as a function of *Re* are shown in figure 12*b*. The mean reactive times $\overline{T}_{\mathcal{AB}}$ vary only modestly with *Re* within each of the decay and the splitting regimes, as do the standard deviations (shown by the error bars).

The results presented in this section were motivated by interest in rare-event pathways and instantons in particular. We observe that reactive trajectories for both decays and splits concentrate around a reactive tube in phase space. This suggests that turbulent fluctuations are dominated by the collective behaviour of trajectories along a most-probable path, which may be an instanton. We observe mild contraction of pathways as we vary *Re* and events become rare (see electronic supplementary material). Such contraction would be expected if the transitions exhibited instanton-like behaviour. At the present time, even using the AMS algorithm, we have not produced sufficient numbers of independent reactive trajectories at very high transition times to draw definite conclusions and more work is needed to relate this behaviour to the Large Deviation theory.

6. Discussion

Determining—or even defining—the threshold for turbulence in wall-bounded shear flows has been an important question since Reynolds' 1883 article [49]. Over time it has become clear that transitional turbulence is typically metastable and that transitions from metastable states play a crucial role in determining the onset of sustained turbulence [50–55]. The culmination of this realization was the study of Avila *et al.* [4] that determined the mean lifetimes for puff decay and puff splitting in pipe flow and showed that these lifetimes cross at a critical value of the Reynolds number Re_c. Although this work involved both numerical simulations and experiments, it was only through experiments that the very long lifetimes associated with Re_c were accessible. This has driven interest in capturing transitions from long-lived metastable states in wall-bounded flows via numerical simulations in order to obtain a clearer theoretical understanding of these events and of their Reynolds number dependence.

We have used the AMS algorithm [7–9] to obtain rare events in plane channel flow. We have specifically analysed transitions from the metastable one-band state to either laminar flow (decay) or to a two-band state (splitting) in tilted-domain simulations of the 3D Navier–Stokes equations with 2×10^7 degrees of freedom. Using AMS on this large system we have been able to obtain mean lifetimes as large as 5×10^6 in advective time units with a gain in computational efficiency of a factor of up to 100 over the standard MC approach. This has permitted us to access timescales near the lifetime crossing point for this flow. With the significant number of rare transitions we obtained, we have been able to construct weak tails in the probability distribution functions for the turbulence fraction. Exploiting ideas by Goldenfeld *et al.* [19] and Nemoto & Alexakis [20,21], we have been able to link directly the super-exponential variation of mean lifetimes with *Re*, for both decay and splitting, to the distribution of fluctuations in the one-band state. Finally, we have examined briefly the reaction pathways for decay and splitting.

Without conducting an extensive companion study in a large untilted domain, we cannot rule out effects of our narrow tilted domain on the transition rates and paths. However, we can cite comparisons of thresholds in the two types of domains. Shimizu & Manneville [29] carried out channel flow simulations in large domains of size $L_{x'} \times L_{z'} = 500 \times 250$ or 1000×500 and obtained a threshold between $Re = 905$ and 984 for one of the two regimes they studied. This is quite close to the crossover at $Re \approx 980$ between the decay and splitting times that we have computed here in a narrow tilted domain via AMS. In plane Couette flow, the threshold for transition to turbulence was estimated to be $Re = 325$ by Shi *et al.* [5] as the decay-splitting lifetime crossing in computations in a narrow tilted domain. This is the same as the value estimated experimentally by Bottin *et al.* [51,52] and numerically by Duguet *et al.* [56], in rectangular domains of size 380×70 and 800×356. An experiment in a much larger domain of size 3927×1500 by Klotz *et al.* [57] yields $Re = 330 \pm 0.5$ as the threshold.

Throughout this study we have focussed on the turbulence fraction as a scalar observable of the state of the system, in large part because it is an easily obtainable quantity of general interest. While turbulence fraction is presumably not a mechanistic driver of either event, it is a very informative observable that is highly correlated to the distance to the targeted states. Our analysis of the super-exponential dependence of mean lifetimes on *Re* is probabilistic and relies heavily on the observed, but unexplained, collapse of rescaled distributions of F_t over what we call the collapse zone.

This approach is complementary to the dynamical-systems approach to turbulence [2,58–60]. It would be useful to connect these approaches and to understand the mechanisms at work within the collapse zone. A particular question is the role played by saddle points or edge states [2,25,61–63] in creating behaviour that can be rescaled and collapse to *Re*-independent form, because this is a key ingredient in how turbulent fluctuations are connected to decay and splitting events. While there is much previous work on decay from a dynamical-systems perspective, there is little to rely upon in the case of splitting.

Our investigation of reaction pathways demonstrates their concentration in phase space for both decay and splitting events. We have also observed a Gumbel distribution for the reaction times. The mild contraction of pathways that we have observed as the transition probability becomes very low resembles an instanton, but is inconclusive. To better support this picture, we would need to quantify the level of the fluctuations of the effective degrees of freedom in the system and how the fluctuation levels depend on the Reynolds number. Following this, we would need to compare the transition-rate dependence on the Reynolds number to what would be expected from the level of fluctuations within Large Deviation theory. This would require us to disentangle the effect of *Re* on turbulent fluctuations from its effect on the potential term, which itself strongly depends on Reynolds number as seen by the parameterization of the PDFs within the one-band state (figures 8 and 10*a*). This approach would thus require the computation of the action minimizer in Large Deviation theory, which is out of the scope of the current study. This fundamental issue is related to the absence of a second parameter that would independently control the level of turbulent fluctuations and thereby allow an approach to a low-noise limit. We note that the states studied here are localized and insensitive to domain length. Hence domain size, the one parameter other than *Re* available in the numerical simulations, does not provide a means to influence the effect of fluctuations on the transitions. We refer the reader to the important studies of Rolland [16,18] on rare events in transitional shear flows.

Finally, while we have succeeded in using the AMS algorithm to compute rare events in the 3D Navier–Stokes equations represented by $O(10^7)$ degrees of freedom, the experience has not been without difficulties. The most notable issues are: (i) the algorithm sometimes stagnates, making very slow progress toward obtaining trajectories reaching the target state and (ii) the variance in the estimated mean lifetimes associated with the AMS realizations is large, thus requiring the costly step of running multiple realizations. The method used here could possibly be improved with the implementation of Anticipated AMS [18]. Most importantly, the score function is well known to be crucial to efficient performance of the algorithm. Finding a score function that targets successful splitting events has been particularly challenging. Although we have obtained a serviceable empirical score function based largely upon the turbulence fraction, a more far-ranging search for appropriate score functions is needed.

Data accessibility. The data that support the findings of this study were generated using the open source code Channelflow [26] and are available from the authors upon request.

Authors' contributions. S.G.: Formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, writing— review and editing; L.T.: funding acquisition, project administration, resources, supervision, writing—original draft, writing—review and editing; D.B.: conceptualization, investigation, methodology, resources, supervision, writing—original draft, writing—review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. Authors declare no competing interests.

Funding. This work was supported by a grant from the Simons Foundation (grant no. 662985, N.G.).

Acknowledgements. The calculations for this work were performed using high performance computing resources provided by the Grand Equipement National de Calcul Intensif at the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS, CNRS) through grant nos A0082A01119 and A0102A01119. We wish to acknowledge Anna Frishman, Tobias Grafke, Freddy Bouchet, Takihiro Nemoto and Alexandros Alexakis for helpful discussions. We also thank Florian Reetz and Alessia Ferraro for their assistance in using Channelflow. This work is dedicated to the memory of our dear friend and colleague Charlie Doering.

References

1. Faisst H, Eckhardt B. 2004 Sensitive dependence on initial conditions in transition to turbulence in pipe flow. *J. Fluid Mech.* **504**, 343–352. (doi:10.1017/S0022112004008134)
2. Eckhardt B, Schneider TM, Hof B, Westerweel J. 2007 Turbulence transition in pipe flow. *Annu. Rev. Fluid Mech.* **39**, 447–468. (doi:10.1146/annurev.fluid.39.050905.110308)
3. Avila M, Willis AP, Hof B. 2010 On the transient nature of localized pipe flow turbulence. *J. Fluid Mech.* **646**, 127–136. (doi:10.1017/S0022112009993296)
4. Avila K, Moxey D, de Lozar A, Avila M, Barkley D, Hof B. 2011 The onset of turbulence in pipe flow. *Science* **333**, 192–196. (doi:10.1126/science.1203223)
5. Shi L, Avila M, Hof B. 2013 Scale invariance at the onset of turbulence in Couette flow. *Phys. Rev. Lett.* **110**, 204502. (doi:10.1103/PhysRevLett.110.204502)
6. Gomé S, Tuckerman LS, Barkley D. 2020 Statistical transition to turbulence in plane channel flow. *Phys. Rev. Fluids* **5**, 083905. (doi:10.1103/PhysRevFluids.5.083905)
7. Cérou F, Guyader A. 2007 Adaptive multilevel splitting for rare event analysis. *Stochastic Anal. Appl.* **25**, 417–443. (doi:10.1080/07362990601139628)
8. Cérou F, Guyader A, Lelievre T, Pommier D. 2011 A multiple replica approach to simulate reactive trajectories. *J. Chem. Phys.* **134**, 054108. (doi:10.1063/1.3518708)
9. Cérou F, Guyader A, Rousset M. 2019 Adaptive multilevel splitting: historical perspective and recent results. *Chaos* **29**, 043108. (doi:10.1063/1.5082247)
10. Rolland J, Simonnet E. 2015 Statistical behaviour of adaptive multilevel splitting algorithms in simple models. *J. Comput. Phys.* **283**, 541–558. (doi:10.1016/j.jcp.2014.12.009)
11. Rolland J, Bouchet F, Simonnet E. 2016 Computing transition rates for the 1-D stochastic Ginzburg–Landau–Allen–Cahn equation for finite-amplitude noise with a rare event algorithm. *J. Stat. Phys.* **162**, 277–311. (doi:10.1007/s10955-015-1417-4)
12. Lestang T, Ragone F, Bréhier C-E, Herbert C, Bouchet F. 2018 Computing return times or return periods with rare event algorithms. *J. Stat. Mech.: Theory Exp.* **2018**, 043213. (doi:10.1088/1742-5468/aab856)
13. Bouchet F, Rolland J, Simonnet E. 2019 Rare event algorithm links transitions in turbulent flows with activated nucleations. *Phys. Rev. Lett.* **122**, 074502. (doi:10.1103/PhysRevLett.122.074502)
14. Simonnet E, Rolland J, Bouchet F. 2021 Multistability and rare spontaneous transitions in barotropic β-plane turbulence. *J. Atmos. Sci.* **78**, 1889–1911. (doi:10.1175/JAS-D-20-0279.1)
15. Lestang T, Bouchet F, Lévêque E. 2020 Numerical study of extreme mechanical force exerted by a turbulent flow on a bluff body by direct and rare-event sampling techniques. *J. Fluid Mech.* **895**, A19. (doi:10.1017/jfm.2020.293)
16. Rolland J. 2018 Extremely rare collapse and build-up of turbulence in stochastic models of transitional wall flows. *Phys. Rev. E* **97**, 023109. (doi:10.1103/PhysRevE.97.023109)
17. Barkley D. 2016 Theoretical perspective on the route to turbulence in a pipe. *J. Fluid Mech.* **803**, P1. (doi:10.1017/jfm.2016.465)
18. Rolland J. 2022 Collapse of transitional wall turbulence captured using a rare events algorithm. *J. Fluid Mech.* **931**, A22. (doi:10.1017/jfm.2021.957)

19. Goldenfeld N, Guttenberg N, Gioia G. 2010 Extreme fluctuations and the finite lifetime of the turbulent state. *Phys. Rev. E* **81**, 035304(R). (doi:10.1103/PhysRevE.81.035304)
20. Nemoto T, Alexakis A. 2018 Method to measure efficiently rare fluctuations of turbulence intensity for turbulent-laminar transitions in pipe flows. *Phys. Rev. E* **97**, 022207. (doi:10.1103/PhysRevE.97.022207)
21. Nemoto T, Alexakis A. 2021 Do extreme events trigger turbulence decay?–a numerical study of turbulence decay time in pipe flows. *J. Fluid Mech.* **912**, A38. (doi:10.1017/jfm.2020.1150)
22. Barkley D, Tuckerman LS. 2005 Computational study of turbulent-laminar patterns in Couette flow. *Phys. Rev. Lett.* **94**, 014502. (doi:10.1103/PhysRevLett.94.014502)
23. Tuckerman LS, Chantry M, Barkley D. 2020 Patterns in wall-bounded shear flows. *Annu. Rev. Fluid Mech.* **52**, 343. (doi:10.1146/annurev-fluid-010719-060221)
24. Tuckerman LS, Kreilos T, Schrobsdorff H, Schneider TM, Gibson JF. 2014 Turbulent-laminar patterns in plane Poiseuille flow. *Phys. Fluids* **26**, 114103. (doi:10.1063/1.4900874)
25. Paranjape CS, Duguet Y, Hof B. 2020 Oblique stripe solutions of channel flow. *J. Fluid Mech.* **897**, A7. (doi:10.1017/jfm.2020.322)
26. Gibson JF. 2012 Channelflow: A Spectral Navier-Stokes Simulator in C++. tech. rep., University of New Hampshire. See Channelflow.org.
27. Pugh J, Saffman P. 1988 Two-dimensional superharmonic stability of finite-amplitude waves in plane Poiseuille flow. *J. Fluid Mech.* **194**, 295–307. (doi:10.1017/S002211208800299X)
28. Barkley D. 1990 Theory and predictions for finite-amplitude waves in two-dimensional plane Poiseuille flow. *Phys. Fluids A* **2**, 955–970. (doi:10.1063/1.857603)
29. Shimizu M, Manneville P. 2019 Bifurcations to turbulence in transitional channel flow. *Phys. Rev. Fluids* **4**, 113903. (doi:10.1103/PhysRevFluids.4.113903)
30. Xiao X, Song B. 2020 The growth mechanism of turbulent bands in channel flow at low Reynolds numbers. *J. Fluid Mech.* **883**, R1. (doi:10.1017/jfm.2019.899)
31. Chantry M, Tuckerman LS, Barkley D. 2016 Turbulent–laminar patterns in shear flows without walls. *J. Fluid Mech.* **791**, R8. (doi:10.1017/jfm.2016.92)
32. Wouters J, Bouchet F. 2016 Rare event computation in deterministic chaotic systems using genealogical particle analysis. *J. Phys. A: Math. Theor.* **49**, 374002. (doi:10.1088/1751-8113/49/37/374002)
33. Bréhier C-E. 2016 Unbiasedness of some generalized adaptive multilevel splitting algorithms. *Ann. Appl. Probab.* **26**, 3559–3601. (doi:10.1214/16-AAP1185)
34. Hof B, De Lozar A, Kuik DJ, Westerweel J. 2008 Repeller or attractor? Selecting the dynamical model for the onset of turbulence in pipe flow. *Phys. Rev. Lett.* **101**, 214501. (doi:10.1103/PhysRevLett.101.214501)
35. Fisher RA, Tippett LHC. 1928 Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Philos. Soc.* **24**, 180–190. (doi:10.1017/S0305004100015681)
36. Manneville P. 2011 On the decay of turbulence in plane Couette flow. *Fluid Dyn. Res.* **43**, 065501. (doi:10.1088/0169-5983/43/6/065501)
37. Shimizu M, Kanazawa T, Kawahara G. 2019 Exponential growth of lifetime of localized turbulence with its extent in channel flow. *Fluid Dyn. Res.* **51**, 011404. (doi:10.1088/1873-7005/aaa73c)
38. Faranda D, Lucarini V, Manneville P, Wouters J. 2014 On using extreme values to detect global stability thresholds in multi-stable systems: the case of transitional plane Couette flow. *Chaos Solitons Fractals* **64**, 26–35. (doi:10.1016/j.chaos.2014.01.008)
39. E W, Vanden-Eijnden E. 2006 Towards a theory of transition paths. *J. Stat. Phys.* **123**, 503.
40. Touchette H. 2011 A basic introduction to large deviations: theory, applications, simulations. (http://arxiv.org/abs/1106.4146)
41. Grafke T, Grauer R, Schäfer T. 2015 The instanton method and its numerical implementation in fluid mechanics. *J. Phys. A: Math. Theor.* **48**, 333001. (doi:10.1088/1751-8113/48/33/333001)
42. Grafke T, Vanden-Eijnden E. 2019 Numerical computation of rare events via large deviation theory. *Chaos* **29**, 063118. (doi:10.1063/1.5084025)
43. Grafke T, Grauer R, Schäfer T. 2013 Instanton filtering for the stochastic Burgers equation. *J. Phys. A: Math. Theor.* **46**, 062002. (doi:10.1088/1751-8113/46/6/062002)

44. Dematteis G, Grafke T, Onorato M, Vanden-Eijnden E. 2019 Experimental evidence of hydrodynamic instantons: the universal route to rogue waves. *Phys. Rev. X* **9**, 041057. (doi:10.1103/PhysRevX.9.041057)
45. Schorlepp T, Grafke T, May S, Grauer R. 2021 Spontaneous symmetry breaking for extreme vorticity and strain in the 3D Navier-Stokes equations. (http://arxiv.org/abs/2107.06153)
46. Woillez E, Bouchet F. 2020 Instantons for the destabilization of the inner solar system. *Phys. Rev. Lett.* **125**, 021101. (doi:10.1103/PhysRevLett.125.021101)
47. Liu T, Semin B, Klotz L, Godoy-Diana R, Wesfreid J, Mullin T. 2021 Decay of streaks and rolls in plane Couette–Poiseuille flow. *J. Fluid Mech.* **915**, A65. (doi:10.1017/jfm.2021.89)
48. Cérou F, Guyader A, Lelievre T, Malrieu F. 2013 On the length of one-dimensional reactive paths. *ALEA, Lat. Am. J. Probab. Math. Stat.* **10**, 359–389.
49. Reynolds O. 1883 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-–982. (doi:10.1098/rstl.1883.0029)
50. Pomeau Y. 1986 Front motion, metastability and subcritical bifurcations in hydrodynamics. *Physica D* **23**, 3–11. (doi:10.1016/0167-2789(86)90104-1)
51. Bottin S, Chaté H. 1998 Statistical analysis of the transition to turbulence in plane Couette flow. *Eur. Phys. J. B* **6**, 143–155. (doi:10.1007/s100510050536)
52. Bottin S, Daviaud F, Manneville P, Dauchot O. 1998 Discontinuous transition to spatiotemporal intermittency in plane Couette flow. *Europhys. Lett.* **43**, 171. (doi:10.1209/epl/i1998-00336-3)
53. Peixinho J, Mullin T. 2006 Decay of turbulence in pipe flow. *Phys. Rev. Lett.* **96**, 094501. (doi:10.1103/PhysRevLett.96.094501)
54. Hof B, Westerweel J, Schneider TM, Eckhardt B. 2006 Finite lifetime of turbulence in shear flows. *Nature* **443**, 59–62. (doi:10.1038/nature05089)
55. Willis AP, Kerswell RR. 2007 Critical behavior in the relaminarization of localized turbulence in pipe flow. *Phys. Rev. Lett.* **98**, 014501. (doi:10.1103/PhysRevLett.98.014501)
56. Duguet Y, Schlatter P, Henningson DS. 2010 Formation of turbulent patterns near the onset of transition in plane Couette flow. *J. Fluid Mech.* **650**, 119–129. (doi:10.1017/S0022112010000297)
57. Klotz L, Lemoult G, Avila K, Hof B. 2022 The phase transition to turbulence in spatially extended shear flows. *Phys. Rev. Lett.* **128**, 014502. (doi:10.1103/PhysRevLett.128.014502)
58. Kerswell R. 2005 Recent progress in understanding the transition to turbulence in a pipe. *Nonlinearity* **18**, R17. (doi:10.1088/0951-7715/18/6/R01)
59. Kawahara G, Uhlmann M, Van Veen L. 2012 The significance of simple invariant solutions in turbulent flows. *Annu. Rev. Fluid Mech.* **44**, 203–225. (doi:10.1146/annurev-fluid-120710-101228)
60. Graham MD, Floryan D. 2021 Exact coherent states and the nonlinear dynamics of wall-bounded turbulent flows. *Annu. Rev. Fluid Mech.* **53**, 227–253. (doi:10.1146/annurev-fluid-051820-020223)
61. Schneider TM, Eckhardt B, Yorke JA. 2007 Turbulence transition and the edge of chaos in pipe flow. *Phys. Rev. Lett.* **99**, 034502. (doi:10.1103/PhysRevLett.99.034502)
62. Duguet Y, Willis AP, Kerswell RR. 2008 Transition in pipe flow: the saddle structure on the boundary of turbulence. *J. Fluid Mech.* **613**, 255–274. (doi:10.1017/S0022112008003248)
63. Chantry M, Schneider TM. 2014 Studying edge geometry in transiently turbulent shear flows. *J. Fluid Mech.* **747**, 506–517. (doi:10.1017/jfm.2014.150)