

Extremely rare collapse and build-up of turbulence in stochastic models of transitional wall flows

Joran Rolland*

Institut du Non-Linéaire de Nice, Université de Nice Sophia Antipolis, UMR CNRS 7010, 1361 route des Lucioles, Valbonne 06560, France

(Received 24 July 2017; revised manuscript received 7 December 2017; published 26 February 2018)

This paper presents a numerical and theoretical study of multistability in two stochastic models of transitional wall flows. An algorithm dedicated to the computation of rare events is adapted on these two stochastic models. The main focus is placed on a stochastic partial differential equation model proposed by Barkley. Three types of events are computed in a systematic and reproducible manner: (i) the collapse of isolated puffs and domains initially containing their steady turbulent fraction; (ii) the puff splitting; (iii) the build-up of turbulence from the laminar base flow under a noise perturbation of vanishing variance. For build-up events, an extreme realization of the vanishing variance noise pushes the state from the laminar base flow to the most probable germ of turbulence which in turn develops into a full blown puff. For collapse events, the Reynolds number and length ranges of the two regimes of collapse of laminar-turbulent pipes, independent collapse or global collapse of puffs, is determined. The mean first passage time before each event is then systematically computed as a function of the Reynolds number r and pipe length L in the laminar-turbulent coexistence range of Reynolds number. In the case of isolated puffs, the faster-than-linear growth with Reynolds number of the logarithm of mean first passage time T before collapse is separated in two. One finds that $\ln(T) = A_p r - B_p$, with A_p and B_p positive. Moreover, A_p and B_p are affine in the spatial integral of turbulence intensity of the puff, with the same slope. In the case of pipes initially containing the steady turbulent fraction, the length L and Reynolds number r dependence of the mean first passage time T before collapse is also separated. The author finds that $T \asymp \exp[L(Ar - B)]$ with A and B positive. The length and Reynolds number dependence of T are then discussed in view of the large deviations theoretical approaches of the study of mean first passage times and multistability, where $\ln(T)$ in the limit of small variance noise is studied. Two points of view, local noise of small variance and large length, can be used to discuss the exponential dependence in L of T . In particular, it is shown how a $T \asymp \exp[L(A'R - B')]$ can be derived in a conceptual two degrees of freedom model of a transitional wall flow proposed by Dauchot and Manneville. This is done by identifying a quasipotential in low variance noise, large length limit. This pinpoints the physical effects controlling collapse and build-up trajectories and corresponding passage times with an emphasis on the saddle points between laminar and turbulent states. This analytical analysis also shows that these effects lead to the asymmetric probability density function of kinetic energy of turbulence.

DOI: [10.1103/PhysRevE.97.023109](https://doi.org/10.1103/PhysRevE.97.023109)**I. INTRODUCTION**

Transitional wall turbulence is an important regime of many geophysical or industrial flows. One fundamental example is the stably stratified planetary boundary layer. This flow manifests itself over ocean (where the stratification is mostly stable) and over land after the late afternoon transition (which corresponds to the stabilization of stratification) [1,2]. In both examples, the stably stratified wall flow will display the feature of simple, academic transitional wall flows such as plane Couette flow and pipe flow [1]. These flows are controlled by the Reynolds number, the ratio of advection over viscosity. In these prototypes of transitional wall turbulence, for a given range of Reynolds number, turbulence can coexist in time and space with laminar flow [3,4]. Turbulent patches (termed puffs in Hagen-poiseuille pipe flow [5]) can split [4,6], extend [5,7–9], or collapse after a time [10,11]. These flows are thus very intermittent and this intermittency is the main obstacle in

understanding and modeling this type of turbulence. This limits the quality of climate models where boundary layer turbulence is parametrized using a simple model because it is not explicitly resolved [2].

Stated more precisely, transitional wall flows undergo *spatiotemporal intermittency* (STI) [10]. The core ingredients of this behavior are the stability of the laminar base flow for all Reynolds numbers [12,13] (or quite beyond the transitional range), transient chaos of small subsections of the flow [14], and the possibility of contagion of neighboring regions [15]. In infinite size domains, turbulence is sustained with a steady turbulent fraction in the transitional range as a consequence of this process [15,16]. One fundamental feature of STI in pipe flow is that the turbulent puffs are transient and metastable and can collapse or split after a *mean first passage time* T [17,18]. This time is often termed the lifetime of the puff in this context. Knowing and understanding the duration of these transients is a necessity if one is to propose models of such flows that are realistic and simple enough so that they can be incorporated in climate models [2], for instance. In pipe flow, extensive data on the passage time is available concerning the low Reynolds number part of the transitional range. One finds that $\ln(T)$ is

*Present address: Institut PPrime, UPR 3346 CNRS, Chasseneuil du Poitou, France; jrolland@phare.normalesup.org

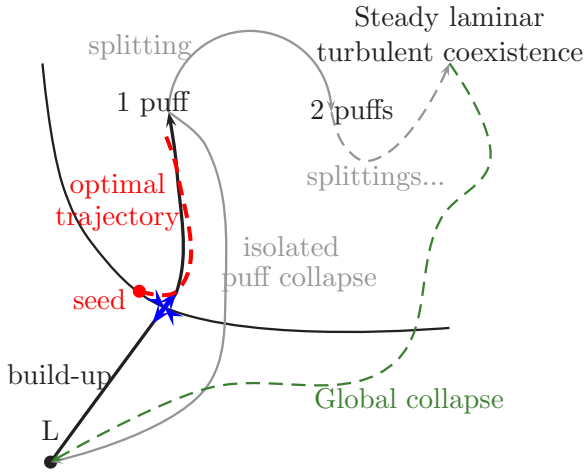


FIG. 1. Sketch of the conceptual phase space of a transitional wall flow, including basin boundary, minimal seed of turbulence, the turbulent coherent structures (1 to n puffs) and the events linking these metastable states: turbulence collapse, puff splitting, build-up of turbulence from the laminar base flow under vanishing perturbations (possibly joining optimal trajectory). L stands for laminar base flow.

faster than linear in Reynolds number [11]. In smaller pipes, one finds that $\ln(T)$ is proportional to pipe length \times Reynolds number [19]. Such a size scaling may have the same origin as other exponentially increasing chaotic transients in extended systems (see [20], Sec. 4). Verification of the proposed scalings has not yet been done in the higher part of the transitional range and in larger pipes. In other flows, the available results are even more fractional. The main limitation in obtaining these results is of course the extreme length of the transients at play. Educing the exact scaling of mean first passage times is fundamental if one wants to discuss the physical mechanism behind turbulence collapse and build-up.

If one is to extend or even obtain the Reynolds number and length dependence of mean first passage time before collapse of transitional wall turbulence, especially in flows with a realistic geometry, one has to propose a strategy adapted to the study of very rare events which is numerically and theoretically tractable. This strategy should be applied in a large enough range of Reynolds number and domain length so that scalings are thoroughly verified. One can define several rare events along with their passage time, which are sketched in Fig. 1. One finds the collapse of puff and laminar-turbulent pipes and the splitting of a turbulent puff. Another class of events is the build-up of turbulence starting from the laminar state, under perturbations of vanishing variance, often passing by optimal perturbations or a minimal seed [21,22]. Fortunately, all three events can be reformulated as a problem of multistability between several metastable states. Multistability has been a very active field of study in statistical physics and probabilities [17,23–27]. Many tools and concepts are available that can be used for the study of transient turbulence. Moreover, many of these tools and concepts originate from kinetic chemistry [17,27], so that they can be presented to a very wide community using references to classical results such as the Arrhenius law [17].

In this paper, I present a joint numerical and theoretical study of the mean first passage time before these events in two models of pipe flow. A spatiotemporal stochastic partial differential equation (SPDE) model of pipe flow was chosen [16]. The extensive numerical study, besides giving insightful physical results, should hopefully propose guidelines and methods for the study of rare events in direct numerical simulations. This study is complemented by a theoretical and numerical analysis of a two degrees of freedom stochastic differential equation (SDE) model [28]. Since the chosen models are stochastic, all the theoretical assertions strictly hold and the orders of approximations are understood. The numerical method used for rare events calculation in these two models is called *adaptive multilevel splitting* (AMS) [23,29]. It is a mutation-selection algorithm [30], arising from kinetic chemistry [27], designed to compute extremely fast, extremely rare multistability events along with the corresponding mean first passage times. The passage times will be computed in a very large range of length and Reynolds number so as to verify precisely scaling laws, and thus discuss physical mechanisms with a sound basis. The numerical results will be discussed and complemented by theory derived from large deviations approaches of the study of rare events [25,31–34]. These approaches concern the asymptotic study of very large mean first passage times T and very small probabilities ρ , under a felt noise of small variance $1/\mathfrak{B}$. Large deviations revolve under the formulation of said passage times, probability density functions (pdf), etc., as a large deviation principle,

$$\lim_{1/\mathfrak{B} \rightarrow 0} \frac{1}{\mathfrak{B}} \ln(T) = I(\mathfrak{X}), \quad \lim_{1/\mathfrak{B} \rightarrow 0} -\frac{1}{\mathfrak{B}} \ln(\rho) = I'(\mathfrak{X}). \quad (1)$$

The functions I and I' are generally called rate functions, are independent of the noise variance $1/\mathfrak{B}$ (as indicated by the diagonal strikethrough \mathfrak{X}), and are of finite amplitude. The rate functions can often be derived using dedicated techniques [24,25,32,33]. Behind many of these computations lies the *instanton*, the most probable trajectory in multistability events in the low variance noise limit. Parametric studies, presentation of numerical results, and comparison with theory will be very convenient using such a framework. In particular, written approximately, one can see that these results generalize the Arrhenius law for mean first passage times $T \asymp \exp(\mathfrak{B}I)$ and the Boltzmann factor for pdf $\rho \asymp \exp(-\mathfrak{B}I')$. This uses the notation \asymp to write in a compact manner that the equalities should be understood in the limit of the logarithm form of Eq. (1) ([25], Sec. 3.2). Earlier successful use of such coupled theoretical and numerical studies is for instance the general study of metastability in the stochastic Ginzburg-Landau equation [24], and similar work has been performed in a model of the zonal turbulent jets of Jupiter [35]. Indeed, geophysical fluid dynamics is another situation where multistability occurs [36], and many parallels can and will be drawn between the two fields of study. An accelerated sampling procedure for computation of mean first passage time before collapse of puffs in a model of pipe flow has also recently been proposed [37]. The sampling is done by reducing or increasing the Reynolds number when the kinetic energy crosses certain thresholds, in order to facilitate turbulence collapse. First passage durations are reconstructed from these samples using a pdf based argument. Empirical criteria give guidelines

for choices of thresholds and Reynolds numbers leading to maximum precision. Large deviation methods have also been applied in a direct manner to compute most probable paths and transition rates between the linearly laminar base flow and Tollmien-Schlichting type waves in low Reynolds number two dimensional plane Poiseuille flow [38]. Improvement of the temporal discretization and of the optimization procedure may lead to calculations in three dimensional plane Poiseuille flow, where wall turbulence can exist.

This study will be presented according to the following plan. The SPDE model of pipe flow and my choice of control parameters will be presented first (Sec. II A). The numerical method, AMS, and its application to the model will be presented next (Sec. II B). Numerical results will then be presented in two parts: first by detailing the feature of selected turbulent collapse, puff splitting, and turbulence build-up trajectories (Sec. III A), second by presenting the parametric study of mean first passage times (Secs. III B and III C). Theory is then used to interpret the parametric dependence of numerically computed passage times. I first discuss the length dependence of mean first passage time in the various regimes of collapse in Sec. IV A. The conceptual two degrees of freedom model is then analyzed analytically to derive the Reynolds number scaling of the mean first passage times before collapse of turbulence in Sec. IV B. The results and perspectives are eventually discussed in the Conclusion (Sec. V).

II. MODEL AND METHOD

A. Stochastic partial differential equations model of transitional Hagen-Poiseuille flow

To a large extent, the complexity of transitional pipe flow can be successfully reduced to one dimension of space (the streamwise position x) models coupling the streamwise velocity u and the intensity of turbulence q [16]. At their core, the models are based on the observation that pipe flow turbulence is very similar to an excitable system, like the action potential in a neuron. In this paper, I use a stochastic partial differential equation (SPDE) model which reads

$$\begin{aligned} \frac{\partial u}{\partial t} &= -(1+U)\frac{\partial u}{\partial x} + \epsilon_1(1-u) - \epsilon_2 u q, \quad (2) \\ \frac{\partial q}{\partial t} + U\frac{\partial q}{\partial x} &= \frac{\partial^2 q}{\partial x^2} + q[u+r-1-(r+\delta)(q-1)^2] \\ &\quad + q\sqrt{\frac{2}{\beta}}\eta. \\ \langle \eta(x,t) \rangle &= 0, \quad \langle \eta(x,t)\eta(x',t') \rangle = \delta(x-x')\delta(t-t'). \end{aligned} \quad (3)$$

The term η is a noise white in time and space. It leads to a multiplicative noise interpreted with an Itô rule (see [18], Sec. 4.2). In this system of stochastic partial equations, r is the Reynolds number, the main control parameter. Following Barkley [16], the other parameters are set as $\epsilon_1 = 0.04$, $\epsilon_2 = 0.2$, $\delta = 0.1$ and the inverse variance of the noise is set at $\beta = 1.5$ so that the model describes the relevant dynamics. One key element is that $q \neq 0$ stimulates $1-u$ and that $1-u$

goes to zeros when $q = 0$. In this manner q leads to $1-u$. The pipe is periodic in x and has a length L .

This model contains the physics of pipe flow, from the fully laminar base flow to higher Reynolds number slugs [39] through the transitional regime [16]. One can consult the mentioned references for more details on the typical regimes and on the internal mechanics of the system of (stochastic) partial differential equations. Note that the origin of the noise in transitional turbulence models is different from that in two dimensional turbulence models, where metastability is also studied [36]. This is not an extrinsic energy injection, so that it is not balanced by dissipation. Instead, it represents the intrinsic fluctuations of turbulence. In particular, such noises terms are absent in direct numerical simulations. In wall turbulence, the injection of energy is actually extraction of energy by turbulence from the laminar base flow mostly through the lift-up process [40]. This effect is taken into account in the deterministic part of the models.

I give here a brief overview of the typical content of the model in the transitional regime. Turbulence [see Figs. 2(a) and 2(b)] exists if $r \gtrsim 1$ and coexists in a steady manner with laminar flow provided $r \lesssim 1.8$ [16] [Fig. 2(c)]. The model realistically reproduces turbulent puffs [16] [Figs. 2(a) and 2(b)]. These coherent structures display localized nonzero turbulence $q \neq 0$ along with a deficit of streamwise velocity $u < 1$. The model contains the asymmetry of u between the sharp trailing edge and the smooth leading edge of the puff.

When studying the collapse of turbulence, one can consider two types of initial conditions. On the one hand, one can consider the collapse of an isolated equilibrium puff [at $r = 1.0$, Fig. 2(a)] [11]. The equilibrium puff is anything but an equilibrium, since it can not only collapse, but also split, until a quasisteady state is reached [Fig. 2(b) at $r = 1.3$]. This second state is hereafter termed a steady turbulent fraction. It is the other starting point for the study of turbulence collapse [10]. One of the reasons of the relevance of this second starting point is that it is well defined no matter the Reynolds number in the transitional regime. Meanwhile isolated puffs become harder and harder to define as r is increased, mainly because of puff splitting. Said puff splittings will be computed using an isolated puff as initial condition. Isolated puffs will also be used as targets for the computation of turbulence build-up.

In order to give a clear view of the meaning of the values of r in terms of regime of turbulence (not sustained, coexistence, uniformly turbulent), the time and space average of the turbulent fraction \mathcal{Q} and the streamwise velocity deficit $1-\mathcal{U}$,

$$\begin{aligned} \mathcal{Q} &\equiv \frac{1}{\mathcal{T}L} \int_{x=0,t=0}^{x=L,t=\mathcal{T}} q(x,t) dx dt, \\ 1-\mathcal{U} &\equiv 1 - \frac{1}{\mathcal{T}L} \int_{x=0,t=0}^{x=L,t=\mathcal{T}} u(x,t) dx dt, \end{aligned} \quad (4)$$

are computed in the range $0.6 \leq r \leq 2.5$ in a pipe of length $L = 1600$ using a duration $\mathcal{T} = 2000$ [Fig. 2(c)]. This indicates that Laminar-turbulent coexistence occurs in the range $1 \lesssim r \lesssim 1.8$. I do not go into detail in the scaling of \mathcal{Q} and \mathcal{U} in r at $r_c \simeq 1.0$ above which a steady nonzero turbulent fraction can exist in the flow. This has already been examined a twin

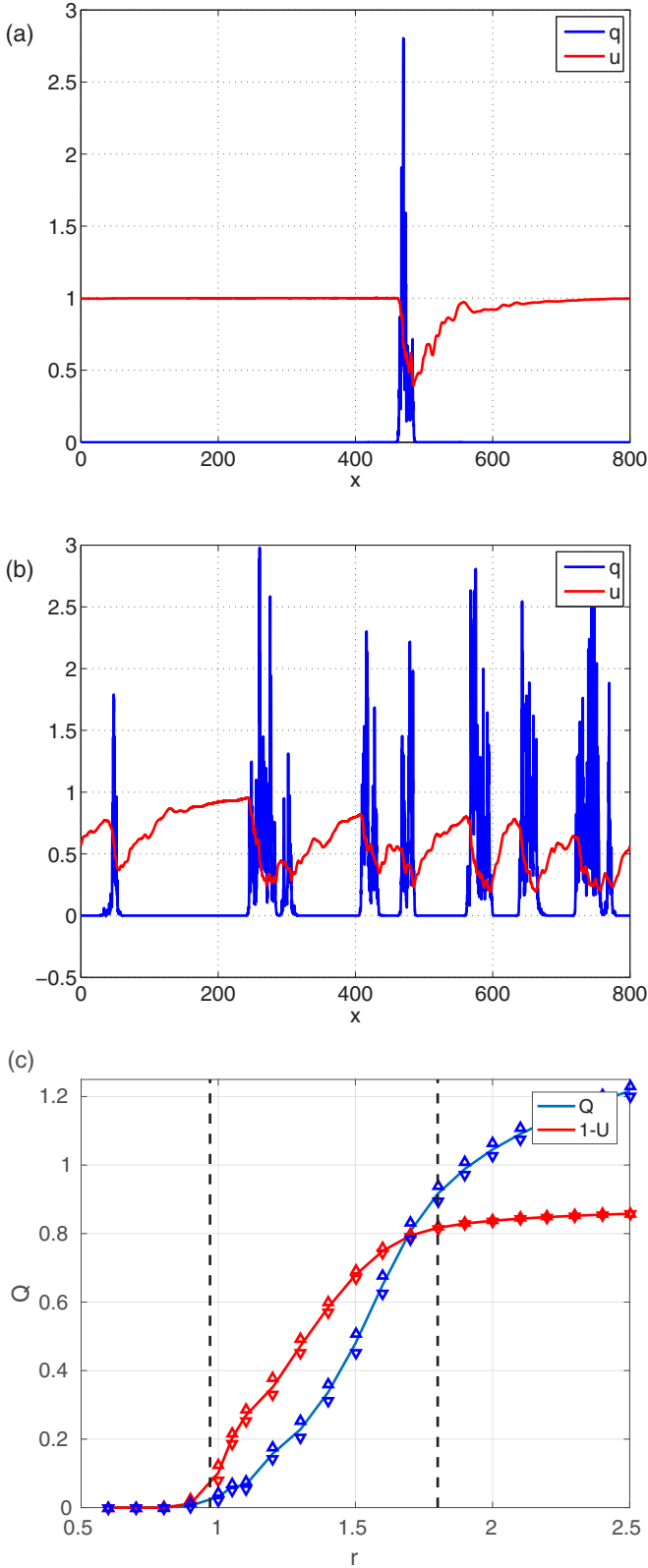


FIG. 2. Streamwise velocity u and turbulence intensity q as a function of space x in two situations: (a) isolated puff, (b) steady turbulent fraction. (c) Space and time average of turbulence intensity Q and streamwise velocity deficit $1 - U$ as a function of the Reynolds number r in the transitional range. The range of laminar-turbulence coexistence, where $0 < 1 - U < 1$, is indicated by black dashed lines.

chaotic model [16]. I will consider collapse of turbulence, puff splitting, and turbulence build-up in the range $0.6 \lesssim r \lesssim 1.8$.

B. Adaptive multilevel splitting

I now explain the principle of the method I use to compute the rare turbulent collapses, puff splittings, and build-ups. The algorithm adaptive multilevel splitting (AMS) will be used to calculate the corresponding trajectories and mean first passage times (i.e., the lifetime of turbulence). For more details, see for instance the monograph by Del Moral [30] for a general presentation of the mathematical framework of mutation-selection algorithms, see [23,41] for the initial presentation of the algorithm and mathematical demonstration of the convergence of such calculations, and see for instance [24] for a thorough presentation of its use coupled to theory to study multistability in a gradient system. Note that Giardana-Kurchan-Lecomte-Tailleur algorithms follow a similar principle of cloning and large deviation theoretical framework, even if they are designed to study atypical trajectories by weighting by Lyapunov exponents [42]. In the context of transitional wall flows, edge states are such atypical trajectories that could be studied using these Lyapunov weighted dynamics [43,44].

1. Principle of the algorithm

AMS is a mutation-selection algorithm, in the family of importance sampling [23,30], which uses N clone dynamics $\{\bar{X}_{1 \leq i \leq N}(t)\}$ to compute N reactive trajectories, the probability for their occurrence, and an estimator of the mean first passage time before they occur. In my study, one has $\bar{X}(t) = (q(x,t), u(x,t))$. First I shall generally define reactive trajectories and first passages. It is very natural to describe this in phase space, a point of view which is commonplace in the study of transitional turbulence. Let us term \mathcal{A} the neighborhood of the starting point \bar{X}_0 [for instance partially turbulent flow Figs. 2(a) and 2(b)] and \mathcal{B} the neighborhood of the arrival point (for instance laminar base flow). They are sketched in Fig. 3(a). I also define \mathcal{C} , a hypersurface that closely surrounds \mathcal{A} . A first passage is the natural dynamics of the system starting from \bar{X}_0 , fluctuating and having excursions out of \mathcal{A} until it reaches \mathcal{B} [Fig. 3(a), black then red line]. The average duration of first passage trajectory is termed the mean first passage time T . This is the lifetime of turbulence in the context of transitional pipe flow. A very important part of the dynamics is the reactive trajectory. This is the portion of the dynamics during which the jump between the two states actually occurs. It is strictly defined as a realization of the dynamics of the system which leaves \mathcal{A} , crosses \mathcal{C} , and then reaches \mathcal{B} before \mathcal{A} [Fig. 3(a), red line]. For a large class of problems and reaction coordinates, AMS calculates specifically these trajectories in an asymptotically unbiased manner along with the mean first passage time T .

I now describe the steps of the algorithm [Fig. 3(b)]. AMS uses N clones dynamics of the system and $\phi : \bar{X} \rightarrow \phi(\bar{X}) \in \mathbb{R}$, a reaction coordinate. The reaction coordinate measures the position of \bar{X} relatively to \mathcal{A} and \mathcal{B} . One often takes $\phi(\partial\mathcal{A}) = 0$ and $\phi(\partial\mathcal{B}) = 1$ on the subsets boundaries. The larger ϕ , the further away the excursion of \bar{X} is. My specific choices of

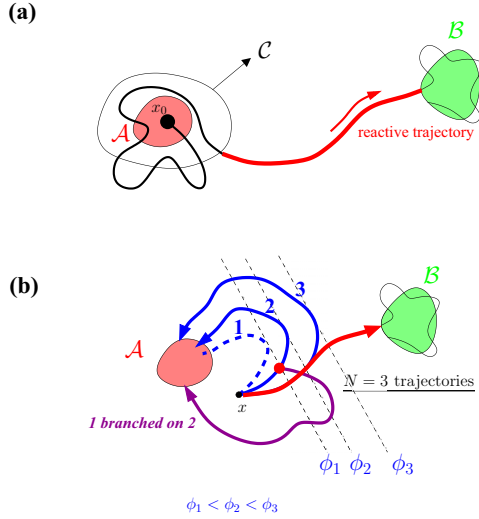


FIG. 3. (a) Sketch of first passages and reactive trajectories in phase space. (b) Sketch of the principle of the adaptive multilevel splitting algorithm, in the case of three clones. The sketch indicates the starting state \mathcal{A} , the arrival state \mathcal{B} , levels of reactions coordinates ϕ , and two examples of branching.

ϕ will be given at the end of the section. The algorithm iteratively computes N reactive trajectories in the following manner [Fig. 3(b)]:

(i) In a stage 0, natural dynamics of the system are run. The $\{\vec{X}_i\}$ start from \mathcal{C} and evolve in time until they either reach \mathcal{A} or \mathcal{B} . If the event is rare, most, if not all, of them reach \mathcal{A} [blue lines in Fig. 3(b)]. The number of iterations is $k = 0$. I then execute the iterative mutation selection stages.

(ii) At each stage, I compute the maximum of ϕ on each trajectory $\Phi_i = \max_t[\phi(\vec{X}_i(t))]$ and order the Φ_i by increasing values [numbers assigned to blue lines in Fig. 3(b)]. The trajectory j which realizes the smallest $\Phi_j = \min_i \Phi_i$ is suppressed [number 1 in Fig. 3(b)]. This is the selection: I favor trajectories which have excursions toward \mathcal{B} . In order to keep the number of trajectories constant, a new trajectory \vec{X}_j is generated by *branching* [number 1' branched on 2 in Fig. 3(b)]. Another trajectory $\vec{X}_{l \neq j}$ is chosen: index l is drawn randomly from $\{1, \dots, j-1\} \cup \{j+1, \dots, N\}$ using a uniform distribution. Since we work with dynamics discretized with time step dt and $\Phi_l > \Phi_j$, there exists $M \in \mathbb{N}$ such that $\phi[\vec{X}_l(mdt)] < \Phi_j, \forall m < M$ and $\phi[\vec{X}_l(Mdt)] > \Phi_j$. This means that Mdt is the first time at which \vec{X}_l crosses the level Φ_j . We set $\vec{X}_j(mdt) = \vec{X}_l(mdt)$ for $0 \leq m \leq M$. Precisely enforcing this constraint is fundamental for unbiasedness of the method [23,45]. The new trajectory $\vec{X}_j(t)$ then follows its natural time discretized dynamics from $(M+1)dt$ on, until it reaches either \mathcal{A} or \mathcal{B} [for instance, the purple or red curves branched respectively on 2 and 3, Fig. 3(b)]. It uses its own realization of the noise, making it different from \vec{X}_l from the branching point on. The number of iterations is increased by 1, $k = k + 1$. This stage is repeated until all trajectories $\vec{X}_i(t)$ reach \mathcal{B} . The final k is used for calculations of passage time.

Each AMS run yields a random number of iteration k . An important intermediate quantity, the probability α of reaching

\mathcal{B} before \mathcal{A} is first calculated. An estimator of α is

$$\alpha = \left(1 - \frac{1}{N}\right)^k,$$

$$\alpha = \left\langle \left(1 - \frac{1}{N}\right)^k \right\rangle_o \simeq \left\langle \exp\left(-\frac{k}{N}\right) \right\rangle_o. \quad (5)$$

The exponential approximation of α is useful to given an order of magnitude of k from α . The \cdot indicates the random output of each AMS run and the average $\langle \cdot \rangle_o$ indicates average over o independent realisations of the algorithm (see Appendix A). When running the algorithm, I will also compute the average duration of reactive trajectories $\underline{\tau}$, the duration of nonreactive trajectories \underline{t}_2 (those that start from \mathcal{C} and go back to \mathcal{A}), and the average duration \underline{t}_1 it takes for the system to go from \vec{X}_0 to the hypersurface \mathcal{C} . It can be demonstrated (see [29]) that the estimator of the mean first passage time is

$$T = \left\langle (\underline{t}_1 + \underline{t}_2) \left(\frac{1}{\alpha} - 1 \right) + (\underline{t}_1 + \underline{\tau}) \right\rangle_o. \quad (6)$$

This can be heuristically understood by noting that there is one out of $n+1$ excursions which realizes a reactive trajectory so that $\alpha \simeq 1/(n+1)$. Indeed, the probability α is numerically defined in direct numerical simulations (DNS) of the type of stage 0 as the number of passing trajectories over N as we take the limit $N \rightarrow \infty$. One then notes that a first passage trajectory consists of $n \simeq 1/\alpha - 1$ “failing” trajectory of duration $t_1 + t_2$ and then a reactive trajectory of duration $t_1 + \tau$. The average over multiple independent realizations of AMS will always be performed before presenting results on mean first passage time T , average duration of trajectories $\tau = \langle \underline{\tau} \rangle_o$.

The acceleration provided by the use of AMS in the computation of reactive trajectories and mean first passage times is *exponential*. If one uses classical direct numerical simulations, the cost is of order T , while using AMS has a cost of order $k \propto \ln(T)$. This is particularly interesting when T increases exponentially with the control parameters of the system. Note that like every numerical procedure, the quantities calculated by AMS converge with numerical parameters. These quantities converge with the time step like \sqrt{dt} [46]. This actually originates from the convergence of hitting times calculated through a numerical discretization with a standard time scheme: no improvement is brought by classical higher order discretizations. Only times calculated using very special discretizations converge faster than \sqrt{dt} [47]. Quantities calculated by AMS also converge with grid size $dx \rightarrow 0$ [24]. The role of the number of clones N in obtaining a precise estimate of α , T , τ is peculiar [23,30,41,45,46]. By precise, I mean that α is within a small and well defined interval of confidence around $\langle \alpha \rangle_o$ with a high probability. I explain the effect of N in details in Appendix A. Once the basic improvements of ϕ are performed, there may still exist a bias in the estimation of α (etc.) by $\langle \alpha \rangle_o$ that goes to 0 when N goes to infinity. Outputs of each single run of AMS are certainly random variables whose variances decrease like $1/\sqrt{N}$. In order to obtain a precise estimate, one should average over o realizations the output of AMS calculations. The larger the N is, the fewer o are necessary. Moreover, the increase of N increases the precision of the outputs of a *single*

AMS calculation. One can therefore consider that the AMS calculations converge as $N, o \rightarrow \infty$. The use of AMS is thus a trade-off between cost ($1/dt, 1/dx, N, o$) and precision. The constraint on ϕ for convergence is that it should be sensible: the better the ϕ , the faster the convergence with N [45,46]. A poorly chosen ϕ leaves very specific traces in the statistics of k over repeated independent realizations of the algorithm, so that it is easily detected [45,46]. In practice, one chooses ϕ after a few trials in a test case which can be confronted to either theory or direct numerical simulation.

2. Reaction coordinates

Now I present the reaction coordinates I use to compute specific events (turbulence collapses, splittings, etc.) with the AMS in the case of the SPDE model of pipe flow.

When computing turbulence collapses, either of an isolated puff [Fig. 2(a)] or of a domain with steady turbulent fraction [Fig. 2(b)], I use the reaction coordinates defined as follows. Let Q_0 be the spatially averaged turbulence intensity of the initial condition. In order to generate collapses, ϕ is defined as

$$Q_0 \equiv \frac{1}{L} \int_0^L q(x, t = 0) dx,$$

$$\phi \equiv 1 - \sqrt{\frac{1}{L Q_0} \int_0^L q(x, t) dx}. \quad (7)$$

The square root extends the range of low Q in ϕ . This is useful in order to treat correctly the last stages of the collapse. In order to generate splittings, we define the turbulent fraction \mathcal{F} (relative area where $q \geq q_{\min}$), and use the reaction coordinate

$$\mathcal{F}(t) \equiv \frac{1}{L} \int_{x=0}^L \theta(q - q_{\min}) dx,$$

$$\phi_{\text{split}} \equiv \frac{\mathcal{F}(t)}{\mathcal{F}(0)} - 1, \quad (8)$$

where θ is the Heaviside step function. I chose $q_{\min} = 0.1$. In order to calculate build-ups of turbulence out of the laminar base flow, I used the reaction coordinate

$$\phi_{\text{seed}} \equiv \sqrt{\frac{1}{L Q_f} \int_0^L q(x, t) dx} - \epsilon, \quad (9)$$

where ϵ is a vanishing positive constant which defines the laminar state \mathcal{A} as a tiny ball around $q(x) = 0 \forall x$. This parameter ϵ has to be small enough so that \mathcal{A} is strictly included in the “basin of attraction” of the laminar state. The constant Q_f is the spatial average of q taken for a typical equilibrium puff [Fig. 2(a)]. In that case, the initial condition is $q = 0$ and $u = 1 \forall x$. Two supplementary independent noises $\sqrt{2/\Sigma} \eta_u$ and $\sqrt{2/\Sigma} \eta_q$, white in time and space, are respectively added to Eqs. (2) and (3). Their variances vanish as we let $\Sigma \rightarrow \infty$. These two noise terms represent the small perturbations that may exist in a controlled though imperfect experimental Hagen-Poiseuille flow. In this paper, I chose them to be white for simplicity and for consistency with the periodic, strongly modeled system I use. If this procedure were to be transposed to a direct numerical simulation, one could for instance localize this noise in space, in a small streamwise range, in order to model perturbations at the pipe inlet or at the

walls, in order to model rugosity of the pipe, external vibration in the experimental facility. In further applications to stably stratified wall flows, this can represent the impact of downward propagating internal gravity waves on the laminar wall flow [2].

Convergence in the limits $N, o \rightarrow \infty$ has been tested (see Appendix A). Precise results at a reasonable cost are obtained if $N = 1000$. I also use $dt = 0.01$ and $dx = 0.2$. Equation (2) is integrated semi-implicitly without introducing artificial drifts ([18], Sec. 4.2).

III. NUMERICAL RESULTS

I now present the numerical computations of collapse splitting and build-up trajectories as well as the corresponding mean first passage times in the SPDE model using the AMS.

A. Visualizations of turbulence collapses, puff splittings, and puff build-ups

I present five typical examples of trajectories in the form of spatiotemporal diagrams of q [Figs. 4(a) and 4(b)], $1 - u$ [Figs. 4(c) and 4(d)], and $\ln(|q|)$ [Fig. 4(e)]. Refer to Fig. 2(c) in order to situate the value of Reynolds number r of each simulation.

The turbulence collapses are computed using reaction coordinate ϕ [Eq. (7)]. I first present the collapse of an isolated puff at $r = 1$ [Fig. 4(a)]. This event occurred in a pipe of length $L = 800$. This is the most studied event in such situations. In the model, the collapse occurs as a block: q decreases relatively homogeneously in the puff with very little shrinking until the puff has collapsed. The mean first passage time before collapse of this puff is $T \simeq 1.3 \times 10^4$. I then present a puff splitting at $r = 0.95$ [Fig. 4(b)]. The splitting was calculated using reaction coordinate ϕ_{split} [Eq. (8)]. This illustrates the well known fact that puff splitting occurs through the extension of the puff on the leading edge and then splitting in the middle (see [4,6] and references within). The splitting occurs after a mean first passage time $T \simeq 2 \times 10^8$. Provided $L \gtrsim 300$, the collapse or splitting of isolated puffs is independent of the pipe length.

I then present events which are harder to sample in direct numerical simulations. They concern the collapse of a pipe starting with the steady turbulent fraction. I first present a peculiar regime which manifests itself in the lower range $1 \lesssim r \lesssim 1.2$ of laminar turbulent coexistence [see Fig. 2(c)], provided the pipe is not too long. A spatiotemporal diagram of $1 - u$ is given in Fig. 4(c). In that case, at $r = 1.1$, the collapse of turbulence occurs puff by puff, quite independently from one another, and takes a relatively long time $\tau \simeq 3000$, while the collapse of each of the puffs has a relatively shorter duration of order $\mathcal{O}(10)$. The mean first passage time before collapse of turbulence in this pipe of length $L = 1200$ at $r = 1.1$ is $T \simeq 10^{17} \gg \tau$. In very long pipes or provided $r \gtrsim 1.2$, the collapse of turbulence in a pipe initially having the steady turbulent fraction is quite generic. Let us present an example of such a collapse with a spatiotemporal diagram of $1 - u$ at $r = 1.2$ and $L = 1200$, for which $T \simeq 10^{32}$ [Fig. 4(d)]. The collapse occurs in a short time $\tau \simeq 35$. In fact, the collapse of all the puffs is simultaneous and no more independent. What most likely changed between $r = 1.1$ [Fig. 4(c)] and $r = 1.2$

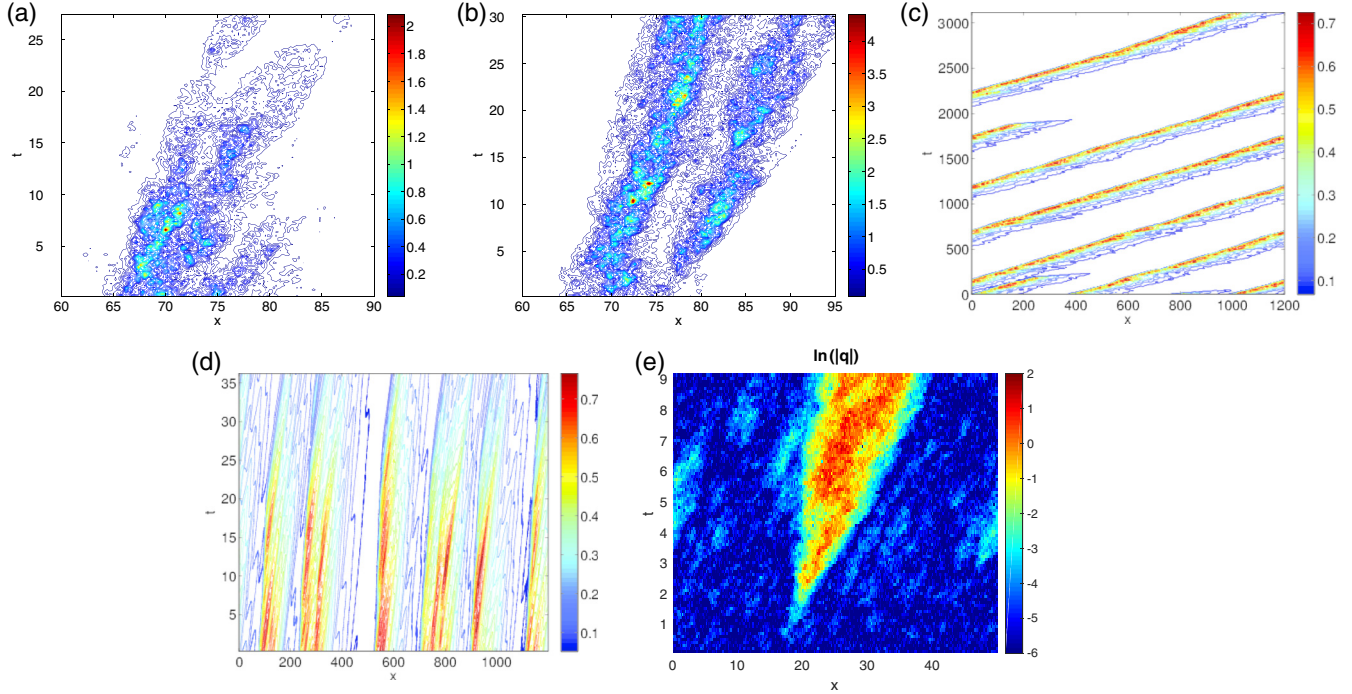


FIG. 4. Spatiotemporal diagrams of reactive trajectories in the SPDE model computed using AMS. (a) Collapse of an isolated puff at $r = 1$. (b) Splitting of a puff at $r = 0.95$. (c) Collapse of a domain of length $L = 1200$, initially with steady turbulent fraction, at $r = 1.1$. (d) Collapse of a domain of length $L = 1200$, initially with steady turbulent fraction, at $r = 1.2$. (e) Color levels of the logarithm of turbulence intensity for a build-up of a turbulent puff out of noise of vanishing variance at $r = 1.15$ and $\Sigma = 25000$.

[Fig. 4(d)] is the increased probability of puff splitting. This makes the long lasting existence of a gap between puffs very unlikely, so that the most probable path for turbulence collapse necessitates that no gap is left. Consequently, the collapse of all the puffs occurs at the same time.

I eventually present the build-up of a turbulent puff, starting from the laminar base flow, out of a noise of vanishing variance. I use the reaction coordinate ϕ_{seed} [Eq. (9)] and I set $\Sigma = 25000$ and $r = 1.15$ in a domain of length $L = 50$. A spatiotemporal diagram of the logarithm of the turbulent intensity $\ln(|q|)$ showing the development of the puff from a laminar state stimulated by noise of very small variance is given in Fig. 4(e). The generation of the puff occurs through the selection of a germ of small spatial length and small but finite amplitude at $t \simeq 1.5$. The germ is created by an extreme realization of the noise of small variance. The germ then grows in amplitude, the spatial maximum of the structure saturates ($t \simeq 3.5$), and it extends in length until it reaches a state similar to that of an equilibrium puff. The logarithm of the mean first passage time out of the laminar state is $\mathcal{O}(10^3)$ at this Σ . The event is *very* rare. Such a large Σ was used to check that I indeed stimulated a germ of turbulence from the laminar state and did not simply push a flow that already contained this germ. Indeed, if the flow already contained the germ, there is no more saddle to pass. The probability of developing a puff would then converge toward a finite value as the variance of the added noise vanishes, $1/\Sigma \rightarrow 0$. The mean first passage time would also converge toward a finite value. Meanwhile, the probability of developing a puff would tend to 0 as $1/\Sigma \rightarrow 0$ and the mean first passage time would diverge exponentially if the saddle had not been crossed and a germ had to be developed

[17,24,25,48]. This is the case of the presented example. The difference in the Σ dependence of T is a confirmation that there exists a saddle point between the laminar state and the one puff state in the model equations (2) and (3).

B. Isolated puffs: Reynolds number dependence of mean first passage times

I now consider the collapse and splitting of isolated puffs [see Figs. 2(a), 4(a), and 4(b)] systematically and quantitatively. I will compute the duration of collapse trajectories of isolated puffs (distribution, average τ , and variance) and the mean first passage time before both events T in the relevant range of Reynolds number [see Fig. 2(c) for comparison] starting from a range of initial conditions.

I first consider the duration of collapse trajectories. I use isolated puffs in pipes of lengths $200 \leq L \leq 1600$ at $r = 1.0$. I perform AMS computations of puff collapse using ϕ [Eq. (7)] and $N = 20000$ clones. Using the 20000 computed reactive trajectories, a sample large enough to faithfully represent the variety of collapses, I compute the empirical distribution of collapse duration $\mu(s)$ as well as the average collapse duration $\tau = 29.2 \pm 0.2$ (for all sizes) and the variance of duration $\sigma = 9.8 \pm 0.1$ (for all sizes). Note that this is the physical distribution of collapse durations. The quantities $\mu(s)$, τ , and σ can in principle be sampled in DNS, at a much higher cost, by accumulating an equivalent number of collapses and using the corresponding durations to perform the statistics. In this case, this is done using approximately 2000 trajectories sampled out of 10^7 trials at $r = 1.0$ in a domain of size $L_x = 200$. The average and variance of trajectory durations calculated

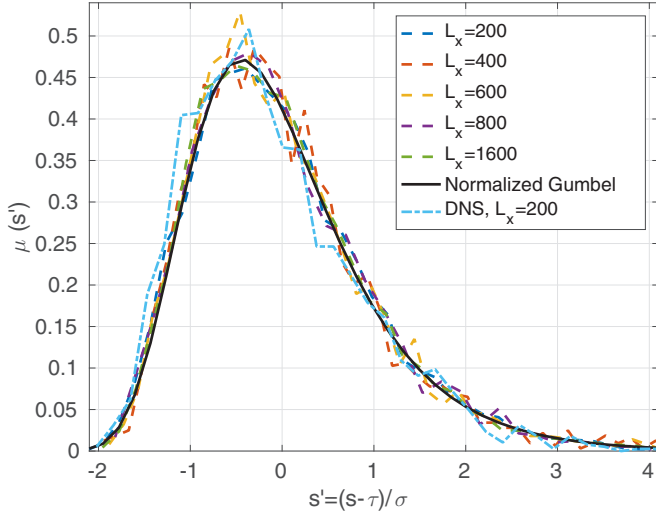


FIG. 5. Normalized distribution of duration of collapses of isolated puffs in pipes of lengths $200 \leq L \leq 1600$ at $r = 1.0$, sampled by mean of AMS and DNS, compared to a normalized Gumbel distribution.

by mean of DNS is equal (up to the uncertainties) to those calculated by AMS.

I display the distributions of durations, originating from AMS and DNS, for all considered lengths as a function of the normalized duration $s' = (s - \tau)/\sigma$ (Fig. 5). On top of that, I add a normalized Gumbel distribution. The normalized Gumbel distribution reads

$$\mu(s') = \frac{\pi}{\sqrt{6}} \exp \left[-\frac{\pi s'}{\sqrt{6}} - \gamma - \exp \left(-\frac{\pi s'}{\sqrt{6}} - \gamma \right) \right], \quad (10)$$

where $\gamma \simeq 0.6$ is the Euler constant. This choice of comparison originates from the fact that it has been demonstrated that for one degree of freedom system, the duration of reactive trajectories follows a Gumbel distribution [26]. Note that this result is demonstrated by considering only the stochastic process, independently of the simulation or sampling procedure. This is numerically verified [46] (for standard behavior). Similarly, the variance of durations σ could be explained by a physical model in the case of random walk of fronts in the Ginzburg-Landau equation [24]. In the case of collapse in this model of pipe flow, the comparison between the distribution of duration of reactive trajectories and the normalized Gumbel is very good. In a spatially extended system, namely the one dimensional Ginzburg-Landau equation, it has been checked that the duration of reactive trajectories followed a distribution very similar to a Gumbel [24]. The comparison was very precise at large durations $s \geq \tau$.

I then compute the mean first passage time before collapse of isolated turbulent puffs as a function of the Reynolds number r in pipes of length $L = 800$ in two types of AMS calculations (both using $N = 1000$ clones).

Using direct numerical simulations of the model, I generated a set of initial conditions which are equilibrium puffs at r_0 in the range of Reynolds number $0.9 \leq r_0 \leq 1.15$. In the first type of AMS calculation, I compute the mean first passage time $T_0(r)$ before collapse of each of these puffs in the range of Reynolds number $0.6 \leq r \leq 1.8$. This covers

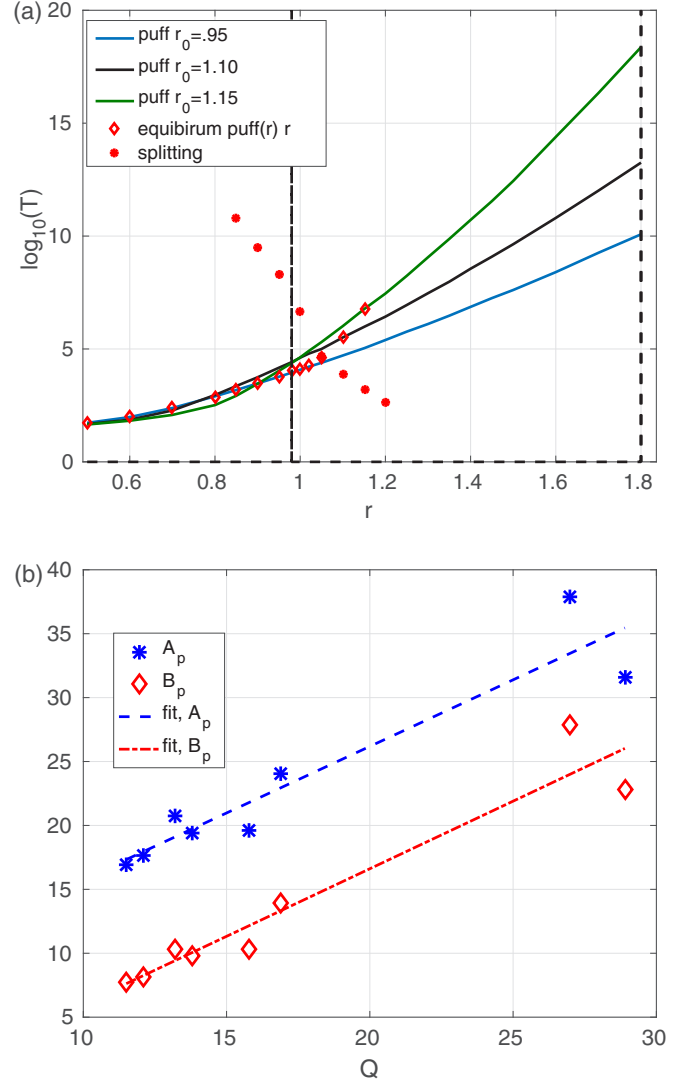


FIG. 6. (a) Decimal logarithm of the mean first passage time of several events as a function of the Reynolds number r in a pipe of length $L = 800$: collapse fixed puffs (lines), collapse of near equilibrium puffs (red lozenges), as well as splitting of near equilibrium puff (red dots). The vertical dashed lines indicate the range of laminar turbulent coexistence [Fig. 2(c)]. (b) Measured slopes A_p (blue asterisks) and ordinate at the origin B_p (red lozenges) for equilibrium puffs at different Reynolds numbers as a function of the total intensity of turbulence of said puffs Q . The dashed lines indicate the affine fits of these data.

the range of r where turbulence cannot be sustained and the range where turbulence coexists with laminar flow with a finite turbulent fraction [Fig. 2(c)]. Since this concerns the collapse of turbulent puffs which are not changed (while the Reynolds number is), these events are termed “collapses of fixed puffs” (in Fig. 6 for instance). Due to the definition of reaction coordinates and of mean first passage times, I condition the dynamics to have a turbulent fraction which is no greater than that of the initial conditions. Such a conditioning is well taken into account theoretically by fixing the necessary boundaries, since we work with stochastic dynamics (see [18], Sec. 5.5 and [25], Sec. 6.1). The numerical studies of the collapse

of isolated puffs actually always include some conditioning, since it is often required that the puffs do not split when mean first passage times are studied in models [16,37]. In this case this is taken one step further in order to investigate how the structure of the phase space of the stochastic model (via the corresponding initial conditions) leads to the dependence of T on r for near equilibrium puffs. From general theoretical argument, one expects that values of T for two initial conditions remain distinct if the control parameters are changed, provided that the initial conditions are not both within one standard deviation of the most probable state. The difference between values of T can actually be written formally in an integral form (see [18], Sec. 5.5 and [25], Sec. 6.1) which is rarely tractable, unless the model is a gradient system. In said gradient systems, the logarithm of the ratio of two conditioned mean first passage times from two distinct initial conditions is actually given by the potential difference between these two initial conditions. This potential difference can grow or decay with control parameters, depending on the system. Performing this type of numerical experiment will thus inform us of how the phase space of the stochastic model changes with r . Provided the revealed dependence on r is simple enough, I may provide a clear description of how the mean first passage time of the near equilibrium puff depends on r .

I present the decimal logarithm $\log_{10}[T_{r_0}(r)]$ for three puffs at equilibrium at $r_0 = 0.95$, $r_0 = 1.10$, and $r_0 = 1.15$ [Fig. 6(a)]. Two regimes of $\log_{10}[T_{r_0}(r)]$ appear. For $r \leq 0.8$, all the curves collapse and $T_{r_0}(r)$ is nearly independent of the initial condition. Only the durations of the collapse trajectories τ , included in T [see Eq. (6)], differ. For $r \geq 1$, the mean first passage time before collapse of these puffs is clearly distinguishable from one initial condition to another. In fact, an affine tendency $\ln[T_{r_0}(r)] = A_p(r_0) \times r - B_p(r_0)$ appears, with $A_p > 0$ and $B_p > 0$. The growth rate $A_p(r_0)$ and ordinate at the origin $B_p(r_0)$ are systematically calculated by affine fits. We find that both $A_p(r_0)$ and $B_p(r_0)$ are both growing affine functions of the initial total turbulent intensity Q_0 of each puff [Fig. 6(b)]. This can be written as

$$\begin{aligned} Q_0(r_0) &\equiv \int_{x=0}^L q_{r_0}(x,0) dx, \\ A_p(r_0) &= A'_p(Q_0(r_0) + A''_p), \\ B_p(r_0) &= B'_p(Q_0(r_0) + B''_p), \quad A'_p \simeq B'_p. \end{aligned} \quad (11)$$

I now present the mean first passage time before collapse and splitting of near equilibrium puffs [Fig. 6(a), red symbols]. This is done for illustrative purpose. The two curves cross near $r = 1.0$, above which splitting becomes more probable than collapse. The slight discrepancy between the Reynolds number threshold of sustained turbulence obtained from the spatial averages [Fig. 2(c)] and the Reynolds number at which $T_{\text{collapse}} = T_{\text{splitting}}$ comes from small errors in both estimations. Since they are not the main focus of this paper, these two Reynolds numbers are estimated using smaller data sets and thus come with an error. The mean first passage time before collapse of equilibrium puffs is actually a subset of the results presented in the former paragraph, since the equilibrium puff at r_0 is only considered as an initial condition for AMS calculation at r_0 . This actually shows how the

former result [Eq. (11)] controls the mean first passage time before collapse of equilibrium puffs. Indeed, if one knows the ensemble average of the total turbulence intensity $Q_0(r)$ of the equilibrium puff at r , one can then deduce the mean first passage time before of equilibrium puffs at this Reynolds number, since $\ln(T) = A'_p(Q_0(r) + A''_p) \times r + A'_p[Q_0(r) + B''_p]$. Performing these two types of experiments showed us that the mean first passage time before collapse of equilibrium puffs could be separated into two parts. First, the reaction of the system to any puff, at a given Reynolds number, gives the affine dependence $\ln(T) = A_p r - B_p$. Second, the amplitude of the equilibrium puff yields the two coefficients A_p and B_p .

The mean first passage time before build-up of turbulence has been considered quantitatively for a range of r and Σ (not shown here). One finds that $\ln(T)/\Sigma \simeq 0.035$. This depends very little on r .

C. Steady turbulent fraction: Length and Reynolds number dependence of mean first passage times

I now consider the collapse of turbulence in pipes initially containing their steady turbulent fraction at r [see Fig. 2(b), examples Figs. 4(c) and 4(d), and dependence of Q on r in Fig. 2(c)]. Unlike the isolated puff, this starting point is always easily defined. The mean first passage time before collapse of turbulence T and the average duration τ of turbulence collapses in these pipes are computed as a function of r and L systematically in the length range $50 \leq L \leq 1600$ and in the Reynolds number range $1.05 \leq r \leq 1.6$, where such a steady turbulent fraction can be sustained in the infinite length limit.

I first present $\ln[T(L)]$ at $r = 1.15$ in Fig. 7(a). The linear growth of $\ln(T)$ with L is visible. A linear fit is performed to calculate the growth rate. The fit is added to the plot in my example. In practice the exponential growth of T with L , $\ln(T) = f(r) \times L$, is found no matter the Reynolds number in the range [1.05; 1.6] [Fig. 7(b)]. Note that the mean first passage time before the collapse of isolated puffs is independent of the length of the pipe provided that it is large enough (for instance $L > 200$ at $r = 1.15$). This is illustrated at $r = 1.15$ by adding the logarithm of the mean first passage time before the collapse of isolated puffs as a function of L in Fig. 7(a) (red curve). I then consider the growth rate $f(r)$ of T as a function of r . It is calculated systematically in the range $1.05 \leq r \leq 1.6$ and displayed in Fig. 7(c). This function $f(r)$ is in good approximation affine with r , $f(r) \simeq Ar - B$ with $A, B > 0$. We can deduce this tendency in a large range of length and Reynolds number in the SPDE model. Note that it had previously been seen in a smaller range of length and Reynolds number in direct numerical simulations [19].

I eventually present the logarithm of the average durations of collapse trajectories as a function of the pipe length in the range of Reynolds number [1.05; 1.6] [Fig. 7(d)]. Error bars are added: they are computed using the variance of τ over the realizations of the algorithm used to compute τ . While there is a slightly larger uncertainty over the value of $\ln(\tau)$ in the intermediate range of $1.2 \leq r \leq 1.3$, these error bars show that the tendencies of $\tau(r, L)$ can be trusted and do not arise from large uncertainties over the estimate. We can distinguish two types of behavior of $\ln(\tau)$. For $r < 1.2$, the average duration of trajectories grows very fast with L after a threshold pipe

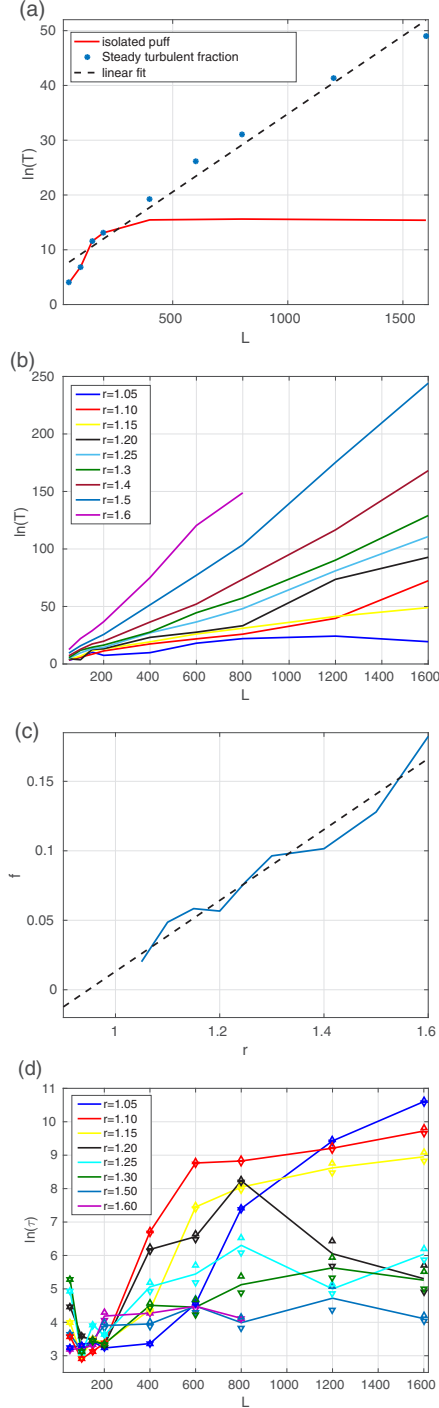


FIG. 7. (a) Logarithm of the mean first passage time before collapse of turbulence as a function of length at $r = 1.15$ in pipes initially containing either the steady turbulent fraction or an isolated puff. The black dashed line is the linear fit of $\ln[T(L)]$. (b) Logarithm of the mean first passage time before collapse of turbulence as a function of length for increasing r in pipes initially containing steady turbulent fraction. (c) Slope $f(r)$ of the linear fit of the logarithm of the mean first passage time before collapse of turbulence as a function of the Reynolds number. (d) Logarithm $\ln(\tau)$ of duration of reactive trajectories of collapse of laminar-turbulent pipes as a function of pipe length for increasing Reynolds number. The error bars are determined using the variance of the estimate of τ over several realizations of AMS computations.

length ($L \geq 200$ for $r \geq 1.1$) and saturates. This threshold pipe length decreases with the Reynolds number. These very long durations are due to the specific type of collapse, which involves independent collapse of each puff [Fig. 4(c)]. The threshold pipe length is simply given by the length above which the pipe can contain two puffs at a given Reynolds number. The duration saturates when the pipe is long enough for splitting to be probable in one of the many holes created by puff collapses ($L \geq 600$ for $r \geq 1.1$). At $r = 1.2$, independent puff collapses occur for intermediate length ($200 \leq L \leq 800$), while global collapses [Fig. 4(d)] become the typical event in long pipes ($L \leq 1200$). At larger Reynolds numbers ($r \geq 1.3$), the collapse duration varies little with length and is two orders of magnitude smaller than what is found at lower Reynolds numbers. The pipe undergoes a global collapse of turbulence. This systematic examination of collapse duration thus gives a quantitative criterion to determine whether the collapses happen independently puff by puff, or globally.

IV. THEORY AND INTERPRETATION

The Reynolds number and length scalings of mean first passage times before collapse of turbulence in the SPDE model can be interpreted and discussed in view of theories of metastability. In Sec. IV A, I consider the collapse of turbulence in pipes initially containing the steady turbulent fraction and I interpret the mean first passage time dependence on length. This will take the point of view of small variance local noise. I show in Sec. IV B how the Reynolds number and size dependence of the mean first passage times $T \asymp \exp[L(Ar - B)]$ can be derived in a two degrees of freedom model. This will take the point of view of large pipe lengths. This will again mostly concern the global collapse of laminar-turbulent pipes, but should also bring some insight on the mean first passage time before collapse of isolated puffs. Relying on these results is entirely consistent since the model I analyzed with AMS is stochastic.

A. Length scaling of mean first passage times before collapse of turbulence in pipe containing the steady turbulent fraction

The first stage of my discussion consists of justifying the persistence of a single scaling in size for T , while collapse trajectories display very different features as r and L are varied. More generally, this discussion also serves as a reminder that independent events, which had been invoked to explain the $\ln(T) \propto L$ scaling, are not the only phenomena which can lead to such a size scaling.

In this section, I use properties of T which are demonstrated in the limit of zero noise variance $\beta \rightarrow \infty$, but nevertheless control T if the noise variance is fixed but relatively small. Indeed, most theoretical results on multistability concerning mean first passage times T are demonstrated in the limit of noise of variance $1/\mathfrak{B} \rightarrow 0$ for a stochastic process with an action \mathcal{S}_t ,

$$\dot{\vec{X}} = \vec{b}(\vec{X}) + \sqrt{\frac{2}{\mathfrak{B}}} a(\vec{X}) \vec{\eta}, \quad \langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t'), \quad (12)$$

$$\mathcal{S}_t = \frac{1}{4} \int_{s=0}^t (\dot{\vec{X}} - \vec{b})^\dagger (a a^\dagger)^{-1} (\dot{\vec{X}} - \vec{b}) ds. \quad (13)$$

Here arrows designate vectors, a is a matrix, and \dagger designate the transpose. The matrix (aa^\dagger) is the correlation matrix of the noise: the total variance of the noise stands as an independent factor $1/\mathfrak{B}$. The term \vec{b} is the deterministic part of the Itô process. This is a formal rewriting of systems such as Eqs. (2) and (3). The results on T take the form of a *large deviation principle*: a probability is exponentially small in the inverse noise variance [25,31]. This is written more rigorously on the logarithm of said mean first passage time in the limit of small variance noise,

$$\lim_{\beta \rightarrow \infty} -\frac{1}{\mathfrak{B}} \ln(T) = I(\mathfrak{X}). \quad (14)$$

This rate function I may very well depend on the control parameters of the system with the strict exception of the felt noise variance. Moreover, this function is not some unknown: it depends in a very regular manner on the properties of the stochastic differential equation of interest. These results are relevant for a broader range of parameter where the noise variance is small $1/\mathfrak{B} \lesssim 1/\mathfrak{B}_c$, that is to say smaller than some minimal threshold. Indeed, many properties on the asymptotic regime, such as the prevalence of instanton-type trajectories, are found in stochastic systems where the inverse noise variance is not varied or turbulent flows where such an inverse noise variance is not a control parameter [49].

There are several ways of calculating the rate function I . A systematic approach can consist in minimizing the action \mathcal{S} of the system [Eq. (13)] over path $\vec{X}(s)$ and duration t (see [24,32,33,38,50], and [25] Sec. 6.1). The action is the time integral of a Lagrangian measuring the amount of noise felt by reactive trajectories [Eq. (13)]. This yields a set of Euler-Lagrange equations, whose integration gives the reactive trajectory. The rate function I is the infimum of the action. Another approach can be using the Gärtner-Ellis theorem: calculating a characteristic function then performing a Legendre transform, an operation similar to using different ensembles in thermodynamics (see [25] Sec. 3.3.1, [34] Secs. 2.3, 6.5, B 1). A less systematic approach can consist of a direct calculation. The rate function $I = \Delta\mathcal{V}$ is given by the difference of a function $\mathcal{V}(\vec{X})$ ($\vec{X} = \{u, q\}$ in this case), termed a quasipotential, between two very specific points \vec{X} of the phase space: the starting point \vec{X}_A of the first passage and \vec{X}_S , the lowest saddle of \mathcal{V} between the starting point and the arrival point \mathcal{B} [31],

$$\Delta\mathcal{V} = \mathcal{V}(\vec{X}_S) - \mathcal{V}(\vec{X}_A). \quad (15)$$

This quasipotential difference can be understood by examining the most probable reactive trajectory, termed the *instanton*, which minimizes the action \mathcal{S} [24,25,32,33]. It first contains a fluctuation path, from the starting point to the saddle, stimulated by noise, on which the action is strictly positive. This yields the quasipotential difference. It then has a relaxation path $\dot{\vec{X}} = b(\vec{X})$, which is deterministic in the zero noise limit, on which the action is zero and which does not contribute to the quasipotential difference.

The name quasipotential arises from the study of multistability in gradient systems, where $\frac{d\vec{X}}{dt} = -\vec{\nabla}V + \sqrt{2/\mathfrak{B}}\vec{\eta}$. In that case the potential $V = \mathcal{V}$ governs the whole deterministic part of the dynamics. In such a gradient system, Eq. (14) is in

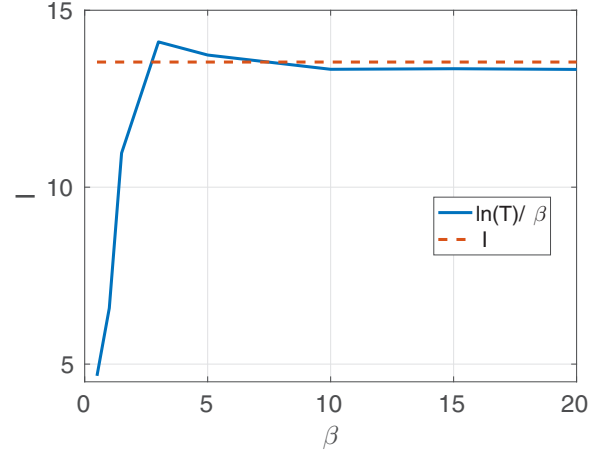


FIG. 8. Rate function I of the mean first passage time before collapse of a laminar-turbulent domain as a function of β [the inverse variance of the multiplicative noise of Eq. (3)], calculated by two manners: ratio of $\ln(T)/\beta$ and slope of the best linear fit of $\ln(T)(\beta)$. The computation is performed using a domain length of $L = 200$ and a Reynolds number of $r = 1.3$.

fact a well known result; it derives from the Eyring-Kramers formula of the mean first passage time [17,48], which is even more widely known as the *Arrhenius law*, owing to its original discovery in kinetic chemistry. The mean first passage time can be entirely calculated in the low noise limit in gradient systems and extensive comparisons with results of AMS calculations can be performed [24]. In the nongradient SPDE model, such a quasipotential would require extensive derivations in order to be analytically readable. However, it is sufficient to know that it exists and to know some of its general properties in order to propose a first discussion of the numerical results.

If we make the loose assumption that the noise locally felt by the puffs has a small variance when $\beta = 1.5$ in Eq. (3), the mean first passage time before collapse of turbulence is governed by the properties of such a quasipotential. This can be checked by performing the artificial experiment where we follow the mean first passage time before collapse of laminar turbulent domain in the model of Eqs. (2) and (3) when β is varied. This is done at $r = 1.3$ and for a domain length of $L = 200$. When examining the logarithm of the mean first passage time as a function of β , I note that this function is very close to linear, hinting toward large deviations of T in the large β limit. We can thus define a rate function I as the best linear fit of $\ln(T)$ function of β . In order to determine how good an approximation of $\ln(T)$ the product $I \times \beta$ is, I display $\ln(T)/\beta$ and I in Fig. 8. This shows that we have $\ln(T) \propto \beta$, for $\beta \gtrsim 2$, with a proportionality factor which has the properties of the quasipotential difference in the large β limit. This justifies the use of results on large deviations of T in the large β limit to give an approximation of T , even if taking this limit is physically artificial. While there can be a 10–20% error on the estimate of $\ln(T)$ by this manner, the variations of I still strongly influence those of $\ln(T)$. This cannot be used if one wants to explain fine changes of T . However, this is acceptable if one wants to discuss changes of $\ln(T)$ by an order of magnitude [Fig. 7(b)]. The length L dependence of the mean first passage time can thus be discussed in view of

results on other systems. In particular, I will invoke the length dependence of potentials and quasipotentials when the reversal of the field is nearly homogeneous in space [Figs. 4(a) and 4(d)] [24]. Such reversals were termed a flipping of the field. In that case, it has been shown analytically and verified numerically that the potential was linear in length $\Delta V \propto L$, in the case of a gradient system. This length scaling originates from the necessity of having a correlated movement of the field at each point of space. This scaling can strongly influence the length dependence of the mean first passage time even if β is finite though relatively small. I can thus explain the three regimes of exponential length dependence, two of which exist for all Reynolds numbers:

(i) If the domain is small and can accommodate only one puff, we have a global collapse of said puff, a flipping of the field, and the quasipotential goes like $\Delta V = L\delta\mathcal{V}$, with $\delta\mathcal{V}$ a density of quasipotential, leading to $\ln(T) \simeq L\beta\delta\mathcal{V}$ at finite β . This regime exists for all Reynolds numbers. The range of pipe length concerned decreases with r as puff density and steady turbulent fraction increase with r [Fig. 2(c)].

(ii) If the domain is larger and can accommodate several puffs *and* if the Reynolds number is not so large that puff splitting is rare enough, then the puffs collapse independently from one another [Fig. 4(c)]. Let us term T_1 the mean first passage time for the collapse of any of the n puffs contained in the domain. The mean first passage time for the collapse of turbulence in the whole pipe is then $T = T_1^n$. Since the collapse of each single puff occurs as a block, one still has $\ln(T_1) \simeq \beta L_{\text{puff}}\delta\mathcal{V}$ in the large β limit. one also has $L = nL_{\text{puff}}$. As a consequence $\ln(T) \simeq \beta n(L/n)\delta\mathcal{V}$ at finite β . One then finds $\Delta V = L\delta\mathcal{V}$. Note that this mechanism was invoked in Ref. [19] for small length systems.

(iii) If the domain is large enough that it can accommodate several puffs and if the Reynolds number is large enough that splitting becomes very probable, then all the puffs collapse together [Fig. 4(d)]. The argument is now the same as in the small scale case; one finds again $\Delta V = L\delta\mathcal{V}$ with the same quasipotential density. This again leads to $T \simeq \exp(L\beta\delta\mathcal{V})$ at finite β .

In any case, one finds that $\ln(T) \simeq \beta\delta\mathcal{V}L$. If one assumes that $\beta\delta\mathcal{V} = (Ar - B)$, we thus have the calculated scaling in length $\ln(T) = L(Ar - B)$ [Figs. 7(a) and 7(b)] of the mean first passage time before collapse of turbulence, throughout all type of collapse.

B. Reynolds number dependence of mean first passage time before collapse of turbulence derived from a quasipotential

Another point of view can be taken in order to discuss the size and Reynolds number scalings of mean first passage times and derive that $T \propto \exp[L(Ar - B)]$. If the Reynolds number r is not so small or the length L is large enough that the turbulence in the pipe collapses as a whole, one can consider the collapse toward zero of quantities like the spatially average intensity of turbulence Q , or the spatially averaged kinetic energy of turbulence. This analysis does not apply to the independent collapse of puffs which occur in the narrow range $1.0 \leq r \leq 1.15$. In that case, instead of invoking somewhat artificial large deviations in the limit of small variance of the *ad hoc* local noise felt by the system

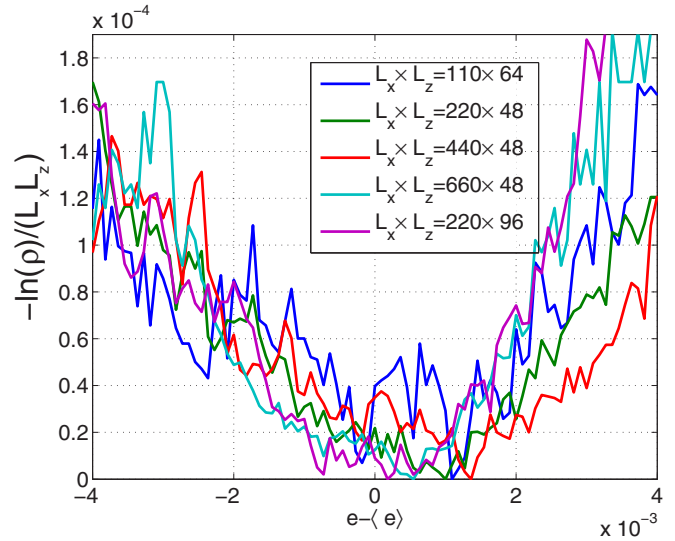


FIG. 9. Rescaled probability density function ρ of the kinetic energy of turbulence e sampled in numerical simulations of plane Couette flow (from the data of [3]).

to approximate T at finite but relatively large β , one can more naturally and more rigorously consider large deviations in the limit of large length [51]. Indeed, one very naturally finds that $\lim_{L \rightarrow \infty} \ln(T)/L = I(r, \mathcal{L})$ in the results of AMS calculations. Such a behavior is not surprising since it has already been reported in the probability density functions ρ of the kinetic energy of turbulence e in plane Couette flow [3]. This is shown again in Fig. 9. This is particularly the case when the PDF are not Gaussians any more, when the length dependence could not be related to central limit behavior, i.e., when the pdfs have exponential tails for $e < \langle e \rangle$ (Fig. 9). This large deviations behavior of the pdfs can be understood with the same quasipotential as for the mean first passage time. Indeed, this quasipotential also governs the probability density function, which has a Boltzmann factor form $\rho \propto \exp(-\beta\mathcal{V})$ [31,52], where β scales the inverse variance (with $\beta \propto L_x L_z$ in plane Couette flow). This motivates us to reduce the quasipotentials in models of wall flows.

After having introduced the general concepts used to study metastability, I now apply them to a simple but enlightening model of transitional wall flow. I will show that the scaling $\lim_{L \rightarrow \infty} \ln(T)/L = AR - B$ can be derived in that case. Following the idea of large deviations in the large length limit, this model should be understood as a model for some spatial averages. While a theoretical analysis of the transitions in the SPDE model may be possible, it would require a thorough technical work which is outside the scope of the current paper. I will proceed in three steps. I first present the model and its basic deterministic properties and show what type of noise should be included to faithfully represent spatially averaged behavior in Sec. IV B 1. I then compute the quasipotential \mathcal{V} which can then be used to determine the probability density functions in Sec. IV B 2. I eventually use the quasipotential to compute analytically the mean first passage times before collapse and the corresponding trajectories, and show how a scaling of the type $T \propto \exp[L(Ar - B)]$ can arise in Sec. IV B 3.

1. Conceptual model

The two degrees of freedom model I use was originally proposed by Dauchot and Manneville [28]. It arises from a projection of principle and truncation of the Navier-Stokes equations on the first wall normal Fourier mode of u_x , the departure to the streamwise base flow, represented by X_1 , and second wall normal Fourier mode of u_y , the wall normal velocity field, represented by X_2 . The two coupled ordinary differential equations read

$$\frac{dX_1}{dt} = s_1 X_1 + X_2 + X_1 X_2, \quad (16)$$

$$\frac{dX_2}{dt} = s_2 X_2 - X_1^2. \quad (17)$$

The first linear terms $s_{1,2} X_{1,2}$ arise from viscous dissipation. For a higher readability of the analysis, I set $s_1 = s_2 = -1/R \leq 0$. The Reynolds number is thus denoted by R . The additive term $+X_2$ comes from advection by the linear base flow and the quadratic terms, from advection by the departure to the base flow. This model contains many of the basic features of transitional wall flows. The fixed point $X_1 = X_2 = 0$, which corresponds to no nontrivial flow on top of the linear flow, is linearly stable for all Reynolds numbers. The linear operator arising from the linearization of the ODEs about $(0,0)$ is highly non-normal, due to $+X_2$ in Eq. (16). This leads to transient growth of the type $t \exp(-t)$ of optimal perturbations to the state $(0,0)$ on both sides of the boundary of the basin of attraction of the laminar state. The model includes lift-up in this manner [40]. If $R \geq 2$, two other fixed points arise from a saddle node bifurcation. They correspond to

$$X_2^\pm = -R X_1^\pm, \quad X_1^\pm = \frac{-1 \pm \sqrt{1 - \frac{4}{R^2}}}{2}, \quad (18)$$

and X_2 explicitly reads

$$X_2^\pm = \frac{-(R - \frac{2}{R}) \pm \sqrt{R^2 - 4}}{2}. \quad (19)$$

The existence of a finite amplitude nontrivial state distinct from the stable linear flow is represented by the $-$ solution and the saddle point in between is represented by the $+$ solution. Note that this saddle node bifurcation can occur because of the non-normal term $+X_2$ in Eq. (16). All the fixed points are illustrated in the bifurcation diagram of Fig. 10. To some extent, X_1 is akin to the spatial average $1 - \mathcal{U}$ and X_2 is akin to the spatial average \mathcal{Q} . I first note that X_1 arises from u_x , like u , while X_2 arises from u_y like q . Also, the stable nontrivial solution X_1^- will saturate quickly, while more and more energy is given to the mode X_2^- , as is seen for the two spatially average variables [Fig. 2(c)]. Given the low dimensionality of the model, these solutions are time independent. Note that in the large Reynolds number limit $R \rightarrow \infty$, a Taylor expansion shows that $X_{1,2}^\pm$ follow simple scaling laws,

$$X_1^- = -1 + \mathcal{O}\left(\frac{1}{R^2}\right), \quad X_2^- = -R + \mathcal{O}\left(\frac{1}{R}\right), \quad (20)$$

$$X_1^+ = -\frac{1}{R^2} + \mathcal{O}\left(\frac{1}{R^4}\right), \quad X_2^+ = -\frac{1}{R^3} + \mathcal{O}\left(\frac{1}{R^5}\right). \quad (21)$$

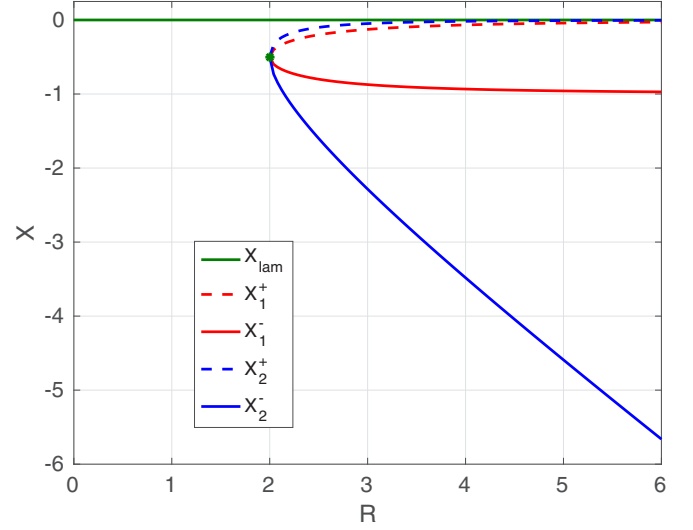


FIG. 10. Bifurcation diagram of the deterministic two degrees of freedom model, as a function of Reynolds number R , including both X_1 and X_2 . The continuous lines indicate the stable fixed points. The dashed lines indicate the unstable fixed points. The green dot indicates the saddle node bifurcation.

These scaling laws are visible in the bifurcation diagram for $R \gtrsim 3$. I recover here another classical property of transitional wall flows. The kinetic energy of perturbations necessary to cross the boundary between turbulent and laminar base flow decreases like a power of R (see [13] and references within). One can find similar models arising from a similar procedure in geophysical fluid dynamics [53]. However, their statistics are more complex than those I consider here [54].

In order to discuss transitions from the nontrivial fixed point $X_{1,2}^-$ to the laminar fixed point $(0,0)$, a measure of complexity, which was integrated out in the model, should be reinserted. In accordance with the SPDE studied numerically, I do so by adding noise to the model, with inverse variance $\bar{\beta}$, which now reads

$$\frac{dX_1}{dt} = -\frac{1}{R} X_1 + X_2 + X_1 X_2 + \sqrt{\frac{2}{\bar{\beta}}} g_a(X_1) \eta_1, \quad (22)$$

$$\frac{dX_2}{dt} = -\frac{1}{R} X_2 - X_1^2 + \sqrt{\frac{2}{\bar{\beta}}} g_a(X_1) \eta_2, \quad (23)$$

$$\langle \eta_i(t) \rangle = 0, \quad \langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t'). \quad (24)$$

I use the variable X_1 in the multiplicative noise led by $\eta_{1,2}$, which are uncorrelated and white in time. There are several reasons for this. First, we will see that X_1 is the fastest variable of the two. This choice of course gives to the noise the property of going to zero when there is no turbulence. More importantly, X_1 remains bounded by 1 and does not grow much with the Reynolds number. As a consequence, this choice also gives the property that the variances of X_1 and X_2 are independent of R , or at the very least grow slowly with R , as is seen in spatially averaged data sampled from direct numerical simulations [3].

I will discuss several types of noise using $g_a(X_1) = |X_1|^a$. When multiplicative, the noise is interpreted with an Itô rule.

The case $a = 0$ is that of simple additive white noise. This case is less relevant in collapses when $X_{1,2}$ are close to 0 (which is then not absorbing). It is however a first, tractable (numerically and analytically) approximation if one is to study the trajectories going from the turbulent state to the saddle: this assumes that the amplitude of turbulence does not vary much on the turbulent side of the saddle. This causes little error, because in the framework of rare events, studying the trajectory from the starting point to the saddle between the starting point and arrival point is sufficient to compute the mean first passage times. The two other cases are $a = 1$ and $a = 1/2$. These two cases correspond to a multiplicative noise going to zero with the amplitude of turbulence. The model is now relevant to describe collapse at all values of $X_{1,2}$. The case $a = 1$ follows the original SPDE. The case $a = 1/2$ follows models arising from reaction diffusion process, like the one describing directed percolation [55]. It is moreover an illustration of a peculiar transition occurring in such model when a is decreased below $\frac{1}{2}$ (see Appendix E).

The noise inverse variance is controlled by the constant $\bar{\beta}$. My theoretical analysis will concern the limit $1/\bar{\beta} \rightarrow 0$. This can either be a crude representation of a turbulent puff, assuming that it feels a relatively low noise variance. More realistically, this can represent the limit $L \rightarrow \infty$, where X_1 and X_2 are space averaged variables, and the Reynolds number is not so small that all turbulence collapse occurs globally [Fig. 4(d)]. These noise terms arise from the spatial integral $\mathcal{Q} = \frac{1}{L} \int_{x=0}^L q\eta$. I can show analytically and confirm numerically in Appendix B that \mathcal{Q} is a delta correlated in time random variable. The average of \mathcal{Q} is 0 and its variance decrease like $1/L$ with a factor which grows like the intensity of turbulence in the flow. I thus choose $\bar{\beta} \propto L$ in my model. This further justifies using a multiplicative noise with $a > 0$. I can then also discuss how nontrivial rescaled pdfs of kinetic energy (see Fig. 9) arise from such models.

2. Educing the quasipotential

As explained in Sec. IV A, the parametric dependence of mean first passage times in stochastic system is very often governed by the quasipotential \mathcal{V} of the system. In order to educate such a quasipotential, the deterministic part of stochastic differential equations of the type of Eqs. (2) and (3), Eq. (12), or Eqs. (22) and (23) can be rewritten through a *transverse decomposition*

$$\dot{\vec{X}} = -aa^\dagger \vec{\nabla} \mathcal{V}(\vec{X}) + \vec{l}(\vec{X}) + \sqrt{\frac{2}{\mathfrak{B}}} a(\vec{x}) \eta. \quad (25)$$

I use \mathfrak{B} as the noise inverse variance and aa^\dagger as the noise correlation matrix. When it exists, this decomposition is unique. There is an orthogonality condition between \vec{l} and $\vec{\nabla} \mathcal{V}$, and \vec{l} cannot be written in a gradient form. One can see the origin of the aa^\dagger factor in Appendix E [Eq. (E1)]. Note that there exist some cases where the existence of such a decomposition does not ensure that \mathcal{V} govern passage times or the steady probability measure (when defined). This happens when there are direct connections between fixed points of the deterministic model [54], for instance. There is no such problem in this model, where fixed points of the deterministic model are isolated. This means that mean first passage time goes like $\exp(\beta \Delta \mathcal{V})$

and probability density functions go like $\exp(-\beta \mathcal{V})$ (when defined). However, this decomposition of the deterministic part of the equation is absolutely not trivial. As a consequence, I will consider numerical solutions of the problem and analytical approximated systems in order to discuss \mathcal{V} .

The quasipotential \mathcal{V} can be displayed by considering the steady pdf $\rho(X_1, X_2)$ of the system Eqs. (22) and (23) in the case $a = 0$. The pdf ρ is obtained by solving numerically the steady Fokker-Planck equation equivalent to the coupled SDEs (see [18] Sec. 5.2),

$$0 = -\frac{\partial}{\partial X_1} \left[\left(-\frac{1}{R} X_1 + X_2 + X_1 X_2 \right) \rho \right] - \frac{\partial}{\partial X_2} \left[\left(-\frac{1}{R} X_2 - X_1^2 \right) \rho \right] + \frac{1}{\bar{\beta}} \left(\frac{\partial^2}{\partial X_1^2} + \frac{\partial^2}{\partial X_2^2} \right) \rho. \quad (26)$$

The partial differential equation is discretized using finite differences. Note that this is not a problem constrained by boundary conditions, but instead by the normalization of ρ to $\int \rho dX_1 dX_2 = 1$. In this paper, this constraint is taken into account in the numerical problem by reformulating the partial differential equation into a variational problem and using a Lagrange multiplier. An alternative method can be to relax the time dependent Fokker-Planck equations [54]. In the low noise limit, the pdf ρ follows large deviations in $1/\bar{\beta}$ and one has

$$-\lim_{1/\bar{\beta} \rightarrow 0} \frac{1}{\bar{\beta}} \ln(\rho) = \mathcal{V}(X_1, X_2), \quad (27)$$

i.e., a Boltzmann factor. We ensure that \mathcal{V} exists by checking that the numerical solution does converge toward such a form in the high $\bar{\beta}$ limit. I give an example of probability density function at $R = 2.1$ (just above the bifurcation of the deterministic system) for $a = 0$ and $\bar{\beta} = 200$ in Fig. 11(a). There is a small maximum near $X_{1,2} = 0$, the ‘‘laminar’’ deterministic stable fixed point. A second more probable maximum is visible, $X_1 \simeq -1/2$ and $X_2 \simeq -1$, near the nontrivial ‘‘turbulent’’ deterministic fixed point. I verify that ρ is governed by a quasipotential in the large $\bar{\beta}$ limit: I display $\min_{X_1} [-(1/\bar{\beta}) \ln(\rho)]$ as a function of X_2 in Fig. 11(b) for increasing values of $\bar{\beta}$. The curves are shifted by a constant so that the minimum is zero, as is common in the study of large deviations. The collapse of the curves is visible. It is very similar to the behavior of the probability density functions sampled in numerical simulations of Couette flow (Fig. 9).

The quasipotential does not have a trivial analytic expression. In particular, it is certainly not a third order polynomial. Hopefully, the dynamics have a very particular structure which will simplify the study. Let us first rescale the equations in $\chi_1 = X_1$, $\chi_2 = X_2/R$, so that $\chi_{1,2} = \mathcal{O}(1)$ through most of the turbulence collapse. The rescaled coupled SDEs read

$$\frac{d\chi_1}{dt} = R \left(\chi_1 \chi_2 + \chi_2 - \frac{1}{R^2} \chi_1 \right) + \sqrt{R} g_a(\chi_1) \sqrt{\frac{2}{\bar{\beta}}} \eta_1, \quad (28)$$

$$\frac{d\chi_2}{dt} = -\frac{1}{R} (\chi_2 + \chi_1^2) + \frac{1}{\sqrt{R}} g_a(\chi_1) \sqrt{\frac{2}{\bar{\beta}}} \eta_2, \quad (29)$$

where $\tilde{\beta} = R\bar{\beta}$. The structure of a slow-fast system appears (see [18] Sec. 8.3). I retained $1/R^2$ for physical reasons

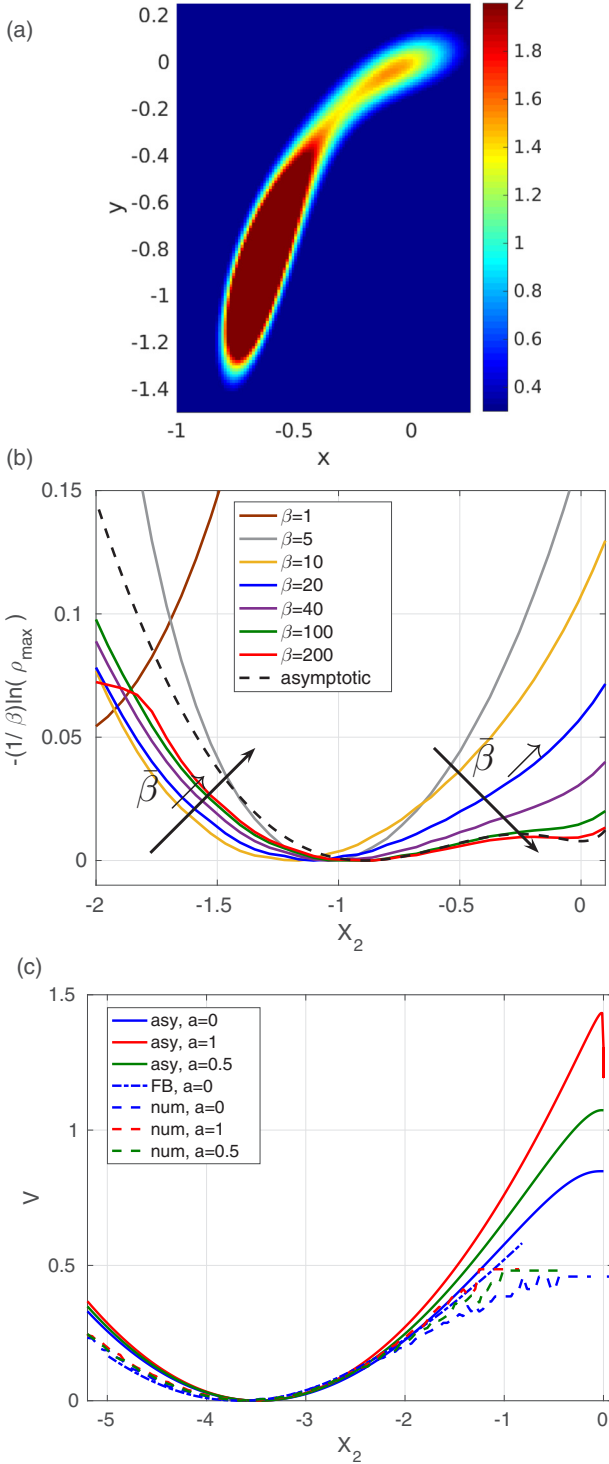


FIG. 11. (a) Color levels of the probability density function $\rho(X_1, X_2)$ obtained by solving the Fokker-Planck equation (26) with $a = 0$ (additive noise), $\bar{\beta} = 200$, $R = 2.1$. (b) $\min_{X_1} [-(1/\bar{\beta}) \ln(\rho)]$ for increasing values of $\bar{\beta}$ at $R = 2.1$ and $a = 0$ (additive noise), along with asymptotic quasipotential. (c) Asymptotic quasipotentials calculated analytically [Eqs. (32), (34), and (36)] for $R = 4$ and all three noise cases (labeled “asy,” $a = 0$, additive, and $a = 1$, $a = \frac{1}{2}$ multiplicative). On top of it is added the quasipotential calculated numerically by solving the Fokker-Planck equation (labeled “FP,” $a = 0$) and the quasipotentials evaluated numerically by integration of the SDEs for all three noise cases (labeled “num”).

discussed throughout the paper. The Reynolds number R controls the time scale separation. The variable χ_1 is slaved to χ_2 because it adjusts very fast to changes of χ_2 , in a time scale of order R . Meanwhile, the variable χ_2 evolves more slowly with a time scale of order $1/R$. This master slave structure of X_1, X_2 is another similarity with the SPDE model, where q leads $1 - u$. Note that the prefactors of the noise in R are also typical of slow-fast dynamics, at fixed $\bar{\beta}$. I introduce the variable ζ_1

$$\zeta_1 \equiv \chi_1 + \frac{\chi_2}{\chi_2 - \frac{1}{R^2}}. \quad (30)$$

This variable will adjust fast to zero and the system will keep $\chi_1 \simeq -\chi_2/(\chi_2 - 1/R^2)$. We can then perform the elimination of the fast variable ζ_1 . This procedure is rather similar to adiabatic elimination of the fast mode in deterministic system (see [56] Sec. 5.1 for the principle of adiabatic elimination and Sec. 5.2 for projection on central manifold), with some subtleties arising from the noise and correlations with the eliminated variable. Note also that this is not a weakly nonlinear analysis near the threshold of a bifurcation, as is often the case in fast variable elimination of deterministic dynamical systems. Such a procedure is commonplace in the study of rare events. However, unlike here, the time scale separation is often governed by $\bar{\beta}$, the large deviation parameter. This leads to a more complex situation [32]. As noted in Sec. II A, this difference with slow-fast systems arising from two dimensional turbulence is caused by the different role of the noise. In this case, it is not an energy injection; the nontrivial state exists even without noise. The noise is instead the trace of the intrinsic turbulent fluctuations. This will not cause problems, since the elimination is based on order identification in the time scale parameter $1/R$, so that we will be able to keep small $1/\bar{\beta}$ and be consistent.

The change of variable to ζ_1 is performed in Appendix C 1. A heuristic elimination in the Langevin equation is performed in Appendix C 2: this procedure is more readable, however it does not strictly ensure the validity of the result in my two parameters $R, \bar{\beta}$ case. Such validity is obtained by performing the elimination in the Fokker-Planck equation (see Appendix C 3), even if the operation is now less readable. When the result is rescaled and terms that will vanish when studying the large deviations are removed, we find the stochastic differential equation for X_2 alone at $a = 0$,

$$\begin{aligned} \dot{X}_2 &= - \left(\frac{1}{R} X_2 + R^2 \frac{X_2^2}{(R X_2 - 1)^2} \right) + \sqrt{\frac{1}{\bar{\beta}}} \eta \\ &= - \frac{d\mathcal{V}}{dX_2} + \sqrt{\frac{1}{\bar{\beta}}} \eta. \end{aligned} \quad (31)$$

An approximation was made to incorporate the saddle and laminar fixed point in the slow system. It makes sense to include the low amplitude $X_1 < 1$, $X_2 < R$ in this procedure, because X_2 still controls X_1 through the linear non-normal lift-up term in this range of amplitude and is thus still the leader in the dynamics. Since it involves only one degree of freedom, one has $l = 0$ when writing the SDE in the form of Eq. (25). The system is necessarily gradient and thus derives from a quasipotential \mathcal{V} . Knowing the noise correlation matrix

as a function of X_2 , we can thus deduce the quasipotential from the deterministic part of my model where X_1 is eliminated, for all three noise types ($a = 0, \frac{1}{2}$, and 1). I do this by rewriting the resulting SDEs for X_2 in the general form of Eq. (25). In the additive noise case, partial fraction decomposition and integration yields the quasipotential

$$\mathcal{V} = \frac{1}{2R} X_2^2 + X_2 + \frac{2}{R} \ln(1 - RX_2) + \frac{1}{R} \frac{1}{1 - RX_2} - \mathcal{V}_0. \quad (32)$$

The quasipotentials, as rate functions, are often shifted by a constant \mathcal{V}_0 chosen so that $\mathcal{V}(X_2^-) = 0$ at the main minimum. I use the constant \mathcal{V}_0 so that $\mathcal{V}(X_2^-) = 0$. I add the analytically calculated \mathcal{V} at $R = 2.1$ to the plot of numerically calculated \mathcal{V} in Fig. 11(b). There is already a good agreement, even if R is small. The minimum near 0 and the saddle is well captured by my approximation of keeping $1/R^2$ in the slow dynamics. Moreover, the approximation will give a more precise value for X_2 of the global minimum of quasipotentials in the whole range of Reynolds number.

The same procedure can be repeated when using a multiplicative noise (see remarks in Appendix C). Note that keeping the X_2 dependence of the multiplicative noise implies an approximation to the fast variable elimination similar to keeping the precise structure of minima and saddle in the quasipotential. If we use $a = 1$, the SDE in which X_1 is eliminated reads

$$\frac{dX_2}{dt} = - \underbrace{\left(\frac{1}{R} X_2 + \frac{X_2^2}{(X_2 - \frac{1}{R})^2} \right)}_{=-[X_2/(X_2-1/R)]^2 d\mathcal{V}/dX_2} + \sqrt{\frac{2}{\beta}} \frac{X_2}{X_2 - \frac{1}{R}} \eta_2. \quad (33)$$

Following the general form of Eq. (25), the quasipotential, as a function of X_2 , reads

$$\mathcal{V} = \frac{1}{2R} X_2^2 + \left(1 - \frac{2}{R^2} \right) X_2 + \frac{1}{R^3} \ln(|X_2|) - \mathcal{V}_0. \quad (34)$$

Note that because of the noise variance vanishing fast enough at $X_{1,2} = 0$, the quasipotential goes to minus infinity at $X_2 = 0$. As a consequence, the function $\exp(-\beta\mathcal{V})$ will diverge extremely fast like $|X_2|^{-\beta}$ at $X_2 = 0$ and β large. Therefore it cannot be normalized and the pdf is not defined as a function. Instead, it is a distribution, a dirac delta function in 0. This divergence of \mathcal{V} at zero arises from the constant part of b and a noise variance going to zero fast enough at zero.

If we choose $a = 1/2$, the SDE where X_1 is eliminated reads

$$\frac{dX_2}{dt} = - \underbrace{\left(\frac{1}{R} X_2 + \frac{X_2^2}{(X_2 - \frac{1}{R})^2} \right)}_{=-[X_2/(X_2-1/R)]d\mathcal{V}/dX_2} + \sqrt{\frac{2}{\beta}} \sqrt{\frac{X_2}{X_2 - \frac{1}{R}}} \eta_2. \quad (35)$$

Following Eq. (25), this yields again the quasipotential

$$\mathcal{V} = \frac{1}{2R} X_2^2 + \left(1 - \frac{1}{R^2} \right) X_2 + \frac{1}{R} \ln(1 - RX_2) - \mathcal{V}_0. \quad (36)$$

Note that now, the quasipotential is finite at 0. This is the case for all $0 \leq a \leq 1/2$. The probability density function will still diverge at 0 like X^{-1} . Note however, that since the system is

discretized numerically, this divergence is slowed down, and this can lead to the laminar state not being absorbing any longer. As discussed in Appendix E, this property should manifest itself only if a is *strictly* under $\frac{1}{2}$.

I compare all three quasipotentials calculated analytically to numerical solutions, either of the Fokker-Planck equation or the coupled SDEs, in Fig. 11(c). The agreement between numerical and analytical estimation of \mathcal{V} is good for $0 > X_2 > X_2^-$, at least within one standard deviation, as shown by the perfect agreement between numerical and analytical results in that range. In particular, the value of the minimum of \mathcal{V} , at $X_2 \gtrsim -R$, is well captured by the approximation performed in the elimination of the fast variable. Some properties are especially worth mentioning. For $X_2 \simeq X_2^-$, the quasipotentials are nearly parabolic. Note however that they remain parabolic away from X_2^- only if $|X_2| > |X_2|^-$. In the range $0 \leq |X_2| \leq X_2^-$, the quasipotentials are nearly linear. This is similar to DNS results (Fig. 9). This leads to exponential tails in the probability density functions. For more than one standard deviation, the numerical results estimated from simulations of the SDEs or solution of the Fokker-Planck equation are not trustworthy, since the corresponding probabilities are very low and thus poorly estimated by standard methods. Nevertheless, one finds both numerically and analytically that for $X_2 > X_2^-$, one has $\mathcal{V}_{a=0} \leq \mathcal{V}_{a=1/2} \leq \mathcal{V}_{a=1}$. In order to check the validity of analytical estimation of \mathcal{V} for more than one standard deviation out of X_2^- and illustrate the analytical calculation of mean first passage time before of turbulence in that model, I will now turn to a comparison between analytical result and AMS numerical results. I will then compare the numerically computed mean first passage times to those predicted by the Arrhenius law $\exp(-\beta\Delta\mathcal{V})$, which is the focus of my study. This will show that the analysis can not only help estimate pdfs and explain their shape but also estimate mean first passage times.

3. Mean first passage times and trajectories

The mean first passage time from state $X_{1,2}^-$ to state $X_{1,2} = 0$ is given analytically by the quasipotential difference between X_2^- and the saddle in all three cases ($a = 0, a = 1$, and $a = 1/2$). This reads $(1/\beta) \ln(T) = \mathcal{V}_{\text{saddle}} - \mathcal{V}(X_2^-)$. This difference can be calculated exactly using the full formula Eqs. (19), (32), (34), and (36) in Appendix D. In the limit of large R , through a Taylor expansion, we find $\Delta\mathcal{V}_{a=0} = \frac{R}{2} - \frac{4\ln(R)}{R} + \mathcal{O}(\frac{1}{R})$, $\Delta\mathcal{V}_{a=1/2} = \frac{R}{2} - \frac{2\ln(R)}{R} + \mathcal{O}(\frac{1}{R})$, and $\Delta\mathcal{V}_{a=1} = \frac{R}{2} - \frac{1}{R} + \mathcal{O}(\frac{1}{R^2})$. This means that in all three cases, at lowest nonzero order, the mean first passage time before the exit of state $X_{1,2}^-$ has the same asymptotic dependence in R and β . In large deviations form and at large R , it reads

$$\lim_{1/\beta \rightarrow 0} \frac{\ln(T)}{\beta} = \Delta\mathcal{V} = \frac{R}{2}. \quad (37)$$

Note that this final asymptotic result is not impacted by the approximation made in the fast variable elimination, since the corrective terms vanish. In all three cases, the actual potential difference is lower than the asymptotic result, since corrections are all negative. The $a = 1$ is quite close to the asymptotic result. Meanwhile, in the $a = 0$ and $a = \frac{1}{2}$ case, the correction, while going to zero at $R \rightarrow \infty$ is decreasing very slowly and is in fact close to a constant of order 1 for $R \lesssim 20$.

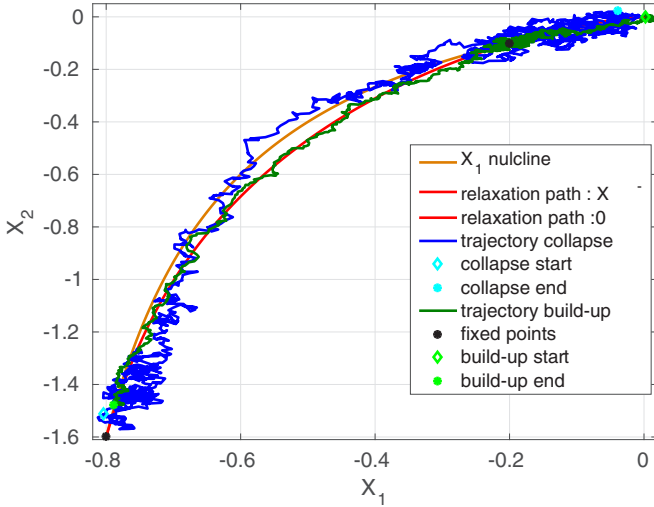


FIG. 12. Reactive trajectories in the two degrees of freedom model at $R = 2.5$, showing build-up of turbulence under additive noise ($\bar{\beta} = 10\,000$), collapse of turbulence under additive noise ($\bar{\beta} = 1500$), the starting and ending points of trajectories, the relaxation paths from the saddle point $X_{1,2}^+$ to the nodes $X_{1,2}^-$ and 0 and the X_1 nullcline $X_1 = [X_2/(X_2 - 1/R)]$.

One can compare this analytical result to AMS calculation of the mean first passage time before reaching $(0, 0)$ in all three cases ($a = 0$, $a = 1$, and $a = 1/2$), and escapes from state $(0, 0)$ in the $a = 0$ case (comparable to build-up of turbulence). I choose the reaction coordinates

$$\begin{aligned} \varphi_{a=0} &= 1 - \frac{X_2}{X_2^-}, \\ \varphi_{a>0} &= (1 + \epsilon_{a>0}) \left(1 - \frac{X_2}{X_2^-} \right), \\ \varphi_{\text{seed}} &= (1 + \epsilon_s) \frac{X_1}{X_1^+} - \epsilon_s. \end{aligned} \quad (38)$$

This distinguishes the cases of additive ($\varphi_{a=0}$) and multiplicative noise ($\varphi_{a>0}$). A small constant $\epsilon_{a>0}$ is added to 1 in $\varphi_{a>0}$, so that AMS stops very slightly before X_2 has entirely collapsed and the noise has entirely vanished, thus ensuring that the final branchings will not be perturbed by a possible absence of noise. In the case of build-up of turbulence a small constant is subtracted to φ_{seed} so as to define \mathcal{A} as a tiny ball around the laminar state. I follow the same procedure as in the study of mean first passage times before collapse of turbulence in the SPDE. I calculate reactive trajectories. I present the collapse of turbulence at $\bar{\beta} = 1500$ and the build-up of turbulence at $\bar{\beta} = 10\,000$, both at $R = 2.5$, in Fig. 12. The collapse trajectory remains very close to the $X_1 = X_2/(X_2 - 1/R)$ line: variable X_2 remains the master on the whole trajectory. The collapse trajectory is stimulated by noise from $X_{1,2}^-$ down to $X_{1,2}^+$ (this is the fluctuation path) and then collapses to 0 deterministically (this is the relaxation path). Meanwhile the build-up trajectory is stimulated by noise from 0 to $X_{1,2}^+$ (this is the fluctuation path) and then reaches the turbulent state $X_{1,2}^-$ deterministically (this is the relaxation path). Even if the trajectories are very close, note that the fluctuation path from $X_{1,2}^-$ to $X_{1,2}^+$ is not the time reverse relaxation path seen in the build-up trajectory,

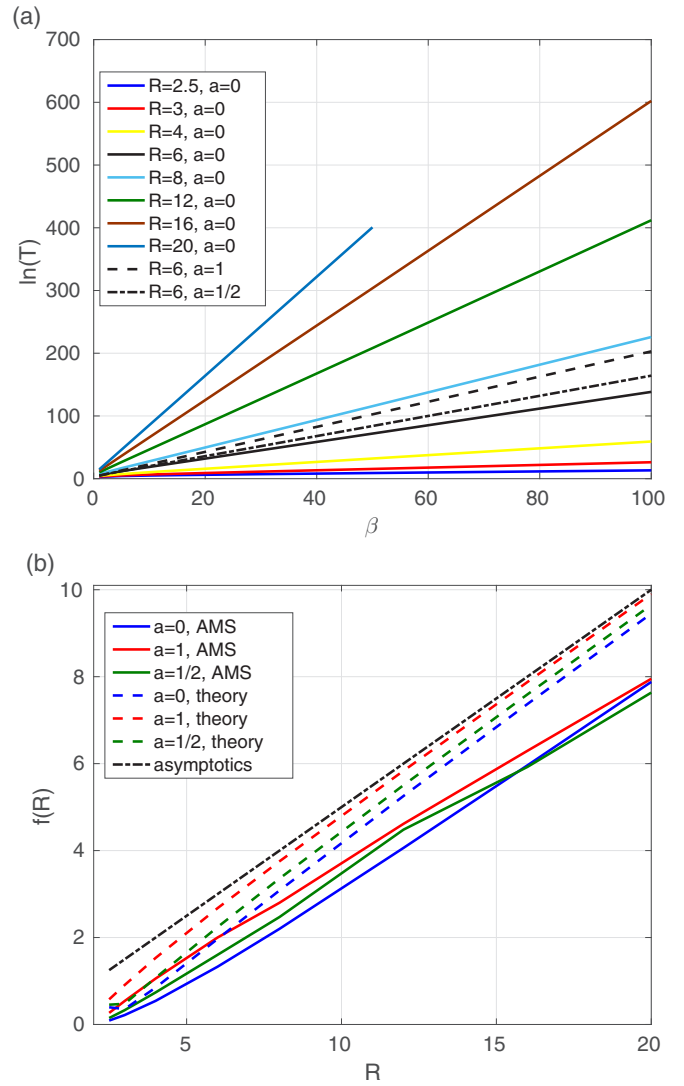


FIG. 13. (a) Logarithm of the mean first passage time before exit of state $X_{1,2}^-$ as a function of $\bar{\beta}$ in the two degrees of freedom model for increasing Reynolds number R in the three noise cases. (b) Growth rate $\bar{f}(R)$ of the mean first passage time before exit of state $X_{1,2}^-$ as a function of Reynolds number R in the two degrees of freedom model for additive noise.

since the system is not gradient. Thus, trajectories of both types are governed by instantons in the limit of low variance noise. Note that both trajectories spend some time fluctuating around the saddle point $X_{1,2}^+$. This decomposition of collapse and build-up trajectories stresses the importance of the saddle points on the separatrix between states or on a minimal germ [see Fig. 4(e)] in the case of build-up of turbulence.

I then calculate T as a function of $\bar{\beta}$ for $R \in [2; 20]$. In the case of additive noise, I display $\ln(T)$ as a function of $\bar{\beta}$ for $2.5 \leq R \leq 20$, as well as $\ln(T)(\bar{\beta})$ at $R = 6$ for the $a = 1$ and $a = 1/2$ multiplicative noise cases in Fig. 13(a). One can see that for all R , $\ln(T)$ grows linearly with $\bar{\beta}$, so that $\bar{f}(R) = \ln(T)/\bar{\beta}$ is independent of $\bar{\beta}$, as calculated analytically [Eq. (37)]. Moreover, the slope of $\ln(T)(R)$ ostensibly grows with the Reynolds number. I perform an affine fit for all Reynolds numbers. This yields $\bar{f}(R)$, for all three noise cases.

I then display $\bar{f}(R)$ for all three noise cases as a function of the Reynolds number, along with the asymptotic prediction [Eq. (37)] and full quasipotential differences [Eqs. (D1), (D3), and (D2)] in Fig. 13(b). The growth rate $\bar{f}_{a=0,1,1/2}(R)$ calculated numerically with AMS is affine in R and is in very good agreement with the asymptotic result: it has a slope within 10% of $\frac{1}{2}$ and within 10% of corresponding quasipotential differences for all three noise cases. The ordinate at the origin is negative in all three cases, for both numerical results and full quasipotential difference. The source of the small discrepancy can be the approximation of the saddle done in the fast variable elimination. The definition of \mathcal{B} can also come into play because the saddle point is very near it. I find analytically and numerically that $\bar{f}_{a=0}(R) < \bar{f}_{a=1/2}(R) < \bar{f}_{a=1}(R)$, as was the case for the quasipotentials [Fig. 11(c)]. All things considered, if one remembers that the noise in the model goes to zero with the length $\bar{\beta} \propto L$, one finds analytically and numerically that this two degrees of freedom model leads to a mean first passage time dependence of the type $T \propto \exp[L(A'R - B')]$. The use of the quasipotential difference also shows the importance of the saddle point in collapse or build-up.

V. CONCLUSION

A. Summary

This paper presented a joint numerical and theoretical study of multistability in two stochastic models of transitional wall flows. The numerical study was focused on turbulence collapse, splitting, and build-up in a SPDE pipe flow model proposed by Barkley [16]. Adaptive multilevel splitting, a mutation selection algorithm dedicated to the fast study of multistability, was used to compute the trajectories of puff splitting, turbulence collapse, and build-up, and corresponding mean first passage times in a reproducible manner. Using AMS exponentially reduced the cost of these computations. Similarly to other rare events or atypical trajectories calculation, AMS modifies the dynamics to sample said rare events. It is indeed a necessary condition for a valid rare events study that said events are sampled and not guessed. An advantage of a method like AMS is that it comes along with a strong mathematical and theoretical background that quantitatively demonstrates, in a increasing number of systems, how the occurrence of observation of rare events is changed. This leads to a precise estimation of mean first passage time. Along with this comes rules for assessment of the precision of estimations and numerical verifications.

The quantitative study was first focused on isolated puffs. The mean first passage time T before puff collapse was computed in the whole range of Reynolds number of transitional turbulence for a wide range of initial conditions (thus including the collapse of equilibrium puffs in the lower range of Reynolds number). Using a fixed initial condition for the computation of mean first passage times before collapse and basing the reaction coordinate on turbulent intensity generalizes the conditioning on the maximal turbulent fraction, which should not be much larger than that of the initial condition. Computing these times in such a conditioned manner while varying R shows that the logarithm of T is affine in the Reynolds number $\ln(T) = A_p r - B_p$. Moreover, the coefficients A_p, B_p depend

in an affine manner on the total turbulence intensity of the initial puff $Q_0 = \int q dx$. Performing such a conditioned study helps uncover the reaction of the phase to given states of isolated puffs in a manner which is well described by theory. Moreover this provides a first separation of the mean first passage time before collapse of turbulence between the reaction of the phase space to a given initial condition and the way the natural initial conditions depend on the Reynolds number. Indeed, knowing $Q_0(r)$ yields the faster than linear growth of $\ln(T)(r)$ for equilibrium puffs. Approximates of the mean first passage times before splitting of equilibrium puffs were computed in order to illustrate the crossover of mean first passage times of collapse and splitting very near the threshold of sustained laminar-turbulent flow.

I could moreover compute turbulence collapse in pipes initially containing their steady turbulent fraction. This showed two regimes of collapse of laminar-turbulent flows: just above the threshold of sustained laminar-turbulent coexistence, the collapse occurs puff by puff through independent events. Meanwhile all puffs collapse in a correlated manner in the rest of the range of laminar-turbulent coexistence, due to increased probability of puff splitting. The length and Reynolds number scaling of the mean first passage time before collapse of turbulence in these laminar-turbulent pipes is independent on the type of collapse and one finds that $T \asymp \exp[L(Ar - B)]$ with positive A and B in the whole transitional range, thus providing a counterpart for results obtained in DNS of short pipes [19].

Large deviation approaches of the study of probabilities and passage times were used to discuss the length and Reynolds number dependence of the mean first passage time before collapse of turbulence. These theories often yield the dependence of said probability and mean first passage times through a quasipotential $(1/\mathfrak{B}) \ln(T) = \Delta\mathcal{V}$ in the limit of zero variance noise $1/\mathfrak{B} \rightarrow 0$ [25]. I discussed the persistence of the same length dependence of mean first passage time before collapse of laminar-turbulent coexistence $\ln(T) \propto L$ in all regimes of collapse. Furthermore, I proposed the computation of the quasipotential in a simple two degrees of freedom model proposed by Dauchot and Manneville [28], which has all the properties one would expect of spatially averaged wall turbulence. This showed that mean first passage times before collapse of turbulence of the type $\exp[L(A'R - B')]$ could be derived in asymptotic limits. This derivation stresses the importance of the saddle between the laminar and the turbulent state for the mean first passage time dependence and the natural trajectories followed by turbulence build-up and collapse. Additional numerical simulations of the simple model showed that these derivations were precise.

B. Discussion

Most of the properties of the two degrees of freedom model are found in the spatially averaged SPDE model and more generally in numerical and laboratory experiment. This means that the derivation of the mean first passage time dependence on L and R in this simple SDE model can enlighten us on the physical mechanisms controlling the mean first passage time before collapse of turbulence initially at its steady turbulent fraction. In all cases, we find master-slave dynamics, where the Q variable leads and the $1 - U$ variable follows. In the average

value of the steady state, the leader Q quantity grows somewhat linearly with the Reynolds number while the follower saturates. The leading variable indicates how much kinetic energy is given to turbulence and how much turbulence has to be collapsed. This amount of energy grows more or less linearly with Reynolds number and this is reflected in the position of the quasipotential minimum. The fact that at fixed length the noise felt by the master variable saturates is reflected in the multiplicative noise depending solely on X_1 . These two combined effects lead to the affine quasipotential difference. The fact that turbulence has to be collapsed as a whole, like “flip” reactive trajectories of gradient systems [24], gives relevance to the study of such spatially averaged variables. This means that the noise variance goes to 0 like 1 over the length and leads to the length dependence of mean first passage time T . The case of the isolated puff is somewhat similar; the increase of length corresponds to the increase of Q_0 .

This two degrees of freedom model can also show us that the physical mechanisms leading to the asymmetric probability density functions of kinetic energy of turbulence (see [3,57]) are the same as those controlling the parametric dependence of mean first passage time before collapse of turbulence. Indeed the mean first passage time before collapse of turbulence is governed by $\Delta\mathcal{V}$ while the pdf is governed by \mathcal{V} . The combined effect of very weak Reynolds number dependence of noise variance and linear growth of kinetic energy of turbulence with Reynolds number leads to the linear quasipotentials in the range $0 \leq e \leq (e)$ and therefore the exponential tails.

C. Perspectives

The success of the combined use of the numerical method for the computation of trajectories and mean first passage times and of the theoretical framework for their interpretation opens a wide range of perspective for the study of transitional turbulence in wall flows. Methods like AMS can be successfully applied to chaotic and turbulent dynamics [58]. This means that the mean first passage time before collapse of turbulence can now be extensively computed both in academic or realistic flow configurations. Mean first passage time scalings with Reynolds number and length and flow configurations can thus be computed extensively, beyond what is deduced from the models studied here or DNS of small systems [19]. One can then check whether the physical mechanisms at work in the models studied in this paper are also relevant in actual pipe flow. Even if the internal dynamics are different (stochastic versus temporal chaos), similar reasonings may be followed, since they are based on the same notions in both cases (comparison of coherent lengths to the domain size, independence of events, etc. [20]). When studying the role of saddles in passage time Reynolds number dependence and turbulence collapse, edge states may very likely play the role of the simple saddle point used in the two degrees of freedom model [43,44].

From another point of view, this also opens a way to estimate optimal evolutions, starting from the laminar base flow, leading to turbulence that goes even beyond computing the minimal seed [21,22]. Indeed, by computing the most probable trajectory traveling from the laminar base flow to turbulence under a noise of vanishing variance, either with an AMS-type method or through action minimization [38],

the trajectory leading to turbulence under a wide range of arbitrarily small external perturbations can be found. This approach does not require us to “teleport” the flow in a state of finite energy since the initial condition is the laminar base flow. The perturbing noise of vanishing variance only plays a role in the fluctuation trajectory from the laminar state to the most probable saddle on such paths: from that point on, theory ensures that the trajectory is a deterministic relaxation path and that it starts from a saddle point [25]. Moreover, this approach requires a single computation: there is no need for a dichotomy on a finite energy of the initial conditions. Such build-up calculations are under way, and should inform us of mechanisms leading to development, collapse, or extension of turbulence in transitional wall flows.

ACKNOWLEDGMENTS

This work was granted access to the HPC and visualization resources of “Centre de Calcul Interactif” hosted by “Université Nice Sophia Antipolis.” The author thanks Eric Simonnet for helpful discussions during the implementation of AMS. Detailed comments by anonymous referees were also helpful for the improvement of the paper and the scientific content. Discussions on the literature with Y. Duguet, suggestions by P. Manneville, and the hospitality of Institute for atmospheric and environmental science, University of Frankfurt, Germany, where the author worked when some of the research was performed, are also acknowledged.

APPENDIX A: EFFECT OF AMS PARAMETERS ON THE PRECISION OF ESTIMATES

The notion of convergence of quantities z estimated by AMS is peculiar, because it involves two parameters: the number of clones N of the computation, and the number o of repetitions of independent AMS computations at a fixed number of clones. Another specificity is that it is probabilistic: α is within a given interval of confidence of the estimate with a given probability. Unless the simulation is extremely expensive, one performs an estimation of $z = \alpha, T, \tau \dots$, etc., by averaging the output of independent AMS computations at fixed clone number N over several to many independent realizations [23,24,45,46],

$$\langle z \rangle_o = \frac{1}{o} \sum_{i=1}^o z_i. \quad (\text{A1})$$

In the paper, I drop the $\langle \cdot \rangle_o$ for readability. The quality of this estimate, depending on N and o , can be quantified using two cumulants: the variance over realizations $\langle (z - \langle z \rangle_o)^2 \rangle_o$ and the difference between the average over realizations and a “true” value estimated either by direct numerical simulation (in order to separate the effect of N from the effect of dt) or by theory: $\langle z \rangle_o - z$. The DNS used for the calculation of α follow the stage 0 of the algorithm and estimates α by the ratio of reactive trajectories by total number of test trajectories, when this number goes to infinity. Mathematical results exist on the variance

$$\sigma_\alpha \equiv \lim_{o \rightarrow \infty} \langle (\alpha - \langle \alpha \rangle_o)^2 \rangle_o \quad (\text{A2})$$

as a function of the number of clones N , using somewhat idealized dynamics [23,45]. Indeed, the estimate of α by AMS has a variance σ_α , no matter the model or the formulation of the algorithm. This variance can be used to compute $\pm 2\sigma_\alpha/\sqrt{o}$ which gives the 95% confidence interval for the estimation of α . It can be demonstrated that

$$\sigma_\alpha \geq \frac{\alpha\sqrt{|\ln(\alpha)|}}{\sqrt{N}}. \quad (\text{A3})$$

There is equality in the ideal case [23]. The relative variance $\sigma_\alpha/(\alpha\sqrt{|\ln(\alpha)|})$ is found to decrease like $1/\sqrt{N}$ even in nonideal cases [24,45,46] (provided ϕ is reasonable). As to $\lim_{o \rightarrow \infty} \langle \alpha \rangle_o$, several situations exist, depending on the formulation of the AMS algorithm, the estimator of α , and the number of degrees of freedom of the model to which AMS is applied. While one can ensure complete unbiasedness in one degrees of freedom models using an estimator slightly improved with respect to Eq. (5) [45,46], a bias may exist between the AMS estimation (averaged over realizations) and the estimate using DNS [23,46]. One empirically finds that the quantity $\Delta\alpha/\alpha_{\text{DNS}} = (\langle \alpha \rangle_o - \alpha_{\text{DNS}})/\alpha_{\text{DNS}}$ is negative or zero: positive values of $\Delta\alpha/\alpha_{\text{DNS}}$ fall quite within a relevant confidence interval around zero and correspond to small sampling errors. In the general case of no systematic bias $\Delta\alpha/\alpha_{\text{DNS}}$ goes to 0 like $1/N$. This means that if the bias is absent or negligible (the number of clones is not so small), obtaining a precise estimate of the probability of observing an event is a tradeoff between the number of clones used in AMS computations and the number of independent realizations over which one averages. If one uses fewer clones in AMS computations (for instance, because of a constraint of memory), one then needs more realizations of AMS computations to average over in order to obtain a precise estimate. If one uses more clones, then fewer realizations are necessary.

No demonstration of the scaling of σ_α exists yet that *mathematically* applies to the use of AMS on a SPDE model like Eqs. (2) and (3). This does not mean that this scaling is false. Indeed, transposing the argument of these demonstrations to SPDEs yields a justification which is certainly not less strong than many physical derivations. This state of fact means that one has to verify scalings and assertions empirically. In order to do so, I choose cases where α can be estimated using direct numerical simulations. I considered collapse events of isolated puffs at $r = 0.85$, $r = 0.95$, $r = 1.0$, and $r = 1.05$ and splitting events at $r = 1.0$, all of them in a domain of length $L = 800$. These cases display relatively small α [down to $\mathcal{O}(10^{-5})$] and span an order of magnitude of α . In all these cases, I estimated α_{DNS} , τ_{DNS} using a large enough number of simulations. The quality of this estimate will be given by 95% confidence intervals. The precision on the estimate of T follows from the precision on the estimate of these quantities. In parallel, I performed AMS computations of these events using a range of clone numbers $10 \leq N \leq 5000$. For each of them, I used a large enough number of realizations (going from thousands when using $N = 10$ clones to tens when using $N = 5000$ clones), so that the statistics of α were estimated with precision.

I first illustrate the convergence of the average of $\langle \alpha \rangle_o$ as o is increased in Fig. 14(a). I used the case of a collapse event at $r = 0.85$. This shows that large errors can be made if one

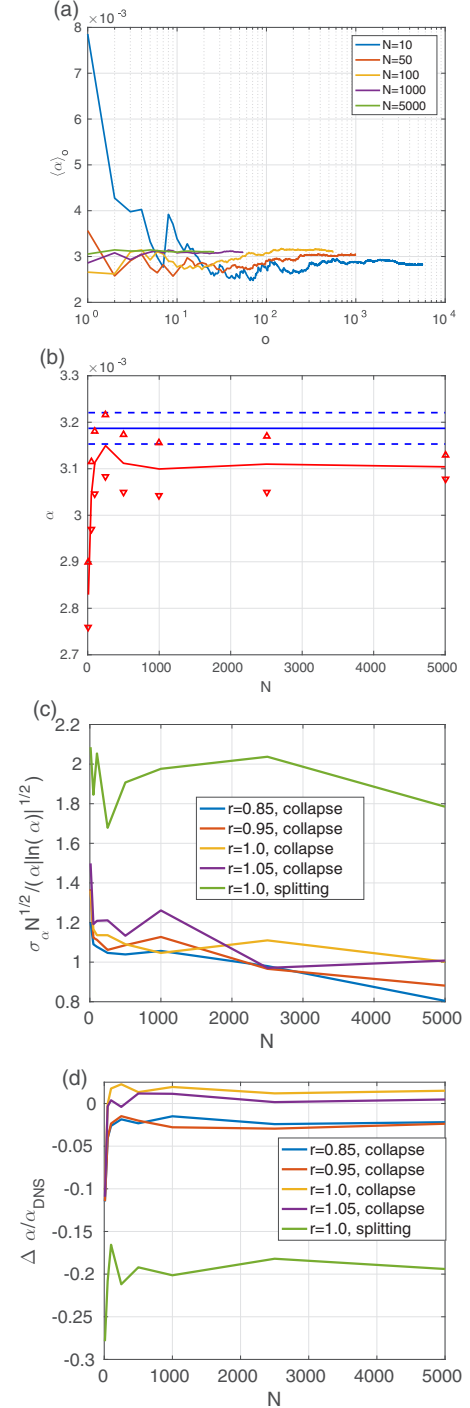


FIG. 14. (a) Probability $\langle \alpha \rangle_o$ of observing a collapse estimated after a sum over the number of independent repetitions of AMS as a function of o at $r = 0.85$, for several N . (b) Estimate of the probability of observing a collapse α (red curve) as a function of the number of clones used in AMS computations, with large o . The blue curve indicates the probability as estimated by DNS α_{DNS} . The error bars indicate the 95% confidence interval computed using the variance over repetitions. The dashed blue lines indicate the 95% confidence interval for estimation of α_{DNS} by DNS. (c) Relative variance of the estimate of the probability of observing a collapse or a splitting as a function of N for several r . (d) Relative difference between the estimate of α by AMS and DNS for collapse and splitting events as a function of N for several r .

uses few clones and few realizations, for instance with $N = 10$. However, one can see that the average over realization always reaches an asymptotic value as the number of realizations is increased. The rate at which this asymptotic value is reached increases as the number of clones is increased. The variance σ_α quantifies this rate. Note that a small but nonzero difference is visible between AMS estimates of α at $N = 10$ clones and AMS estimates with higher numbers of clones. At small N , this difference is larger when using the exponential approximate estimator of α [Eq. (5)]. One can then illustrate the effect of the number of clones on $\langle \underline{\alpha} \rangle_o$ as a function of N , displayed on top of the estimate of α using DNS [Fig. 14(b)]. I used a o large enough for the asymptotic value to be reached. One can see that this estimate reaches an asymptotic value as N is increased. One also see that even if a small difference appears between AMS and DNS computation, one still has an overlap of the 95% confidence interval as soon as $N \geq 50$, ensuring the probabilistic quality of the estimate. Thus, if the variance of α follows the ideal scaling of Eq. (A3), one can know how precise an estimate of α is, given the number of clones in AMS computations and the number of independent realizations. In order to verify this scaling and quantify how many o are generally need to estimate $\langle \underline{\alpha} \rangle_o$, I display the rescaled variance $((\alpha - \langle \underline{\alpha} \rangle_o)^2 \sqrt{N}) / (\alpha \sqrt{|\ln(\alpha)|})$ as a function of the number of clones N [Fig. 14(c)]. One can see that for the collapse events, the variance of α is actually very close to the lower boundary of the idealized case, since the rescaled variance is nearly 1 for all r and N . Meanwhile, one finds that the rescaled variance is larger by a factor of 2 for the splitting events, even if it follows the predicted scalings in α and N . This indicates that splitting events are harder to simulate and that a larger number of clones or repetitions may be needed in order to obtain a given precision. There may be room for improvement of the reaction coordinate. Figure 14(c) overall indicates that even if the estimate of the variance is not yet demonstrated to follow the scaling of Eq. (A3) in the case of SPDE [Eqs. (2) and (3)], such a proposition may very well be true. It is therefore reasonable to undertake such a demonstration. I eventually consider the relative difference between the estimate of α using AMS and the estimate of α using DNS: $\Delta\alpha/\alpha_{\text{DNS}} \equiv (\langle \underline{\alpha} \rangle_o - \alpha_{\text{DNS}})/\alpha_{\text{DNS}}$ [Fig. 14(d)]. Note that this quantity is the typical bias in the uncertainty when considering the logarithm of α [or similar quantities like $\ln(T)$]. One can see that in all cases, the quantity $\Delta\alpha/\alpha_{\text{DNS}}$ reaches an asymptotic value as N is increased. In the case of collapse events, this value is very close to 0 [typically 10^{-2} with my estimates; this uncertainty should be compared to values of order $\mathcal{O}(10)$]. There are small differences in the asymptotic large N value of $\Delta\alpha/\alpha_{\text{DNS}}$ as r is changed. My samples show that for $r = 0.85$ and $r = 0.095$, the same asymptotic value $\lim_{N \rightarrow \infty} \Delta\alpha/\alpha_{\text{DNS}} \simeq -0.025$ is reached. As r is increased, the samples show that one has $\lim_{N \rightarrow \infty} \Delta\alpha/\alpha_{\text{DNS}} \simeq 0.01$ at $r = 1.0$ and an asymptotic value of this relative difference which cannot be distinguished from 0 at $r = 1.05$ (when a small confidence interval on α_{DNS} is taken into account). This would indicate that the quality of the estimate of α by AMS is improving as r is increased. The *a priori* is that the quality of the estimate decreases as α decreases. It is possible that some changes in the structure of the reactive trajectories as r is increased make them easier to compute by AMS. In the case of splitting events, while the

relative difference converges toward a relatively small value (10^{-1} , to be compared to logarithms of α of order 10), this asymptotic value is not zero, hinting again toward the higher difficulty of estimating the splitting probability.

Other AMS estimates can be compared to DNS estimates. Mathematical results exist only for α , so that there are only empirical guidelines for other physical quantities [46]. I tested the case of the average duration of trajectories, because it is physically well defined and because it is easily computed by both AMS and DNS. The average duration of trajectories τ computed by AMS is close to τ_{DNS} . Indeed, when considering the five cases of Fig. 14 (collapse at $r = 0.85$, $r = 0.95$, $r = 1.0$, and $r = 1.05$, and splitting at $r = 1.0$), one find that $|\langle \underline{\tau} \rangle_o - \tau_{\text{DNS}}|/\tau_{\text{DNS}}$ is smaller than 10^{-2} for collapse events at all four Reynolds numbers, when using $10 \leq N \leq 5000$ clones in computations. Unlike $|\langle \underline{\alpha} \rangle_o - \alpha_{\text{DNS}}|/\alpha_{\text{DNS}}$, no tendency with N appears: this difference is always small, even when using $N = 10$ clones. Behind this lies the precision in the estimate of reactive trajectories themselves by AMS. If one wishes to sample them in order to examine the physics of transitions, using few clones N and many repetition o is *always* a good strategy. The relative difference is slightly larger for splitting events at $r = 1.0$, $|\langle \underline{\tau} \rangle_o - \tau_{\text{DNS}}|/\tau_{\text{DNS}} = \mathcal{O}(10^{-1})$. It is smaller than the relative difference of splitting probabilities [Fig. 14(d)]. Note also that the rather large durations of collapses puff by puff [Figs. 4(c) and 7(d), $1.0 \leq r \leq 1.15$] is not likely a numerical artifact but quite probably a physical effect. Indeed, no effect of the number of clones was seen in these estimates of τ . Moreover, when using an imperfect reaction coordinate, AMS calculation may in fact overselect short duration trajectories, which are actually improbable, over longer duration trajectories, which are more probable and which imply spending a long time around local probability maxima [45,46]. When this occurs, AMS outputs leave specific traces in the estimates, such as fat tails in the distributions of k , a large, nonvanishing as $N \rightarrow \infty$, skewness in the distribution of $\underline{\alpha}$. This makes the problem easy to detect. Collapses puff by puff are such long duration trajectories displaying bottlenecks near local probability maxima. For this reason, the long durations displayed for $1.05 \leq r \leq 1.15$ in Fig. 7(d) are very doubtfully numerical artifacts.

Note that when using AMS to study transitions in the two degrees of freedom model Eqs. (22) and (23), effects comparable to those see in Fig. 14(d) while using the estimator [Eq. (5)] remain much smaller, while $\ln(T)$ spans more than two orders of magnitude.

APPENDIX B: LENGTH SCALING OF SPATIALLY AVERAGED MULTIPLICATIVE NOISE

The spatial average of the noise term of Eq. (3) reads

$$\Omega \equiv \sqrt{\frac{2}{\beta}} \frac{1}{L} \underbrace{\int_{x=0}^L q(x,t) \eta(x,t) dx}_{\equiv I}. \quad (\text{B1})$$

This comes from the SPDE interpreted with an Itô rule, so that the average of $q(x,t)$ is uncorrelated from $\eta(x,t)$: “ η happens

dt later than q in the time discretization.” This means that

$$\langle \Omega \rangle = \sqrt{\frac{2}{\beta}} \frac{1}{L} \int_{x=0}^L \underbrace{\langle q(x,t)\eta(x,t) \rangle}_{=0} dx. \quad (\text{B2})$$

Let us now consider the lagged variance of Ω . I will work with discretized time, since this is how Itô processes are mathematically defined. In that case, one has that

$$\langle \eta(x,t)\eta(x',t') \rangle = \delta(x-x')\delta_{t,t'} \frac{1}{\sqrt{dt}}. \quad (\text{B3})$$

Note the distinction between the Dirac delta function $\delta(x-x')$ and the Kronecker symbol $\delta_{t,t'}$ arising from the distinction between continuous space and time discretized with a time step dt . The time covariance of Ω is

$$\begin{aligned} & \langle [\Omega(t) - \langle \Omega \rangle][\Omega(t') - \langle \Omega \rangle] \rangle \\ &= \langle \Omega(t)\Omega(t') \rangle = \frac{2}{\beta L^2} \langle I(t)I(t') \rangle. \end{aligned} \quad (\text{B4})$$

We could simplify using Eq. (B2). This is a first consequence of having an Itô process. Let us now consider the lagged covariance

$$\langle I(t)I(t') \rangle = \int_{x=0}^L \int_{x'=0}^L \langle q(x,t)q(x',t')\eta(x,t)\eta(x',t') \rangle dx dx'. \quad (\text{B5})$$

Without loss of generality, let us set that $t' \geq t$. In Eq. (B5), we have two cases for the two points two times covariance of q and η

$$\begin{aligned} & \langle q(x,t)q(x',t')\eta(x,t)\eta(x',t') \rangle \\ &= \langle q(x,t)\eta(x,t)q(x',t') \rangle \langle \eta(x',t') \rangle, \quad \text{if } t' > t, \end{aligned} \quad (\text{B6})$$

$$\begin{aligned} & \langle q(x,t)q(x',t')\eta(x,t)\eta(x',t') \rangle \\ &= \langle q(x,t)q(x',t) \rangle \langle \eta(x,t)\eta(x',t) \rangle, \quad \text{if } t' = t. \end{aligned} \quad (\text{B7})$$

This again comes from the fact that we are considering an Itô process (see [18] Sec. 4.2). The spatial averages and correlations are not made explicit at this point. Since $\eta(x',t')$ can be extracted in the case $t' > t$ and we have $\langle \eta(x',t') \rangle = 0$, this leads to $\langle I(t)I(t' > t) \rangle = 0$. The lagged correlation of Eq. (B5) can be rewritten with a Kronecker delta symbol $\delta_{t,t'}$. I also use that $\langle \eta(x',t)\eta(x,t) \rangle = \frac{1}{dt} \delta(x-x')$ in discretized equations. We can thus rewrite the lagged correlation

$$\begin{aligned} \langle I(t)I(t') \rangle &= \frac{1}{dt} \int_{x=0}^L \int_{x'=0}^L \langle q(x,t)q(x',t) \rangle \delta(x-x') \delta_{t,t'} dx dx' \\ &= \frac{1}{dt} \delta_{t,t'} \int_{x=0}^L \langle q(x,t)q(x,t) \rangle dx. \end{aligned} \quad (\text{B8})$$

I performed a first integration on x' which uses the Dirac δ function. Due to the translational invariance in the periodic pipe, the second order moment $\langle q(x,t)^2 \rangle = C_q(t)$ is independent on space. Note however, that this function is not independent on time, since it contains the information on the collapse of turbulence after a mean first passage time T . This corresponds to having $q = 0$ on ensemble average after T . One typically

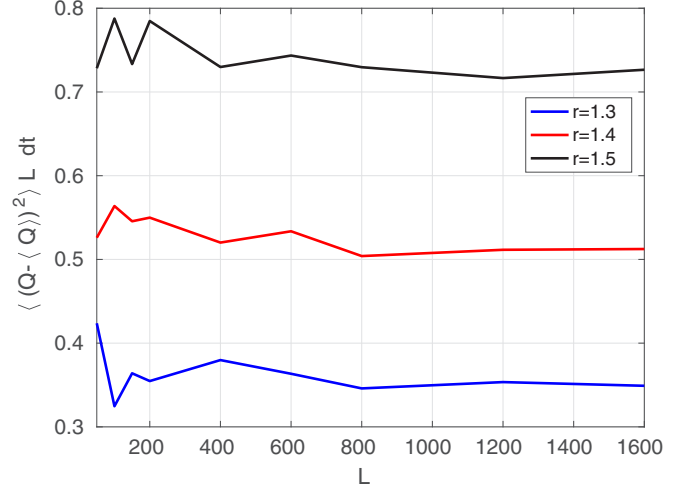


FIG. 15. Normalized variance $\langle (\Omega - \langle \Omega \rangle)^2 \rangle L dt$ sampled in direct numerical simulations of the model (while no collapse occurred) as a function of the domain size, for several Reynolds numbers.

has $C_q(t) = \mathcal{O}(1)$ if $t \lesssim T$ and $0 \leq C_q(t) \ll 1$ if $t \gtrsim T$. One important fact is that $C_q(t)$ is expected to become independent on size as size is increased, since the spatial correlation extends within one puff or up to another puff at most. Note that $C_q(t)$ does however contain some information about the spatial averages of $1-u$ and q . We thus have the lagged covariance of the multiplicative noise term

$$\begin{aligned} & \langle [\Omega(t) - \langle \Omega \rangle][\Omega(t') - \langle \Omega \rangle] \rangle \\ &= \frac{2}{\beta L^2 dt} \delta_{t,t'} L C_q(t) = \frac{2}{\beta L dt} \delta_{t,t'} C_q(t). \end{aligned} \quad (\text{B9})$$

The average and lagged variance of Ω are therefore the same as that of a noise term of the type $\sqrt{\frac{1}{2\beta}} f_a(q,u)\eta$, with $\bar{\beta} \propto L$, provided C_q brings little size effect, as argued.

In order to illustrate this and show that C_q indeed adds very little size dependence, I compute $\langle [\Omega(t) - \langle \Omega \rangle]^2 \rangle L dt$ from direct numerical simulations of the model at $r = 1.3, 1.4$, and 1.5 (Fig. 15). Time series of $\sqrt{2/\beta}\eta q$ are sampled over a duration of 50 000 time unit (too short for a likely collapse at these Reynolds numbers). I compute the cumulants. We can first see that C_q brings no size dependence to the variance of the noise, since the rescaled variance is independent of size. The increase of the variance of this noise with the Reynolds number also encourages to use a multiplicative noise in the model, since it follows the increase of \mathcal{Q} and $1-\mathcal{U}$ with Reynolds. I also computed the relative skewness. It remains mostly under 10^{-2} in absolute value and has no definite sign, for all sizes. The decrease of the absolute value of the skewness with the sample duration indicates us that most asymmetry in the distribution of $q\eta$ is the consequence of the finite duration of the sample. This further hints toward Gaussianity of this noise term. Note that some ensemble averaging would be necessary to infinitely increase the size of the sample, while enforcing the condition that no collapse occurs.

APPENDIX C: CHANGE OF VARIABLE AND FAST VARIABLE ELIMINATION

1. Slow-fast system

I change variables in the system of SDEs [Eqs. (22) and (23)] to use the new variable ζ_1 , Eq. (30). Due to the nonlinearity of the change of variable, care must be taken. I first calculate the derivatives of ζ_1 ,

$$\begin{aligned}\frac{\partial \zeta_1}{\partial \chi_1} &= 1, & \frac{\partial^2 \zeta_1}{\partial \chi_1^2} &= 0, \\ \frac{\partial \zeta_1}{\partial \chi_2} &= -\frac{1}{R^2} \frac{1}{\left(\chi_2 - \frac{1}{R^2}\right)^2}, & \\ \frac{\partial^2 \zeta_1}{\partial \chi_2^2} &= \frac{2}{R^2} \frac{1}{\left(\chi_2 - \frac{1}{R^2}\right)^3}.\end{aligned}\quad (\text{C1})$$

I then perform the change of variables using the Itô lemma (Ref. [18] Sec. 4.3.4) in the case $a = 0$. The new system of SDEs now reads

$$\begin{aligned}\frac{d\zeta_1}{dt} &= R\left(\chi_2 - \frac{1}{R^2}\right)\zeta_1 + \sqrt{R}\sqrt{\frac{2}{\beta}}\eta_1 + \frac{1}{R^3\left(\chi_2 - \frac{1}{R^2}\right)^2} \\ &\times \left(\chi_2 + \frac{\chi_2^2}{\left(\chi_2 - \frac{1}{R^2}\right)^2} - \frac{2\zeta_1\chi_2}{\chi_2 - \frac{1}{R^2}} + \zeta_1^2\right) \\ &+ \frac{2}{\beta R^3\left(\chi_2 - \frac{1}{R^2}\right)^3} + \frac{\sqrt{2}}{\sqrt{\beta}R^{\frac{5}{2}}}\frac{1}{\left(\chi_2 - \frac{1}{R^2}\right)^2}\eta_2,\end{aligned}\quad (\text{C2})$$

$$\begin{aligned}\frac{d\chi_2}{dt} &= -\frac{1}{R}\left(\chi_2 + \frac{\chi_2^2}{\left(\chi_2 - \frac{1}{R^2}\right)^2}\right) + \frac{1}{\sqrt{R}}\sqrt{\frac{2}{\beta}}\eta_2 \\ &- \frac{1}{R}\left(\frac{2\zeta_1\chi_2}{\chi_2 - \frac{1}{R^2}} + \zeta_1^2\right).\end{aligned}\quad (\text{C3})$$

In Eq. (C2), the second line and the second term of the third line arise from $d\chi_2/dt$. Meanwhile the first term of the third line is specific to changes of variables in stochastic differential equations and takes into account particular correlations between variables.

2. Langevin equations

I first perform a heuristic elimination working only with the Langevin equations. This approach is more readable, however, I will perform more approximations. Fortunately, this will have little impact on the leading order result. From the Langevin equation at leading order for ζ_1 (neglecting $1/R^3$ drift and the $1/R^5$ variance noise), if the time scale separation is large enough, one can assume that χ_2 is a constant. Thus, taking the initial condition $\zeta_1 = 0$, I obtain by variation of the constant

$$\zeta_1(t) \simeq \int_{s=0}^t \exp\left(-R\left|\chi_2 - \frac{1}{R^2}\right|(t-s)\right) \sqrt{\frac{2R}{\beta}} \eta_1(s) ds.\quad (\text{C4})$$

This tells us that $\langle \zeta_1 \rangle = 0$ and that

$$\begin{aligned}\langle \zeta_1(t)\zeta_1(t') \rangle &= \frac{1}{\beta} \left[\exp\left(-R\left|\chi_2 - \frac{1}{R^2}\right||t-t'|\right) \right. \\ &\quad \left. - \exp\left(-R\left|\chi_2 - \frac{1}{R^2}\right|(t+t')\right) \right].\end{aligned}\quad (\text{C5})$$

Inserted in Eq. (C3), I therefore find that $h_1 = 2\zeta_1\chi_2/[R(\chi_2 - \frac{1}{R^2})]$ is a noise term (*a priori* not Gaussian) whose variance is of order $\mathcal{O}(\frac{1}{R^2\beta})$, negligible with respect to $\sqrt{2/(R\beta)}\eta_2$ in the R expansion. Meanwhile $h_2 = \zeta_1^2/R$ contains a deterministic drift term of order $1/(R\beta)$ which will not be neglected in the slow-fast separation. However, it is a deterministic term of order $1/\beta$ which will vanish the $1/\beta \rightarrow 0$ limit. Let us term $(1/\beta)\mathcal{V}_1$ a quasipotential containing the corrections brought to the deterministic part by h_2 . If one has the one degree of freedom process $dx/dt = -\nabla[\mathcal{V}_0 + (1/\beta)\mathcal{V}_1] + \sqrt{2/\beta}\eta$, the pdf and mean first passage time will be proportional to $g = \exp(-\beta\mathcal{V}_0 + \mathcal{V}_1)$, so that in the limit, one finds that $\lim_{\beta \rightarrow \infty} -\frac{1}{\beta} \ln(g) = \mathcal{V}_0$. As a consequence, we will study the process

$$\frac{d\chi_2}{dt} = -\frac{1}{R}\left(\chi_2 + \frac{\chi_2^2}{\left(\chi_2 - \frac{1}{R^2}\right)^2}\right) + \frac{1}{\sqrt{R}}\sqrt{\frac{2}{\beta}}\eta_2,\quad (\text{C6})$$

for comparison with ρ and T .

As I will explain more formally in Appendix C3, fast variable eliminations are not always that simple. If the scale separation is also the large deviations parameter (see [32]), a parameter like $1/\beta$ remains finite, and the effect of h_1 and h_2 have to be taken into account in full detail. Taking this effect into account with the procedures used in this paper would not be possible. They are caused by very large deviations of χ_2 out of its starting point. The corrections can quite dramatically impact the large deviations of the slow variable, especially in the case where their moments build up to create the noise felt by the slow variable and even the multiple equilibria.

A similar procedure can be used to work out the case of multiplicative noise. $\eta_{1,2}$ will come with $g[|\chi_2/(\chi_2 - 1/R)|]$ factors and corrections in ζ_1 will arise from the expansion of g . Again, they will have high orders in $1/\beta$ and will be negligible in the $1/\beta \rightarrow 0$ limit.

3. Elimination in the Fokker-Planck equation

I now perform the elimination in the Fokker-Planck equations equivalent to the two coupled SDEs [Eqs. (C2) and (C3)]. I follow the principle of the elimination presented in [18] Sec. 8.3.3, where the time scale separation and operator properties are very similar. Note that some difference in the treatment will arise from appearing drift terms and very high order in $1/R$ which will trivially be eliminated. This approach is less readable in terms of the effect of each term in the dynamics. However it shows more clearly how and why terms are eliminated, especially since there are two small parameters

$1/R$ and $1/\bar{\beta}$. I define some differential operators

$$\mathcal{L}_1 f \equiv \frac{\partial}{\partial \zeta_1} \left[\left(\chi_2 - \frac{1}{R^2} \right) \zeta_1 f \right] + \frac{1}{\bar{\beta}} \frac{\partial^2 f}{\partial \zeta_1^2}, \quad (C7)$$

$$\mathcal{L}_2 f \equiv \zeta_1^2 \frac{\partial f}{\partial \chi_2}, \quad (C8)$$

$$\mathcal{L}'_2 f \equiv -2\zeta_1 \frac{\partial}{\partial \chi_2} \left(\frac{\chi_2}{\chi_2 - \frac{1}{R^2}} f \right), \quad (C9)$$

$$\mathcal{L}_3 f \equiv \frac{\partial}{\partial \chi_2} \left[\left(\chi_2 + \frac{\chi_2^2}{\left(\chi_2 - \frac{1}{R^2} \right)^2} \right) f \right] + \frac{1}{\bar{\beta}} \frac{\partial^2 f}{\partial \chi_2^2}, \quad (C10)$$

$$\mathcal{L}_4 f \equiv \frac{\partial}{\partial \zeta_1} \left[\frac{f}{\chi_2 - \frac{1}{R^2}} \left(\frac{2}{\beta \left(\chi_2 - \frac{1}{R^2} \right)} + \chi_2 + \frac{\chi_2^2}{\left(\chi_2 - \frac{1}{R^2} \right)^2} - \frac{2\zeta_1 \chi_2}{\chi_2 - \frac{1}{R^2}} + \zeta_1^2 \right) \right], \quad (C11)$$

$$\mathcal{L}_5 f \equiv \frac{1}{\beta} \frac{\partial^2}{\partial \zeta_1^2} \left(\frac{f}{\chi_2 - \frac{1}{R^2}} \right). \quad (C12)$$

Note that I made an approximation in including Reynolds number dependence in \mathcal{L}_3 . This is done in order to give a better description of the quasipotential minima and saddle. Even if the time scales of $X_{1,2}$ are not separated by the same scalings if of low amplitude, X_2 is still leader through the non-normal linear lift up term. The Fokker-Planck equation for pdf $P(\zeta_1, \chi_2, t)$ equivalent to the Langevin equations (C2) and (C3) reads

$$\frac{\partial P}{\partial t} = R\mathcal{L}_1 P + \frac{1}{R}\mathcal{L}_3 P + \frac{1}{R}\mathcal{L}_2 P + \frac{1}{R}\mathcal{L}'_2 P + \frac{1}{R^3}\mathcal{L}_4 P + \frac{1}{R^5}\mathcal{L}_5 P. \quad (C13)$$

One finds ν the steady pdf of ζ_1 , where the slow dynamics of χ_2 is eliminated, from the leading order $\mathcal{L}_1 \nu = 0$. This normalized pdf reads

$$\nu_{\chi_2}(\zeta_1) = \sqrt{\frac{\bar{\beta} \left| \chi_2 - \frac{1}{R^2} \right|}{2\pi}} \exp \left(-\bar{\beta} \left| \chi_2 - \frac{1}{R^2} \right| \frac{\zeta_1^2}{2} \right). \quad (C14)$$

Using ν , I define the projector of a pdf on statistics where the fast steady dynamics of ζ_1 and the slow time dependent dynamics of χ_2 are independent,

$$\Pi f(\chi_2, \zeta_1, t) = \nu_{\chi_2} \int f(\chi_2, \zeta_1, t) d\zeta_1. \quad (C15)$$

In particular, applied to the pdf P , this yields $\Pi P = \nu_{\chi_2}(\zeta_1) \times p(\chi_2)$, where $p(\chi_2)$ is the pdf of χ_2 when the fast variable is eliminated. The goal of the procedure is to obtain a partial differential equation of the Fokker-Planck type for $p(\chi_2)$. We have that $\mathcal{L}_1 \Pi f = 0$, since ν is in the kernel of \mathcal{L}_1 and the integral is independent on ζ_1 and thus factors out of the derivatives. Meanwhile $\Pi \mathcal{L}_1 f = 0$, since this is the integral of derivatives both with respect to ζ_1 : it yields boundary terms at infinity which are zero. I define the probability density function of $\nu(\chi_2, \zeta_1, t)$ in which the two variables are made independent due to time scale separation as well as another function w ,

$$v \equiv \Pi P, \quad w = (1 - \Pi)P, \quad \Pi w = 0. \quad (C16)$$

We will work with Laplace transforms

$$\hat{h} = \int h(t) \exp(st) dt, \quad (C17)$$

so that the time derivatives correspond to a multiplication by s for Laplace transforms. Using \hat{v} , \hat{w} , and applying operator Π and $1 - \Pi$ to Eq. (C13) yields respectively

$$s\hat{v} = \frac{1}{R} \Pi \mathcal{L}_3 \hat{v} + \frac{1}{R} \Pi (\mathcal{L}_2 \hat{v} + \mathcal{L}'_2 \hat{v}) + \frac{1}{R} \Pi (\mathcal{L}_3 + \mathcal{L}_2 + \mathcal{L}'_2) \hat{w} + \frac{1}{R^3} \Pi \left(\mathcal{L}_4 + \frac{1}{R^2} \mathcal{L}_5 \right) \hat{P}, \quad (C18)$$

$$s\hat{w} = R\mathcal{L}_1 \hat{w} + \frac{1}{R} (1 - \Pi) (\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3) \hat{w} + \frac{1}{R} (1 - \Pi) (\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3) \hat{v} + \frac{1}{R^3} (1 - \Pi) \left(\mathcal{L}_4 + \frac{1}{R^2} \mathcal{L}_5 \right) \hat{P}. \quad (C19)$$

From Eq. (C19) we can write that

$$\hat{w} = \frac{1}{R} \left[\frac{s}{R} + \mathcal{L}_1 + \frac{1}{R^2} (1 - \Pi) (\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3) \right]^{-1} \circ \left[\frac{1}{R} (1 - \Pi) (\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3) \hat{v} + \frac{1}{R^3} (1 - \Pi) \left(\mathcal{L}_4 + \frac{1}{R^2} \mathcal{L}_5 \right) \hat{P} \right]. \quad (C20)$$

We can then insert it in Eq. (C18); this yields

$$\begin{aligned}
s\hat{v} &= \frac{1}{R}\Pi\mathcal{L}_3\hat{v} + \frac{1}{R}\Pi(\mathcal{L}_2\hat{v} + \mathcal{L}'_2\hat{v}) + \frac{1}{R^3}\Pi(\mathcal{L}_3 + \mathcal{L}_2 + \mathcal{L}'_2) \circ \left[\frac{s}{R} + \mathcal{L}_1 + \frac{1}{R^2}(1 - \Pi)(\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3) \right]^{-1} \\
&\circ [(1 - \Pi)(\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3)\hat{v}] + \frac{1}{R^3}\Pi(\mathcal{L}_3 + \mathcal{L}_2 + \mathcal{L}'_2) \circ \left[\frac{s}{R} + \mathcal{L}_1 + \frac{1}{R^2}(1 - \Pi)(\mathcal{L}_2 + \mathcal{L}'_2 + \mathcal{L}_3) \right]^{-1} \\
&\circ (1 - \Pi)\left(\mathcal{L}_4 + \frac{1}{R^2}\mathcal{L}_5\right)\hat{P} + \frac{1}{R^3}\Pi\left(\mathcal{L}_4 + \frac{1}{R^2}\mathcal{L}_5\right)\hat{P}. \tag{C21}
\end{aligned}$$

Since the time scale at which $p(\chi_2)$ lives is R and it is the time dependence of v , we find that \hat{v} takes non-negligible values when $s = \mathcal{O}(1/R)$. The procedure works by order identification; the relevant order for $p(\chi_2)$ is $1/R$, where we have

$$s\hat{v} = \frac{1}{R}\Pi\mathcal{L}_3\hat{v} + \frac{1}{R}(\Pi\mathcal{L}_2\hat{v} + \mathcal{L}'_2\hat{v}). \tag{C22}$$

In all these eliminations, the quality of the result depends on how much the time scale is actually concentrated around $1/R$. Note that Π and \mathcal{L}_3 do not commute, however, we have

$$\Pi\mathcal{L}_3\hat{v} = v_{\chi_2} \int_{\zeta'_1} [\mathcal{L}_3(v_{\chi_2})\hat{p}(\chi_2)]d\zeta'_1 = v_{\chi_2}\mathcal{L}_3 \left(\underbrace{\left(\int_{\zeta'_1} v_{\chi_2}(\zeta'_1) d\zeta'_1 \right)}_{=1} p(\chi_2) \right) = v_{\chi_2}(\zeta_1)\mathcal{L}_3\hat{p}(\chi_2). \tag{C23}$$

The integral can be passed within the differential operators and the probability density function $v_{\chi_2}(\zeta_1)$ is normalized. We then note that $\Pi\mathcal{L}'_2\hat{v} = 0$, since the integrand is odd. Meanwhile, we pass ζ_1 and the integral in the ∂_{χ_2} derivative and perform the Gaussian integral in $\Pi\mathcal{L}_2\hat{v}$; this yields

$$\Pi\mathcal{L}_2\hat{v} = v_{\chi_2}(\zeta_1) \frac{1}{\tilde{\beta}} \frac{\partial}{\partial \chi_2} \left(\frac{1}{\chi_2 - \frac{1}{R^2}} \hat{p}(\chi_2) \right).$$

This gives rise to a deterministic term $1/[\tilde{\beta}(\chi_2 - 1/R^2)]$, since it appears in a first order derivative of the pdf $p(\chi_1)$. This is actually the same term as the average in the fast variable elimination in the SDE. Again, we will not take this term into account in the final result. Indeed, as explained in Sec. C2, while this term remains in the slow dynamics, it will disappear when we consider large deviations in the $1/\tilde{\beta}$ limit. This can be shown more formally by performing the elimination not on the probability density function, but on a large deviation function S such that $P \propto \exp(-\tilde{\beta}S)$. One then does the order identification, on R and on $\tilde{\beta}$ (similarly to [32], where the scale separation and the large deviation parameter are the same). In this case the “ \mathcal{L}_2 ” terms are indeed retained at order $\mathcal{O}(1/R)$: they are actually the type of terms which bring the subtle effects in the slow-fast separation of Grafke *et al.* [32], with one small parameter. However, they will be one order too low in $\tilde{\beta}$ to be retained in the differential equation for the large deviation function, and thus will not perturb the large deviations of χ_2 .

Once these two terms are neglected, we find that

$$s v_{\chi_2}(\zeta_1)\hat{p} = v_{\chi_2}(\zeta_1) \frac{1}{R} \mathcal{L}_3\hat{p}. \tag{C24}$$

We can factor $v_{\chi_2}(\zeta_1)$ and take the inverse Laplace transform. This yields the Fokker-Planck equivalent to Langevin equation (C6).

APPENDIX D: QUASIPOTENTIAL DIFFERENCES

I give the quasipotential difference between the saddle X_2^+ and the minimum X_2^- , which appears in mean first passage times. If the noise is additive, $a = 0$, we have

$$\begin{aligned}
\mathcal{V}(X_2^+) - \mathcal{V}(X_2^-) &= \Delta\mathcal{V} = \frac{R}{2} \sqrt{1 - \frac{4}{R^2}} - \frac{4 \ln(R)}{R} + \frac{2}{R} \sqrt{1 - \frac{4}{R^2}} \\
&- \frac{2}{R} \left[\ln \left(\frac{\frac{1}{R^2} + \frac{(1 - \frac{2}{R^2}) + \sqrt{1 - \frac{4}{R^2}}}{2}}{1 + R \frac{(R - \frac{2}{R}) - R\sqrt{1 - \frac{4}{R^2}}}{2}} \right) \right]. \tag{D1}
\end{aligned}$$

In the first case of multiplicative noise $a = \frac{1}{2}$, we have

$$\begin{aligned}
\Delta\mathcal{V} &= \frac{R}{2} \sqrt{1 - \frac{4}{R^2}} - \frac{2 \ln(R)}{R} \\
&- \frac{1}{R} \left[\ln \left(\frac{\frac{1}{R^2} + \frac{(1 - \frac{2}{R^2}) + \sqrt{1 - \frac{4}{R^2}}}{2}}{1 + R \frac{(R - \frac{2}{R}) - R\sqrt{1 - \frac{4}{R^2}}}{2}} \right) \right]. \tag{D2}
\end{aligned}$$

In the second case of multiplicative noise $a = 1$, we have

$$\begin{aligned}
\Delta\mathcal{V} &= \left(\frac{R}{2} - \frac{1}{R} \right) \sqrt{1 - \frac{4}{R^2}} \\
&- \frac{1}{R^3} \ln \left(\frac{(R - \frac{2}{R}) + R\sqrt{1 - \frac{4}{R^2}}}{(R - \frac{2}{R}) - R\sqrt{1 - \frac{4}{R^2}}} \right). \tag{D3}
\end{aligned}$$

APPENDIX E: NONABSORBING LAMINAR STATE FOR $a < 1/2$

In the $a = 1$ case, the pdf $\rho(X_1, X_2)$ [solution of Eq. (26)] is not defined as a function, but instead as a distribution. In the

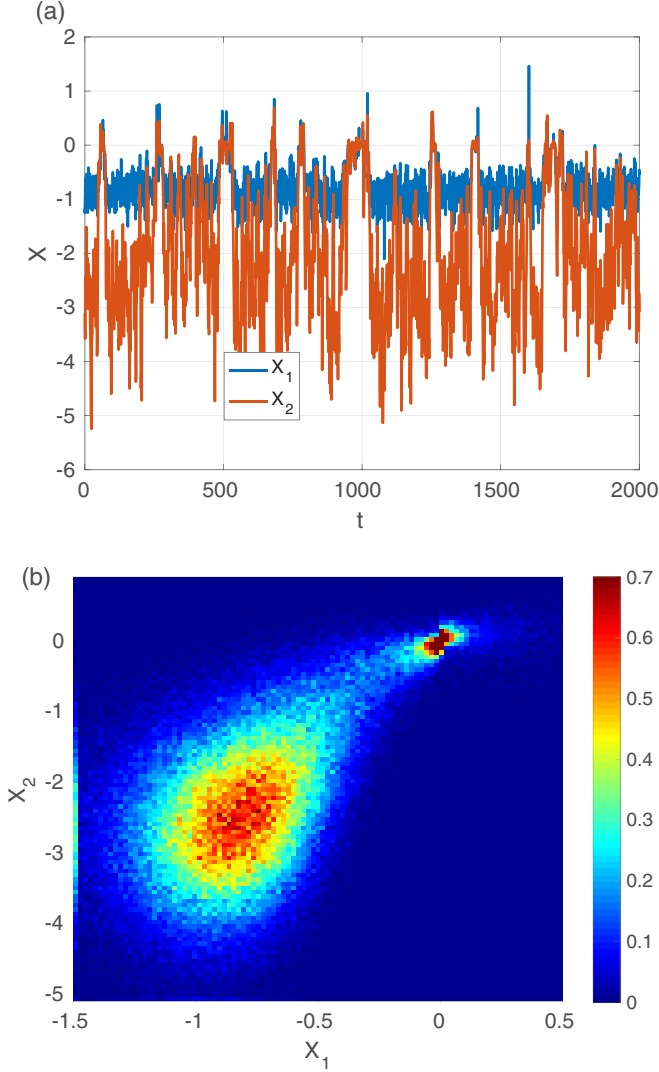


FIG. 16. (a) Time series of X_1, X_2 obtained from numerical integration of the coupled SDEs in the $a = \frac{1}{2}$ case (square root multiplicative noise) at $\bar{\beta} = 5$ and $R = 3$. (b) Color levels of probability density function $\rho_{a=\frac{1}{2}}(X_1, X_2)$ sampled numerically by integration of the SDEs for $R = 3$ and $\bar{\beta} = 5$.

$a < \frac{1}{2}$ cases, this is quite different and the pdf is a function. This is also the case for the empirical distribution sampled numerically in the $a = \frac{1}{2}$ case. The pdf can be normalized and defined as a function in that case. This can be illustrated by numerical integration of the coupled SDEs at $R = 3$ and $\bar{\beta} = 5$ and $a = \frac{1}{2}$ [Fig. 16(a)]. The time series show that the $X_{1,2} = 0$, the *a priori* absorbing point, can be visited quite often, but that

the system still manages to avoid turbulence collapse. This is in agreement with the fact that while the probability of being in the neighborhood of the laminar state is large in the steady state, the probability of being in the turbulent state, while smaller, is nonzero. This is seen in the steady probability density function of $\rho_{a=\frac{1}{2}}(X_1, X_2)$, sampled numerically [Fig. 11(b)]. The slow divergence at zero is not captured numerically due to the finite duration of the sample and binning, but $\rho_{\frac{1}{2}}(0,0)$ is still quite large.

This can be understood by calculating the probability density function ρ of X_2 , when the fast variable is eliminated and the quasipotential is known. The pdf is the solution of the steady Fokker-Planck equation,

$$0 = -\frac{\partial}{\partial X_2} \left\{ -\left[\frac{d\mathcal{V}}{dX_2} + \mathcal{O}\left(\frac{1}{\beta}\right) \right] \left(\frac{X_2}{X_2 - \frac{1}{R}} \right)^{2a} \rho \right\} + \frac{1}{\bar{\beta}} \frac{\partial}{\partial X_2} \left(\underbrace{\left(\frac{X_2}{X_2 - \frac{1}{R}} \right)^{2a}}_{=\mu} \rho \right). \quad (E1)$$

Note that the quasipotential is already introduced, since I include a $(\frac{X_2}{X_2 - \frac{1}{R}})^{2a} / (\frac{X_2}{X_2 - \frac{1}{R}})^{2a}$ factor in the first order derivative. I solve for μ and compute ρ ,

$$\rho = C \left(\frac{X_2}{X_2 - \frac{1}{R}} \right)^{-2a} \exp[-\beta\mathcal{V} + \mathcal{O}(1)] = C\rho'. \quad (E2)$$

The pdf ρ can be defined as a function if it is normalizable, if $C = 1 / \int \rho' dX_2$ is defined. This integral is obviously not defined if $a \geq 1$, due to the divergence of the quasipotential and thus, the exponential. If $a < \frac{1}{2}$, not only the quasipotential is finite, but the integral of the prefactor is also defined. Indeed,

$$0 \leq \frac{1}{C} \leq \int_{X_2=-\infty}^{-1} \left(1 + \frac{1}{R} \right)^{2a} \exp[-\bar{\beta}\mathcal{V} + \mathcal{O}(1)] dX_2 + \int_{X_2=-1}^0 \left(\underbrace{\frac{1}{RX_2}}_{\leq (1+1/R)/|X_2|} \right)^{2a} \times \max_{X_2 \in [-1;0]} \exp[-\bar{\beta}\mathcal{V} + \mathcal{O}(1)] dX_2. \quad (E3)$$

The first integral is finite because the integral is a Gaussian; meanwhile the second integral is bounded above by $[1/(1-2a)](1+1/R)^{2a} \max_{X_2 \in [-1;0]} \exp[-\bar{\beta}\mathcal{V} + \mathcal{O}(1)]$ if $a < \frac{1}{2}$, otherwise it is not defined.

[1] C. Ansonge and J. P. Mellado, *Boundary-Layer Meteorol.* **153**, 89 (2014).
 [2] L. Mahrt, *Annu. Rev. Fluid Mech.* **46**, 23 (2014).
 [3] J. Rolland, *Eur. Phys. J. B* **88**, 66 (2015).
 [4] D. Moxey and D. Barkley, *Proc. Natl. Acad. Sci. USA* **18**, 8091 (2010).

[5] I. J. Wygnanski and F. H. Champagne, *J. Fluid Mech.* **59**, 281 (1973).
 [6] I. Wygnanski, M. Sokolov, and D. Friedman, *J. Fluid Mech.* **69**, 283 (1975).
 [7] A. Lundbladh and A. V. Johansson, *J. Fluid Mech.* **229**, 499 (1991).

- [8] Y. Duguet, O. Le Maître, and P. Schlatter, *Phys. Rev. E* **84**, 066315 (2011).
- [9] M. Couliou and R. Monchaux, *Phys. Rev. E* **93**, 013108 (2016).
- [10] S. Bottin and H. Chaté, *Eur. Phys. J. B* **6**, 143 (1998).
- [11] B. Hof, A. de Lozar, D. J. Kuik, and J. Westerweel, *Phys. Rev. Lett.* **101**, 214501 (2008).
- [12] V. A. Romanov, *Funkt. Anal. Appl.* **7**, 137 (1973).
- [13] G. Lemoult, J.-L. Aider, and J. E. Wesfreid, *Phys. Rev. E* **85**, 025303 (2012).
- [14] A. Schmiegél and B. Eckhardt, *Phys. Rev. Lett.* **79**, 5250 (1997).
- [15] H. Chaté and P. Manneville, *Phys. Rev. Lett.* **58**, 112 (1987).
- [16] D. Barkley, *J. Phys.: Conf. Ser.* **318**, 032001 (2011); *J. Fluid Mech.* **803**, 1 (2016).
- [17] P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).
- [18] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Science* (Springer, Heidelberg, 2009).
- [19] T. M. Schneider and B. Eckhardt, *Phys. Rev. E* **78**, 046310 (2008).
- [20] T. Tél and Y.-C. Lai, *Phys. Rep.* **460**, 245 (2008).
- [21] S. M. E. Rabin, C. P. Caulfield, and R. R. Kerswell, *J. Fluid Mech.* **712**, 244 (2012).
- [22] A. Monokrousos, A. Bottaro, L. Brandt, A. Di Vita, and D. S. Henningson, *Phys. Rev. Lett.* **106**, 134502 (2011).
- [23] F. Cérou and A. Guyader, *Stochastic Anal. Appl.* **25**, 417 (2007).
- [24] J. Rolland, F. Bouchet, and E. Simonnet, *J. Stat. Phys.* **162**, 277 (2016).
- [25] H. Touchette, *Phys. Rep.* **478**, 1 (2009).
- [26] F. Cérou, A. Guyader, T. Lelièvre, and F. Malrieu, *Lat. Am. J. Probab. Math. Stat.* **10**, 359 (2013).
- [27] M. Rosenbluth and A. Rosenbluth, *J. Chem. Phys.* **23**, 356 (1955).
- [28] O. Dauchot and P. Manneville, *J. Phys.* **2**, 371 (1997).
- [29] F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier, *J. Chem. Phys.* **134**, 054108 (2011).
- [30] P. Del Moral, *Feynman-Kac formulae. Genealogical and Interacting Particle Systems with Applications* (Springer, New York, 2004).
- [31] F. Bouchet and J. Reygner, *Ann. Henri Poincaré* **17**, 3499 (2016).
- [32] T. Grafke, F. Bouchet, T. Tangarife, and E. Vanden Eijnden, *J. Stat. Phys.* **162**, 793 (2016).
- [33] J. Laurie and F. Bouchet, *New J. Phys.* **17**, 015009 (2015).
- [34] T. Tangarife, Ph.D. thesis, ENS de Lyon, 2015.
- [35] E. Simonnet and F. Bouchet (private communication).
- [36] F. Bouchet and E. Simonnet, *Phys. Rev. Lett.* **102**, 094504 (2009).
- [37] T. Nemoto and A. Alexakis, *Phys. Rev. E* **97**, 022207 (2018).
- [38] X. Wan and H. Yu, *J. Comput. Phys.* **331**, 209 (2017).
- [39] B. Song, D. Barkley, B. Hof, and M. Avila, *J. Fluid Mech.* **813**, 1045 (2017).
- [40] F. Waleffe, *Phys. Fluid* **9**, 883 (1997).
- [41] E. Simonnet, *Stat. Comput.* **26**, 211 (2016).
- [42] T. Laffargue, K.-D. Nguyen, T. Lam, J. Kurchan, and J. Tailleur, *J. Phys. A: Math. Theor.* **46**, 254002 (2013).
- [43] T. M. Schneider, B. Eckhardt, and J. A. Yorke, *Phys. Rev. Lett.* **99**, 034502 (2007).
- [44] A. P. Willis and R. R. Kerswell, *J. Fluid Mech.* **619**, 213 (2008).
- [45] C.-E. Bréhier, M. Gazeau, L. Goudenège, T. Lelièvre, and M. Rousset, *Ann. Appl. Probab.* **26**, 3559 (2016).
- [46] J. Rolland and E. Simonnet, *J. Comput. Phys.* **283**, 541 (2015).
- [47] E. Gobet, *Stochastic Process Appl.* **87**, 167 (2000).
- [48] J. S. Langer, *Phys. Rev. Lett.* **21**, 973 (1968).
- [49] M. Berhanu, R. Monchaux, S. Fauve, N. Mordant, F. Pétrélis, A. Chiffaudel, F. Daviaud, B. Dubrulle, L. Marié, F. Ravelet, M. Bourgoïn, P. Odier, J.-F. Pinton, and R. Volk, *Europhys. Lett.* **77**, 59001 (2007).
- [50] W. E. W. Ren and E. Vanden-Eijnden, *Commun. Pure Appl. Math.* **57**, 1 (2004).
- [51] B. Derrida, *J. Stat. Mech.* (2007) P07023.
- [52] R. Graham and T. Tél, *J. Stat. Phys.* **35**, 729 (1984).
- [53] E. Simonnet and H. A. Dijkstra, *J. Phys. Oceanogr.* **32**, 1747 (2002).
- [54] F. Bouchet and H. Touchette, *J. Stat. Mech.* (2012) P05028.
- [55] P.-A. Rey Täuber and J. L. Cardy, *J. Phys. A: Math. Gen.* **32**, 1585 (1999).
- [56] P. Manneville, *Dissipative Structures and Weak Turbulence* (Academic Press, San Diego, London, 1990).
- [57] D. Faranda, V. Lucarini, P. Manneville, and J. Wouters, *Chaos Solitons Fractals* **64**, 26 (2014).
- [58] J. Wouters and F. Bouchet, *J. Phys. A* **49**, 374002 (2016).