Article:
Hollerbach, R, Kim, E-J and Schmitz, L (2020) Time-dependent probability density functions and information diagnostics in forward and backward processes in a stochastic prey-predator model of fusion plasmas. Physics of Plasmas, 27. 102301. ISSN 1070-664X

https://doi.org/10.1063/5.0011473

Time-dependent probability density functions and information diagnostics in forward and backward processes in a stochastic prey-predator model of fusion plasmas

Rainer Hollerbach[1], Eun-jin Kim[2,a)] and Lothar Schmitz[3]

[1] *Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK*

[2] *Fluid and Complex Systems Research Centre, Coventry University, Coventry CV1 2TT, UK*

[3] *Department of Physics and Astronomy, University of California Los Angeles, Los Angeles, CA 90095-7799, USA*

[a)] *Author to whom correspondence should be addressed: ejk92122@gmail.com*

Abstract

Forward and backward processes associated with the Low-to-High (L-H) transition in magnetically confined fusion plasmas are investigated by using a time-dependent Probability Density Function (PDF) approach and information length diagnostics. Our model is based on the extension of the deterministic prey-predator-type model (Kim and Diamond, Phys. Rev. Lett. **91**, 185006, 2003) to a stochastic model by including two independent, short-correlated Gaussian noises. The 'forward' process consists of ramping up the input power linearly in time so that zonal flows self-regulate with turbulence after their initial growth from turbulence. The 'backward' process ramps the power down again, by starting at time $t = t_*$ when the input power is switched to $Q(t) = Q(2t_* - t)$ for $t > t_*$, linearly decreasing with time until $t = 2t_*$. Using three choices for $Q(t)$, with differing ramping rates, the time-dependent PDFs are calculated by numerically solving the appropriate Fokker-Planck equation, and several statistical measures including the information length for the forward and backward processes are investigated. The information length $\mathcal{L}_x(t)$ and $\mathcal{L}_v(t)$ for turbulence and zonal flows, respectively, are path-dependent dimensionless numbers, representing the total number of statistically different states that turbulence and zonal flows evolve through in time t. In particular, PDFs are shown to be strongly non-Gaussian with convoluted structures and multiple peaks, intermittency in zonal flows playing a key role in turbulence regulation. The stark difference between the forward and backward processes is captured by time-dependent PDFs of turbulence and zonal flows and corresponding information length diagnostics. The latter are shown to give us a useful insight in understanding the correlation and self-regulation, and transitions to the self-regulatory dithering phase.

I. INTRODUCTION

The need for a proper statistical theory for understanding fusion plasmas has grown significantly over the last few decades, with experiments and simulations revealing ample evidence for strong, non-Gaussian fluctuations, anomalous transport, or intermittency [1–6]. The latter question the validity of the mean-field-type theory based on small, Gaussian fluctuations, the concepts of transport coefficients, or the utility of the different moments/cumulants (mean value, variance, skewness, kurtosis, etc.) while highlighting the importance of the calculation of an entire probability density function (PDF) [7, 8]. In particular, rare but large-amplitude events can mediate large transport and contribute to the tails of PDFs whose effects are not easily captured by a few low-order moments. Although PDF tails were successfully predicted by using non-perturbative methods such as instantons (the extreme of a path-integral) in different types of fusion plasma turbulence (Hasegawa-Mima, ion temperature gradient, edge turbulence, etc.) [9–12], these studies tend to focus on stationary PDFs. For plasma turbulence out of equilibrium, the prediction of an entire, time-dependent PDF is in order. We have initiated a time-dependent PDF approach in various non-equilibrium processes [13–24]. The main purpose of this paper is to demonstrate the importance of such a time-dependent PDF approach and intermittency in understanding the Low-to-High (L-H) transition and the backward H-L transition in fusion plasmas.

The L-H transition constitutes one of the most interesting examples of self-organization, in which plasmas organize themselves into an ordered, high-confinement (H) mode from the low-confinement (L) mode at a critical power threshold [25–46], accompanied by the regulation between structures (zonal flows, mean flows) and turbulence [47–49]. Given their critical implications for future burning plasmas and commercial power plants as well as the success of the ITER project [28, 35] where the H-mode is a basic scenario, the L-H transition and zonal flows have been of primary interest from the perspectives of theory, computer simulations, and experiments. Much attention has been paid to the issues surrounding the causality relation and correlation among different players, triggering mechanisms, hysteresis, the threshold power scaling, and the effects of density, diverter/magnetic geometry and neutrals [35], most of which still remain controversial. Statistical methods like Fourier/wavelet analysis, bicoherence, phase relation, turbulence-flow energy transfer, fluxes, turbulent coefficients, and moments (mean value, variance, etc.) have often been adopted to study these

issues.

Recently, we have extended the deterministic version of the prey-predator model of the L-H transition to a stochastic model [29] by incorporating the overall (incoherent) effect of the unresolved scales [24], which are not included for the deterministic model, and investigated the time-dependent PDFs. We note that there have been experimental observations of mini-avalanches and the power flow from the core has fast transients, stochastic noise being clearly part of the physics. In this stochastic model, while turbulence is mainly determined by the input power, it is also weakly driven by a small stochastic noise. As a result, for a fixed turbulence energy required for the L-H transition, the contribution from this stochastic noise to turbulence slightly lowers the required power threshold. Furthermore, we have shown that time-dependent PDFs offer a new insight into the L-H transition that is simply inaccessible otherwise [24]. In particular, we demonstrated strongly non-Gaussian PDFs and the interesting possibility that intermittency – thought to be important for enhancing transport (e.g. by blobs, streamers, etc.) – can also play a key role in the regulation of turbulence by zonal flows in the L-H transition.

Here, we extend this work to elucidate the effects of different input power ramp rates (heating power) and stochasticity on hysteresis. Specifically, we ramp up an input power linearly in time so that zonal flows self-regulate with turbulence (undergoing the dithering phase) after their initial growth from turbulence, which we call the forward process. Before the completion of the L-H transition, we start decreasing the input power again to induce the transition back to the L-mode, which we call the backward process. The faster the ramping rate, the further from a stationary state the system is; we therefore investigate the consequences of different ramping rates. We compute the time-dependent PDFs in the entire forward and backward sequence, and calculate different statistical measures including the information length (see §II, §IV-V), entropy, and Fisher information [50] (see §IV-V). Here, entropy measures uncertainty, disorder, or the lack of information, while the Fisher information is a concept opposite to the entropy, measuring the amount of information or order associated with a PDF (see §III for more details).

Our principal results include that i) the mean-field-type theory is invalid, with the limited utility of mean value, variance, stationary PDFs; ii) the L-H transition can involve a strongly non-Gaussian PDF and intermittency; iii) faster ramping yields more deviation from Gaussian PDFs and less duration of the dithering phase; iv) enhanced right tails for rare large

zonal flows can play a crucial role in turbulence regulation; v) the stark difference between the forward and backward processes is captured by time-dependent PDFs of turbulence and zonal flows; vi) the larger the stochasticity, the more asymmetry there is between forward and backward processes. Furthermore, we show that the information geometric tool (information length) provides us with a useful index to understand correlation and self-regulation, as well as forecasting the transitions to the dithering phase.

The remainder of this paper is organized as follow. §II introduces the information length and key expressions. §III presents our stochastic model by extending [29], and §IV summarizes the definitions of various statistical measures that are analyzed. Results are presented in §V for different choices of input power and for the forward and backward processes. Discussion and conclusions are found in §VI. Appendices A and B contain some detailed derivations of the key properties of the information length, while Appendix C provides the derivation for our main Fokker-Planck equation (12) from the Langevin equations (9)-(10) and the statistical properties of the noise terms (11).

II. INFORMATION LENGTH AND INFORMATION PHASE PORTRAIT

A key characteristic of non-equilibrium processes is a PDF which changes with time. For simplicity, we consider the case of one stochastic variable x which has a time-dependent PDF $p(x,t)$. In the following, we use the concept of relative entropy which is different from entropy. Specifically, relative entropy is a way of comparing two PDFs such that its value is zero for two identical PDFs while becoming larger as the difference between the PDFs increases. For a PDF which continuously changes with time, the main interest is the comparison of two (temporally) adjacent PDFs along the evolution path, say, $p(x,t)$ and $p(x,t+\delta t)$ in the limit $\delta t \to 0$. We thus calculate an (infinitesimal) relative entropy between $p(x,t)$ and $p(x,t+\delta t)$ in the limit $\delta t \to 0$. Summing the square-root of the infinitesimal relative entropy ($\tau(t)^{-1}$) along the path, we define the (dimensionless) information length $\mathcal{L}(t)$ [14–17, 19–21] (see also Appendix A)

$$\mathcal{L}(t) = \int_0^t \frac{dt_1}{\tau(t_1)} \quad = \int_0^t dt_1 \sqrt{\int dx \frac{1}{p(x,t_1)} \left[\frac{\partial p(x,t_1)}{\partial t_1}\right]^2}, \tag{1}$$

where

$$\frac{1}{\tau(t)^2} = \int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t}\right]^2 \equiv \mathcal{E}(t). \tag{2}$$

The unit of τ in Eq. (2) is time, since the unit of $p(x,t)$ is x^{-1}; $\tau(t)$ is referred to as a dynamical time unit at time t for information change, representing the characteristic time scale of a PDF. Since $\tau(t)$ depends on time in general, $\mathcal{L}(t)$ in Eq. (1) is obtained by measuring the clock time in (instantaneous time) unit $\tau(t)$; $\mathcal{L}(t)$ quantifies the total number of statistically different states that a variable x passes through between time 0 and t as $p(x,t)$ evolves from its initial PDF $p(x,0)$. $\mathcal{L}$ is dimensionless, and we can check that Eqs. (1)-(2) are invariant under a time-independent change of variable (see Appendix B) unlike other entropy-based statistical measures (e.g. entropy, relative entropy, Jensen divergences, etc.). Note that Eq. (1) is different from the statistical distance invoked by Wooters [51] which represents the shortest distance between two given PDFs, where t is a parameter that labels a path between these two PDFs.

In principle, $\mathcal{L}(t)$ depends on the parameters in the initial PDF $p(x,0)$ and in the model. However, when all the parameters are fixed, $\mathcal{L}(t)$ is a function of time only, representing the cumulative change to the PDF along the path, starting from a given initial PDF $p(x,0)$ at $t=0$. For instance, $\mathcal{L}(t=0)$ compares $p(x,0)$ and $p(x,0)$, and is thus always zero since there is no difference between two (initially) identical PDFs. $\mathcal{L}(t)$ increases with time unless $p(x,t)$ approaches a stationary state where there is no temporal change in $p(x,t)$. Also, if $\mathcal{L}(t=T)=0$, the system is stationary with no change in the PDF for all time $t=[0,T]$. However, $p(x,0)=p(x,T)$ does not mean $\mathcal{L}(t=T)=0$ since the PDF at some intermediate time $t\in(0,T)$ can be different from $p(x,0)=p(x,T)$. The larger $\mathcal{L}(T)$, the more change in the PDF along the path $t=(0,T)$.

By using this path-dependence and by varying one parameter in the initial condition or model, we can study how $\mathcal{L}(t)$ depends on this parameter at a fixed time. For instance, we can map out an attractor structure in a relaxation problem by measuring $\mathcal{L}_\infty=\mathcal{L}(t\to\infty)$ against the mean position x_0 of the initial PDF since $\mathcal{L}(t)$ approaches a finite value [14, 15] in this limit $t\to\infty$ due to the relaxation of an initial PDF into a stationary PDF around $x=0$ [13, 14, 16, 18, 22]. Some of the interesting findings include that the relation between $\mathcal{L}_\infty$ and x_0 is linear for a linear Ornstein-Uhlenbeck (O-U) process regardless of its initial width and the strength of stochastic noise, illustrating that $\mathcal{L}_\infty$ captures the linear geometry of a linear process [14, 18]. However, for a nonlinear relaxation problem, $\mathcal{L}_\infty$ against x_0 is not linear but a power-law, the power-law index depending on the width of an initial PDF and the strength of a stochastic noise [14, 18]. In both cases, $\mathcal{L}_\infty$ increases monotonically

with x_0. In contrast, for a chaotic system, $\mathcal{L}_\infty$ abruptly changes with x_0, revealing the sensitive dependence of $\mathcal{L}_\infty$ on the initial condition (like the Lyapunov exponent) [13, 14]. The path-dependence of $\mathcal{L}(t)$ was also utilized to understand hysteresis involved in phase transitions such as Ginzburg-Landau model [19], and will be studied further in this paper.

In general, for a system with m variables x_i $(i = 1, 2, ..m)$, we extend Eq. (2) to

$$\mathcal{L}_{x_i}(t) = \int_0^t \frac{dt_1}{\tau_{x_i}(t_1)}, \tag{3}$$

$$\frac{1}{[\tau_{x_i}(t)]^2} = \int dx_i \frac{1}{p(x_i,t)} \left[\frac{\partial p(x_i,t)}{\partial t}\right]^2 \equiv \mathcal{E}_{x_i}, \tag{4}$$

where $p(x_i,t) = \int \Pi_{j\neq i}(dx_j)\, p(x_1, x_2, ..., x_m)$ is a marginal PDF of x_i. Eqs. (3)-(4) are again invariant under time-independent changes of variables x_i (see Appendix B). As a dimensionless number, $\mathcal{L}_{x_i}(t)$ gives a useful proxy for the evolution of x_i in its statistical state in time, permitting us to compare the dynamics of different x_i, which have different units as well as quantifying the correlation among them. In particular, τ_{x_i} and $\mathcal{L}_{x_i}$ depend on the path of x_i, and the correlation or causality between x_i and x_j $(i \neq j \in [1, m])$ can be inferred by comparing τ_{x_i} and τ_{x_j}. Furthermore, the 2-D plane of $\tau_{x_i}^{-1}$ against $\tau_{x_j}^{-1}$ – an *information phase portrait* – is useful for capturing self-regulation between x_i and x_j.

III. MODEL

We take advantage of the property of self-organising systems that statistical properties tend to be robust across different models in the absence of constraints (e.g. conserved quantities/ideal invariants) that depend on the dimensionality of the system. For instance, a similar PDF of a self-organized shear flow was shown in low-order, 0- and 1-dimensional models and in 2-D fluid model [47], see also [20]. A low-dimensional deterministic, prey-predator model of the L-H transition [29] and its extension [30–34] have proven to be valuable in gaining the key insight into results from numerical simulations of more complicated models and experiments [28, 37, 44]. Furthermore, given the challenges in calculating time-dependent PDFs from plasma turbulence in regards to computational efforts and data analysis, it is valuable to consider a low-dimensional model to facilitate the calculation of an exact time-dependent PDF. Thus, we investigate a stochastic extension of [29] in the following.

A. Previous deterministic model [29]

For completeness, we briefly explain the prey-predator model [29] which is governed by

$$\frac{\partial \epsilon}{\partial t} = N\epsilon - a_1\epsilon^2 - a_2 V^2\epsilon - a_3 V_{ZF}^2\epsilon, \tag{5}$$

$$\frac{\partial V_{ZF}}{\partial t} = b_1\epsilon\frac{V_{ZF}}{1+b_2V^2} - b_3 V_{ZF}, \tag{6}$$

$$\frac{\partial N}{\partial t} = -c_1\epsilon N - c_2 N + Q. \tag{7}$$

Here, ϵ, V_{ZF} and N represent turbulence amplitude, zonal flow and density gradient, respectively; a_i, b_i and c_i are non-negative constants; Q is an external heating (input power); $V = dN^2$ is the mean flow with $d > 0$. In Eq. (5), the right side represents the linear growth of turbulence by the density gradient and turbulence damping due to turbulence nonlinear interaction, mean flow and zonal flow, respectively. The right side of Eq. (6) represents the zonal flow growth from turbulence, subject to the mean flow damping $(1+b_2V^2)$, and linear (collisional) damping. The right side of Eq. (7) represents the damping of the density gradient due to turbulence and neo-classical/collisional effects, and the density gradient growth due to an external heating Q. In [34], the periodic perturbation in Q was shown to help the L-H transition. Eqs. (5)-(7) support the L-H transition either with or without going through limit-cycle oscillations (I-phase) depending on parameter values and Q, dithering appearing for a slow ramping of $Q(t)$. We note that in this model [29], the H-mode is a quiescent H-mode where both zonal flow and turbulence are zero. As in [24, 52], we employ the following approximation of Eq. (7) as

$$N = \frac{Q}{c_1\epsilon + c_2}. \tag{8}$$

Note that different approximations of Eq. (8) were also used, e.g. such as $N = Q/c_2$ in [32] or the neglect of the mean flow in [53] to elucidate intermittency and bistability of zonal flows and geodesic acoustic modes, respectively. See §V for the discussion about $Q(t)$.

B. Stochastic model

The deterministic model in Eqs. (5)-(8) is a mean-field model where the overall (incoherent) contribution from unresolved scales is ignored, as noted in Section I. We incorporate such contribution by adding the two independent δ-correlated Gaussian stochastic noises ξ

and η in Eqs. (5) and (6) respectively [54] and consider the following stochastic equivalents of Eqs. (5)-(6):

$$\frac{dx}{dt} = f + \xi, \qquad f = \frac{1}{2}\Big[N - a_1 x^2 - a_2 V^2 - a_3 v^2\Big]x, \tag{9}$$

$$\frac{dv}{dt} = g + \eta, \qquad g = \frac{b_1 x^2 v}{1 + b_2 V^2} - b_3 v, \tag{10}$$

where $x = \pm\sqrt{\epsilon}$, $v = V_{ZF}$, and N is given by Eq. (8) (with $\epsilon = x^2$). The noise terms satisfy

$$\langle \xi(t)\xi(t')\rangle = 2D_x\delta(t-t'), \qquad \langle \eta(t)\eta(t')\rangle = 2D_v\delta(t-t'),$$
$$\langle \xi(t)\eta(t')\rangle = 0, \qquad \langle \xi\rangle = \langle \eta\rangle = 0. \tag{11}$$

Here, the angular brackets denote averages over ξ and η. D_x and D_v are the amplitudes of the stochastic noise ξ and η, affecting x and v respectively. It is worth noting that we work with x instead of ϵ in order to implement the boundary conditions that a PDF vanishes as $x \to \pm\infty$ and $v \to \pm\infty$, and avoid having numerical boundaries at $\epsilon = 0$ (see §III.C). In our stochastic model, turbulence is excited not only by the instability but also by the weak stochastic noise ξ (small D_x) as can be seen in Eq. (9). Thus, the instability/weak stochasticity causes turbulence, driving the zonal flows and the transition to the dithering.

C. Fokker-Planck equation

In order to obtain exact PDFs and to avoid handling noisy data, we solve the following Fokker-Planck equation [54] for a joint PDF $p = p(x, v, t)$ corresponding to Eqs. (10)-(11) (see Appendix C, [54, 55])

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}\Big[f(x,v)p\Big] - \frac{\partial}{\partial v}\Big[g(x,v)p\Big] + D_x\frac{\partial^2 p}{\partial x^2} + D_v\frac{\partial^2 p}{\partial v^2}, \tag{12}$$

instead of performing stochastic simulations of Eqs. (10)-(11) (e.g. [47]). As noted above, using x instead of ϵ enables us to use natural boundary conditions $x \in (-\infty, \infty)$ and $v \in (-\infty, \infty)$, and $p(x, v, t) \to 0$ as $x, v \to \pm\infty$. The numerical solution of (12) involves second-order finite-differencing, with grid spacings as small as 10^{-3} in both x and v. The time-stepping is second-order Runge-Kutta, with time-steps as small as $2 \cdot 10^{-5}$. In principle solutions could be investigated that are asymmetric in either x or v, but here we only consider symmetric solutions satisfying $p(x, v, t) = p(x, -v, t) = p(-x, v, t)$. This allows the computational domain to be restricted to a single quadrant of the (x, v) plane and

appropriate symmetry conditions imposed, with a fourfold savings in computer time. Taking a box size with $x_{max} = v_{max} = 2$ is sufficiently large to be a good approximation to $x, v \to \infty$; that is, the total probability $\iint p\, dx\, dv$ remains conserved within 10^{-4} or better for all runs presented here.

IV. STATISTICAL MEASURES

Before presenting the results in §V, we summarize different statistical measures here.

A. Information length diagnostics

For our model, we have the two variables $x_1 = x$ and $x_2 = v$. Thus, from $p(x, v, t)$ we calculate the marginal PDFs $p(x, t)$ and $p(v, t)$ as

$$p(x,t) = \int dv\, p(x,v,t), \quad p(v,t) = \int dx\, p(x,v,t). \tag{13}$$

From Eqs. (3)-(4), we then have

$$\mathcal{E}_x = \frac{1}{[\tau_x(t)]^2} = \int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t}\right]^2, \tag{14}$$

$$\mathcal{E}_v = \frac{1}{[\tau_v(t)]^2} = \int dv \frac{1}{p(v,t)} \left[\frac{\partial p(v,t)}{\partial t}\right]^2, \tag{15}$$

$$\mathcal{L}_x(t) = \int_0^t \frac{dt_1}{\tau_x(t_1)}, \;\; \mathcal{L}_v(t) = \int_0^t \frac{dt_1}{\tau_v(t_1)}. \tag{16}$$

We note that since Eqs. (3)-(4) and (14)-(16) are invariant under the independent change of variables, as noted previously (see also Appendix B), τ_x calculated from $p(x, t)$ is guaranteed to be the same as τ_ϵ calculated from $p(\epsilon, t)$ (where we recall ϵ and x are related by $\epsilon = x^2$).

We also define $\mathcal{L}$, $\mathcal{E}$ and τ from the joint PDF $p(x, v, t)$ directly as

$$\mathcal{L}(t) = \int_0^t \frac{dt_1}{\tau(t_1)}, \tag{17}$$

$$\mathcal{E} = \frac{1}{[\tau(t)]^2} = \int dx dv \frac{1}{p(x,v,t)} \left[\frac{\partial p(x,v,t)}{\partial t}\right]^2, \tag{18}$$

where the quantity without the subscript x or v denotes those calculated from the joint PDFs directly. For the two independent variables x and v with $p(x, v, t) = p(x, t)p(v, t)$, $\mathcal{L}_x + \mathcal{L}_v = \mathcal{L}$.

B. Entropy

As noted in §I, entropy measures disorder or the lack of information [50]. We define entropy S_x, S_v and S from the marginal PDFs $p(x,t)$ and $p(v,t)$ and the joint PDF $p(x,v,t)$ as follows

$$S_x = -\int dx\, p(x,t) \ln\left(p(x,t)\right), \tag{19}$$

$$S_v = -\int dv\, p(v,t) \ln\left(p(v,t)\right), \tag{20}$$

$$S = -\int dxdv\, p(x,v,t) \ln\left(p(x,v,t)\right). \tag{21}$$

Unlike the information length diagnostics, entropy changes under the coordinate transformation (e.g. under $\epsilon \to x$), while being independent of a linear translation (e.g. $x \to x +$ const). The latter property, in particular, makes entropy less useful when quantifying the movement of a PDF as happens in the L-H transition. Note that $S_x + S_v - S \geq 0$ represents the mutual entropy, the equality $S = S_x + S_v$ holding when $p(x,v,t) = p(x,t)p(v,t)$.

C. Fisher information

The Fisher information deals with a PDF and is a way of measuring the amount of information or order associated with a PDF [50] as noted in §I. Qualitatively, the narrower (broader) a PDF, the larger (smaller) the Fisher information and the smaller (larger) the entropy. We calculate Fisher information [50] F_{xx} and F_{vv} from the marginal PDFs $p(x,t)$ and $p(v,t)$

$$F_{xx} = 4\int dx\, q(x,t) \left[\frac{\partial q(x,t)}{\partial x}\right]^2, \tag{22}$$

$$F_{vv} = 4\int dv\, q(v,t) \left[\frac{\partial q(v,t)}{\partial v}\right]^2, \tag{23}$$

$$\tag{24}$$

where $q(x,t) = \sqrt{p(x,t)}$ and $q(v,t) = \sqrt{p(v,t)}$. Note that Eqs. (22)-(23) are given in terms of q – the square-root of a PDF – instead of a PDF itself to avoid the (superficial appearance of) singularity in the denominator. Like entropy, Fisher information changes under the coordinate transformation (e.g. under $\epsilon \to x$) while being independent of the linear translation (e.g. $x \to x +$ const), and is therefore again less useful to understand time-dependent PDFs.

D. Cross-correlation

The cross-correlation C_{xy} between fluctuating x and v, their normalised version c_{xy} and standard deviations σ_x and σ_v of x and v are defined as

$$C_{xv} = \langle (x - \langle x \rangle)(v - \langle v \rangle) \rangle = \int_0^\infty dx \int_0^\infty dv \, (x - \langle x \rangle)(v - \langle v \rangle) \, p(x, v, t), \tag{25}$$

$$c_{xv} = \frac{C_{xy}}{\sigma_x \sigma_v}, \tag{26}$$

$$\sigma_x^2 = \langle (x - \langle x \rangle)^2 \rangle = \int_0^\infty dx \int_0^\infty dv \, (x - \langle x \rangle)^2 \, p(x, v, t), \tag{27}$$

$$\sigma_v^2 = \langle (v - \langle v \rangle)^2 \rangle = \int_0^\infty dx \int_0^\infty dv \, (v - \langle v \rangle)^2 \, p(x, v, t), \tag{28}$$

where $\langle x \rangle = \int dx dv \, x p(x, v, t)$ and $\langle v \rangle = \int dx dv \, v p(x, v, t)$. Note that the average $\langle \cdot \rangle$ refers to the mean value over the first quadrant $x, v > 0$ only, that is, $\langle f \rangle \equiv \int_0^\infty \int_0^\infty f \, p \, dx \, dv$.

V. RESULTS

Our key interest in this paper is not on the exploration of all possible cases but on demonstrating the consequences of different forms of $Q(t)$ (fast vs slow ramping) on time-dependent PDFs and information geometry in the forward and backward transitions and resulting hysteresis. Here, hysteresis refers to a lag between the input and output in our system upon the reversal of time direction (forward vs backward processes). To this end, we increase $Q(t)$ linearly at a certain rate to induce the transition to dithering up to time $t = t_*$, and then decrease $Q(t)$ to simulate the back transition to the L-mode. $Q(t)$ is chosen to be symmetric around t_* so that for $t > t_*$, $Q(t) = Q(2t_* - t)$; the forward and backward processes refer to the phase where $Q(t)$ increases and decreases, respectively. The different evolution in the forward and backward processes despite the symmetry in $Q(t) = Q(2t_* - t)$ is called hysteresis in this paper. There has been the experimental observation of hysteresis between the forward (L-mode to dithering) and the backward (dithering to L-mode) processes reported in [44].

We note that our $Q(t)$ is not chosen to represent an input power ramping in other experiments, but to demonstrate how our method works in general and inform us of their effects on time evolution and statistical properties of the transitions. For instance, our method works for the cases of different heating scenarios (e.g. [44–46]). In particular, while the

characteristic time scale for coupled auxiliary power to flow from the core to edge is on the order of the global energy confinement time, the local power balance in the edge that is relevant to the evolution of limit cycles can be governed by intermittent turbulence dynamics and avalanches that occur on turbulence time scales $t \sim a/c_s \leq 100\mu$s. Here, a is the minor plasma radius and c_s is the local ion sound speed. These nonlinear events occur predominantly in the edge due to the high saturated levels of edge turbulence. Also, plasmas can be strongly driven out of equilibrium, e.g. due to a strong (overpowered) beam pulse and then backing off the power [44], with hysteresis between forward and backward processes. A reduction of the limit cycle frequency has been experimentally observed with higher input power. These experimental observations are reproduced in our model as discussed below.

Furthermore, given a time-dependent problem, we do not concern ourselves with bifurcation analysis [30], which would have only a limited validity for a time-dependent $Q(t)$. Specifically, we consider the following three different cases

- Set 1: $Q = 0.03t + 0.1$, ramping back down at $t = t_* = 40$;
- Set 2: $Q = 0.05t$, ramping back down at $t = t_* = 20$;
- Set 3: $Q = 0.1t$, ramping back down at $t = t_* = 12$.

We note that for Set 1, we choose $Q(t = 0) = 0.1$ instead of $Q(t = 0) = 0$ so that the system evolves to the dithering phase quicker. That is, $Q(t = 0) = 0$ would just have lengthened the duration of the L-mode without much effect on the subsequent evolution, the system spending a lot of time in the L-mode. In comparison, for the faster ramping for Sets 2-3, the system enters the dithering phase quicker without costing much extra computational efforts.

We use the same parameter values $a_1 = 0.2$, $a_2 = a_3 = 0.7$, $b_1 = 1.5$, $b_2 = b_3 = 1$, $c_1 = 1$, $c_2 = 0.5$, and $d = 1$ as in [29]. For the results presented here, we use an initial condition $p(x, v, 0) \propto \exp[-((|x| - 0.5)^2 - v^2)/5 \cdot 10^{-3}]$; other initial conditions with small values of x and v yielded similar results. For the noise terms D_x and D_v, we considered different combinations of $D_x = [1, 2, 4, 8, 16, 32] \cdot 10^{-4}$ and $D_v = [1, 2] \times 10^{-4}$. Since results did not change much for different D_v, in the following, we present results for $D_x = [1, 4, 16] \cdot 10^{-4}$ and $D_v = 10^{-4}$ only. In [24], we investigated the forward process up to time $t = 50$ by using $Q = 0.03t + 0.1$ (the same as Set 1 here) which covers the transition to the dithering phase and then the approach to the H-mode. Here we take Set 1 only up to $t = t_* = 40$ before

ramping $Q(t)$ back down again, thereby switching back to the L-mode. If Q becomes too large before the ramping down is started, the system does not revert to the L-mode, but instead continues evolving toward the H-mode ($x, v \to 0$). Sets 2 and 3, where Q is increased more quickly, are then also ramped down sooner, so that all three sets have broadly similar $Q_{\max}$ values. We are then interested in comparing the forward and backward processes under different ramping speeds, and exploring the different statistical measures. It is important to note that for the time duration of forward and backward processes in this paper, turbulence regulation is mainly by zonal flows, which will be our focus in this paper.

A. Set 1: $Q = 0.03t + 0.1$ and $t_* = 40$

1. *Mean, standard deviation, cross-correlation and phase portrait*

Fig. 1 shows the average quantities $\langle x \rangle$, $\langle v \rangle$ (column 1), the standard deviations $\sigma_x = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$ and $\sigma_v = \sqrt{\langle (v - \langle v \rangle)^2 \rangle}$ (column 2), the (normalized) cross-correlation $\langle (x - \langle x \rangle)(v - \langle v \rangle) \rangle / (\sigma_x \sigma_v)$ (column 3), and the phase-portrait $\langle x \rangle$ against $\langle v \rangle$ (column 4). We note that the average $\langle \cdot \rangle$ refers to the mean value over the first quadrant $x, v > 0$ only, that is, $\langle f \rangle \equiv \int_0^2 \int_0^2 f\, p\, dx\, dv$ since the computational domain is $x, v = [0, 2]$. To distinguish the forward and backward processes, we use blue and red colors to mark the forward ($t = [0, 40]$) and backward ($t = [40, 80]$) processes, respectively, in the time histories (for turbulence x) and in the phase portrait. Note that the same black line is used for zonal flows v in the time history. If the backward process had undergone the same evolution as the forward process, the behavior for $t = [40, 80]$ would be a mirror image of the behavior for $t = [0, 40]$. This is clearly not the case, indicating that the forward and backward processes are substantially different from one another.

In column 1, the rapid growth of $\langle v \rangle$ from $\langle x \rangle$ is seen up to $t \approx 11$ for all D_x, followed by the dithering I-phase where $\langle x \rangle$ and $\langle v \rangle$ oscillate around each other. The smaller D_x, the more prominent these oscillations are. Note that if the ramping-up had been continued for $t > t_*$, the dithering phase would have ended when $\langle x \rangle$ and $\langle v \rangle$ both start collapsing back towards zero, which is the H-mode. As noted above though, if the system is allowed to evolve all the way to that point, then ramping Q down again does not cause the system to revert to the L-mode, but instead remains in the H-mode. If instead Q is ramped down

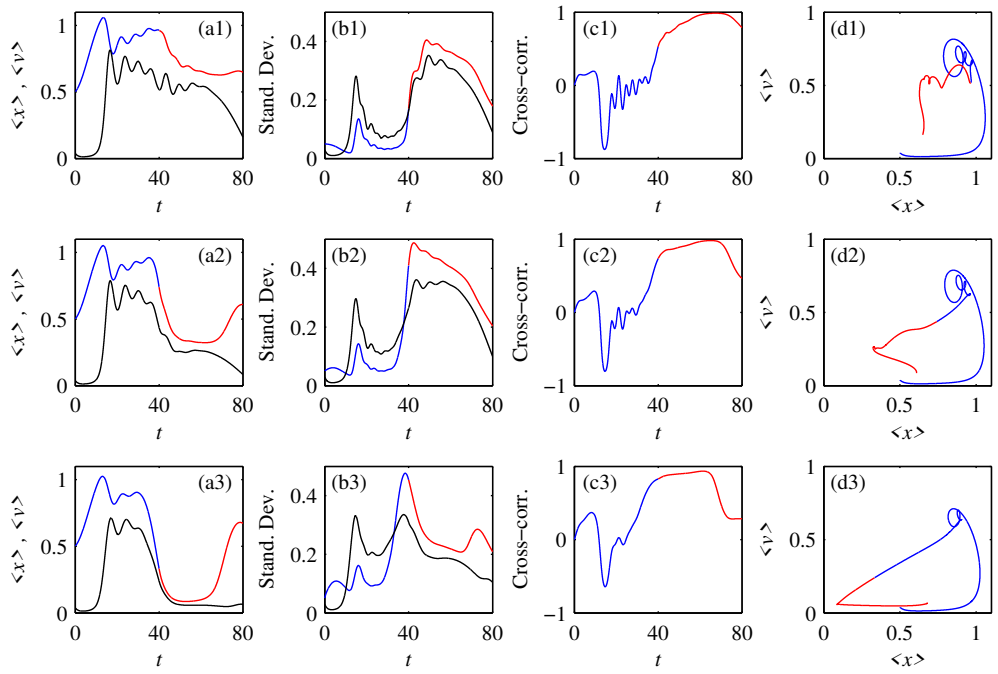

FIG. 1: From top to bottom, the three rows (1,2,3) show results for Set 1 at $D_x = 10^{-4}$, $4 \cdot 10^{-4}$ and $16 \cdot 10^{-4}$. Column 1 (a1,a2,a3) shows $\langle x \rangle$ and $\langle v \rangle$ as functions of time, with $\langle x \rangle$ in blue/red (with the switch in color at the switch time $t_* = 40$) and $\langle v \rangle$ in black. Column 2 (b1,b2,b3) shows the associated standard deviations σ_x and σ_v, with the same color-coding. Column 3 (c1,c2,c3) shows the cross-correlation $c_{xv} = \langle (x - \langle x \rangle)(v - \langle v \rangle) \rangle / (\sigma_x \sigma_v)$. Column 4 (d1,d2,d3) shows the phase portrait in the $(\langle x \rangle, \langle v \rangle)$ plane, with blue again corresponding to $Q(t)$ increasing and red to $Q(t)$ decreasing.

before the transition to the H-mode is complete, then the ramp-down during $t = [40, 80]$ induces the transition to the L-mode as $\langle x \rangle$ and $\langle v \rangle$ evolve towards $\langle x \rangle \sim 0.5$ and $\langle v \rangle \sim 0$, albeit via non-monotonic fluctuations.

Specifically, for $D_x = 10^{-4}$, $\langle x \rangle$ undergoes some weak oscillations with an overall decrease before increasing again, while $\langle v \rangle$ exhibits mostly oscillatory behavior up to $t \sim 60$ before collapsing rapidly. The backward evolution for $t = [40, 80]$ is clearly not the mirror image of the forward evolution $t = [0, 40]$, showing the disparity between the forward and backward transitions (e.g. $\langle x \rangle(t = 80) \neq \langle x \rangle(t = 0)$, $\langle v \rangle(t = 80) \neq \langle v \rangle(t = 0)$). In fact, the evolution of $\langle v \rangle$ for the entire $t = [0, 80]$ seems reminiscent of the L-H transition despite the power ramping down for $t = [40, 80]$ with the extended region of the dithering. This is probably

because of the inertia of x and v which requires a sufficient time to respond to the change introduced by a time-varying input power.

A stochasticity D_x introduces irreversibility. Thus, the larger D_x is, the larger the disparity between forward and backward processes. Specifically, as D_x increases to $D_x = [4, 16] \times 10^{-4}$, the backward evolution of $\langle x \rangle$ and $\langle v \rangle$ becomes more different from their forward evolution, with their fast decay around $t \sim 40$. Note that D_x was shown to help entering the H-mode at earlier time, and thus at smaller power Q ($= 0.03t + 0.1$) in [24]. That is, a larger stochastic noise helps the L-H transition by reducing the required power threshold. We note that $\langle x \rangle$ increases during the backward process as the system approaches the L-mode where turbulence has a finite amplitude.

The standard deviation σ_x and σ_v in column 2 in Fig. 1 shows the large increase in σ_x and σ_v at the beginning of the dithering phase around $t \sim 20$, and an even larger increase around $t \sim 50$, $t \sim 42$ and $t \sim 38$ for $D_x = [1, 4, 16] \times 10^{-4}$, respectively. The larger D_x is, the earlier the second peak of σ_x and σ_v appears. Another prominent feature is that σ_x and σ_v tend to take larger values in the backward process than in the forward process. Also, the values of σ_x and σ_v can exceed those of $\langle x \rangle$ and $\langle v \rangle$ in the backward process, implying strongly out-of-equilibrium dynamics. Furthermore, σ_x tends to be larger than σ_v in the backward process, naively suggesting broader PDFs of x. However, in the case of a multiple-peak PDF, the meaning of PDF width is unclear, and a careful examination of the form of the actual PDFs is needed to understand the evolution. In fact, as shown later, these behaviors are due to the formation of multi-peak PDFs in the dithering phase and their persistent evolution in the backward process.

Fig. 1 column 3 shows that the self-regulation in the dithering phase in the forward process is reflected by the negative sign of the cross-correlation σ_{xv} starting around $t \approx 15$, when fluctuating x and v alternate in sign. In comparison, the sign of σ_{xv} is always positive in the backward process. It is thus tempting to conclude the absence of self-regulation in the backward process. However, as noted above, the literal interpretation of the sign of cross-correlation requires some caution for broad, non-Gaussian PDFs (especially ones with multiple peaks). This is because the mean values for multiple peak PDFs do not represent the mostly likely values. Consequently, fluctuations which measures the deviation from the mean values can also be misleading and thus cross-phase can be misleading for multi-peak PDFs (see below).

The phase-portrait $\langle v \rangle$ against $\langle x \rangle$ in Fig. 1 column 4 shows the asymmetry between the forward (in blue) and backward (in red) processes. In particular, the forward process starting with $\langle x \rangle = 0.5$ and $\langle v \rangle = 0$ undergoes the increases in $\langle x \rangle$ and $\langle v \rangle$ before showing the circular trajectory during the dithering; the backward process shows very different evolution, with almost no overlap with the forward process trajectory, and ends at $\langle x \rangle(t = 80) \neq \langle x \rangle(t = 0)$ and $\langle v \rangle(t = 80) \neq \langle v \rangle(t = 0)$. The disparity between the forward and backward processes increases as D_x increases.

2. Information length diagnostics

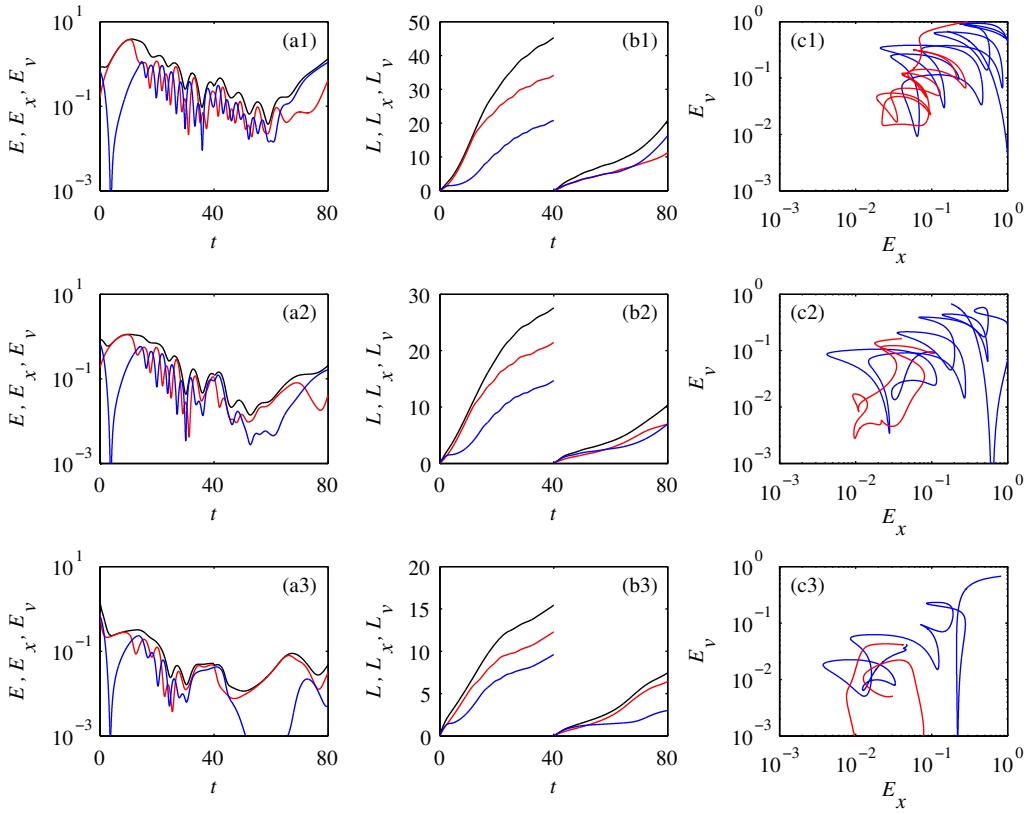

FIG. 2: The various information length diagnostics associated with the results in Fig. 1. Column 1 (a1,a2,a3) shows $\mathcal{E}$ (black), $\mathcal{E}_x$ (red), and $\mathcal{E}_v$ (blue) as functions of time. Column 2 (b1,b2,b3) shows $\mathcal{L}$ (black), $\mathcal{L}_x$ (red), and $\mathcal{E}_v$ (blue); note how these quantities are reset to zero at the start of the Q-decreasing phase. Column 3 (c1,c2,c3) shows the phase portrait in the $(\mathcal{E}_x, \mathcal{E}_v)$ plane, with blue corresponding to $Q(t)$ increasing and red to $Q(t)$ decreasing.

The information diagnostics calculated from the time-dependent PDFs are shown in Fig. 2. As in Fig. 1, the three rows are for the different values of D_x, increasing from the top to bottom as $D_x = [1, 4, 16] \times 10^{-4}$. Column 1 shows $\mathcal{E}$ (black), $\mathcal{E}_x$ (red), and $\mathcal{E}_v$ (blue) as functions of time. We observe that $\mathcal{E}$ calculated from the joint PDF $p(x, v, t)$ is always larger than $\mathcal{E}_x$ and $\mathcal{E}_v$ calculated from the marginal PDFs $p(x, t)$ and $p(v, t)$. This is because the averaging over x or v reduces the information content in the marginal PDFs. ($\mathcal{E} - (\mathcal{E}_x + \mathcal{E}_v)$ has a rather complicated time evolution, with a larger fluctuation in the forward process than in the backward process.) For the same reason, the values of $\mathcal{E}$, $\mathcal{E}_x$ and $\mathcal{E}_v$ decrease as D_x increases from the top to the bottom row.

In the first column in Fig. 2, the appearance of the maxima in $\mathcal{E}_x$ (red) and $\mathcal{E}_v$ (blue) forecast the transition to the dithering phase earlier (better) than corresponding mean values in Fig. 1; the maximum in $\mathcal{E}_x$ occurs at earlier times (e.g. $t \approx 10.5$ for $D_x = 10^{-4}$, $t \approx 9.5$ for $D_x = 16 \cdot 10^{-4}$) than $\langle x \rangle$ (at $t \approx 13.5$). This is followed by a series of oscillations during dithering when $\mathcal{E}_x$ compete and oscillate around each other, alternating in which is larger. The crossing between $\mathcal{E}_x$ (red) and $\mathcal{E}_v$ (blue) signifies the matching of the time scales of $p(x, t)$ and $p(v, t)$ ($\tau_x = \tau_v$). $\tau_x = \tau_v$ can be viewed as a resonance in the statistical state and signifies a strong correlation between x and v. The crossing between $\mathcal{E}_x$ and $\mathcal{E}_v$ extends up to around $t \sim 60$ for $D_x = 10^{-4}$, suggesting that the self-regulation between zonal flows and turbulence continues in the backward process for a while although such behavior is difficult to infer from the phase portraits in Fig. 1.

Time matching is clearly seen in the information phase portrait of τ_v against τ_x in column 3 where $\mathcal{E}_x$ and $\mathcal{E}_v$ oscillate around a straight line $\mathcal{E}_x = \mathcal{E}_v$ ($\tau_x = \tau_v$) not only in the forward process (in blue) but also in the backward process (in red). The oscillation of $\mathcal{E}_x$ and $\mathcal{E}_v$ around $\mathcal{E}_x = \mathcal{E}_v$ represents the competition between x and v, with the larger of $\mathcal{E}_x$ and $\mathcal{E}_v$ dominating the competition. As D_x increases, the number of crossings between the two decreases.

For the information length shown in column 2, we reset its value to zero at the start of the backward process so that we can compare the total cumulative information change in the forward and backward processes. $\mathcal{L}$ (black) calculated from the joint PDF $p(x, v, t)$ is always larger than $\mathcal{L}_x$ (red), and $\mathcal{L}_v$ (blue) since $\mathcal{E} > \mathcal{E}_x, \mathcal{E}_v$. $\mathcal{L}$ resembles the behavior of the greater of $\mathcal{L}_x$ and $\mathcal{L}_v$. Compared with $\mathcal{E}$ ($\mathcal{E}_x$ and $\mathcal{E}_v$), $\mathcal{L}$ ($\mathcal{L}_x$ and $\mathcal{L}_v$) shows much smoother time evolution, with the tendency of a slower increase during dithering. In fact,

for $D_x = 10^{-4}$ (top row), $\mathcal{L}_x \sim \mathcal{L}_v$ during the dithering can be seen at $t \sim (40, 65)$ while the slope of $\mathcal{L}_x$ is similar to $\mathcal{L}_v$ at $t \sim (20, 40)$. That is, the self-regulation between x and v synchronizes and reduces the rate at which the information length of x and v increases.

3. *Entropy and Fisher information*

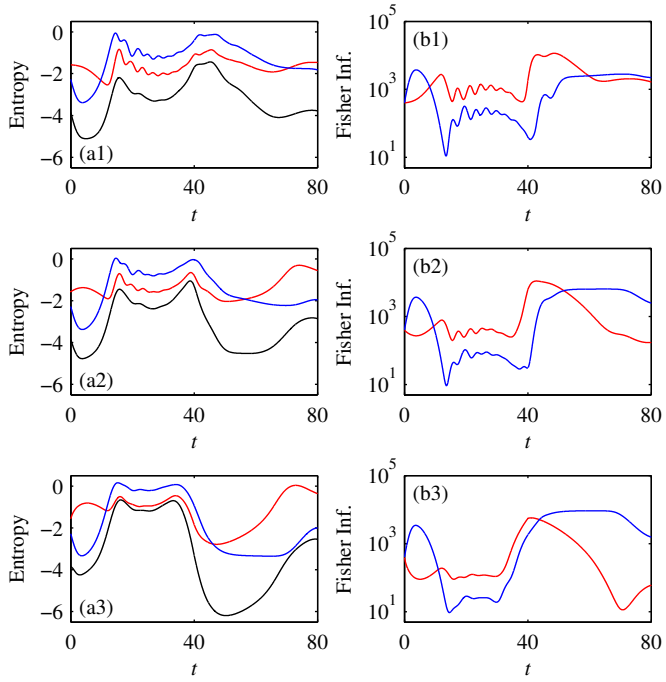

FIG. 3: Entropy and Fisher information associated with the results in Figs. 1 and 2. Column 1 (a1,a2,a3) shows the entropies S (black), S_x (red), and S_v (blue) as functions of time. Column 2 (b1,b2,b3) shows the Fisher information F_{xx} associated with the marginal PDF $p(x,t)$ in red, and the Fisher information F_{vv} associated with the marginal PDF $p(v,t)$ in blue. As in Figs. 1 and 2, the three rows (1,2,3) correspond to $D_x = 10^{-4}$, $4 \cdot 10^{-4}$ and $16 \cdot 10^{-4}$.

As noted in §IV, entropy and Fisher information do not detect the movement of a PDF. As a result, entropy or Fisher information is less informative compared with the information length diagnostics discussed above. To demonstrate this, the results corresponding to Figs. 1 and 2 are shown in Fig. 3. Specifically, column 1 in Fig. 3 shows the evolution of the entropies S (black), S_x (red), and S_v (blue). We note first that S from the joint PDF $p(x,v,t)$ is always smaller than S_x and S_v calculated from the marginal PDFs. This is because the averaging

needed for calculating marginal PDFs increases uncertainty and thus entropy. The local maxima of S, S_x and S_v appear to coincide with those of σ_x and σ_v in Fig. 1, although the ordering between S_x and S_v does not follow that between σ_x and σ_v. What is clear though is that the ordering between S_x and S_v is opposite to that between Fisher information F_{xx} (in red) and F_{vv} (blue) shown in column 2. This is because the Fisher information increases when the uncertainty decreases (e.g. when a PDF becomes narrow) while entropy decreases. Also, comparing with Fig. 2, we notice that the dithering appears only in the forward process and that the effect of different D_x is much less noticeable in Fig. 3. That is, the information length diagnostics are more sensitive in capturing the dynamics of the system (e.g. self-regulation/dithering, the effect of different parameters).

4. Joint PDFs

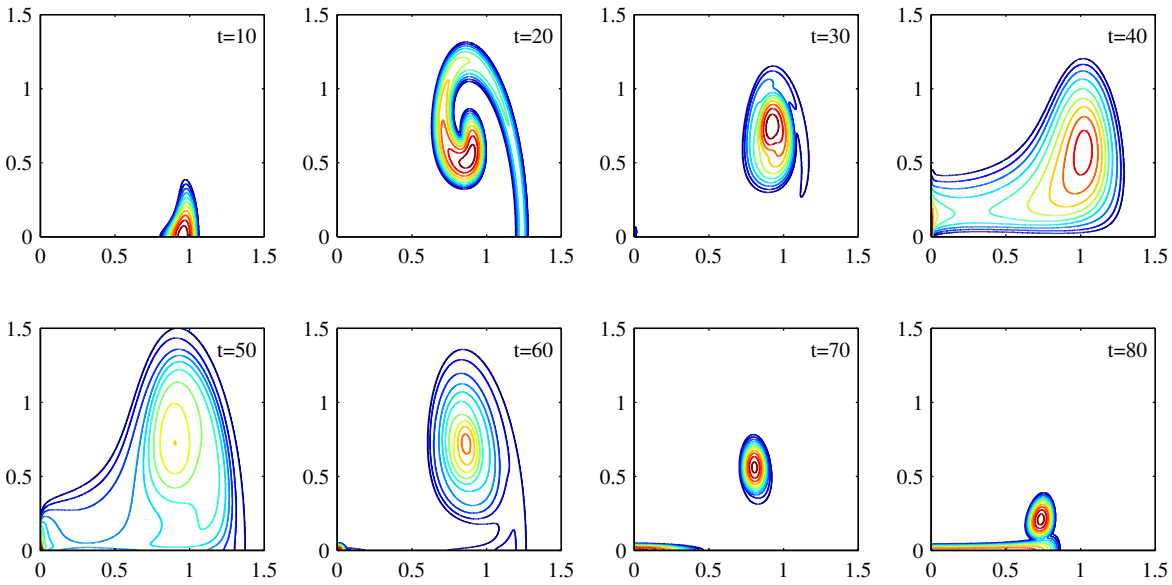

FIG. 4: Contour plots of the $D_x = 10^{-4}$ joint PDFs $p(x, v, t)$ at the times labelled in each panel, x on the horizontal axis, and v on the vertical axis. The top row corresponds to the Q-increasing phase, and the bottom row to the Q-decreasing phase. Contour intervals are on a logarithmic scale, with the smallest value at 10^{-3}, then $10^{-2.5}$, 10^{-2}, etc.

Fig. 4 shows the snapshot of joint PDFs $p(x, v, t)$ in the (x, v) plane, taken at times $t = 10, 20, 30, ...80$ for $D_x = 10^{-4}$. While the peak of the PDF (in red) tends to follow the evolution of $\langle x \rangle$ and $\langle v \rangle$ in Fig. 1, the joint PDF actually exhibits a quite complicated

evolution with convoluted structures. First of all, it is immediately clear that the PDF is (strongly) non-Gaussian and develops multiple peaks during its evolution. Note that since our results are shown only for $x, v \geq 0$, PDFs actually have multiple (more than two) peaks in $x, v = (-\infty, \infty)$. Specifically, around $t = 40$, the second peak appears around $x = v = 0$; up to $t = 60$, the original peak remains mainly in the same position while the second peak around $x = 0$ gradually becomes stronger. At $t = 70$, the original peak appears to move towards $v = 0$ in the v direction while it looks as if the two peaks got separated in the x direction due to the development of a deep valley between them. The final PDF at $t = 80$ is seen to be stretched in x.

This complicated evolution of the PDF involving its overall movement and the change in the number of PDF peaks highlight the significant limitations in using conventional methods such as just mean value, standard deviation, entropy, etc. to understand dynamical changes in general. In this regards, it is useful to recall again that for a PDF with more than one peak, mean value does not give us the mostly likely value, and similarly standard deviation does not give the PDF width, and their interpretation therefore requires caution. We note also that in our previous work [24], the final collapse to $x, v \to 0$ did not proceed by a simple motion of the peak toward the origin; instead, a secondary peak developing at the origin overtakes the original peak which remains largely in the same position.

5. Marginal PDFs

The marginal PDFs $p(x, t)$ and $p(v, t)$, including the case that corresponds to Fig. 4 ($D_x = 10^{-4}$), are shown in Figs. 5 and 6, respectively; the three different values of $D_x = [1, 4, 16] \times 10^{-4}$ are denoted by the black, blue and red curves in each panel. Prominent features are strongly non-Gaussian, asymmetric PDFs, multiple peaks being observed at $t = 20$ in both figures. For instance, the formation of a peak around the origin is seen for $D_x = 16 \times 10^{-4}$ (red) at $t = 30$ already, while this happens later at $t = 40$ for $D_x = 10^{-4}$. This is consistent with the previous results [24] where the stochasticity promotes the transition. Of interest is that $p(v, t)$ develops a more pronounced right tail around its main peak around $v \sim 0.5$, even with the appearance of a local maximum around $v \sim 1.2$ at $t = 20$. This means that rare, strong zonal flows v associated with this stretched right tail $v > 0.5$ play a key role in regulating turbulence at the initial stage of the dithering. Also, the right tail of

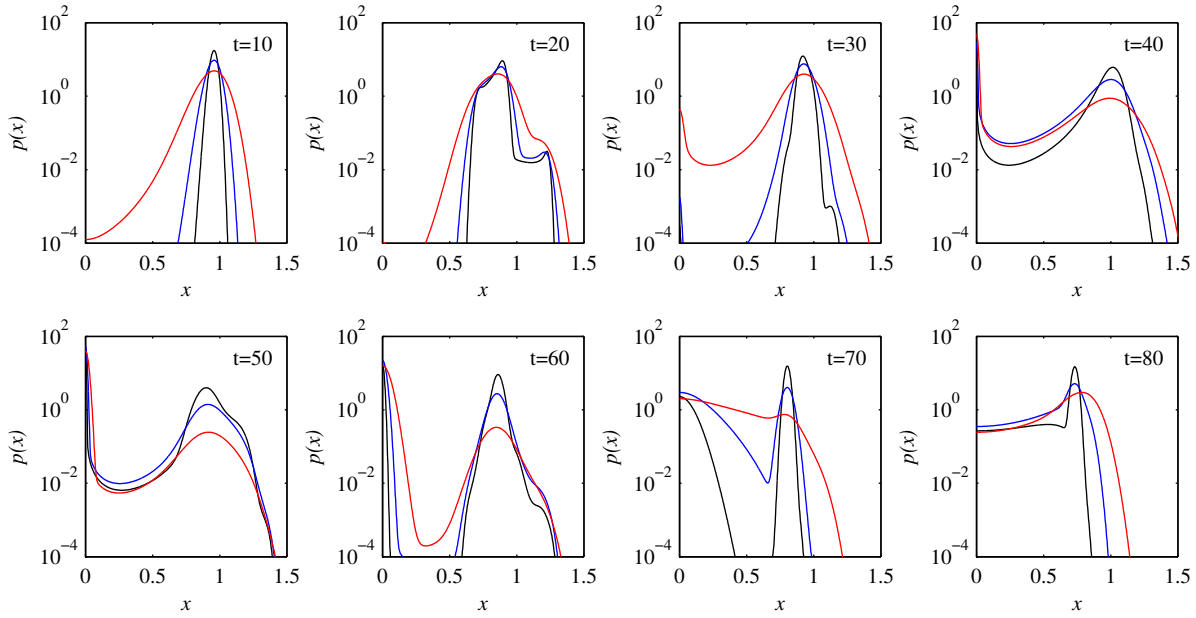

FIG. 5: The black curves in each panel show the marginal PDFs $p(x,t)$ for $D_x = 10^{-4}$, corresponding to the joint PDFs shown in Fig. 4. The blue curves show equivalent results at $D_x = 4 \cdot 10^{-4}$, and red at $D_x = 16 \cdot 10^{-4}$.

$p(v,t)$ is more pronounced than that of $p(x,t)$, highlighting the importance of intermittent zonal flows. That is, intermittency, which is often thought to cause anomalous transport, can actually reduce transport due to the enhanced population of a strong zonal flows (reflected in the right tail of $p(v,t)$).

B. Set 2: $Q = 0.05t$ and $t_* = 20$

We now consider the case of a faster ramping as $Q = 0.05t$ compared with Set 1 above. Because of a shorter time for the system to adjust to the change in Q, the system is driven further from equilibrium, as experimentally observed in [44]. Specifically, this results in the reduction in the dithering phase as well as in the difference across results for different D_x, as shown below. (As noted above, higher ramp rates slowing the cycle was shown in [44].) Figs. 7-12 are equivalent to Figs. 1-6 with the only difference in $Q = 0.05t$ and $t_* = 20$ instead of $Q = 0.03t + 0.1$ and $t_* = 40$. Qualitatively, quite similar behaviors are observed, and thus, the following focuses on pointing out the main differences between Figs. 7-12 and Figs. 1-6.

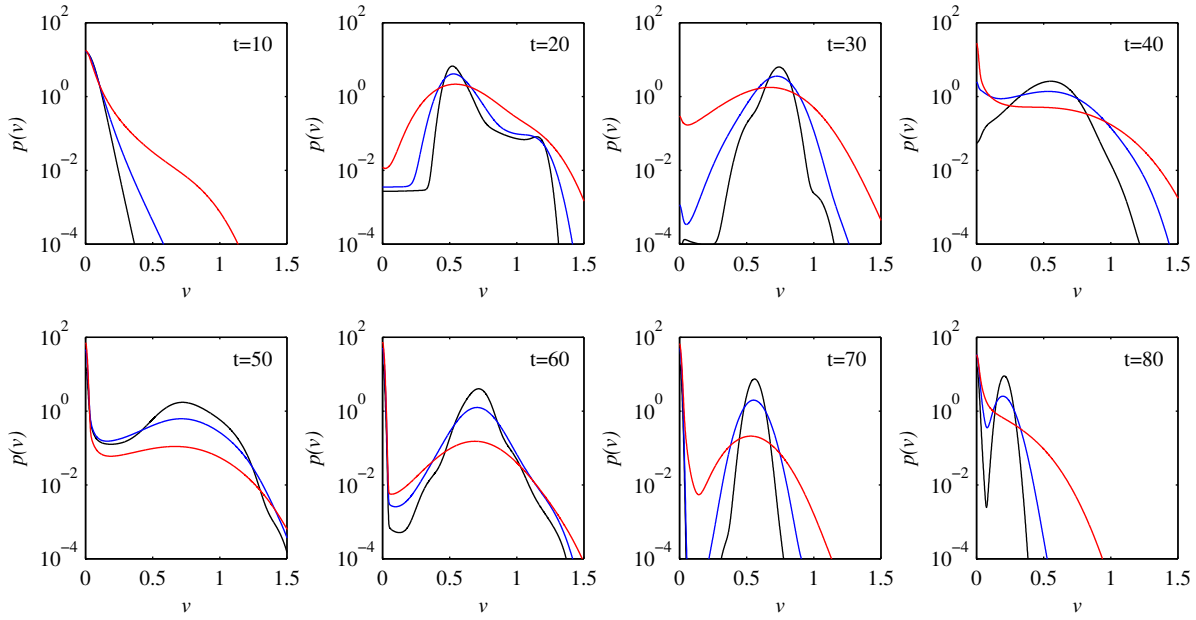

FIG. 6: As in Fig. 5, but now showing the marginal PDFs $p(v,t)$ rather than $p(x,t)$. The black curves is for $D_x = 10^{-4}$, again corresponding to the joint PDFs in Fig. 4; blue denotes $D_x = 4 \cdot 10^{-4}$ and red $D_x = 16 \cdot 10^{-4}$.

1. Mean, standard deviation, cross-correlation and phase portrait

In comparison with Fig. 1, the dithering phase is more clearly seen in the backward process in Fig. 7. Also, the effect of different D_x seems much less pronounced, with quite similar evolutions in the different rows ($D_x = [1, 4, 16] \times 10^{-4}$); the standard deviation (column 2) exhibits less difference between the forward and backward processes.

2. Information length diagnostics

In Fig. 8, we again observe that the oscillations between $\mathcal{E}_x$ and $\mathcal{E}_v$ are more pronounced than those between $\langle x \rangle$ and $\langle v \rangle$ in Fig. 7. The rapid power ramp-down leads to the overall reduction in the values of the information length diagnostics in comparison with Fig. 2 as a result of being strongly driven out of equilibrium. Of note is that $\mathcal{L}_x$ (red) tends to be larger than $\mathcal{L}_v$ (blue) in the forward process, while the opposite holds in the backward process (see Fig. 2). This can imply that the forward process is mainly driven by turbulence x while the backward process by zonal flow v. Although the time-histories of $\langle x \rangle$ and $\langle v \rangle$ in Fig. 7 appear quite similar for different D_x, the information length $\mathcal{L}$, $\mathcal{L}_x$ and $\mathcal{L}_v$ in Fig. 8 show

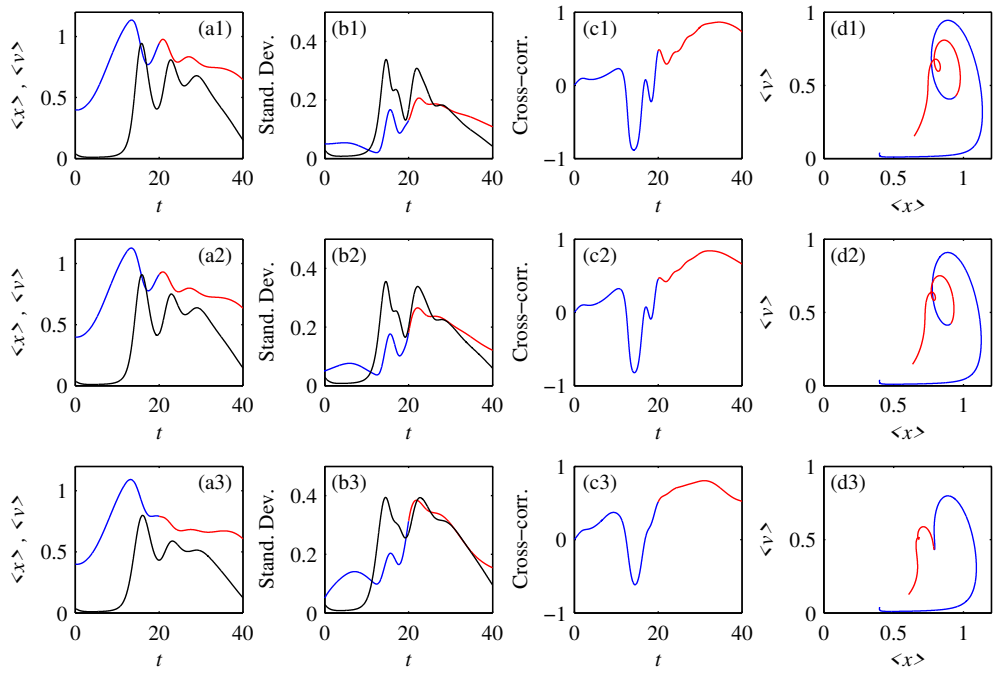

FIG. 7: From top to bottom, the three rows (1,2,3) show results for Set 2 at $D_x = 10^{-4}$, $4 \cdot 10^{-4}$ and $16 \cdot 10^{-4}$. Column 1 (a1,a2,a3) shows $\langle x \rangle$ and $\langle v \rangle$ as functions of time, with $\langle x \rangle$ in blue/red (with the switch in color at $t_* = 20$) and $\langle v \rangle$ in black. Column 2 (b1,b2,b3) shows the associated standard deviations σ_x and σ_v, with the same color-coding. Column 3 (c1,c2,c3) shows the cross-correlation $\langle (x - \langle x \rangle)(v - \langle v \rangle) \rangle / (\sigma_x \sigma_v)$. Column 4 (d1,d2,d3) shows the phase portrait in the $(\langle x \rangle, \langle v \rangle)$ plane, with blue again corresponding to $Q(t)$ increasing and red to $Q(t)$ decreasing.

clear differences, for instance, their overall values decreasing for a larger D_x.

3. Entropy and Fisher information

In comparison with Fig. 3, the forward and backward processes in Fig. 9 are more symmetric around $t = t_* = 20$. In particular, the two local maxima in entropies appear much closer with each other near $t_* = 20$ in Fig. 9. Also, the difference across different rows is less pronounced than in Fig. 3.

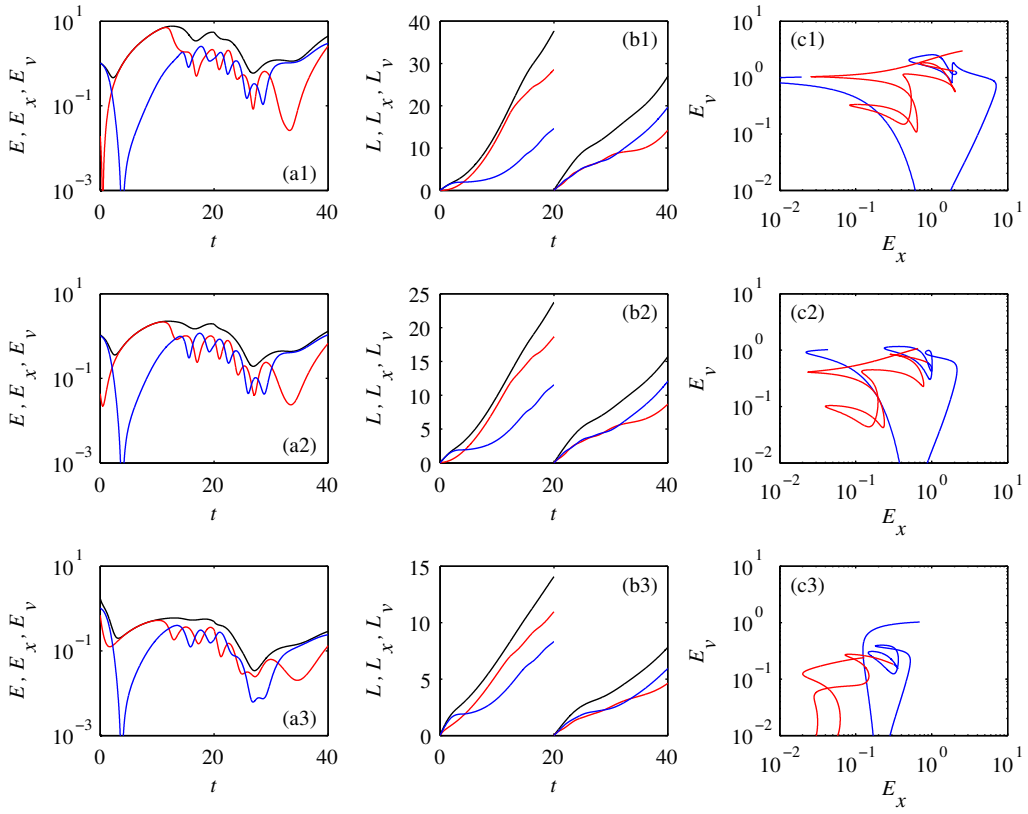

FIG. 8: The various information length diagnostics associated with the results in Fig. 7. Column 1 (a1,a2,a3) shows $\mathcal{E}$ (black), $\mathcal{E}_x$ (red), and $\mathcal{E}_v$ (blue) as functions of time. Column 2 (b1,b2,b3) shows $\mathcal{L}$ (black), $\mathcal{L}_x$ (red), and $\mathcal{E}_v$ (blue); note how these quantities are reset to zero at the start of the Q-decreasing phase. Column 3 (c1,c2,c3) shows the phase portrait in the $(\mathcal{E}_x, \mathcal{E}_v)$ plane, with blue corresponding to $Q(t)$ increasing and red to $Q(t)$ decreasing.

4. Joint PDFs

Fig. 10 uses the same color coding as that in Fig. 4. Thus, in comparison with Fig. 4, we can conclude that $p(x, v, t)$ in Fig. 10 contains more convoluted structures in space and is less localized. In particular, a curly structure that appears at $t = 15$ persists through the backward process without showing the breakup into the two peaks. This extended structure manifests higher uncertainty in predicting x and v values and can explain the overall smaller values in Fig. 8 in comparison with Fig. 2.

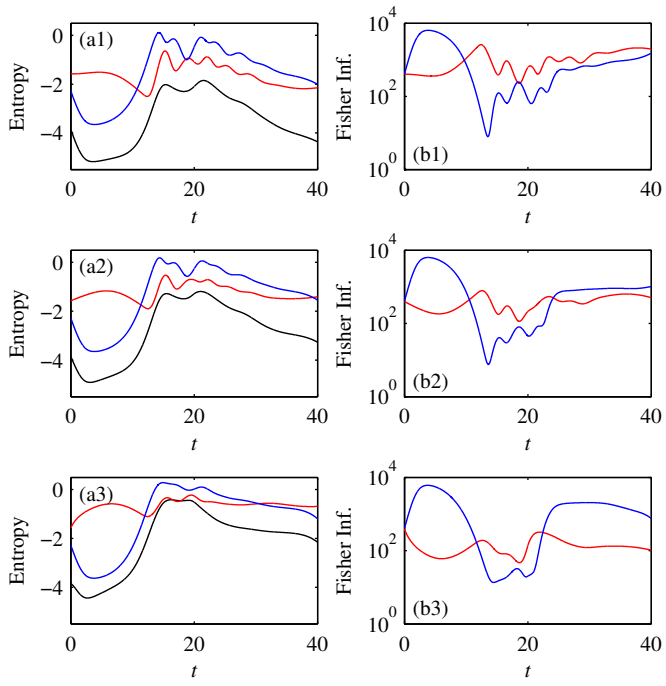

FIG. 9: Entropy and Fisher information diagnostics associated with the results in Figs. 7 and 8. Column 1 (a1,a2,a3) shows the entropies S (black), S_x (red), and S_v (blue) as functions of time. Column 2 (b1,b2,b3) shows the Fisher information F_{xx} associated with the marginal PDF $p(x,t)$ in red, and the Fisher information F_{vv} associated with the marginal PDF $p(v,t)$ in blue. As in Figs. 7 and 8, the three rows (1,2,3) correspond to $D_x = 10^{-4}$, $4 \cdot 10^{-4}$ and $16 \cdot 10^{-4}$.

5. Marginal PDFs

The marginal PDFs $p(x,t)$ in Fig. 11 and $p(v,t)$ in Fig. 12 show in more detail the extended structures of the PDFs noted above. In particular, at $t = 15$, $D_x = 10^{-4}$ (black) shows the three local maxima of $p(x,t)$ at $x \neq 0$ in Fig. 11, implying the total six peaks of $p(x,t)$ in the entire $x = [-2, 2]$ range. On the other hand, at $t = 15$, $D_x = 10^{-4}$ (black) shows an almost flat top $p(v,t)$ in Fig. 12.

C. $Q = 0.1t$ and $t = t_* = 12$.

When we increase the ramping rate further to $Q = 0.1t$, we observe that the tendency of less difference between the forward and backward processes (noted in §V.B) persists. As examples, here we show only the three Figs. 13, 14 and 15 that correspond to Figs. 7, 8 and

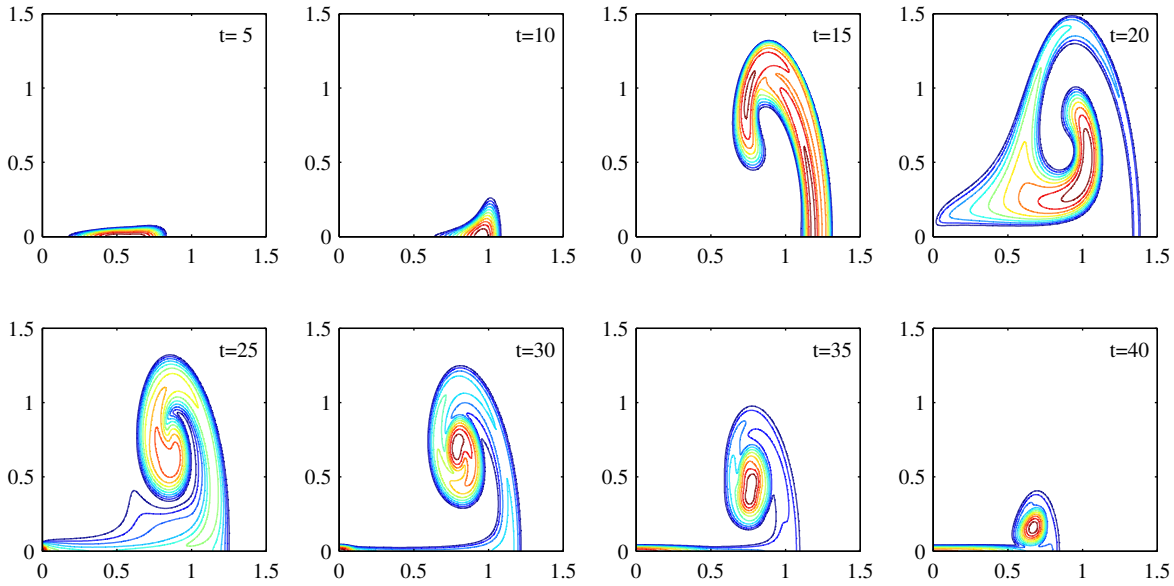

FIG. 10: Contour plots of the $D_x = 10^{-4}$ joint PDFs $p(x, v, t)$ at the times labelled in each panel, x on the horizontal axis, and v on the vertical axis. The top row corresponds to the Q-increasing phase, and the bottom row to the Q-decreasing phase. Contour intervals are on a logarithmic scale, with the smallest value at 10^{-3}, then $10^{-2.5}$, 10^{-2}, etc.

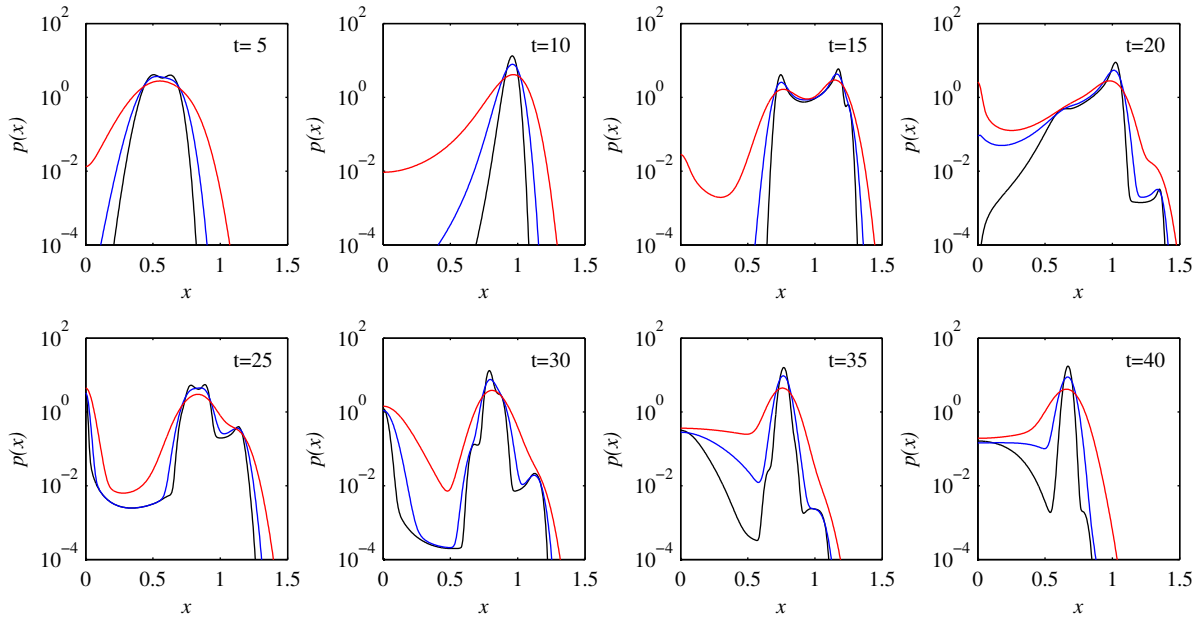

FIG. 11: The black curves in each panel show the marginal PDFs $p(x, t)$ corresponding to the joint PDFs shown in Fig. 10, which are again at $D_x = 10^{-4}$. The blue curves show equivalent results at $D_x = 4 \cdot 10^{-4}$, and red at $D_x = 16 \cdot 10^{-4}$.

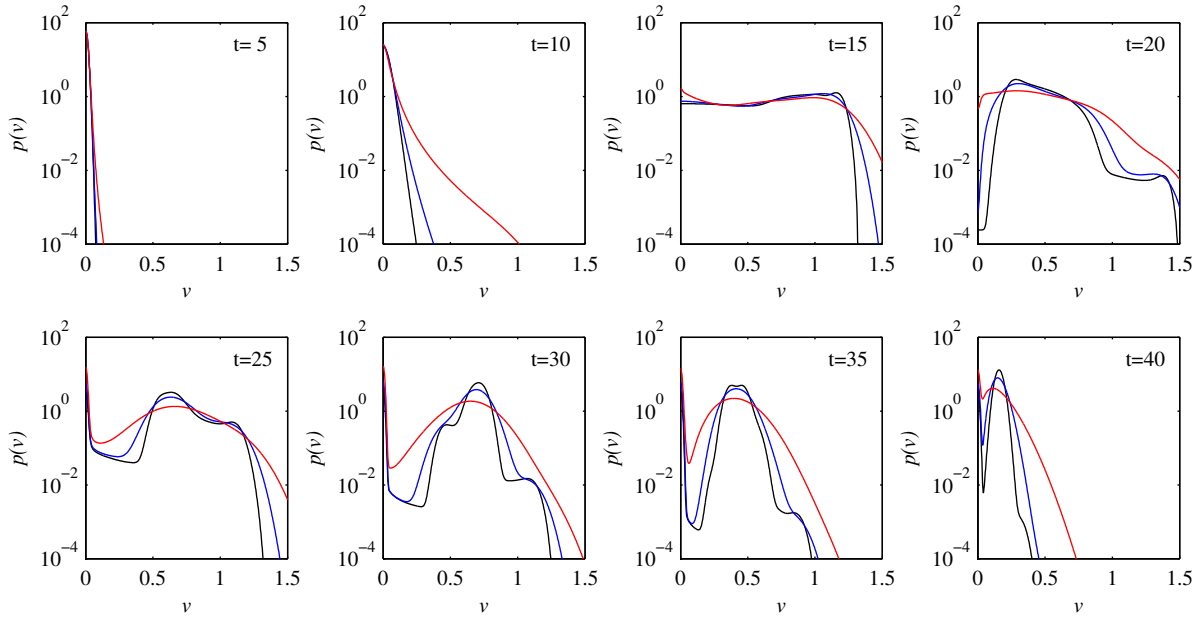

FIG. 12: As in Fig. 11, but now showing the marginal PDFs $p(v,t)$ rather than $p(x,t)$. The black curves again correspond to the joint PDFs in Fig. 10, then blue denotes $D_x = 4 \cdot 10^{-4}$ and red $D_x = 16 \cdot 10^{-4}$.

10, respectively.

It is interesting to observe also that there is virtually no difference across different rows ($D_x = [1, 4, 16] \times 10^{-4}$) in Fig. 13. That is, there is reduced sensitivity to noise at higher power ramp rate. As noted above, this curtailing of the limit cycle at high power/ramp rate is similar to what was found experimentally at high power input, where only one half of one cycle is executed (see Fig. 2(e) in [44]). In comparison, the information length diagnostics in Fig. 14 show the reduction in their values as D_x increases. Finally, Fig. 15 shows beautifully convoluted structures in the joint PDF $p(x, v, t)$.

VI. DISCUSSIONS AND CONCLUSIONS

We have investigated the stochastic, prey-predator L-H transition model by calculating time-dependent PDFs in the forward and backward processes where the input power increases and decreases, respectively. From the time-dependent PDFs, we calculated various statistical measures including the information length $\mathcal{L}_x(t)$ and $\mathcal{L}_v(t)$ for turbulence and zonal flows. Our principal results include that i) the mean-field-type theory is invalid with

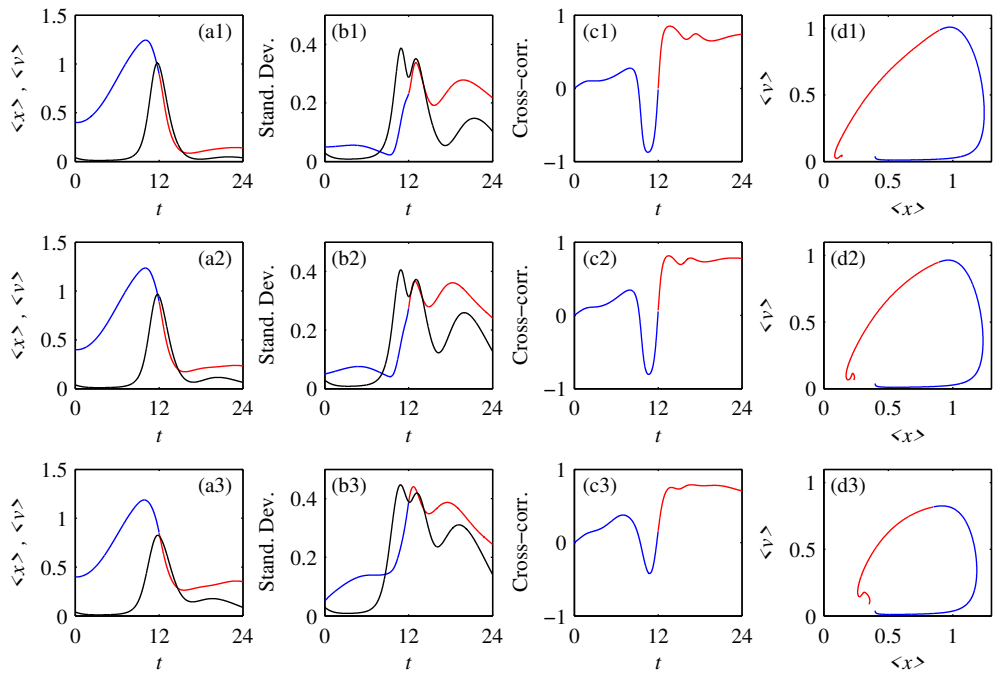

FIG. 13: From top to bottom, the three rows (1,2,3) show results for Set 3 at $D_x = 10^{-4}$, $4 \cdot 10^{-4}$ and $16 \cdot 10^{-4}$. Column 1 (a1,a2,a3) shows $\langle x \rangle$ and $\langle v \rangle$ as functions of time, with $\langle x \rangle$ in blue/red (with the switch in color at the transition time $t = 40$) and $\langle v \rangle$ in black. Column 2 (b1,b2,b3) shows the associated standard deviations σ_x and σ_v, with the same color-coding. Column 3 (c1,c2,c3) shows the cross-correlation $\langle (x - \langle x \rangle)(v - \langle v \rangle) \rangle / (\sigma_x \sigma_v)$. Column 4 (d1,d2,d3) shows the phase portrait in the $(\langle x \rangle, \langle v \rangle)$ plane, with blue again corresponding to $Q(t)$ increasing and red to $Q(t)$ decreasing.

the limited utility of mean value, variance, stationary PDFs; ii) the L-H transition can involve a strongly non-Gaussian PDF and intermittency; iii) the faster the ramping is, the greater the deviation from Gaussian PDFs and less duration of the dithering; iv) enhanced right tails for rare, large zonal flows can play a crucial role in turbulence regulation; v) the stark difference between the forward and backward processes is captured by time-dependent PDFs of turbulence and zonal flows; vi) the larger the stochasticity, the more asymmetry there is between forward and backward processes. Furthermore, the information geometric tool (information length) was shown to be a useful index to understand correlation and self-regulation, the transitions as well as forecasting the transitions.

While ensembles with a given statistical property (as used in this paper) facilitate analysis, it always raises the question how relevant those ensembles are to a real system (which

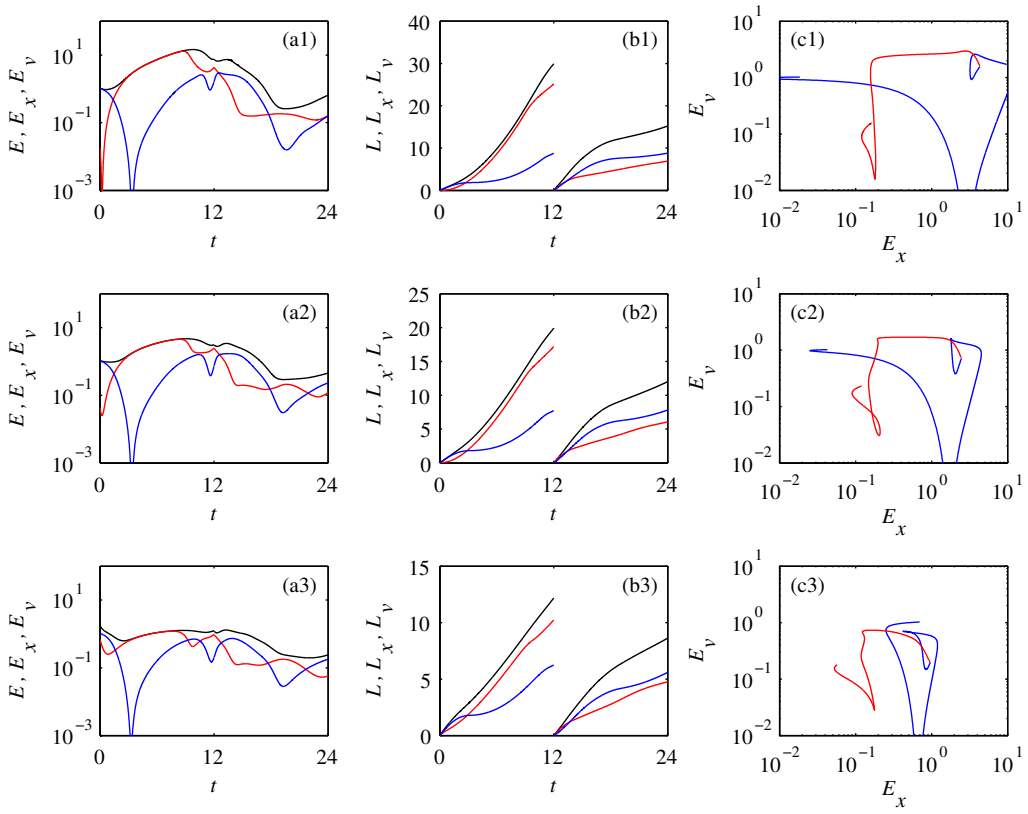

FIG. 14: The various information length diagnostics associated with the results in Fig. 1. Column 1 (a1,a2,a3) shows $\mathcal{E}$ (black), $\mathcal{E}_x$ (red), and $\mathcal{E}_v$ (blue) as functions of time. Column 2 (b1,b2,b3) shows $\mathcal{L}$ (black), $\mathcal{L}_x$ (red), and $\mathcal{E}_v$ (blue); note how these quantities are reset to zero at the start of the Q-decreasing phase. Column 3 shows the phase portrait in the $(\mathcal{E}_x, \mathcal{E}_v)$ plane, with blue corresponding to $Q(t)$ increasing and red to $Q(t)$ decreasing.

does not have multiple copies for ensembles). Based on ergodicity, time or spatial samplings are often used for ensembles assuming that there are many similar copies of the system at different times or spatial locations. This thus requires high-resolution data from numerical simulations or future experiments for good statistics. For instance, we can utilise different (temporal and/or spatial) samplings of data from high-resolution numerical simulations of fluid models or gyrokinetic models (which are computationally very expensive) for statistics. Even if high spatial resolution data might be difficult to obtain, time-dependent PDFs can still be calculated by sampling the data in the time-series of different variables (fluctuating density, electric field, etc.) from high temporal resolution data by using moving-time windows (e.g. see [21]). High-resolution temporal data would be necessary to ensure that there are a sufficient number of samples within the sliding windows to represent ensembles for a

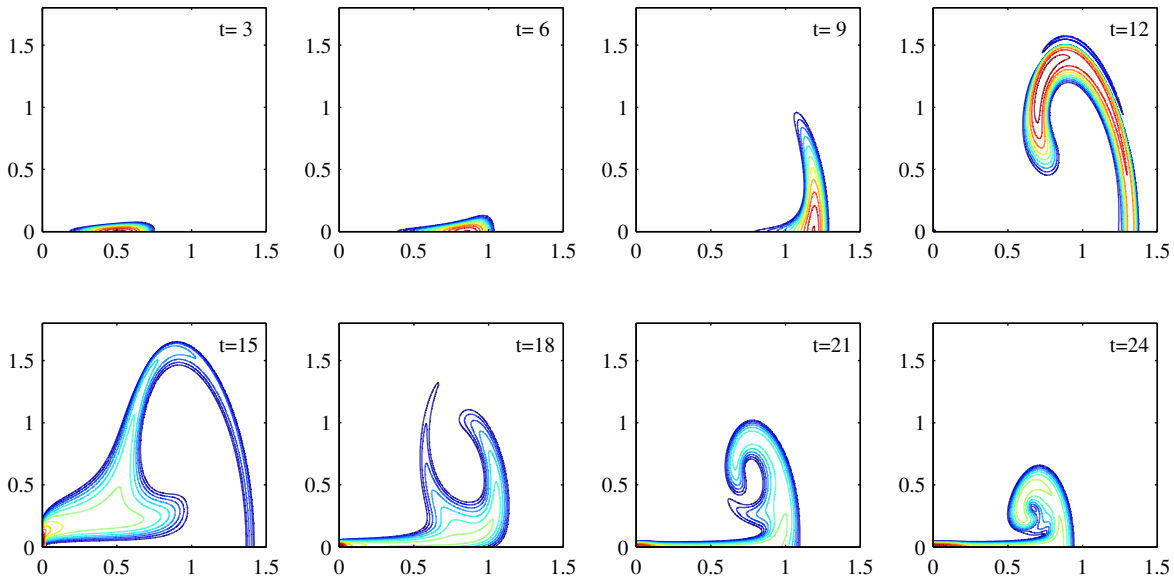

FIG. 15: Contour plots of the $D_x = 10^{-4}$ joint PDFs $p(x, v, t)$ at the times labelled in each panel, x on the horizontal axis, and v on the vertical axis. The top row corresponds to the Q-increasing phase, and the bottom row to the Q-decreasing phase. Contour intervals are on a logarithmic scale, with the smallest value at 10^{-3}, then $10^{-2.5}$, 10^{-2}, etc.

non-stationary process. Ensembles are often used to analyse non-stationary time-series for forecasting, etc., e.g. [58–60].

The latter method was in fact used in our recent work of the analysis of the Hasegawa-Wakatani turbulence [21] where the information diagnostics were shown to be a novel methodology of assessing the effects of coherent structures and turbulent dynamics in plasmas. In particular, a sudden change in the system, like intermittent transport by a coherent structure, increases $\mathcal{E}$ to a large value. Also, correlation/decorrelation of the flux between different spatial positions due to coherent structure was captured by similar/disparate time evolution of information length $\mathcal{L}$ and $\mathcal{E}$. (We note that the strong correlation between the two interacting species in [22] was captured by the same evolution of $\mathcal{L}$ and $\mathcal{E}$.) Similar analysis can be performed to interpret simulation results of the L-H transition models by using data at different spatial points for different variables to infer the correlation between different spatial points (e.g. plasma edge, core, etc.) as well as elucidating the effects of zonal and mean flows. High-resolution data in time and space will be highly desirable from experiments.

Since our methodology is much more sensitive to dynamical changes such as entering the

dithering phase than other traditional methods (mean values, variance), we may be able to identify undesirable plasmas events (large ELMs, major disruption, etc.) [56, 57] well before other methods can. This will then give more time to control or avoid the occurrence of such events. We note that the importance of a sufficient warning time to control major disruption was highlighted in [57].

In summary, the sensitivity of our information length diagnostics is welcome in addressing the main challenge in fusion research in regards to the prediction and control of anomalous transport or undesirable large plasmas events [56, 57], facilitating the control of such events. Finally, our future work will include extension to the 3D Fokker-Planck equation by treating N dynamically, different power rampings, more general stochastic forcings, extension to the fluid model to analyse the temporal-spatial dynamics in other L-H and H-L transition models or ELMs, experimental data, and the prediction of plasma eruption and mitigation. In the abstract mathematical sense, it will be of interest to understand the information budget [23] in space as well as its time-evolution.

Appendix A: Relation between $\mathcal{L}$ and relative entropy

We show that $\tau(t)$ and $\mathcal{L}(t)$ in Eqs. (2)-(6) are related to the relative entropy (Kullback–Leibler divergence) [15, 16] by considering two nearby PDFs $p_1 = p(x, t_1)$ and $p_2 = p(x, t_2)$ at times $t = t_1$ and t_2. We consider the limit of a very small $\delta t = t_2 - t_1$ and Taylor expand $D[p_1, p_2] = \int dx\, p_2 \ln(p_2/p_1)$ by using

$$\frac{\partial}{\partial t_1} D[p_1, p_2] = -\int dx\, p_2 \frac{\partial_{t_1} p_1}{p_1}, \tag{A1}$$

$$\frac{\partial^2}{\partial t_1^2} D[p_1, p_2] = \int dx\, p_2 \left\{ \frac{(\partial_{t_1} p_1)^2}{p_1^2} - \frac{\partial_{t_1}^2 p_1}{p_1} \right\}, \tag{A2}$$

$$\frac{\partial}{\partial t_2} D[p_1, p_2] = \int dx \left\{ \partial_{t_2} p_2 + \partial_{t_2} p_2 \big[\ln p_2 - \ln p_1\big] \right\}, \tag{A3}$$

$$\frac{\partial^2}{\partial t_2^2} D[p_1, p_2] = \int dx \left\{ \partial_{t_2}^2 p_2 + \frac{(\partial_{t_2} p_2)^2}{p_2} + \partial_{t_2}^2 p_2 \big[\ln p_2 - \ln p_1\big] \right\}. \tag{A4}$$

In the limit $t_2 \to t_1 = t$ ($p_2 \to p_1 = p$), Eqs. (A1)–(A4) give us

$$\begin{aligned} \lim_{t_2 \to t_1} \frac{\partial}{\partial t_1} D[p_1, p_2] &= \lim_{t_2 \to t_1} \frac{\partial}{\partial t_2} D[p_1, p_2] = \int dx \partial_t p = 0, \\ \lim_{t_2 \to t_1} \frac{\partial^2}{\partial t_1^2} D[p_1, p_2] &= \lim_{t_2 \to t_1} \frac{\partial^2}{\partial t_2^2} D[p_1, p_2] = \int dx \frac{(\partial_t p)^2}{p} = \frac{1}{\tau^2}. \end{aligned} \tag{A5}$$

Up to $O((dt)^2)$ ($dt = t_2 - t_1$), Eq. (A5) and $D(p_1, p_1) = 0$ lead to

$$D[p_1, p_2] = \frac{1}{2} \left[\int dx \frac{(\partial_t p(x,t))^2}{p(x,t)} \right] (dt)^2, \tag{A6}$$

and thus the infinitesimal distance $dl(t_1)$ between t_1 and $t_1 + dt$ as

$$dl(t_1) = \sqrt{D[p_1, p_2]} = \frac{1}{\sqrt{2}} \sqrt{\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p(x, t_1)}}\, dt. \tag{A7}$$

By summing $dt(t_i)$ for $i = 0, 1, 2, \ ..., n-1$ (where $n = t/dt$) in the limit $dt \to 0$, we have

$$\lim_{dt \to 0} \sum_{i=0}^{n-1} dl(idt) = \lim_{dt \to 0} \sum_{i=0}^{n-1} \sqrt{D[p(x, idt), p(x, (i+1)t)]}\, dt \propto \int_0^t dt_1 \sqrt{\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p(x, t_1)}} = \mathcal{L}(t), \tag{A8}$$

where $\mathcal{L}(t)$ is the information length in Eq. (1). We note that unlike a path-dependent $\mathcal{L}$, the relative entropy $D[p(x, 0), p(x, t)]$ depends only on PDFs at time 0 and t for a finite t and thus does not tell us about intermediate states between initial and final states.

Appendix B: Invariance of Eqs. (1)-(2) and (3)-(4)

To show the invariance of Eq. (1)-(2) under a change of variables, let us consider $y = F(x)$. Then, we have

$$p(y,t) = p(x,t)\left|\frac{dx}{dy}\right| = p(x,t)\frac{1}{|F'(x)|}, \tag{B1}$$

where $F'(x) = \frac{dF(x)}{dx}$. Since $F'(x)$ is independent of time t, we have $\partial_t p(y,t) = [\partial_t p(x,t)]\frac{1}{|F'(x)|}$. Using this and $dy = dx|F'(x)|$, we have

$$\int dy \frac{1}{p(y,t)}\left[\frac{\partial p(y,t)}{\partial t}\right]^2 = \int dx \frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2. \tag{B2}$$

Thus, we obtain the identical $\tau(t)$ from $p(x,t)$ and $p(y,t)$.

Similarly, for Eqs. (3)-(4), by considering the independent changes of variables as $y = F(x)$ and $z = F(v)$ that are independent of time t, we can show that $\tau_y = \tau_x$ and $\tau_z = \tau_v$.

Appendix C: Derivation of the Fokker-Planck Equations in Eq. (12)

In order to derive the Fokker-Planck equation (12) from the Langevin equation (9)-(11), it is useful to introduce a generating function Z (see also [54, 55])

$$Z = e^{i(\lambda x(t) + \gamma v(t))}. \tag{C1}$$

Then, by definition of 'average', the average of Z is related to the PDF, $p(x,v,t)$, as

$$\langle Z\rangle = \int dxdv\, Z\, p(x,v,t) = \int dxdv\, e^{i(\lambda x(t)+\gamma v(t))}\, p(x,v,t). \tag{C2}$$

Thus, we see that $\langle Z\rangle$ is the Fourier transform of $p(x,v,t)$. The inverse Fourier transform of $\langle Z\rangle$ then gives $p(x,v,t)$:

$$p(x,v,t) = \left(\frac{1}{2\pi}\right)^2 \int d\lambda d\gamma\, e^{-i(\lambda x+\gamma v)}\, \langle Z\rangle. \tag{C3}$$

We note that Eq. (C3) can be written as

$$p(x,v,t) = \left\langle \left(\frac{1}{2\pi}\right)^2 \int d\lambda d\gamma\, e^{-i[\lambda(x-x(t))+\gamma(v-v(t))]} \right\rangle = \langle \delta(x(t)-x)\delta(v(t)-v)\rangle\,, \tag{C4}$$

which is another form of $p(x,v,t)$. To obtain the equation for $p(x,v,t)$, we first derive the equation for $\langle Z\rangle$ and then take the inverse Fourier transform as summarized in the following.

We differentiate Z with respect to time t and use Eqs. (9)-(10) to obtain

$$\partial_t Z = i[\lambda \partial_t x Z + \gamma \partial_t v Z] = i[\lambda(f + \xi(t)) + \gamma(g + \eta(t))]Z. \tag{C5}$$

The formal solution to Eq. (C5) is

$$Z(t) = \int_0^t dt_1 \, i[\lambda(f(t_1) + \xi(t_1)) + \gamma(g(t_1) + \eta(t_1))]Z(t_1). \tag{C6}$$

The average of Eq. (C5) gives

$$\partial_t \langle Z \rangle = i \left[\lambda \langle f Z \rangle + \gamma \langle g Z \rangle\right] + i\lambda \langle \xi(t) Z(t) \rangle + i\gamma \langle \eta(t) Z(t) \rangle. \tag{C7}$$

To find $\langle \xi(t) Z(t) \rangle$, we use Eq. (C6) as follows:

$$\begin{aligned} \langle \xi(t) Z(t) \rangle &= i \langle \xi(t) \int_0^t dt_1 \left[\lambda(f(t_1) + \xi(t_1)) + \gamma(g(t_1) + \eta(t_1))\right] Z(t_1) \rangle \\ &= i\lambda^2 D_x \langle Z(t) \rangle. \end{aligned} \tag{C8}$$

Here we used the independence of $\xi(t)$ and $Z(t_1)$ for $t_1 < t$, $\langle \xi(t) Z(t_1) \rangle = \langle \xi(t) \rangle \langle Z(t_1) \rangle = 0$, Eq. (11), $\int_0^t dt_1 \, \delta(t - t_1) = 1/2$, $\langle \xi \rangle = \langle \eta \rangle = 0$ and $\langle \xi(t) \eta(t') \rangle = 0$. Note that $\langle \xi(t) Z(t_1) \rangle = \langle \xi(t) \rangle \langle Z(t_1) \rangle = 0$ for $t_1 < t$ is an exact result since the stochastic noise changes very rapidly due to its short correlation time, and thus its value at any future time t evolves independently of the previous value of $Z(t_1)$ at the earlier time t_1. Similarly, we can show that

$$\langle \eta(t) Z(t) \rangle = i\gamma^2 D_v \langle Z(t) \rangle. \tag{C9}$$

By substituting Eqs. (C8)-(C9) into Eq. (C7) we obtain

$$\partial_t \langle Z \rangle = i \left[\lambda \langle f Z \rangle + \gamma \langle g Z \rangle\right] - [\lambda^2 D_x + \gamma^2 D_v] \langle Z(t) \rangle. \tag{C10}$$

The inverse Fourier transform of Eq. (C10) then gives us

$$\frac{\partial}{\partial t} p(x, v, t) = -\frac{\partial}{\partial x}\Big[f p(x, v, t)\Big] - \frac{\partial}{\partial v}\Big[g p(x, v, t)\Big] + \left[D_x \frac{\partial^2}{\partial x^2} + D_v \frac{\partial^2}{\partial v^2}\right] p(x, v, t). \tag{C11}$$

which is Eq. (12) in the text.

Alternative ways of deriving Eq. (12) are also given by [54, 55, 61]. Note further that the Fokker-Planck Eq. (12) corresponding to the Langevin Eqs. (9)-(11) agrees with the general expression for the Fokker-Planck equation (4.94)-(4.95) corresponding to the Langevin equation (3.110)-(3.113) for the multi-variables in [54], which has also been used extensively in the literature (e.g. Eqs. (14)-(15) in the recent paper [62]).

Acknowledgement EK acknowledges the Leverhulme Trust Research Fellowship (RF-2018-142-9). We thank Y. Andrew, Y. Zheng, G. McKee and DIII-D team for helpful discussion.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

---

[1] S. J. Zweben, J. A. Boedo, O. Grulke, C. Hidalgo, B. LaBombard, R. J. Maqueda, P. Scarin, and J. L. Terry, "Edge turbulence measurements in toroidal fusion devices," Plasma Phys. Contr. Fusion **49**, S1-23 (2007).

[2] P. A. Politzer, "Observation of avalanche-like phenomena in a magnetically confined plasma," Phys. Rev. Lett. **84**, 1192-1195 (2000).

[3] P. Beyer, S. Benkadda, X. Garbet, and P. H. Diamond, "Nondiffusive transport in tokamaks: Three-dimensional structure of bursts and the role of zonal flows," Phys. Rev. Lett. **85**, 4892-4895 (2000).

[4] J. F. Drake, P. N. Guzdar, and A. B. Hassam, "Streamer formation in plasma with a temperature gradient," Phys. Rev. Lett. **61**, 2205-2208 (1988).

[5] G. Y. Antar, S. I. Krasheninnikov, P. Devynck, R. P. Doerner, E. M. Hollmann, J. A. Boedo, S. C. Luckhardt, and R. W. Conn, "Experimental evidence of intermittent convection in the edge of magnetic confinement devices," Phys. Rev. Lett. **87**, 065001 (2001).

[6] B. A. Carreras, C. Hidalgo, E. Sanchez, M. A. Pedrosa, R. Balbin, I. Garcia-Cortes, B. van Milligen. D. E. Newman, and V. E. Lynch, "Fluctuation-induced flux at the plasma edge in toroidal devices," Phys. Plasmas **3**, 2664-2672 (1996).

[7] J. Anderson and P. Xanthopoulos, "Signature of a universal statistical description for drift-wave plasma turbulence," Phys. Plasmas **17**, 110702 (2010).

[8] E. Kim and J. Anderson, "Structure based statistical theory of intermittency," Phys. Plasmas **15**, 114506 (2008).

[9] E. Kim and P.H. Diamond, "Intermittency in drift-wave turbulence: Structure of the momentum flux probability distribution function," Phys. Rev. Lett. **88**, 225002 (2002).

[10] E. Kim, P. H. Diamond, M. Malkov, T. S. Hahm, K. Itoh, S.-I. Itoh, S. Champeaux, I. Gruzinov, O. Gurcan, C. Holland *et al.*, "Non-perturbative models of intermittency in drift-

wave turbulence: towards a probabilistic theory of anomalous transport," Nuclear Fusion **43**, 961-968 (2003).

[11] J. Anderson and E. Kim, "The momentum flux probability distribution function for ion-temperature-gradient turbulence," Phys. Plasmas **15**, 052306 (2008).

[12] J. Anderson and E. Kim, "Nonperturbative models of intermittency in edge turbulence," Phys. Plasmas, **15**, 122303 (2008).

[13] S.B. Nicholson and E. Kim, "Investigation of the statistical distance to reach stationary distributions," Phys. Lett. A **379**, 83-88 (2015).

[14] E. Kim, "Intermittency and self-organisation in turbulence and statistical mechanics," Entropy **21**, 574 (2019).

[15] J. Heseltine and E. Kim, "Novel mapping in non-equilibrium stochastic processes," J. Phys. A **49**, 175002 (2016).

[16] E. Kim, U. Lee, J. Heseltine, and R. Hollerbach, "Geometric structure and geodesic in a solvable model of nonequilibrium process," Phys. Rev. E **93**, 062127 (2016).

[17] E. Kim and R. Hollerbach, "Time-dependent probability density function in cubic stochastic processes," Phys. Rev. E **94**, 052118 (2016).

[18] E. Kim and R. Hollerbach, "Signature of nonlinear damping in geometric structure of a nonequilibrium process," Phys. Rev. E **95**, 022137 (2017).

[19] E. Kim and R. Hollerbach, "Geometric structure and information change in phase transitions," Phys. Rev. E **95**, 062107 (2017).

[20] E. Kim, Q. Jacquet, and R. Hollerbach, "Information geometry in a reduced model of self-organised shear flows without the uniform coloured noise approximation," J. Stat. Mech., 023204 (2019).

[21] J. Anderson, E. Kim, B. Hnat, and T. Rafiq, "Elucidating plasma dynamics in Hasegawa-Wakatani turbulence by information geometry," Phys. Plasmas **27**, 022307 (2020).

[22] J. Heseltine and E. Kim, "Comparing information metrics for a coupled Ornstein-Uhlenbeck process," Entropy **21**, 775 (2019).

[23] E. Kim, J. Heseltine, and H. Liu, "Information length as a useful index to understand variability in the global circulation," Mathematics **8**, 299 (2020).

[24] E. Kim and R. Hollerbach, "Time-dependent probability density functions and information geometry of the low-to-high confinement transition in fusion plasma," Phys. Rev. Res. **2**,

023077 (2020).

[25] F. Wagner, G. Becker, K. Behringer, D. Campbell, A. Eberhagen, W. Engelhardt, G. Fussmann, O. Gehre, J. Gernhardt, G. von Gierke *et al.*, "Regime of improved confinement and high-beta in neutral-beam-heated divertor discharges of the ASDEX tokamak," Phys. Rev. Lett. **49**, 1408-1412 (1982).

[26] P. H. Diamond, S.-I. Itoh, K. Itoh, and T. S. Hahm, "Zonal flows in plasma – a review," Plasma Phys. Contr. Fusion **47**, R35-161 (2005).

[27] C. S. Chang, S. Ku, G. R. Tynan, R. Hager, R. M. Churchill, I. Cziegler, M. Greenwald, A. E. Hubbard, and J. W. Hughes, "Fast low-to-high confinement mode bifurcation dynamics in a tokamak edge plasma gyrokinetic simulation," Phys. Rev. Lett. **118**, 175001 (2017).

[28] L. Schmitz, "The role of turbulence-flow interactions in L- to H-mode transition dynamics: recent progress," Nuclear Fusion **57**, 025003 (2017).

[29] E. Kim and P. H. Diamond, "Zonal flows and transient dynamics of the L-H transition," Phys. Rev. Lett. **90**, 185006 (2003).

[30] M. A. Malkov and P. H. Diamond, "Weak hysteresis in a simplified model of the L-H transition," Phys. Plasmas **16**, 012504 (2009).

[31] H. Zhu, S. C. Chapman, and R. O. Dendy, "Robustness of predator-prey models for confinement regime transitions in fusion plasmas," Phys. Plasmas **20**, 042302 (2013).

[32] N. H. Bian, "On-off intermittent regulation of plasma turbulence," Phys. Plasmas **17**, 044501 (2010).

[33] K. Miki, P. H. Diamond, O. D. Gurcan, G. R. Tynan, T. Estrada, L. Schmitz, and G. S. Xu, "Spatio-temporal evolution of the L $\rightarrow$ I $\rightarrow$ H transition," Phys. Plasmas **19**, 092306 (2012).

[34] S. Douglas, M. Mohamed, and E. Kim, "Confinement improvement by fluctuating input power," Phys. Plasmas **20**, 114504 (2013).

[35] C. F. Maggi, E. Delabie, T. M. Biewer, M. Groth, N. C. Hawkes, M. Lehnen, E. de la Luna, K. McCormick, C. Reux, F. Rimini *et al.*, "L-H power threshold studies in JET with Be/W and C wall," Nuclear Fusion **54**, 023007 (2014).

[36] T. Nishizawa, A. F. Almagri, J. K. Anderson, W. Goodman, M. J. Pueschel, M. D. Nornberg, S. Ohshima, J. S. Sarff, P. W. Terry, and Z. R. Williams, "Direct measurement of a toroidally directed zonal flow in a toroidal plasma," Phys. Rev. Lett. **122**, 105001 (2019).

[37] L. Schmitz, L. Zeng, T. L. Rhodes, J. C. Hillesheim, E. J. Doyle, R. J. Groebner, W. A. Peebles,

K. H. Burrell, and G. Wang, "Role of zonal flow predator-prey oscillations in triggering the transition to H-mode confinement," Phys. Rev. Lett. **108**, 155002 (2012).

[38] T. Kobayashi, K. Itoh, T. Ido, K. Kamiya, S.-I. Itoh, Y. Miura, Y. Nagashima, A. Fujisawa, S. Inagaki, K. Ida, and K. Hoshino, "Spatiotemporal structures of edge limit-cycle oscillation before L-to-H transition in the JFT-2M tokamak," Phys. Rev. Lett. **111**, 035002 (2013).

[39] J. Cheng, J. Q. Dong, K. Itoh, L. W. Yan, M. Xu, K. J. Zhao, W. Y. Hong, Z. H. Huang, X. Q. Ji, W. L. Zhong *et al.*, "Dynamics of low-intermediate-high-confinement transitions in toroidal plasmas," Phys. Rev. Lett. **110**, 265002 (2013).

[40] G. S. Xu, H. Q. Wang, M. Xu, B. N. Wan, H. Y. Guo, P. H. Diamond, G. R. Tynan, R. Chen, N. Yan, D. F. Kong *et al.*, "Dynamics of L-H transition and I-phase in EAST," Nuclear Fusion **54**, 103002 (2014).

[41] Z. Yan, G. R. McKee, R. Fonck, P. Gohil, R. J. Groebner, and T. H. Osborne, "Observation of the L-H confinement bifurcation triggered by a turbulence-driven shear flow in a tokamak plasma," Phys. Rev. Lett. **112**, 125002 (2014).

[42] J. J. Rasmussen, A. H. Nielsen, J. Madsen, V. Naulin, and G. S. Xu, "Numerical modeling of the transition from low to high confinement in magnetically confined plasma," Plasma Phys. Contr. Fusion **58**, 014031 (2016).

[43] T. I. Wagner, "The history of research into improved confinement regimes," Europ. Phys. J. H **43**, 523-549 (2018).

[44] L. Schmitz, L. Zeng, T. L. Rhodes, J. C. Hillesheim, W. A. Peebles, R. J. Groebner, K. H. Burrell, G. R. McKee, Z. Yan, G. R. Tynan, P. H. Diamond, J. A. Boedo, E. J. Doyle, B. A. Grierson, C. Chrystal, M. E. Austin, W. M. Solomon, and G. Wang, "The role of zonal flows and predator–prey oscillations in triggering the formation of edge and core transport barriers," Nuclear Fusion **54**, 073012 (2014).

[45] C. C. Petty and the DIII-D Team, "DIII-D research towards establishing the scientific basis for future fusion reactors," Nuclear Fusion **59**, 112002 (2019).

[46] P. Gohil, T. E. Evans, M. E. Fenstermacher, J. R. Ferron, T. H. Osborne, J. M. Park, O. Schmitz, J. T. Scoville, and E. A. Unterberg, "L–H transition studies on DIII-D to determine H-mode access for operational scenarios in ITER," Nuclear Fusion **51**, 103020 (2011).

[47] A. Newton, E. Kim, and H. Liu, "On the self-organizing process of large scale shear flows," Phys. Plasmas **20**, 092306 (2013).

[48] E. Kim and I. Movahedi, "Effect of enhanced dissipation by shear flows on transient relaxation and probability density function in two dimensions," Phys. Plasmas **24**, 112306 (2017).

[49] E. Kim, "Consistent theory of turbulent transport in two-dimensional magnetohydrodynamics," Phys. Rev. Lett. **96**, 084504 (2006).

[50] B. R. Frieden, *Science from Fisher Information* (Cambridge University Press, 2004).

[51] W. K. Wootters, "Statistical distance and Hilbert-space," Phys. Rev. D **23**, 357-362 (1981).

[52] M. Dam, M. Brons, J. J. Rasmussen, V. Naulin, and G. S. Xu, "Bifurcation analysis and dimension reduction of a predator-prey model for the L-H transition," Phys. Plasmas **20**, 102302 (2013).

[53] K. Miki and P. H. Diamond, "Novel states of pre-transition edge turbulence emerging from shearing mode competition," Nuclear Fusion **51**, 103003 (2011).

[54] H. Risken, *The Fokker-Planck Equation: Methods of Solutions and Applications* (Springer, 2013).

[55] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon Press, 2002).

[56] P. C. De Vries, M. F. Johnson, B. Alper, P. Buratti, T. C. Hender, H. R. Koslowski, V. Riccardo, JET-EFDA Contributors, "Survey of disruption causes at JET," Nuclear Fusion **51**, 053018 (2011).

[57] J. Kates-Harbeck, A. Svyatkovskiy and W. Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," Nature **568**, 527 (2019).

[58] J. L. McCauley, *Stochastic Calculus and Differential Equations for Physics and Finance*, Chap. 13.2-13.4 (Cambridge University Press, 2013).

[59] J. L. McCauley, "Time versus ensemble averages for nonstationary time series," Physica A **387**, 5518-5522 (2008).

[60] H. Allende and C. Valle, *Ensemble Methods for Time Series Forecasting.* In: Seising R., Allende-Cid H. (eds) Claudio Moraga: A Passion for Multi-Valued Logic and Soft Computing. Studies in Fuzziness and Soft Computing, Vol. 349 (Springer, 2017).

[61] P. Hänggi, "Path integral solutions for non-Markovian processes," Z. Physik B **75**, 275-281 (1989).

[62] A. Medved, R. Davis and P. A. Vasquez, "Understanding fluid dynamics from Langevin and Fokker-Planck equations," Fluids **5**, 40 (2020).