Article:
Kim, E and Hollerbach, R orcid.org/0000-0001-8639-0967 (2022) Non-equilibrium Statistical Properties, Path-dependent Information Geometry, and Entropy Relations in Edge-Localized Modes in Fusion Plasmas. Physics of Plasmas, 29 (11). 112302. ISSN 1070-664X

https://doi.org/10.1063/5.0109257

Non-equilibrium Statistical Properties, Path-dependent Information Geometry, and Entropy Relations in Edge-Localized Modes in Fusion Plasmas

Eun-jin Kim[1,a)] and Rainer Hollerbach[2]
[1] *Fluid and Complex Systems Research Centre, Coventry University, Coventry CV1 2TT, UK*
[2] *Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK*

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

We investigate time-varying turbulence statistical properties of edge-localized modes in fusion plasmas. By utilizing a simplified stochastic model, we calculate a time-dependent probability density function (PDF) and various entropy-related quantities such as entropy, entropy production, entropy flux, mutual information, and information flow and path-dependent information geometry. A thorough analysis is performed to elucidate the effects on ELM dynamics (evolution, suppression, mitigation, etc.) of different values of stochastic noise and different forms of a time-varying input power. Furthermore, the time-irreversibility and hysteresis are investigated through the employment of forward and back processes where a time-varying input power varies mirror-symmetrically in time. Among all the statistical quantities, the path-dependent information geometry is shown to be a robust diagnostic for quantifying hysteresis and self-regulation as well as for an early detection of subtle changes in ELM dynamics, e.g., caused by a sudden change in the input power.

I. INTRODUCTION

Plasmas in fusion devices are volatile and constitute one of the prototypical examples of non-equilibrium systems [1] involving large fluctuations. Excited on a broad range of scales (e.g., micro instability for electron/ion transport), various instabilities cause anomalous transport/loss of energy/particles that are larger than what is expected from collisional values, significantly degrading the confinement. In a typical low-confinement mode (L-mode), confinement degrades as the plasmas become hotter with increasing input power. The tendency to more complex dynamics as the free energy is increased is in fact common to many systems such as in Rayleigh-Bénard convection, where an increase in temperature contrast across the system leads to turbulent convection. Understanding anomalous transport is critical for achieving the well-controlled extraction of fusion energy in magnetic fusion, with significant practical implications (ITER, future power plants).

At the same time, the volatile characteristics of fusion plasmas provides us with a valuable theoretical framework in which one can investigate non-equilibrium processes and test statistical methods to disentangle complex dynamics (e.g., interaction, correlation, causality among different players involved in transport) [2, 3]. To address this, various statistical methods [3–7] including bi-coherence spectral analysis, phase-space portrait, Reynolds stress analysis, wavelet analysis, cross-phase analysis etc. have been valuable. However, the exploration of non-perturbative statistical methods using Probability Density Functions (PDFs) taking into account strong fluctuations [8–16] will help us characterize rare, large-amplitude events in PDF tails [17–22], whose early prediction can be vital for ensuring enough time for applying controls. Furthermore, a PDF method will enable us to employ and test different statistical diagnostics based on information theory (e.g. transfer entropy, etc.) or information geometry. The latter refers to the application of differential geometry to probability and statistics [23–25].

This paper aims to touch on these practical and theoretical issues through stochastic modelling of Edge-Localized Modes (ELMs) [6, 26–31] that occur in the high-confinement mode (H-mode) plasmas. The H-mode is the enhanced confinement regime [32] that is attained when plasmas bifurcate from the L-mode at the critical input power P_{cr}. The latter, the so-called L-H transition, is caused by the spontaneous formation of the $E \times B$ flow velocity shear layer in the edge plasma accompanied by the reduction in the turbulence and transport and the steepening of pressure profile [7, 33–46].

ELMs manifest as sudden, quasi-periodic oscillations/bursts caused by instabilities of pressure/current gradient in the transport barrier for a sufficiently high input power in the H-mode. In particular, rare, bursty, large amplitude (Type I) ELMs impose a large heat load, or even damage to the fusion container walls. Therefore, great attention has been paid to ELM mitigation or suppression using resonant magnetic perturbations or a pellet injection, e.g., in DIII-D, JET, ASDEX-U, EAST tokamaks [30, 31, 47]. Notably, the potential for stochastic magnetic field to reduce the pressure gradient by electron transport was suggested more than a few decades ago [48]. There have been more recent works (e.g., [49, 50] and references therein) demonstrating the importance of magnetic perturbations in transport at kinetic level, for instance, those arising internally (due to a disruption) or externally (due to external coils) acting in such a way as to inhibit runaway electrons (avalanches). Thus, a proper stochastic modelling of ELMs will be beneficial instead of utilizing averages or moments.

In this paper, we investigate non-perturbative statistical diagnostics in a stochastic model of ELMs. These include: i) the entropy-related quantities such as entropy, entropy production, entropy flux, mutual information, and information flow; ii) information geometry. In particular, statistical properties are not necessarily stationary in fusion plasmas and can undergo significant temporal changes, especially near the onset of bursts, abrupt events, etc. Thus, we focus on a time-dependent PDF method [25, 45, 46, 52–60] to capture time-varying statistical properties and path-dependent information geometry.

Specifically, for a stochastic variable x with a PDF $p(x,t)$, our path-dependent information geometry captures the evolution from $p(x,t_0)$ to $p(x,t_F)$ along a particular path between the times $t=(t_0,t_F)$ among an infinite number of trajectories between the two. This has the advantage over using other metrics such as the Kullback-Leibler (K-L) divergence (see Appendix A) that compares two given PDFs $p(x,t_0)$ and $p(x,t_F)$, which does not tell us anything about the evolution.

Furthermore, far from equilibrium, the time-evolution of a system does not obey time-symmetry (time-irreversibility) even when a model parameter is symmetric in time. In order to quantify this time-dependence, we are interested in how a PDF changes with time in the forward and back processes for a time-varying parameter which is mirror-symmetric in time. For our stochastic ELM model, the model parameter would be an input power, as discussed in §II.

In order to perform a thorough investigation of evolution paths of ELMs under different condition, we take advantage of a simple computational model. Given the simplicity of our model, our study will be quantitative with the focus on identifying overall tendencies and generic properties. We aim to answer the following questions:

- What are the effects of stochastic noise on ELM dynamics?
- What are the effects of a non-constant model parameter (input power)?
- How do we characterize irreversibility and hysteresis and how are they affected by stochasticity and the temporal form of an input power?
- What is a robust diagnostic capturing a sudden change in model parameters?
- What is a robust diagnostic that detects subtle changes to the system, the onset of bursts due to a sudden internal event (i.e., disruption, explosive event, burst, instability) or external perturbation, which is important for predicting ELMs?
- How do we characterize self-regulation among different variables that have different physical units?

The remainder of this paper is organized as follows. We provide a stochastic ELM model in Section II. Sections III and IV present the summary of our main entropy-related statistical diagnostics, including thermodynamic and path-dependent information geometry diagnostics. Sections V and VI contain our numerical methods and results. We conclude in Section VII. Appendices A-D contain some background material including some of the details of the deterministic model of ELMs and the results that are not included in the main text of the paper.

II. STOCHASTIC ELM MODEL.

To elucidate the effects of stochastic noises on ELMs, we utilize a minimal deterministic ODE model of ELM dynamics proposed and analyzed in [26], which provided detailed theoretical explanations of the model's connection to plasma physics and nuclear fusion. This model involves the evolution of the pressure gradient, magnetic fluctuation amplitude together with the ion radial-force balance relation that the electric fields contain the contributions from the pressure gradient as well as the poloidal velocity [26].

For the input power P_{in} far above the critical power-threshold P_{cr} [61], diamagnetic velocity dominates over the poloidal velocity so that the electric field is mainly driven by the pressure gradient. This is the limit we will focus on for our comprehensive computational study. Due to the absence of the contribution from the poloidal flow to $E\times B$ shear flows, the L-mode bifurcates into the ELMy H-mode without an ELMy free H-mode gap. The resulting ODE model for the dimensionless pressure gradient p and magnetic fluctuations E_M is given in Eqs. (C1)-(C3) in Appendix C where some fixed point solutions are described to make this paper more self-contained for readers.

A. Langevin model

We extend the deterministic ODE model (C1)-(C3) by adding the two independent, short-correlated Gaussian stochastic noises ξ and η to incorporate uncertainty in the ODE model. Here, ξ and η are assumed to be of the strength Q_x and Q_y, respectively, with the following statistical properties

$$\langle \xi(t)\xi(t')\rangle = 2Q_x\delta(t-t'),\ \langle \eta(t)\eta(t')\rangle = 2Q_y\delta(t-t'),\\ \langle \xi(t)\eta(t')\rangle = 0,\ \ \langle \xi\rangle = \langle \eta\rangle = 0, \tag{1}$$

where $\delta(t-t')$ means that the memory time of ξ and η is shorter than any other characteristic time scales (e.g., ELM period) in the system. Specifically, we use the variables $x=p$ and $y=\sqrt{E_M}$ for convenience to represent the dimensionless pressure gradient p and the magnetic fluctuations E_M. Our stochastic ODE model then takes

the following form:

$$\frac{dx}{dt} = \Phi - \tilde{D}(x)x - xy^2 + \xi = f + \xi, \tag{2}$$

$$\frac{dy}{dt} = \frac{1}{2}\lambda(x-1)y + \eta = g + \eta, \tag{3}$$

$$\tilde{D} = d_0 + d(x - c^2x^4)\Theta(\tilde{P} - x). \tag{4}$$

Here, $\Theta(x)$ is the Heaviside function with $\Theta(x) = 1$ for $x \geq 0$ and $\Theta(x) = 0$ for $x < 0$; $\tilde{P}$ is the critical pressure gradient x for the complete suppression of turbulence due to the shear; $c \equiv \tilde{P}^{-3/2}$. Φ is the control parameter that represents the energy flux (input power) by assuming a constant temperature. λ and $d_0 \ll d$ are non-negative constants.

We note that the Heaviside function Θ is discontinuous, but the additional factor $(x - c^2x^4)$ ensures that $\tilde{D}$ in Eq. (4) is nevertheless continuous (but not differentiable). Since the Fokker-Planck Eq. (5) below involves $\frac{d}{dx}\tilde{D}$, one might wonder whether this lack of differentiability affects the solutions or not. Various numerical experiments were done smoothing out the discontinuity in Θ over a small but finite region, and indicate that using a true Heaviside function or a slightly smoothed one yield the same results.

Physically, ξ, η in Eqs. (2)-(3) represent the contributions from the internal (incoherent unresolved-scale dynamics) and any external perturbations that are not included in the ODE model. These include fluctuating energy flux of unresolved scales, the outward energy flux at the edge (e.g. [8, 62]), pellet pacing [30], mini-avalanches [46], stochastic magnetic fields induced by internal disruption, kinetic instabilities (due to the runaway electrons) [51], or external magnetic coils, etc. The accurate quantification of the values of Q_x and Q_y is infeasible since they would depend on many factors (different turbulence models, experimental conditions, plasma configurations, etc.). Thus, our aim is to elucidate the generic effects of stochastic noise and time-dependent input power Φ on ELM dynamics, as noted previously. To this end, we investigate how different values/forms of Q_x and Q_y and $\Phi(t)$ affect statistical properties and dynamics of ELMs.

Finally, we note that Eqs. (2)-(4) are non-dimensionalised such that time t is in the units of $[(c_s/\rho_s)k\rho_s(\Delta_c^4/\rho_s^2L_p^2)]^{-1}$, where $c_s = \sqrt{T_e/m_i}$ is the ion sound speed, $\rho_s = c_s/\omega_{ci}$, ω_{ci} is the ion cyclotron frequency, and k and Δ_c are the poloidal wave number and radial correlation length of the turbulence (see [26, 63] for details).

B. Joint PDF $p(x, y, t)$

In the stochastic model with non-zero ξ and η, the trajectories of x are random and its probability taking certain value is described by a PDF. A joint PDF $p(x, y, t)$ represents the likelihood of x and y taking different values around what are expected from the deterministic model. Since x and y can take any real number, their values can be considered as the spatial coordinates. The time-evolution of $p(x, y, t)$ is then governed by the following Fokker-Planck equation [64], which is effectively a partial differential equation in the two (spatial) + one (time) coordinates since $x, y \in \mathbb{R}$;

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(f\,p) - \frac{\partial}{\partial y}(g\,p) + Q_x\frac{\partial^2 p}{\partial x^2} + Q_y\frac{\partial^2 p}{\partial y^2} \equiv -\partial_x J_x - \partial_y J_y. \tag{5}$$

In Eq. (5)

$$f \equiv \Phi - \tilde{D}(x)x - xy^2, \qquad g \equiv \frac{\lambda}{2}(x-1)y, \tag{6}$$

$$J_x = fp - Q_x\partial_x p, \qquad J_y = gp - Q_y\partial_y p, \tag{7}$$

where J_x and J_y are the probability currents of x and y. Eq. (5) represents the local conservation of probability density. That is, the local temporal change of a PDF at particular x and y is caused by the probability current flowing in or out of that region. The probability current consists of the two parts due to the deterministic force (f or g) and diffusion due to the stochastic noise. Consequently, for larger stochastic noise (Q_x, Q_y), a PDF spreads more and becomes more widely distributed (in x, y).

We note that an alternative method to the Fokker-Planck method is to perform multiple, stochastic simulations of Eqs. (2)-(4) [65]. This would however require dealing with various issues in processing noisy data to calculate time-dependent PDFs and information diagnostics accurately, which is circumvented in the Fokker-Planck method.

C. Marginal PDFs

Integrating Eq. (5) over y gives us the Fokker-Planck equation for the marginal PDF $p(x, t) = \int dy\, p(x, y, t)$:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\int dy\ (f\,p) + Q_x\frac{\partial^2 p}{\partial x^2} \equiv -\partial_x \tilde{J}_x. \tag{8}$$

Here, we used the boundary condition $J_y(x, y \to \pm\infty, t) = 0$ ($p(x, y \to \pm\infty, t) = 0$) and defined the effective probability current $\tilde{J}_x$ as

$$\tilde{J}_x = \int dy\, J_x = \int dy\, [f(x, y)p(x, y, t) - Q_x\partial_x p(x, t)]. \tag{9}$$

By using $p(x, y, t) = p(y|x, t)p(x, t)$ where $p(y|x, t)$ is a conditional PDF of y for given x at time t, we have

$$\tilde{J}_x = \int dy\, [f(x, y)p(y|x, t)p(x, t) - Q_x\partial_x p(x, t)] \equiv \tilde{f}p(x, t) - Q_x\partial_x p(x, t), \tag{10}$$

where $\tilde{f} \equiv \int dy\, f(x,y)p(y|x,t)$. We note that $\tilde{f}$ cannot be calculated without knowing $p(y|x)$ or $p(x,y,t)$.

Similarly, the x integral of Eq. (5) and $J_y(x, y \to \pm\infty, t) = 0$ ($p(x \to \pm\infty, y, t) = 0$) yields ($p(y,t) = \int dx\, p(x,y,t)$)

$$\frac{\partial p(y,t)}{\partial t} = -\frac{\partial}{\partial y}(g\, p) + Q_y \frac{\partial^2 p}{\partial y^2} \equiv -\partial_y \tilde{J}_y. \tag{11}$$

Here,

$$\begin{aligned}\tilde{J}_y &= \int dx\, J_y = \int dx\, [g(x,y)p(x,y,t) - Q_y \partial_y p(y,t)] \\ &= \int dx\, [g(x,y)p(x|y,t)p(y,t) - Q_y \partial_y p(y,t)] \\ &\equiv \tilde{g} p(y,t) - Q_y \partial_y p(y,t),\end{aligned} \tag{12}$$

where $\tilde{g} \equiv \int dx\, g(x,y)p(x|y,t)$. We note that $\tilde{f}$ and $\tilde{g}$ in Eqs. (10) and (12) are formally introduced to help the discussions in the following section III.

III. ENTROPY-RELATED STATISTICAL DIAGNOSTICS

Entropy measures disorder or the lack of information [25, 66–70]. While in equilibrium it is well-defined with a clear thermodynamic meaning of the heat flow measured in units of the absolute temperature, its meaning is less clear in non-equilibrium. Nevertheless, entropy and its related concepts like entropy production, entropy flow, mutual information, and information flow are among the most popular and often invoked for characterizing plasma states and dynamics in non-equilibrium processes. These quantities are defined in §III A-III B and used for our analysis in §V.

Note that in practical settings, there have been various attempts to estimate entropy directly from time-series [73, 74] when a PDF is not available. For instance, [75] utilized permutation entropy to analyze non-stationary fusion plasma data for characterizing plasma states and the detection of changes in plasmas.

A. Entropy, entropy production, entropy flow

For PDFs, we calculate differential entropies S_x, S_y, and S from the marginal PDFs $p(x,t)$, $p(y,t)$, and the joint PDF $p(x,y,t)$, respectively, as follows:

$$S_x = -\int dx\, p(x,t) \ln\left(p(x,t)\right), \tag{13}$$

$$S_y = -\int dy\, p(y,t) \ln\left(p(y,t)\right), \tag{14}$$

$$S = -\int dxdy\, p(x,y,t) \ln\left(p(x,y,t)\right). \tag{15}$$

It is useful to note that the values of S, S_x and S_y change under the coordinate transformation (e.g. under $E_M = y^2 \to y$), while being independent of a linear translation (e.g. $y \to y + \text{const}$). This means that entropy cannot be used to detect large events associated with the sudden change in the mean value or advection (such as the emergence of large vortices). Furthermore, they take both negative and positive values, and only their relative value is meaningful.

The time-derivatives of Eqs. (13)-(15) relate entropy to the total entropy production rate $\dot{S}_T$ and flow rate $\dot{S}_m$. Specifically, we use Eq. (8) [$\dot{p} = -\partial_x J_x - \partial_y J_y$, $J_x = fp - Q_x \partial_x p$ and $J_y = gp - Q_y \partial_y p$] and integration by parts to express $\frac{d}{dt} S(x,y,t) = \dot{S}$ as

$$\begin{aligned}\dot{S} &= -\int dxdy \dot{p} \ln p = \int dxdy [\partial_x J_x \ln p + \partial_y J_y \ln p] \\ &= -\int dxdy [J_x \partial_x \ln p + J_y \partial_y \ln p] \\ &= \dot{S}_T - \dot{S}_m,\end{aligned} \tag{16}$$

where

$$\begin{aligned}\dot{S}_T &= \int dxdy \left(\frac{1}{Q_x p} J_x^2 + \frac{1}{Q_y p} J_y^2\right), \\ \dot{S}_m &= \int dxdy \left(\frac{1}{Q_x} J_x f + \frac{1}{Q_y} J_y g\right).\end{aligned} \tag{17}$$

Note that $\dot{S}_T \geq 0$ represents the rate of entropy production due to internal processes. It is non-negative by definition and serves as a proxy of irreversibility. S_m is entropy flux to the environments (heat bath with the temperature Q_x and Q_y). $\dot{S}_T$ and $\dot{S}_m$ consist of the contributions from J_x, f and J_y, g, respectively:

$$\dot{S}_T = \frac{dS_T}{dt} \equiv \dot{S}_{Tx} + \dot{S}_{Ty}, \tag{18}$$

$$\dot{S}_m = \frac{dS_m}{dt} \equiv \dot{S}_{mx} + \dot{S}_{my}, \tag{19}$$

where

$$\dot{S}_{Tx} = \int dxdy \frac{1}{Q_x p} J_x^2, \;\; \dot{S}_{Ty} = \int dxdy \frac{1}{Q_y p} J_y^2, \tag{20}$$

$$\dot{S}_{mx} = \int dxdy \frac{1}{Q_x} J_x f, \;\; \dot{S}_{my} = \int dxdy \frac{1}{Q_y} J_y g. \tag{21}$$

$\dot{S}_{Tx}$ and $\dot{S}_{Ty}$ represent the entropy production rates in x and y, respectively, while $\dot{S}_{mx}$ and $\dot{S}_{my}$ represent the entropy flow rates from x and y to its heat bath, respectively. For instance, $\dot{S}_{mx}$ is positive when the entropy flows from the system x to its environment (at temperature Q_x) while negative when the entropy flows from the environment to the system, and similarly for $\dot{S}_{my}$. In non-equilibrium stationary states where PDFs do not change in time, $\dot{S} = 0$ and $\dot{S}_T = \dot{S}_m$. S_T was often used as a proxy for irreversibility (see [67, 68]).

While Eq. (16) gives $\dot{S}_T = \dot{S}_m + \dot{S}$, a similar relation does not hold in general for the individual x or

y components due to the interaction between x and y. To see this, we calculate $\dot{S}_x$ from the marginal PDF $p(x,t) = \int dy\, p(x,y,t)$ by using Eqs. (8)-(10)

$$\dot{S}_x = -\int dx \partial_t p(x,t) \ln p(x,t) = \int dx\, \partial_x \tilde{J}_x \ln p(x,t)$$
$$= -\int dx\, \tilde{J}_x \partial_x \ln p(x,t), \tag{22}$$

where we recall $\tilde{J}_x = \int dy J_x$. Similarly, from the marginal PDF $p(y,t) = \int dx\, p(x,y,t)$, we obtain

$$\dot{S}_y = -\int dy \partial_t p(y,t) \ln p(y,t) = \int dy\, \partial_y \tilde{J}_y \ln p(y,t)$$
$$= -\int dy\, \tilde{J}_y \partial_y \ln p(y,t). \tag{23}$$

Now, using Eq. (9) in Eq. (22) leads to

$$\dot{S}_x = \int dx \left[\frac{\tilde{J}_x^2}{Q_x p(x,t)} - \frac{\tilde{J}_x \tilde{f}}{Q_x} \right], \tag{24}$$

and similarly for $\dot{S}_y$.

Obviously, the right-hand side of Eq. (24) cannot be written as $\dot{S}_{Tx} - \dot{S}_{mx}$ (see Eqs. (20)-(21)) since $J_x \neq \tilde{J}_x$ in general. It is only when f is independent of y as $f = f(x)$ that $J_x = \tilde{J}_x$ and thus $\dot{S}_x = \dot{S}_{Tx} - \dot{S}_{mx}$. Similarly, it is only when g is independent of y as $g = g(y)$ that $\dot{S}_y = \dot{S}_{Ty} - \dot{S}_{my}$ holds. The importance of the coupling between x and y through $f(x,y)$ or $g(x,y)$ is quantified through mutual information in §III B.

B. Mutual information, information flow

For two independent variables x and y, $p(x,y,t) = p(x,t)p(y,t)$, leading to the equality $S = S_x + S_y$ in Eqs. (13)-(15). In general, $S_x + S_y - S \equiv I \geq 0$, where I is the mutual information [69, 70] defined by

$$I = \int dxdy\, p(x,y,t) \ln \frac{p(x,y,t)}{p(x,t)p(y,t)}. \tag{25}$$

Taking the time derivative of Eq. (25) gives us

$$\frac{dI}{dt} = \int dxdy\, \partial_t p(x,y,t) \ln \frac{p(x,y,t)}{p(x,t)p(y,t)}, \tag{26}$$

since the time-derivative of the term including the logarithm vanishes due to the total probability conservation $\int dxdy\, p(x,y,t) = 1$. Then, Eq. (26) can be recast as follows by using $\partial_t p(x,y,t) = -\partial_x J_x - \partial_y J_y$:

$$\frac{dI}{dt} = T_{y\to x} + T_{x\to y}. \tag{27}$$

Here, $T_{y\to x}$ and $T_{x\to y}$ are information flows from y to x and from x to y, respectively, defined by

$$T_{y\to x} = \int dxdy\, \ln \left[\frac{p(x,t)p(y,t)}{p(x,y,t)} \right] \partial_x J_x(x,y,t)$$
$$= -\int dxdy\, J_x(x,y,t) \partial_x \ln \left[\frac{p(x,t)}{p(x,y,t)} \right]$$
$$(= \partial_\tau I(x(t+\tau), y(t))|_{\tau\to 0}), \tag{28}$$
$$T_{x\to y} = -\int dxdy\, J_y(x,y,t) \partial_y \ln \left[\frac{p(y,t)}{p(x,y,t)} \right]$$
$$(= \partial_\tau I(x(t), y(t+\tau))|_{\tau\to 0}). \tag{29}$$

To relate Eq. (28) to $\dot{S}_{Tx}, \dot{S}_{mx}, \dot{S}_x$, we rewrite it as

$$T_{y\to x} = -\int dxdy\, J_x(x,y,t) \partial_x \ln \left[\frac{p(x,t)}{p(x,y,t)} \right]$$
$$= -\int dxdy\, J_x(x,y,t) \left[\frac{\partial_x p(x,t)}{p(x,t)} - \frac{\partial_x p(x,y,t)}{p(x,y,t)} \right]. \tag{30}$$

Then, using

$$\partial_x p(x,y,t) = \frac{1}{Q_x}[f p(x,y,t) - J_x], \tag{31}$$

and Eqs. (22), (20), and (21), we recast Eq. (30) as

$$T_{y\to x} = \dot{S}_x + \int dxdy\, \frac{J_x}{Q_x}[fp - J_x]$$
$$= \dot{S}_x + \dot{S}_{mx} - \dot{S}_{Tx}, \tag{32}$$

where $\dot{S}_x$ is given in Eq. (22). Thus, Eq. (32) gives us the expression showing how $\dot{S}_{Tx}$ is affected by the interaction between x and y through $T_{y\to x}$:

$$\dot{S}_{Tx} = \dot{S}_x + \dot{S}_{mx} - T_{y\to x}. \tag{33}$$

A similar analysis gives

$$\dot{S}_{Ty} = \dot{S}_y + \dot{S}_{my} - T_{x\to y}, \tag{34}$$

where $\dot{S}_y$ is given in Eq. (23). From Eqs. (18)-(19), we have

$$\dot{S}_T = \dot{S}_{Tx} + \dot{S}_{Ty}$$
$$= (\dot{S}_x + \dot{S}_y) + (\dot{S}_{mx} + \dot{S}_{my}) - [T_{y\to x} + T_{x\to y}]$$
$$= (\dot{S}_x + \dot{S}_y) + (\dot{S}_{mx} + \dot{S}_{my}) - \frac{d}{dt} I(x,y)$$
$$(= \dot{S} + \dot{S}_m), \tag{35}$$

where we used Eqs. (25) and (27) and Eqs. (18)-(21).

In the following, we numerically calculate $\dot{S}_{Tx}$, $\dot{S}_{mx}$, $\dot{S}_{Ty}$, $\dot{S}_{my}$, S, S_x, S_y, and I under different conditions and present some of the results in §V. In particular, special attention is paid to the effect of a sudden change in a model parameter Φ on the entropy production rate and the comparison of their values in forward and back processes.

IV. PATH-DEPENDENT INFORMATION GEOMETRIC DIAGNOSTICS

Our path-dependent information geometry quantifies the temporal change in a PDF by a dimensionless distance along the evolution path. To clarify its meaning, one-variable systems are first considered in §IV A, followed by the multi-variable case in §IV B below.

A. One-variable system

For a stochastic variable x which has a time-dependent PDF $p(x,t)$, the statistical state (x) takes a continuous value, making it difficult to count the number of different states among two PDFs. A convenient way of doing this is to calculate the information rate Γ, and then its time-integral to obtain the information length $\mathcal{L}(t)$ as follows [45, 52–60]

$$\Gamma^2(t) = \int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t}\right]^2,$$
$$\mathcal{L}(t) = \int_0^t dt_1 \Gamma(t_1). \tag{36}$$

We recall that Γ can also be calculated from the infinitesimal symmetric relative entropy between $p(x,t)$ and $p(x,t+\delta t)$ in the limit $\delta t \to 0$, and that the unit of Γ^{-1} in Eq. (36) is time, representing a dynamical time unit for information change. Alternatively, Γ represents the rate of change in information; the faster the temporal change in a PDF, the larger the information rate Γ is.

$\mathcal{L}(t)$ can be interpreted as the total clock time measured in units of Γ^{-1}, or the total number of statistically different states that x passes through between time 0 and t. By definition, $\mathcal{L}(t) = 0$ at $t = 0$. In simple terms, $\mathcal{L}(t)$ quantifies the cumulative change in $p(x,t)$ taking into account the uncertainty in measuring the mean value due to a finite width of $p(x,t)$. Furthermore, Γ and $\mathcal{L}$ are *invariant under (time-independent) change of variables*, which can thus be directly compared with each other unlike physical variables having different units.

For a Gaussian PDF, Γ^2 can be rewritten in terms of a metric tensor (see Appendix B). The simplest interpretation of Γ and $\mathcal{L}$ can be made in the special case of a Gaussian PDF whose width does not change with time, keeping a constant standard deviation σ. In this limit, $\Gamma(t)$ basically gives us the rate of change of a mean value $\mu(t)$ measured in units of σ as $\Gamma(t) = \left|\frac{1}{\sigma}\frac{d\mu}{dt}\right|$. Thus, $\mathcal{L}(t_F) = \frac{\mu(t_F)-\mu(t=0)}{\sigma}$ for $\mu(t) > 0$ for $t = (0, t_F)$, quantifying the total change in the mean value as a dimensionless number. The key point is thus that the evolution path involving a widely distributed PDF will involve smaller information length as a large uncertainty in the broad distribution obscures the information change.

This concept has been applied to different types of problems. In particular, in the relaxation problem where $p(x,t)$ settles into an equilibrium PDF, say, $p(x, t \to \infty)$. In that case, we can quantify the total change in information by $\mathcal{L}(t \to \infty) = \mathcal{L}_\infty$. Since the latter will take different values when the initial condition or parameters change, we can investigate, for instance, how $\mathcal{L}_\infty$ depends on the different initial conditions for the same parameter values in the long time limit to quantify attractor structures (e.g. stable or chaotic attractors) of dynamical system [25, 58].

Finally, for a Gaussian process $p(x,t)$, Γ^2 was shown to be related to $\dot{S}_T$ and $\dot{S}$ as follows [25]:

$$\Gamma^2 = \frac{D}{\sigma^2}\dot{S}_T + \dot{S}^2. \tag{37}$$

Eq. (37) becomes simplified for constant σ as $\Gamma^2 = \frac{1}{\sigma^2}(\partial_t\langle x\rangle)^2 = \frac{D}{\sigma^2}\dot{S}_T$, with a direct proportionality between the entropy production $\dot{S}_T$ and the information geometric diagnostics Γ^2.

B. Multi-variable system

For a system with m variables x_i $(i = 1, 2, .., m)$, we can extend Eq. (36) to [25, 46]

$$\mathcal{L}_{x_i}(t) = \int_0^t dt_1 \Gamma_i(t_1), \tag{38}$$
$$\Gamma^2_{x_i}(t) = \int dx_i \frac{1}{p(x_i,t)} \left[\frac{\partial p(x_i,t)}{\partial t}\right]^2, \tag{39}$$

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. Note that Γ_{x_i} and $\mathcal{L}_{x_i}$ depend on the path of x_i, and the correlation or causality among different variables can be inferred by comparing Γ_{x_i} for different x_i, as shown in [45, 46].

Specifically, for $m = 2$ with the joint PDF $p(x, y, t)$, we obtain

$$\mathcal{L}(t) = \int_0^t dt_1 \Gamma(t_1), \tag{40}$$
$$\Gamma^2(t) = \int dxdy \frac{1}{p(x,y,t)} \left[\frac{\partial p(x,y,t)}{\partial t}\right]^2. \tag{41}$$

From the marginal PDFs $p(x,t)$ and $p(y,t)$ (Eq. (38) for $x_i = x, y$), we have

$$\mathcal{L}_x(t) = \int_0^t dt_1 \Gamma_x(t_1), \tag{42}$$
$$\mathcal{L}_y(t) = \int_0^t dt_1 \Gamma_y(t_1), \tag{43}$$
$$\Gamma^2_x(t) = \int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t}\right]^2, \tag{44}$$
$$\Gamma^2_y(t) = \int dy \frac{1}{p(y,t)} \left[\frac{\partial p(y,t)}{\partial t}\right]^2. \tag{45}$$

For two independent variables x and y, $\Gamma^2 = \Gamma^2_x + \Gamma^2_y$.

For more than a one-variable system, comparing the time-evolution of Γ or $\mathcal{L}$ can give us a measure of correlation. For instance, the strong correlation between the two switching species was captured by their same evolution of $\mathcal{L}(t)$ [56]. Furthermore, self-regulatory dithering in the L-H transition was manifested by the similar evolution of $\mathcal{L}$'s of zonal flows and turbulence [45, 46]; Γ's calculated from marginal PDFs of zonal flows and turbulence compete with each other. Also, the path-dependence of $\mathcal{L}(t)$ is desirable for measuring hysteresis involved in phase transitions [46, 53] such as the L-H transition.

V. NUMERICAL METHODS AND SETUP

As noted previously, in this paper we solve the Fokker-Planck equation as given in Eqs. (5)-(7). We use second-order finite differences to discretize x and y [45, 46], using grid spacings as small as 10^{-3} in both x and y. For the time-stepping, we use second-order Runge-Kutta with time steps as small as 2×10^{-5}.

Given our definition $E_M = y^2$, y can be interpreted as magnetic field fluctuation. Since both $y > 0$ and $y < 0$ will have a similar effect on ELM dynamics with the same E_M value, we look for the solution $p(x, y, t) = p(x, -y, t)$. This leads to the symmetry condition $\frac{\partial p}{\partial y} = 0$ at $y = 0$. Thus, the Fokker-Planck equation is solved in a half-plane $y \geq 0$. Our computational domain is a 2D-box in $x = [x_{min}, x_{max}]$ and $y = [0, y_{max}]$. We choose/adjust the values of x_{min}, x_{max}, and y_{max} to ensure that the solution becomes sufficiently small at the boundaries of the 2D-box (e.g., $p(x_{max}, y, t) = p(x_{min}, y, t) = p(x, y_{max}, t) = 0$) so that our computational solution from a finite-size box is a good approximation of the solution obtained from the true infinite domain. Specifically, we check that the total probability maintains $\iint p(x, y, t)\, dxdy = 1$ to within 10^{-4}.

As an initial condition, all the processes use a narrow Gaussian PDF with the mean values $\langle x(0) \rangle = 1.2$ and $\langle y(0) \rangle = 0.2$ and standard deviations $\sigma_x(0) = \sigma_y(0) = 0.04$. We fix the parameters $d_0 = 10^{-3}$, $d = 0.1$, $\tilde{P} = 1.05$, and $\lambda = 5$ and consider Φ/d in the range $[0.4, 1.2]$ in order to capture the most interesting dynamics (see Appendix C). Note that in the deterministic model with $\xi = \eta = 0$, as Φ/d increases from 0.4 to 1.2, ELMs gradually change from giant ELMs to grassy ones with the shortening of oscillation period. For noise levels, we explored the range $Q_x, Q_y = 10^{-5} - 10^{-3}$ and here present results for only $Q_y = 10^{-5}$ and $Q_x = 10^{-5}, 3 \times 10^{-5}, 10^{-4}, 3 \times 10^{-4}$.

We have considered the different time-dependences of Φ/d summarized in Table 1. **CT1, CT2, CT3** are the three cases where Φ/d is kept constant at $\Phi/d = 0.4, 0.8, 1.2$, respectively, for time $t = [0, 100]$. In the case of a time-dependent Φ/d, its value is varied between $[0.4, 1.2]$ for time $t = [0, 50]$ or $[0, 100]$. **RU/RD** refers to ramp-up/down where Φ increases/decreases linearly in time, as indicated in Table 1; **RU1** and **RU2** (**RD1** and **RD2**) differ in the rate of the increase (decrease) of Φ/d in time. In comparison, **JU/JD** refers to jump-up/down where the change in Φ occurs in steps; **JU1** and **JU2** (**JD1** and **JD2**) involves one and five step changes, respectively, in going from $\Phi/d = 0.4$ (1.2) to $\Phi/d = 1.2$ (0.4). We note that **JU/JD** cases are intended to mimic a step-wise increase in neutral beam power in experiments.

TABLE I. Summary of different cases

Cases	$\Phi(t)/d$ time-dependence and time interval
CT1-3	Constant at 0.4, 0.8, 1.2 for $t = [0, 100]$
Forward	Increase over time as $0.4 \rightarrow 1.2$
RU1	$0.4 + 0.8t/100$ for $t = [0, 100]$
RU2	$0.4 + 0.8t/50$ for $t = [0, 50]$
JU1	0.4 for $t = [0, 50]$, then 1.2 for $t = [50, 100]$
JU2	$0.4, 0.6, 0.8, 1.0, 1.2$ for ten time units for $t = [0, 50]$
Back	Decrease over time as $1.2 \rightarrow 0.4$
RD1	$1.2 - 0.8t/100$ for $t = [0, 100]$
RD2	$1.2 - 0.8t/50$ for $t = [0, 50]$
JD1	1.2 for $t = [0, 50]$, then 0.4 for $t = [50, 100]$
JD2	$1.2, 1.0, 0.8, 0.6, 0.4$ for ten time units for $t = [0, 50]$

For our conventions, the four cases where Φ increases with time are referred to as *forward processes* (say, denoted by a PDF p_F) while the four cases where Φ decreases with time are *back processes* (denoted by a PDF p_B). $\Phi(t)$ is mirror-symmetric in time such that $\Phi(t)$ for the forward process is the same as $\Phi(t_m)$ for the back process where t_m is the mirror image time.

VI. RESULTS

In the following subsections §VI A-VI D and Appendix D, we selectively discuss results for different diagnostics. Specifically, the results **CT1-3** with constant Φ are shown in Figures 14-16 in Appendix D while some of the highlights from time-dependent Φ are presented in §VI A-VI D and Appendix D.

It is important to note that unlike a deterministic model, stochastic noise causes random trajectories and consequently phase-mixing with the loss of phase information over time. This leads to the decrease in ELM oscillation amplitude over time – the so-called amplitude death in nonlinear dynamical systems (see Figures 14-16). One consequence is that for constant Φ, $p(x, y, t)$ starting with a given initial PDF will completely lose the phase information after a sufficiently long time and settle into a stationary PDF (whose spatial form is determined by parameter values). This signifies the complete ELM suppression. The time it takes to reach a stationary PDF depends on Q_x and Q_y. In addition, the frequency of ELM oscillation is also affected by stochastic noise and

becomes shorter in the early evolution for $\Phi/d = 0.4$ in Figure 14, reminiscent of ELM mitigation.

A. JU1 and JD1

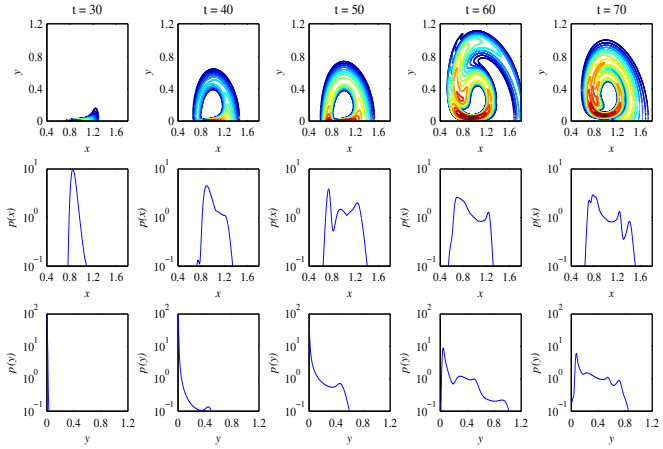

FIG. 1. **JU1**: Joint PDFs, marginal PDFs at different times. The top row shows the joint PDFs $p(x, y, t)$ at the times indicated above each panel. Contours are on a logarithmic scale, starting at 10^{-2} and doubling with each successive contour line. The second row shows the corresponding marginal PDFs $p(x, t)$, and the third row $p(y, t)$.

To highlight the impact of a sudden change in Φ on time-dependent PDFs, we start with the cases with the most abrupt change in Φ/d, **JU1** and **JD1**. Figures 1-2 present the snapshots of joint PDF $p(x, y, t)$ and marginal PDFs $p(x, t)$ and $p(y, t)$ for $Q_x = 10^{-5}$ at each ten time units $t = 30, 40, 50, 60, 70$. The latter is chosen to highlight the evolution of PDFs around $t = 50$ when the sudden step change in Φ occurs. In the contour plots shown in the first row, the red (blue) color denotes high (low) probability. The PDF evolution between $t = 30$ and $t = 50$ is rather monotonic in both figures, involving the gradual spreading of a PDF due to phase-mixing caused by random trajectories while Φ is kept constant. In contrast, much more complicated PDF shapes are seen at $t = 60$ caused by a step change in Φ at $t = 50$. The latter tentatively drives a system further into non-equilibrium, causing large fluctuations or intermittency. Such behavior is reflected in time-series in Figures 3-4, as shall be seen below.

If the system were to equilibrate quickly at any time to a quasi-stationary state, we would expect $p_F(t = 30) = p_B(t = 70)$, $p_F(t = 40) = p_B(t = 60)$, and $p_F(t = 50) = p_B(t = 50)$ with no hysteresis, the back process being a mirror image of the forward process. However, a quick glance at Figures 1-2 immediately reveals that this is not the case. Furthermore, the marginal PDFs $p(x, t)$ and $p(y, t)$ in Figures 1-2 behave very differently. For instance, the forward process in Figure 1 involves much more complicated PDF structures with

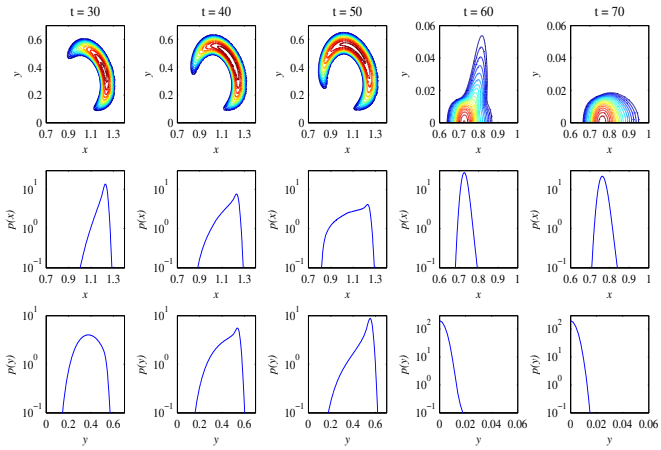

FIG. 2. **JD1**: As in Figure 1, joint and marginal PDFs at different times as indicated above the columns.

multiple peaks and widely distributed (signifying large fluctuations) while PDFs in Figure 2 are narrower with a single peak. (Note that the x and y ranges in Figure 2 are smaller than in Figure 1.)

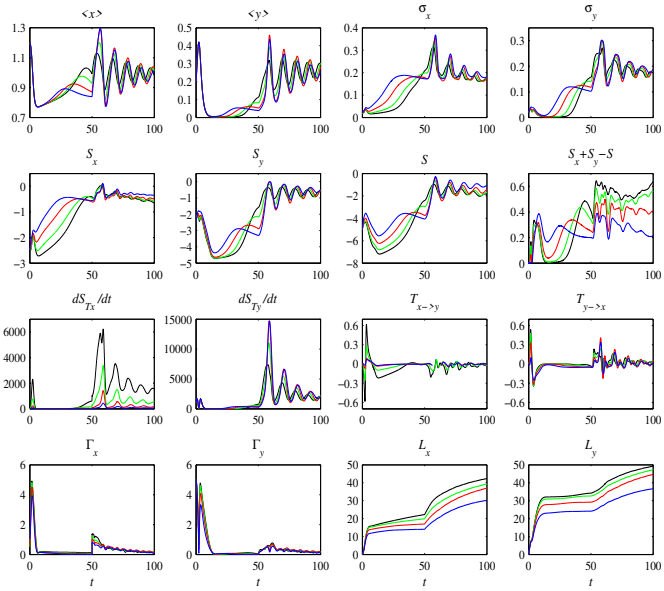

FIG. 3. **JU1**: $\Phi = 0.04$ for $t = [0, 50]$, then $\Phi = 0.12$ for $t = [50, 100]$. Time-traces of $\langle x\rangle$, $\langle y\rangle$, σ_x, σ_y, S_x, S_y, S, $I = S_x + S_y - S$, $\dot{S}_{Tx}$, $\dot{S}_{Ty}$, $T_{x\to y}$, $T_{y\to x}$, Γ_x, Γ_y, $\mathcal{L}_x$, $\mathcal{L}_y$ as labelled above each panel. $Q_x = 10^{-5}$, 3×10^{-5}, 10^{-4}, 3×10^{-4} color-coded as black, green, red, blue, respectively.

Corresponding to Figures 1-2, respectively, Figures 3-4 show time-traces of $\langle x\rangle$, $\langle y\rangle$, σ_x, σ_y in the first row, S_x, S_y, S, $I = S_x + S_y - S$ [Eqs. (13)-(15), (25)] in the second row, $\dot{S}_{Tx}$, $\dot{S}_{Ty}$, $T_{x\to y}$, $T_{y\to x}$ [Eqs. (18), (20), (28)-(29)] in the third row, and Γ_x, Γ_y, $\mathcal{L}_x$, $\mathcal{L}_y$ [Eqs. (42)-(45)] in the last row. Different colors are used for different $Q_x = 10^{-5}$ (black), 3×10^{-5} (green), 10^{-4} (red), and 3×10^{-4} (blue).

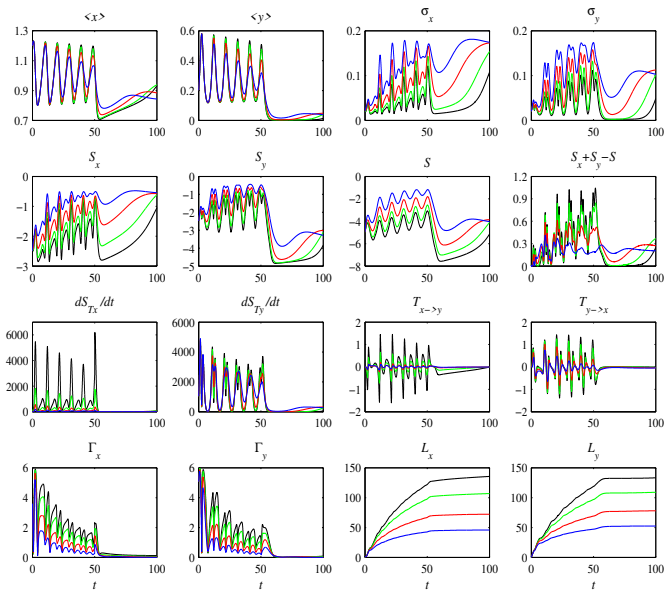

FIG. 4. **JD1**: $\Phi = 0.04$ for $t = [0, 50]$, then $\Phi = 0.12$ for $t = [50, 100]$. Labelling and color-coding as in Figure 3.

Overall, the larger Q_x is, the larger σ_x, σ_y, S_x, and S_y and the smaller $\langle x\rangle$, $\langle y\rangle$, and I become. This means: i) even though Q_x acts directly only on x, it effectively increases stochasticity in y through the coupling between x and y; ii) the phase-mixing by stochastic noise obviously makes the trajectories more unpredictable, with large uncertainties in x and y and smaller I (less correlation) while damping the amplitude of oscillations. Furthermore, $\langle y\rangle$, σ_x, σ_y, S_x, S_y ($\dot{S}_{mx}$, $\dot{S}_{my}$ - not shown), $\dot{S}_{Tx}$, $\dot{S}_{Ty}$ tend to be all in phase while $\langle y\rangle$ and $\langle x\rangle$ are out of phase, the peak of $\langle x\rangle$ appearing before that of $\langle y\rangle$. This is due to the linear instability of y driven by x.

Now, the step increase in Φ at $t = 50$ is seen to increase not only the oscillation amplitude but also the fluctuation level, causing the peaking up of σ_x, σ_y, S_x, and S_y around the time $t \sim 60$. Interestingly, in all cases that we investigated, $\dot{S}_{Tx}$, $\dot{S}_{mx}$, $\dot{S}_{Tx}$, and $\dot{S}_{Ty}$ are significantly larger than $T_{x\to y}$ and $T_{y\to x}$. That is, $\dot{S}_{Tx}$ and $\dot{S}_{mx}$ ($\dot{S}_{Ty}$ and $\dot{S}_{my}$) take quite similar values that are much larger than $T_{x\to y}$. For this reason, the results for $\dot{S}_{mx}$ and $\dot{S}_{my}$ are not included in the figures in this paper.

$T_{x\to y}$ and $T_{y\to x}$ exhibit rapid temporal changes when I is large. Their signs tend to be opposite. Note that a negative sign of $T_{x\to y}$ means the information flows from y to x, which is equivalent to a positive sign of $T_{y\to x}$. By definition, $T_{x\to y}$ and $T_{y\to x}$ sensitively depend on entropy with no dependence on $\langle x\rangle, \langle y\rangle$. Consequently, they do not detect any changes occuring in mean values [52], for instance, Note also that $T_{x\to y}$ and $T_{y\to x}$ take zero values at $t = 0$ due to the lack of correlation in the initial condition, and thus does not reveal the instantaneous statistical property at $t = 0$.

Now, comparing Figures 1-2, we see that the forward process (towards instability) is driven further from equilibrium by perturbation, with larger $\sigma_x, \sigma_y, S_x, S_y$, etc. In particular, $\dot{S}_{Ty}$ shoots up to a much larger value (~ 15000) in **JU1** compared with **JD1** (~ 6000). However, the opposite tendency is observed for the values of $T_{x\to y}$ and $T_{y\to x}$ which are larger in **JD1** than **JU1**. This can be understood in terms of stronger correlation in the back process than the forward process.

Another useful way of quantifying the hysteresis is to compare the total numbers of different statistical states ($\mathcal{L}_x$ and $\mathcal{L}_y$) in the forward and back processes. The last row shows that **JU1** involves the total change of $\mathcal{L}_x, \mathcal{L}_y$ up to ~ 50. In comparison, **JD1** undergoes the increases of $\mathcal{L}_x, \mathcal{L}_y$ over 130. These demonstrate that the forward and back processes undergo very different changes (evolution) in PDFs, e.g., the former involving more widely distributed PDFs.

Finally, unlike $T_{x\to y}(t = 0) = T_{y\to x}(t = 0) = 0$, $\Gamma_x(t = 0) \neq 0$ and $\Gamma_y(t = 0) \neq 0$ quantify how quickly the PDFs change at $t = 0$ due to an initial non-equilibrium state. Notably, among all different diagnostics, it is Γ_x, Γ_y that detect Φ change earliest, with almost vertical line seen at $t = 50$. This is because the information rate is very sensitive to the temporal change of a PDF caused by Φ change. It is also notable that despite the different evolution of x and y in all other diagnostics, including a PDF, $\mathcal{L}_x$ and $\mathcal{L}_y$ behave quite similarly. This will be further elaborated in §VI D.

B. JU2 and JD2

We now consider the second most drastic cases of **JU2** and **JD2**. The snapshots of joint PDFs $p(x, y, t)$ and marginal PDFs $p(x, t)$ and $p(y, t)$ are shown at time $t = 10, 20, 30, 40, 50$ in Figures 5-6 at the times of the step changes in Φ. Again, the clear differences between forward and back processes are seen as $p_F(t = 10) \neq p_B(t = 40)$, etc. Nevertheless, the disparity in PDFs between the forward and back processes is smaller compared with the cases of **JU1** and **JD1** in Figures 1-2, suggesting less hysteresis. As shall be seen later, a similar trend persists when Φ undergoes less dramatic changes.

Corresponding to Figures 5-6, respectively, Figures 7-8 show time-traces of $\langle x\rangle$, $\langle y\rangle$, σ_x, σ_y in the first row, S_x, S_y, S, $I = S_x + S_y - S$ [Eqs. (13)-(15), (25)] in the second row, $\dot{S}_{Tx}$, $\dot{S}_{Ty}$, $T_{x\to y}$, $T_{y\to x}$ [Eqs. (18), (20), (28)-(29)] in the third row, and Γ_x, Γ_y, $\mathcal{L}_x$, $\mathcal{L}_y$ [Eqs. (42)-(45)] in the last row, following the same format as in Figures 3-4. Again, different colors are used for different $Q_x = 10^{-5}$ (black), 3×10^{-5} (green), 10^{-4} (red), and 3×10^{-4} (blue).

Compared with Figures 1-2, the overall magnitudes of diagnostics measuring randomness σ_x, σ_y, S_x, S_y, $\dot{S}_{Tx}$, $\dot{S}_{Ty}$ are smaller due to less dramatic change in Φ. For instance, while $\dot{S}_{Ty}$ reaches up to 15,000 for **JU1**, its maximum is less than 10,000 for **JU2**.

Among all the diagnostics, it is again the information

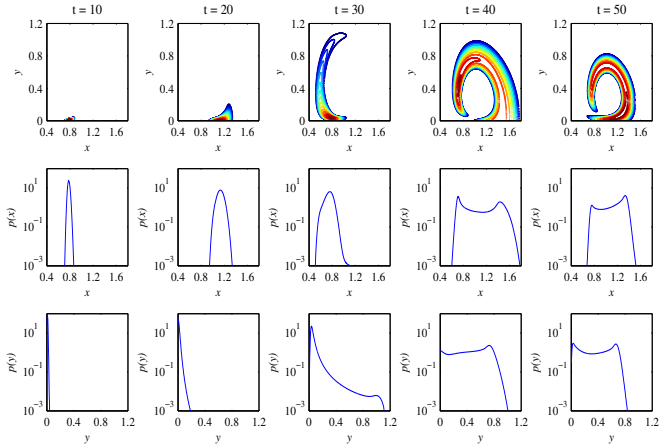

FIG. 5. **JU2**: Joint PDFs, marginal PDFs at different times.

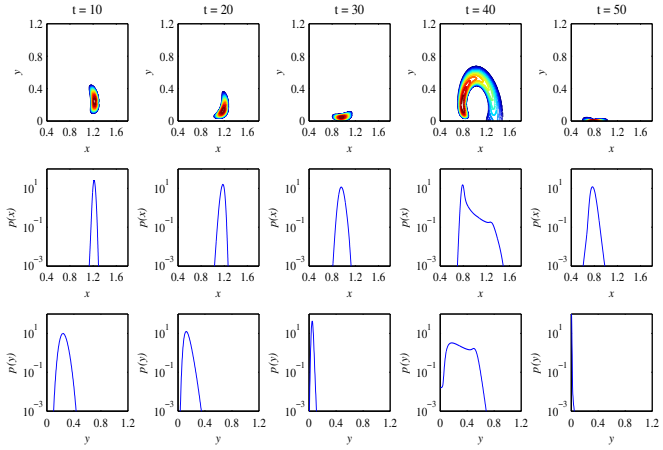

FIG. 6. **JD2**: Joint PDFs, marginal PDFs at different times.

rate (especially, Γ_x at $t = 10, 20, 30$ for **JU2**) that captures the sudden step changes in Φ earliest and most clearly. $\mathcal{L}_x, \mathcal{L}_y$ increase up to $\sim 55, 95$, in Figures 7 and 8, respectively, with about 40 difference. This is much smaller than what was observed for **JU1-JD1**. This is another manifestation of the reduced hysteresis due to less dramatic change in Φ. Finally, we again observe rather remarkably similar evolutions of $\mathcal{L}_x$ and $\mathcal{L}_y$ in both processes.

C. RU1 and RD1

The next dramatic case is the linear ramp-up/down case for time $t = [0, 50]$, whose results are presented in Appendix E. In this subsection, we briefly go over the least dramatic case of a linear ramp-up/down for time $t = [0, 100]$ to highlight the main points. Figures 9-10 show the time-series of various diagnostics in the same format as in Figures 3-4. Overall, we observe similar trends such as similar evolutions of $\mathcal{L}_x$ and $\mathcal{L}_y$, larger values of $\dot{S}_{Tx}$ and $\dot{S}_{Ty}$ in the forward process than those in back process. The fluctuation-related diagnostics (standard deviations and entropy-related diagnostics) are seen to be all reduced compared with the more dramatic cases in §VI A-VI B.

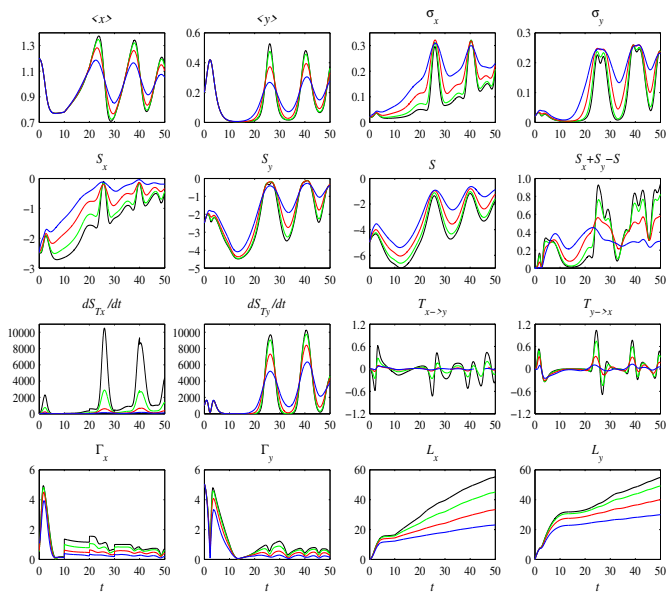

FIG. 7. **JU2**: $\Phi = 0.4, 0.6, 0.8, 1.0, 1.2$, each for ten time units for $t = [0, 50]$

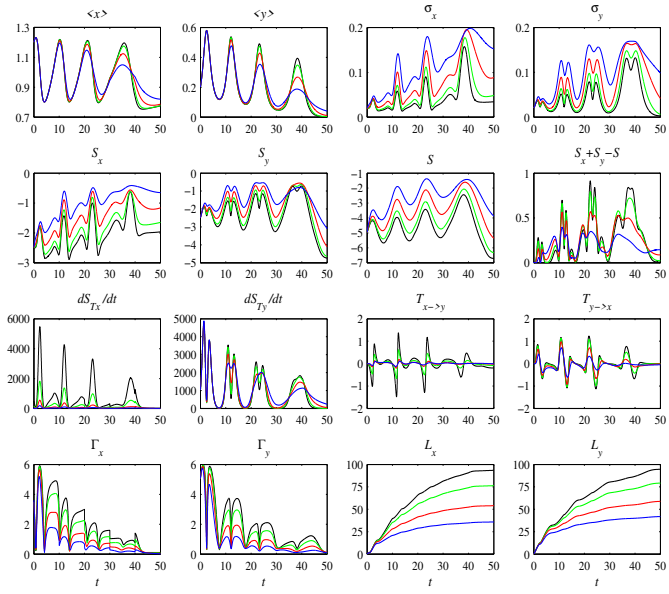

FIG. 8. **JD2**: $\Phi = 1.2, 1.0, 0.8, 0.6, 0.4$, each for ten time units for $t = [0, 50]$

D. Information phase-portrait: Γ_x vs Γ_y

.

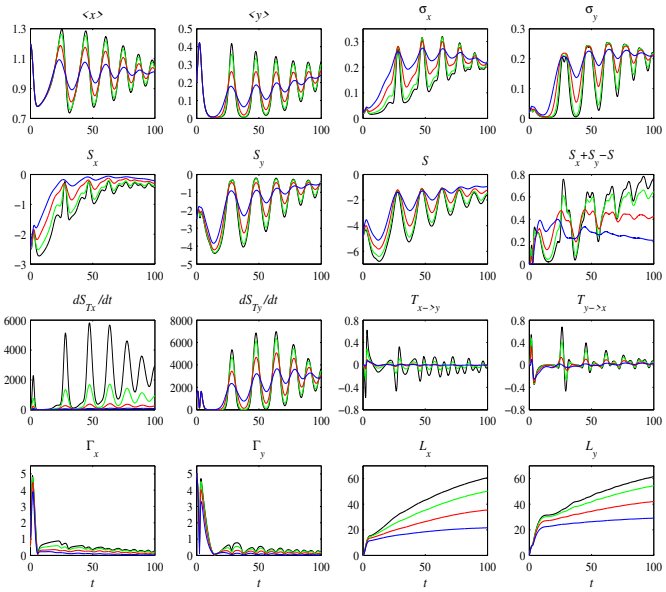

FIG. 9. **RU1**: $\Phi = 0.04 + 0.08t/100$, so increasing from 0.04 to 0.12 over $t = [0, 100]$

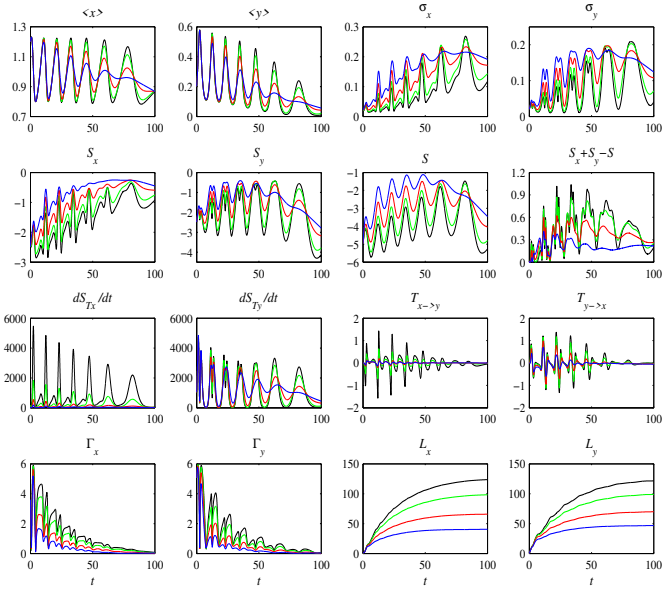

FIG. 10. **RD1**: $\Phi = 0.12 - 0.08t/100$, so decreasing from 0.12 to 0.04 over $t = [0, 100]$

In §VI A-VI C, similar evolution of $\mathcal{L}_x$ and $\mathcal{L}_y$ was noted. Given that all other diagnostics show quite different behavior, this is not a trivial result. It highlights that the coupling (correlation) between x and y can be well captured by information geometry, that is, by the number of different statistical states that each variable passes through despite the differences in the details of the time-evolutions of the two variables [56].

As noted earlier, the forward process is more out of equilibrium and expected to be less correlated compared

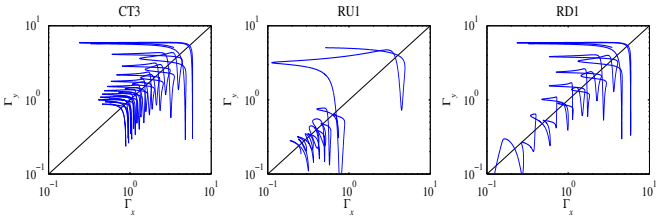

FIG. 11. Γ_x against Γ_y for $\Phi/d = 1.2$ (**CT3**), **RU1**, and **RD1**, from left to right.

with back process. To demonstrate this, Figure 11 shows the information phase-portrait which plots Γ_y against Γ_x. The diagonal line $\Gamma_x = \Gamma_y$ is overplotted in black solid line along which the time-scales of the change in $p(x,t)$ and $p(y,t)$ coincide. The oscillation around $\Gamma_x = \Gamma_y$ reveals regulatory interaction between x and y as a result of overshooting (due to inertial) and restoring forces (due to interaction). In particular, when Γ_x and Γ_y cross each other, the time-scales of x and y match with a perfect balance. In Figure 11, such regulatory behaviour is most prominent in the case of **CT3** for a constant $\Phi/d = 1.2$, then **RD1**. Quite large deviation from $\Gamma_x = \Gamma_y$ is observed for **RU1** (forward process) due to intermittency.

E. Information rate vs entropy-related diagnostics

In §IV A, it was noted that for one Gaussian variable, the information rate can be related to entropy-related diagnostics as in Eq. (37). In general for a coupled nonlinear system, it is unknown what relationship might hold. To gain some insight, it is useful to see how well or badly Eq. (37) would work. To look into this, we use numerically obtained results Γ_x, $\dot{S}_{Tx}$, $\dot{S}_x$ from $p(x,t)$ and Γ_y, $\dot{S}_{Ty}$, $\dot{S}_y$ from $p(y,t)$ and compare the left and right sides of the following Eqs. (46) and (47):

$$\Gamma_x^2 \stackrel{?}{=} \frac{Q_x}{\sigma_x^2}\dot{S}_{Tx} + \dot{S}_x^2, \tag{46}$$

$$\Gamma_y^2 \stackrel{?}{=} \frac{Q_y}{\sigma_y^2}\dot{S}_{Ty} + \dot{S}_y^2. \tag{47}$$

Figure 12 shows the results for the case of **RU2** and **RD2**, with $Q_x = 10^{-5}$. The top row shows results for **RU2**, the bottom row for **RD2**. The first panel in each row shows the x quantities from Eq. (46), and the second panel the y quantities from Eq. (47), as labelled also above each panel. Γ_x^2 and Γ_y^2 are in blue, and the right-hand sides of Eqs. (46) and (47) are in red. In the initial evolution stage, the red and blue lines overlap reasonably, suggesting that Eqs. (46) and (47) are good approximation. The disparity between the left-hand and right-hand sides of Eqs. (46) and (47) gradually increases over time, with the overall tendency of the entropy-related quantities on the right-hand side exceeding the other.

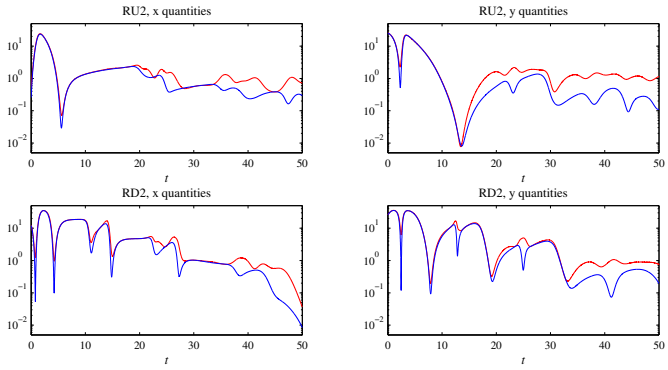

FIG. 12. **RU2-RD2**: The top row shows results for **RU2**, the bottom row for **RD2**. The first panel in each row shows the x quantities from Eq. (46), and the second panel the y quantities from Eq. (47), as labelled also above each panel. Γ_x^2 and Γ_y^2 are in blue, and the right-hand sides of Eqs. (46) and (47) are in red.

VII. CONCLUSION

We presented the detailed investigation into how statistical properties change with time in a stochastic model of ELMs under different conditions. We employed a time-dependent PDF method and various entropy-related quantities (entropy, entropy production, entropy flux, mutual information, and information flow) and path-dependent information geometry (information rate, information length). To quantify hysteresis, we considered the forward and back processes for an input power which is mirror-symmetric in time.

Our principal findings are

i) Stochastic noise has non-trivial effects on ELM dynamics, altering the amplitude and period of oscillations and thereby leading to different evolutions.

ii) Time-dependent model parameter (input power Φ) has a non-trivial effect: Compared with back process, forward processes involve larger fluctuations and are further from equilibrium, with widely distributed PDFs possibly with multiple peaks; alternatively, back processes involve stronger correlation between x and y.

iii) As consequences of ii), standard deviation, entropy, entropy production tend to be larger in the forward than back processes while $T_{x\to y}$ and $T_{y\to x}$ (measuring the interaction) tend to be larger in back than forward processes.

iv) A temporal change in Φ tentatively drives a system further into non-equilibrium and causes large fluctuations and entropy production, leading to intermittency and the amplification of oscillation amplitude.

v) The impact of iv) disappears after a sufficiently long time due to the adjustment of the system via the dissipation, nonlinear interactions, etc. in the system.

vi) The severity of iv) is determined by how quickly Φ changes in time, with the most (instantaneous) dramatic effect expected from **JU1**/**JD1**, then **JU2**/**JD2**, **RU2**/**RD2**, and **RU1**/**RD1**.

vii) Among all different diagnostics, it is information rate (Γ_x, Γ_y) that detects Φ change earliest. This is a reflection of the fact that the information rate is very sensitive to the temporal change of a PDF caused by Φ change.

viii) Information phase-portrait (Γ_x and Γ_y) reveals that a strong coupling between x and y can be measured by the oscillation of Γ_x and Γ_y around $\Gamma_x = \Gamma_y$, suggesting the time competition in statistical space.

ix) The coupling (correlation) between x and y and their self-regulation can be well captured by information geometry, that is, by the number of different statistical states that each variable passes through despite the differences in the details of the time-evolutions of the two variables [56].

Some of the important experimental implications of our findings are as follows:

First, the different evolution of ELMs depending on stochastic noises, together with the effect of stochastic noises on the L-H transition in [45, 46], implies the possibility that stochasticity might be one of the important factors (hidden variables) that cause scatters in experimental data. Specifically, the stochastic noise caused by different sources (e.g., due to fluctuating energy flux of unresolved scales, intermittency such as the outward energy flux at the edge or mini-avalanches, stochastic magnetic fields induced by internal disruption, kinetic instabilities or external magnetic coils, etc.) will differ among experiments because of different configurations (e.g., divertor, q-profile, etc.). To understand those scatters, it would be valuable to perform (time-dependent) statistical analysis of experimental data.

Second, non-trivial effects of the temporal form of the input power call for attention on the importance of time scheduling of an input power, e.g., to minimize irreversible heat generation/loss or hysteresis.

Third, high-resolution experimental data would be valuable for statistical analysis as time-dependent PDFs and information geometric diagnostics can be calculated by sampling different (temporal and/or spatial) selections of data, e.g., by moving-time windows, kernel density estimation, etc. [55, 57, 59, 60].

It would be of interest to extend the work further to investigate if similar conclusions can be drawn in other models, e.g. by including a poloidal flow contribution to $E \times B$ flows, coherent magnetic structure [6], or by considering different models with non-zero turbulence

in the H-mode [77]. This will require the investigation of higher-dimensional models including three or more stochastic variables, which will need much more expensive computational resources and time. For instance, the Langevin equation for three stochastic variables can be simulated by solving the 3+1 partial differential equation (Fokker-Planck equation). An alternative method of direct stochastic simulations (e.g. in [65]) will be further developed to investigate these higher-dimensional models. It will also be interesting to explore other information theoretical tools such as transfer entropy (e.g. see [3] and references therein) or else apply our methods to analyze experimental data [59, 60] or computational data from fluid turbulence models [55, 57].

In conclusion, we provided a theoretical framework to understand the effects of uncertainty in ODE models while demonstrating non-trivial effects of stochastic noise and time-varying model parameter and potential utility of the path-dependent information geometry in understanding ELM dynamics in fusion plasmas. As is often the case, it seems ironic that while a desirable, H-mode for fusion plasma is a quiet state (low turbulence), it comes with the cost of setting off instabilities which possibly cause significant damage. The mitigation/suppression of such instabilities via stochastic perturbations and consequent increase in turbulence suggests that an ideal fusion operation might not be a too quiet state, but instead a moderately turbulent state which compromises between instabilities (developed in a quiet state) and turbulence activity, with a subtle balance between the two. The latter might be described as *self-organised* state which evolves dynamically maintaining such a subtle balance.

Appendix A: K-L divergence

The Kullback-Leibler (K-L) divergence – relative entropy – has proven to be useful in understanding irreversibility and non-equilibrium thermodynamic inequality relations (e.g, see [68] and references therein). K-L divergence between the two PDFs [66] is defined by

$$\mathcal{K}(p_1|p_2) = \int dx\, p_1 \ln\left(\frac{p_1}{p_2}\right). \tag{A1}$$

A symmetric version of K-L divergence is $\frac{1}{2}\left[\mathcal{K}(p_1|p_2) + \mathcal{K}(p_2|p_1)\right]$. When a parameter (say, $\Phi(t)$) of a system changes in time, at any instant of time t_1, a time-dependent PDF $p(x, y, t = t_1)$ is not necessarily the same as what is expected from a stationary solution (say, $p(x, y, t \to \infty)$ obtained by using the fixed value of $\Phi(t_1)$ in the long time limit). This non-equilibrium characteristics contributes to irreversibility and hysteresis.

Furthermore, for a time-varying parameter which is mirror-symmetric in time, with the forward and back PDFs, $p_F(t)$ and $p_R(t)$), respectively, the dissipated work was shown to be related to the relative entropy

$$W_D = D\mathcal{K}[p_F(t)|p_R(t_m))]. \tag{A2}$$

(e.g., see [76]), where t_m is the mirror image time such that $\Phi(t)$ for the forward process is the same as $\Phi(t_m)$. This manifests that the disparity between $p_R(t)$ and $p_R(t_m)$ is linked to irreversibility (dissipation) and hysteresis.

Appendix B: Γ for Gaussian PDFs

Γ can be related to Fisher information [66] for a Gaussian PDF which is fixed by the two parameters, the mean value μ and standard deviation σ. In terms of these parameters $\lambda_i = (\mu, \sigma)$, Γ can be shown to be

$$\Gamma^2 = g_{ij}\partial_t\lambda_i\partial_t\lambda_j. \tag{B1}$$

Here, g_{ij} is the metric tensor defined in the statistical space defined by λ_i

$$g_{ij}(t) = \int dx p(x,t)\partial_{\lambda_i}(\ln p)\,\partial_{\lambda_j}(\ln p) = \frac{1}{\sigma^2}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}. \tag{B2}$$

The non-zero components of g_{ij} are inversely proportional to the variance σ^2. That is, a broad PDF with large σ leads to smaller $\Gamma(t)$ (and $\mathcal{L}(t)$).

Appendix C: ODE model of ELMs

The ODE ELM model in [26] far above the threshold power consists of the evolution of the pressure gradient $P = x$ and magnetic fluctuation amplitude E_M:

$$\frac{dP}{\partial t} = \Phi - \tilde{D}(P)P - E_M P, \tag{C1}$$

$$\frac{dE_M}{\partial t} = \lambda(P-1)E_M, \tag{C2}$$

$$\tilde{D} = d_0 + d(P - c^2P^4)\Theta(\tilde{P} - P). \tag{C3}$$

Here $c = \tilde{P}^{-\frac{3}{2}}$ is constant; $\Theta(x)$ is the Heaviside function with $\Theta(x) = 1$ for $x \geq 0$ and $\Theta(x) = 0$ for $x < 0$; Φ is the control parameter representing the input power; λ and $d_0 \ll d$ are non-negative constants. $\tilde{P}$ is the critical pressure gradient for a complete suppression of turbulence due to shear. We note that no chaotic solutions were found in the original ODE ELM model in [26] unlike in [6].

Fig. 13 shows the time-evolution of x and y from Eqs. (C1)-(C3) using larger x ranges for Φ because the period of ELMs decreases with Φ. $\Phi/d = 0.4$ supports giant ELMs with a long period and a short duration of magnetic activity. They become more grassy for larger Φ/d with a shorter period and a longer duration of magnetic activity. The period of oscillation and the duration

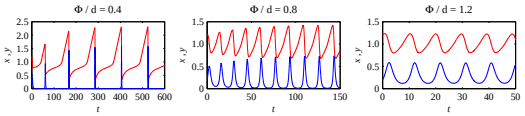

FIG. 13. Deterministic solution for $\Phi/d = 0.4, 0.8, 1.2$, from left to right. x is red, y is blue.

of magnetic activity become similar for $\Phi/d = 0.12$. Alternatively, for $\Phi/d = 0.4$, y stays nearer the unstable L-mode solution ($y = 0$), followed by large deviation in a large burst.

The three different types of the fixed points were reported in [26] depending on the value of $\tilde{P}$ as the control parameter Φ changes.

i) Case 1: When $\tilde{P} > \left(\frac{5}{2}\right)^{1/3}$, there are two stable states: a) the L-mode with $E_M = 0$ if $\Phi < d_0 + d(1 - \tilde{P}^{-3})$, b) the L-mode with $E_M \neq 0$ if $\Phi > d_0 + d(1 - \tilde{P}^{-3})$. This corresponds to $\Phi/d > 0.146d$ for our parameter values $\tilde{P} = 1.05$, $d_0 = 0.001$, $d = 0.1$. [At Φ_{max} and $E_M = 0$, $\Phi = \tilde{D}p_0 = (d_0 + d(P_0 - c^2 P_0^4))$. For $P_0 = 1$, $\Phi_{max} = d_0 + d(1 - c^2) = d_0 + d(1 - \tilde{P}^{-3})$.]

ii) Case 2: a) For $1 < \tilde{P} \leq \left(\frac{5}{2}\right)^{1/3}$, stable L-mode state with $E_M = 0$ if $\Phi < [d_0 + \left(\frac{2}{5}\right)^{4/3} \tilde{P}d] \left(\frac{2}{5}\right)^{1/3} \tilde{P}$. However, this criterion seems approximate. For instance, for the parameter values to be used in this paper $\tilde{P} = 1.05, d_0 = 0.001d = 0.1$, this gives $\Phi/d = 0.247$ for the L-mode exists while the L-mode persisted up to at least $\Phi/d = 0.33d$ from our simulations. b) the H-mode state with $p_0 = 1$ and $E_M \neq 0$ if $\Phi > [d_0 + \left(\frac{2}{5}\right)^{4/3} \tilde{P}d] \left(\frac{2}{5}\right)^{1/3} \tilde{P}$. When $d_0 \ll d$, the H-mode has MHD activity.

iii) Case $\tilde{P} \leq 1$, no ELMs; the L-mode, marginally stable H-mode with $E_M \neq 0$, and the stable H-mode without MHD acitivity for $d_0 > \Phi > d_0 \tilde{P}$.

We will focus on the Case 2 with $E_M \neq 0$ in this paper.

These three different cases above are based on how the nonlinear flux/diffusion $(\tilde{D} - d_0)P \equiv F(P)$ in Eq. (C4) behaves as

- $F(P) = d(P - c^2 P^3)P = 0$ when $P = \tilde{P} = c^{-2/3}$
- $F(P) \neq 0$ between $0 < P < c^{-2/3}$
- The maximum of F occurs when $\frac{dF}{dP}|_{P=P_*} = 0$. If $P_* > 1$, as P increases from $P = 1$ (one of the equilibrium point for $E_M \neq 0$), a larger damping $F(p)$ brings P back to $P \sim 1$. Since $\frac{dF}{dP} = d(2P - 5c^2 P^4) = 0$ occurs at $P_*^3 = \frac{2}{5c^2} = \frac{2}{5}\tilde{P}^3$, $\tilde{P} > \left(\frac{5}{2}\right)^{1/3}$.
- If $\frac{dF}{dP} = d(2P - 5c^2 P^4) = 0$ occurs for $P_*^3 = \frac{2}{5c^2} = \frac{2}{5}\tilde{P}^3 < 1$. Thus, $\tilde{P} < \left(\frac{5}{2}\right)^{1/3}$. In addition, a nonzero $\Theta(\tilde{P} - P)$ for $P = 1$ requires $\tilde{P} \geq 1$. Thus, $1 \leq \tilde{P} < \left(\frac{5}{2}\right)^{1/3}$.
- If the equilibrium point $P = 1$ is larger than $\tilde{P}$ ($\tilde{P} < 1$), no nonlinear diffusion at $P = 1$.
- When $\tilde{D} = d_0$, the fixed points satisfy

$$\Phi = (d_0 + E_M)P, \quad (P - 1)E_M = 0 \tag{C4}$$

Thus, $P = 1$ or $E_M = 0$. When $P = 1$, $E_M = \Phi - d_0$. When $E_M = 0$, $P = \frac{\Phi}{d_0}$ (which may not be stable).

Appendix D: CT1-3 with constant $\Phi/d = 0.4, 0.8, 1.2$

Figures 14-16 show the three cases of constant $\Phi/d = 0.4, 0.8, 1.2$, respectively.

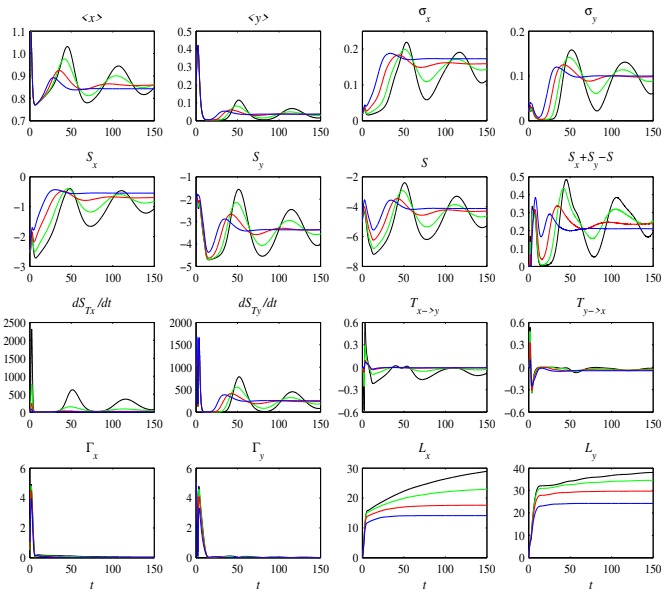

FIG. 14. Constant $\Phi = 0.04$, and plotted up to $t = 100$.

From Figures 14-16, we can see $\langle y \rangle$, σ_x, σ_y, S_x, S_y, $\dot{S}_{mx}$, $\dot{S}_{my}$, $\dot{S}_{Tx}$, $\dot{S}_{Ty}$ are all in phase, etc.

The mutual information $I = S_x + S_y - S > 0$ when these quantities are maximum with the strongest magnetic activity while $I < 0$ when magnetic activity is minimum. That is, magnetic activity promotes the information change between x and y. When I is maximum, the information flow $T_{x\to y}$ and $T_{y\to x}$ exhibit rapid temporal changes. Their signs tend to be opposite. Note that a negative sign of $T_{x\to y}$ means the information flows from y to x, which is equivalent to a positive sign of $T_{y\to x}$. By definition, $T_{x\to y}$ and $T_{y\to x}$ sensitively depend on entropy with no dependence on $\langle x \rangle, \langle y \rangle$. Both $T_{x\to y}$ and $T_{y\to x}$ take zero values at $t = 0$ due to the lack of correlation.

$\dot{S}_{Tx}$ and $\dot{S}_{Ty}$ represent how far the system is from equilibrium. The larger their values, further from the equilibrium. $\dot{S}_{mx} > 0$ and $\dot{S}_{my} > 0$ the entropy flow always from the system (x and y) to its environment (due to ξ and η).

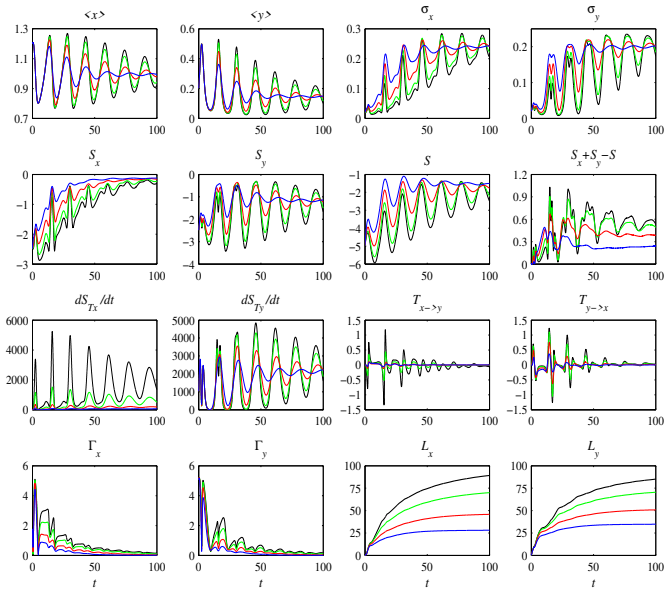

FIG. 15. Constant $\Phi = 0.08$, and plotted up to $t = 100$.

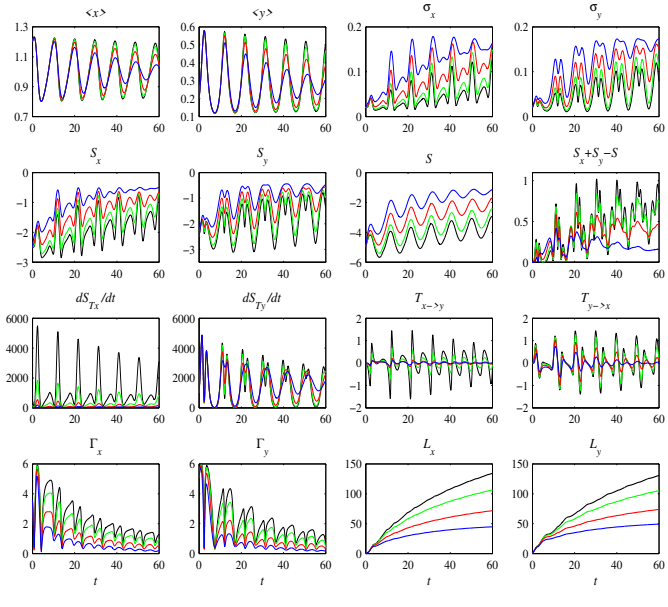

FIG. 16. Constant $\Phi = 0.12$, and plotted up to $t = 100$.

Appendix E: RD2 and RD2

We now look at the case of the linear ramping cases **RD2** and **RD2** in a shorter time interval $t = [0, 50]$ (than **RD2** and **RD2**). That is, the ramping up/down rates are faster than the cases shown in Figures 9-10. Figures 17-18 show the time-series of various diagnostics in the same format as in Figures 17-18. Figure 17 is the case of **RD1** where $\Phi = 0.12 + 0.08t/100$, so decreasing from 0.12 to 0.04 over $t = [0, 50]$; Figure 18 is the case of **RD2** where $\Phi = 0.12 - 0.08t/100$, so decreasing from 0.12 to 0.04 over $t = [0, 50]$.

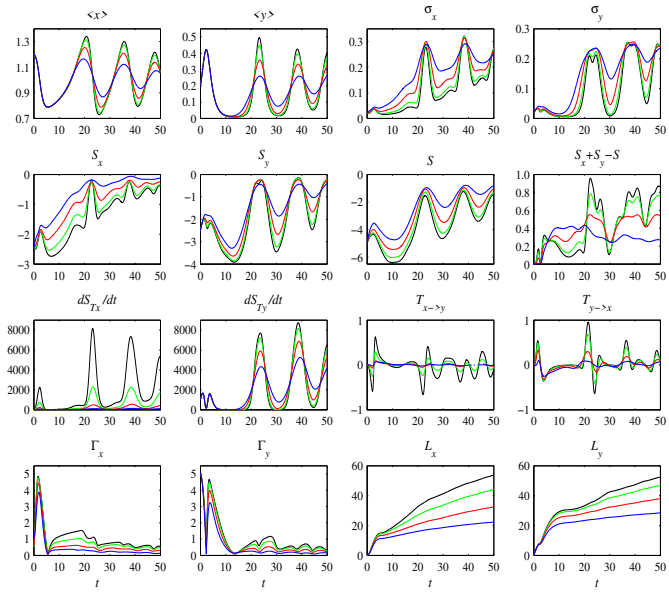

FIG. 17. **RU2**: $\Phi = 0.04 + 0.08t/50$, so increasing from 0.04 to 0.12 over $t = [0, 50]$

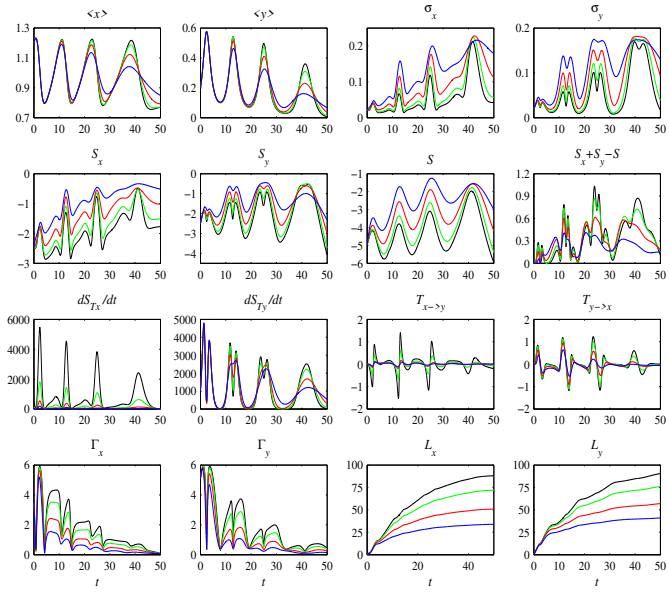

FIG. 18. **RD2**: $\Phi = 0.12 - 0.08t/50$, so decreasing from 0.12 to 0.04 over $t = [0, 50]$

CONFLICT OF INTEREST: The authors have no conflicts to disclose.

ACKNOWLEDGEMENTS: We thank Dr Yasmin Andrews for helpful discussions.

DATA AVAILABILITY STATEMENTS: The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] B. Kadomtsev and E.W. Laing, *Tokamak Plasma: A Complex Physical System* (IOP Publishing Ltd., 1992).

[2] C. Hildago, M.A. Pedrosa, and B. Goncalves, Fluctuations, sheared radial electric fields and transport interplay in fusion plasmas, New J. Phys. **4**, 51 (2002).

[3] B. P. van Milligen, B. A. Carreras, I. Voldiner, U. Losada, C. Hidalgo, and TJ-II Team, Causality, intermittence, and crossphase evolution during confinement transitions in the TJ-II stellarator, Phys. Plasmas **28**, 092302 (2021).

[4] K. Itoh, Y. Nagashima, S.-I. Itoh, P. H. Diamond, A. Fujisawa, M. Yagi, and A. Fukuyama, On the bicoherence analysis of plasma turbulence, Phys. Plasmas **2**, 102301 (2005).

[5] B.Ph. van Milligen, C. Hidalgo, E. Sánchez, M. A. Pedrosa, R. Balbín, and I. García-Cortés, Statistically robust linear and non-linear wavelet analysis applied to plasma edge turbulence, Rev. Sci. Instrum. **68**, 967 (1997).

[6] A. Thyagaraja, F. A. Haas, and D. J. Harvey, A nonlinear dynamic model of relaxation oscillations in tokamaks Phys. Plasmas **6**, 2380 (1999).

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

[8] Y. Sarazin and P. Ghendrih, Intermittent particle transport in two-dimensional edge turbulence, Phys. Plasmas **5**, 4214 (1998).

[9] K. Itoh, S.-I. Itoh, A. Fukuyama, and M. Yagi, Theory of plasma turbulence and structural formation – nonlinearity and statistical view, J. Plasma Fusion Res. **79**, 608 (2003).

[10] J. Wang, G.R. Tyne, R. Hong, L. Nie, Y. Chen, R. Ke, T. Wu, T. Long, P. Zheng, M. Xu, and HL-2A Team, Edge turbulence evolution and intermittency development near the density limit on the HL-2A tokamak, Phys. Plasmas **26**, 092303 (2019).

[11] 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).

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

[13] 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).

[14] 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).

[15] 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).

[16] 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).

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

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

[19] 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).

[20] 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).

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

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

[23] A. M. Kowalski, M. T. Martin, A. Plastino, O. A. Rosso, and M. Casas, Distances in probability space and the statistical complexity setup, Entropy **13**, 1055 (2011).

[24] M. Oizumi, N. Tsuchiya, and S. Amari, Unified framework for information integration based on information geometry, Proc. Nat. Acad. Sci. **113**, 14817 (2016).

[25] E. Kim, Information geometry, fluctuations, non-equilibrium thermodynamics, and geodesics in complex systems, Entropy **23**, 1393 (2021).

[26] V. B. Lebedev, P. H. Diamond, I. Gruzinov, and B. A. Carreras, A minimal dynamical model of edge localized mode phenomena, Phys. Plasmas **2**, 3345 (1995).

[27] H. K. Park, J. Lee, G. S. Yun, X. Q. Xu, and M. Bécoulet, BOUT++ nonlinear simulation for a comparative study with the measured 2D ELM structures in the KSTAR H-mode plasma, Phys. Plasmas **26**, 052502 (2019).

[28] Y. M. Wang, X. Q. Xu, Z. Yan, G. R. Mckee, B. A. Grierson, T. Y. Xia, and X. Gao, Simulation of density fluctuations before the L-H transition for hydrogen and deuterium plasmas in the DIII-D tokamak using the BOUT++ code, Nuclear Fusion **58**, 026026 (2018).

[29] A. W. Leonard, Edge-localized-modes in tokamaks, Phys. Plasmas **21**, 090501 (2014).

[30] P. T. Lang, B. Alper, R. Buttery, K. Gal, J. Hobirk, J. Neuhauser, M. Stamp and JET-EFDA contributors, ELM triggering by local pellet perturbations in type-I ELMy H-mode plasma at JET, Nuclear Fusion **47**, 754 (2007).

[31] F. Orain, M. Hoelzl, F. Mink, M. Willensdorfer, M. Bécoulet, M. Dunne, S. Günter, G. Huijsmans, K. Lack-

ner, S. Pamela, W. Suttrop, E. Viezzer, ASDEX Upgrade Team, and EUROfusion MST Team, Non-linear modeling of the threshold between ELM mitigation and ELM suppression by resonant magnetic perturbations in ASDEX upgrade, Phys. Plasmas **26**, 042503 (2019).

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

[33] 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).

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

[35] 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).

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

[37] 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).

[38] Y. Andrew, J.-P. Bähner, R. Battle, and T. Jirman, H-mode power threshold studies on MAST, Plasma **2**, 328 (2019).

[39] 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).

[40] 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).

[41] J. Cheng, 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).

[42] 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).

[43] 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).

[44] 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).

[45] 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).

[46] R. Hollerbach, E. Kim, and L. Schmitz, Time-dependent probability density functions and information diagnostics in forward and backward processes in a stochastic prey-predator model of fusion plasmas. Phys. Plasmas **27**, 102301 (2020).

[47] Y. Sun, Y. Liang, Y. Q. Liu, S. Gu, X. Yang, W. Guo, T. Shi, M. Jia, L. Wang, B. Lyu *et al.*, Nonlinear transition from mitigation to suppression of the edge localized mode with resonant magnetic perturbations in the EAST tokamak, Phys. Rev. Lett. **117**, 115001 (2016).

[48] A. B. Rechester and M. N. Rosenbluth, Electron heat transport in a tokamak with destroyed magnetic surfaces, Phys. Rev. Lett. **40**, 38 (1978).

[49] J. Mlynar, O. Ficker, E. Macusova, T. Markovic, D. Naydenkova, G. Papp, J. Urban, M. Vlainic, P. Vondracek, V. Weinzettl *et al.*, Runaway electron experiments at COMPASS in support of the EUROfusion ITER physics research, Plasma Phys. Control. Fusion **61**, 014010 (2018).

[50] L. Zeng, Z. Y. Chen, Y. B. Dong, H. R. Koslowski, Y. Liang, Y.P. Zhang, H.D. Zhuang, D.W. Huang and X. Gao, Runaway electron generation during disruptions in the J-TEXT tokamak, Nuclear Fusion **57**(4), 046001 (2017).

[51] P. Svensson, O. Embreus, S. L. Newton, K. Särkimäki, O. Vallhagen, and T. Fülöp, Effects of magnetic perturbations and radiation on the runaway avalanche, J. Plasma Phys. **87**, 905870207 (2021).

[52] A. J. Guel-Cortez and E. Kim, Information geometric theory in the prediction of abrupt changes in system dynamics, Entropy **23**, 694 (2021).

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

[54] 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).

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

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

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

[58] E. Kim, Investigating information geometry in classical and quantum systems through information length, Entropy **20**, 574 (2018).

[59] H. R. Chamorro, A. Guel-Cortez, E. Kim, F. Gonzalez-Longatt, A. Ortega, and W. Martinez, Information length quantification and forecasting of power systems kinetic energy, IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2022.3146314 (2022).

[60] Y. Andrew, E. Kim, D. Kologi, L. Schmitz, and Z. Wei, Time-dependent PDF analysis of DIII-D L-H transition, Royal Society Scientific meeting – H-mode Transition and Pedestal Studies in Fusion, poster (2022).

[61] This corresponds to $P_{in}/P_{cr} \gg (R\rho_s L_p^2)(\rho_s/L_p)^{1/3}$, the pressure gradient (the diamagnetic velocity) becomes sufficiently large to dominate the contribution from the poloidal velocity. Here, R, $\rho_s = c_s/\omega_{ci}$, $c_s = \sqrt{T_e/m_i}$, ω_{ci}, L_p, and a represent the major radius, the ion Larmor

radius calculated from the ion sound speed, the sound speed, the ion cyclotron frequency, the length-scale of the pressure gradient, and the minor radius, and ion sound speed, respectively [26].

[62] H. Y. W. Tsui, A. J. Wootton, J. D. Bell, R. D. Bengston, D. Diebold, J. H. Harris, N. Hershkowitz, C. Hidalgo, J. C. Ingraham, S. J. Kilpatrick *et al.*, A comparison of edge turbulence in tokamaks, stellarators, and reversed-field pinches, Phys. Fluids B **5**, 2491 (1993).

[63] Obviously, parameter values depend on the characteristics of plasmas and turbulence. For MAST edge plasmas [71] with $B \sim 0.5$ T, $T_e \sim 50$ eV, $\rho_s \sim 0.5$ cm, $\Delta_c \sim 1$ cm, and $L_p \sim (1-10)\rho_s$ the time unit is $\sim (1-30)\mu$secs. For DIII-D edge plasmas with $B \sim 2$ T, $T_e \sim 200$ eV, $\rho_s \sim 0.2$ cm, $k\rho_s \sim 0.1$, $\Delta_c \sim (5-10)\rho_s$[72], and $L_p \sim 5$ cm, the time unit is $O(1-10)\mu$secs.

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

[65] A. A. Thiruthummal and E. Kim, Monte Carlo Simulation of Stochastic Differential Equation to Study Information Geometry, Entropy **24**(8), 1113 (2022).

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

[67] A. Ignacio, G. B. Martínez, M. H. Jordan, and M. R. P. Juan, Inferring broken detailed balance in the absence of observable currents, Nat. Commun. **10**, 3542 (2019).

[68] É., Roldán, J. Barral, P. Martin, J. M. R. Parrondo, F. Jülicher, Quantifying entropy production in active fluctuations of the hair-cell bundle from time irreversibility and uncertainty relations, New J. Phys. **23**, 083013 (2021).

[69] T. Sagawa, *Thermodynamics of Information Processing in Small Systems* (Springer, 2012).

[70] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[71] N. Walkden, F. Riva, J. Harrison, F. Militello, T. Farley, J. Omotani, and B. Lipschultz, The physics of turbulence localised to the tokamak divertor volume, Comm. Phys. **5**, 139 (2022).

[72] T. L. Rhodes, J.-N. Leboeuf, R. Sydora, E. J. Doyle, R. A. Moyer, C. L. Rettig, K. H. Burrell, D. M. Thomas, G. R. McKee, and W. A. Peebles, Turbulent radial correlation length in the DIII-D tokamak, United States. https://www.osti.gov/servlets/purl/766721 (2000).

[73] E. Dudewicz, L. Gyorfi, and E. van der Meulen, Nonparametric entropy estimation: an overview, Int. J. Sci. Res. Math. Stat. Sci. **6**, 17 (1997).

[74] C. Bandt and B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. **88**, 174102 (2002).

[75] J. F. Guerrero Arnaiz, A. Dinklage, J. Geiger, M. Hirsch, U. Höfel, C. Brandt, H. Thomsen, J. Schilling, K. Rahbarnia, U. Neuner, M. Zanini, and the W7-X Team, Fast characterization of plasma states in W7-X with permutation entropy, Plasma Phys. Control. Fusion **64**, 084005 (2022).

[76] J. M. R. Parrondo, C. van den Broeck, and R. Kawai, Entropy production and the arrow of time, New J. Phys. **11**, 073008 (2009). 555

[77] G. Y. Antar, M. Tsalas, E. Wolfrum, V. Rohde and the ASDEX Upgrade Team, Turbulence during H- and L-mode plasmas in the scrape-off layer of the ASDEX Upgrade tokamak, Plasma Phys. Control. Fusion **50** (9), 095012 (2008).