Information Geometry of Non-Equilibrium Processes in a Bistable System with a Cubic Damping

Rainer Hollerbach *

Department of Applied Mathematics,

University of Leeds, Leeds LS2 9JT, UK;

* Correspondence: rh@maths.leeds.ac.uk

Eun-jin Kim

School of Mathematics and Statistics,

University of Sheffield, Sheffield, S3 7RH, UK

Abstract

A probabilistic description is essential for understanding the dynamics of stochastic systems far from equilibrium, given uncertainty inherent in the systems. To compare different Probability Density Functions (PDFs), it is extremely useful to quantify the difference among different PDFs by assigning an appropriate metric to probability such that the distance increases with the difference between the two PDFs. This metric structure then provides a key link between stochastic systems and information geometry. For a non-equilibrium process, we define an infinitesimal distance at any time by comparing two PDFs at times infinitesimally apart and sum these distances in time. The total distance along the trajectory of the system quantifies the total number of different states that the system undergoes in time, and is called the information length. By using this concept, we investigate the information geometry of non-equilibrium processes involved in disorder-order transitions between the critical and subcritical states in a bistable system. Specifically, we compute time-dependent PDFs, information length, the rate of change in information length, entropy change and Fisher information in disorder-to-order and order-to-disorder transitions, and discuss similarities and disparities between the two transitions. In particular, we show that the total information length in order-to-disorder transition is much larger than that in disorder-to-order transition, and elucidate the link to the drastically different evolution of entropy in both transitions. We also provide the comparison of the results with those in the case of the transition between the subcritical and supercritical states and discuss implications for fitness.

keywords: order-disorder transitions; fokker-planck equation; fisher information

I. INTRODUCTION

The spontaneous emergence of order out of disorder is one of the most fascinating phenomena in nature and laboratory experiments, attracting ever-increasing interest. Important examples include phase transition/critical phenomena in cosmology, elementary particle theory, condensed matter, chemistry, biology, and social-economic movement [1–5]. Order is usually quantified by a non-zero value of a macroscopic observable (global mode). While triggered by an external parameter (such as temperature) or spontaneously, a macroscopic observable often does not simply evolve passively, but undergoes an indispensable interaction with fluctuations (microscopic variables). The self-regulation between macroscopic and microscopic variables leads to a dynamical equilibrium (self-organisation), which involves fluctuations as an essential part [6]. There has been accumulating evidence for relevance and important role of self-organisation in different systems such as shear flows or vortices in fluids or plasmas, pattern formation in chemical oscillators, homeostasis in biosystems and even traffic flows [4, 7–16]. In particular, self-organised shear (zonal) flows are now believed to play a crucial role in stabilising laboratory plasmas, beneficial for extracting fusion energy [7]. Due to large fluctuations involved in order-disorder transition or self-organising systems, it is essential to use statistical tools to describe these systems.

The aim of this paper is to understand order-disorder transition from the perspective of information change associated with transition and uncover a geometrical structure in a statistical space, which can be utilised to understand ever-increasing experimental/observational data. To this end, we investigate a bistable stochastic system which is often invoked as a canonical model of self-regulating systems e.g. in electric circuits [17], in various cellular processes such as cycles, differentiation and apoptosis, regulation of heart, brain, etc. [18–23]. In this model, we calculate time-dependent Probability Density Function (PDF) and the total number of statistically different states that the system undergoes in time. The latter is defined by the dimensionless information length [24–28]:

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

where $p(x, t)$ is a time dependent PDF for a stochastic variable x. In Eq. (1), $\tau(t)$ is the

2

time-varying 'time-unit':

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

$\tau(t)$ in Eq. (2) has dimensions of time, and quantifies the correlation time over which the (dimensionless) information changes, thereby serving as the time unit for information change. Note that in equilibrium where $\frac{\partial p}{\partial t} = 0$, $\tau \to \infty$. Measuring the total elapsed time in units of τ between time $t = 0$ and t gives the information length in Eq. (1). The latter thus establishes a distance between the initial and final PDFs in the statistical space.

We note that our information length is based on Fisher information (c.f. [29]) and is a generalisation of statistical distance mainly used in equilibrium or near-equilibrium systems [30–39] to non-equilibrium systems [24–28]. In particular, the linear geometry of a linear Ornstein-Uhlenbeck (O-U) process was captured by the linear relation $\mathcal{L}_\infty \propto x_0$, while the power-law geometry of a nonlinear (cubic) stochastic process was revealed by power-law scalings $\mathcal{L}_\infty \propto x_0^n$ ($n \sim 1.5 - 1.9$) [28]. Furthermore, interesting geodesic solutions were found [27] by time-periodic modulation of the model parameters in an O-U process which by itself does not support a geodesic solution without the modulation of parameters. As a geodesic is a unique path along which a system undergoes the minimum number of changes in the statistical states given the initial and final conditions, it is beneficial to a system where the change is costly. It is thus important to elucidate the key characteristics of a stochastic system that permits or facilitates the existence of a geodesic in general (without the modulation of model parameters). Finally, we emphasise that Eq. (1) can be applied to any data; [25] constructed time-dependent PDFs from MIDI files of musics and elucidated the information change in music by $\mathcal{L}$ and τ.

In this paper, in order to gain a key insight, we focus on a zero-dimensional (0D) model which has one control parameter and propose an on-quenching experiment by a sudden change of a control parameter from the critical to subcritical and from the subcritical to critical values to trigger disorder-to-order and order-to-disorder transitions, respectively. A pair of disorder-to-order and order-to-disorder transitions with suitable choice of initial conditions then provides a simple model in which a continuous switch between ordered and disordered states can be studied in great detail. Each pair of disorder-to-order and

order-to-disorder transitions models a burst (e.g. a burst in the gene expression consisting of a pair of induction and repression). Since an initial condition represents the 'resting' state of a stochastic system in between the two bursts, it is important to understand the effect of different initial conditions on information change. In particular, we aim to elucidate what might be an optimal initial 'resting' state that minimizes the information change, sustaining a robust geodesic solution. To this end, we provide detailed comparison of on-quenching processes in this paper with off-quenching processes, where the control parameter changes between subcritical and supercritical, reported in [41]. While some aspects are similar between the two processes, there are also important differences that will be presented and discussed here. The remainder of this paper is organized as follows. Section II presents our model, and Section III provides details on the time evolution of PDFs. We discuss information length in Section IV and differential entropy and Fisher information in Section V. We conclude in Section VI. Appendices contain the derivation of equations used in the main text.

II. MODELS

We consider the following 0D Ginzburg-Landau model [40] for a stochastic variable x:

$$\frac{dx}{dt} = F(x) + \xi = -\lambda x - \mu x^3 + \xi. \tag{3}$$

Here, $F(x) = -\lambda x - \mu x^3$ is a deterministic force; ξ is a white noise with the following statistical property:

$$\langle \xi(t)\xi(t') \rangle = 2D\delta(t - t'), \tag{4}$$

where D is the strength of the forcing and the angular brackets denote the average over ξ. In our study, λ is a control parameter. In the numerical computations we will fix the value of $\mu \, (= 1)$ with no loss of generality, while keeping track of μ in analytical calculations for clarity. λ represents the deviation from the critical value (e.g. $\lambda \propto T - T_c$ for the temperature T where T_c is the critical temperature). That is, our system is subcritical for $\lambda < 0$, supercritical for $\lambda > 0$, and critical at $\lambda = 0$. Eq. (3) is the extension of our recent work [28, 42] to a cyclic transition. The Fokker-Planck equation [43, 44] corresponding to Eqs. (3)-(4) is as follows:

$$\frac{\partial}{\partial t}p(x,t) = \frac{\partial}{\partial x}\left[-F(x) + D\frac{\partial}{\partial x}\right]p(x,t). \tag{5}$$

In this paper, we consider the transition between the critical state and subcritical state by changing λ between 0 and $-\gamma$ ($\gamma > 0$). Here, we note that we are using $\gamma > 0$ to explicitly represent a positive growth rate in the subcritical state $\lambda < 0$. To this end, we induce the phase transition from the initial state $\lambda = 0$ by a sudden change in λ. This mimics the "on-quenching" experiment where the quenching occurs at the critical state; disorder-to-order transition (order-to-disorder transition) represents the transition from the critical to subcritical (the subcritical to critical) states. This thus contrasts to the case of the transition from the supercritical to subcritical state studied in [41] where the quenching occurs off the critical state.

Specifically, our model is described as follows:

- Forward Process (FP): $\lambda = -\gamma < 0$: at $t = 0$, a unimodal PDF with a peak at $x = 0$, which evolves into a bimodal PDF with peaks at $x = \pm\sqrt{\gamma/\mu} \neq 0$ as $t \to \infty$;

- Backward Process (BP): $\lambda = 0$: at $t = 0$, a bimodal PDF with peaks at $x = \pm\sqrt{\gamma/\mu} \neq 0$, which evolves into a unimodal PDF with a peak at $x = 0$ as $t \to \infty$.

FP and BP have the following stationary distributions $p_F(x)$ and $p_B(x)$, respectively:

$$p_F(x) \;\propto\; e^{-\frac{\mu}{4D}\left(x^2 - \frac{\gamma}{\mu}\right)^2}, \tag{6}$$

$$p_B(x) \;\propto\; e^{-\frac{\mu}{4D}x^4}. \tag{7}$$

That is, FP has a stationary bimodal distribution peaked at $x = \pm\sqrt{\frac{\gamma}{\mu}}$, which can be approximated as a double Gaussian for small D as follows:

$$p_F(x) \;\sim\; \frac{\sqrt{\beta_F}}{2\sqrt{\pi}}\left[e^{-\beta_F\left(x+\sqrt{\frac{\gamma}{\mu}}\right)^2} + e^{-\beta_F\left(x-\sqrt{\frac{\gamma}{\mu}}\right)^2}\right], \tag{8}$$

where $\beta_F = \frac{\gamma}{D}$. Eq. (8) represents the sum of the two Gaussians (double Gaussian) with the peak at $\pm\sqrt{\frac{\gamma}{\mu}}$ and variance $\sigma_F = \langle x^2 \rangle - \langle x \rangle^2 = \frac{D}{2\gamma} = \frac{1}{2\beta_F}$. In comparison, BP has a unimodal quartic exponential PDF centered around $x = 0$ in equilibrium. To model a cyclic transition between ordered and disordered states, we use $p_B(x)$ as an initial distribution for FP and $p_F(x)$ for BP, respectively. Consequently, the initial PDFs in both FP and BP are strongly out of equilibrium. We investigate time-dependent PDFs and information length during this transient relaxation. In particular, we are interested in how $\mathcal{L}$ depends on the deviation from the critical value (γ) and the strength of the stochastic noise (D). Table 1

summarizes the value of λ in Eq. (3) and initial conditions for FP and BP, together with the variance $\sigma = \langle x^2 \rangle - \langle x \rangle^2$ of the initial and final PDFs, where the angular brackets denote the average over the stochastic noise ξ. We note that for small D, the equilibrium variance for FP, $D/2\gamma$, is much smaller than that for BP, $2\sqrt{D/\mu}\,\Gamma\left(\frac{3}{4}\right)/\Gamma\left(\frac{1}{4}\right)$.

Case	λ	$p(x,0)$	$p(x, t \to \infty)$	σ_0	σ_F
FP	$-\gamma$	$p_B(x)$	$p_F(x)$	$2\sqrt{D/\mu}\,\Gamma\left(\frac{3}{4}\right)/\Gamma\left(\frac{1}{4}\right)$	$D/2\gamma$
BP	0	$p_F(x)$	$p_B(x)$	$D/2\gamma$	$2\sqrt{D/\mu}\,\Gamma\left(\frac{3}{4}\right)/\Gamma\left(\frac{1}{4}\right)$

TABLE I: Summary of FP, BP: FP and BP have equilibrium PDFs (in the limit of $t \to \infty$), p_F and p_B, respectively. FP and BP start with the initial PDFs, $p(x,0) = p_B$ and $p(x,0) = p_F$, respectively and reach the equilibrium PDFs, $p(x, t \to \infty) = p_F$ and p_B, respectively. p_F and p_B have equilibrium variances, $D/2\gamma$ and $2\sqrt{D/\mu}\,\Gamma\left(\frac{3}{4}\right)/\Gamma\left(\frac{1}{4}\right)$, respectively. σ_0 and σ_F are the initial and final variances at $t = 0$ and $t \to \infty$, respectively, for each process.

III.   TIME-EVOLUTION OF PDFS

To solve Eq. (5) numerically, we note first that γ, D and t can always be rescaled such that $\mu = 1$, thereby reducing the number of parameters that need to be explored to only γ and D. Any numerical solution also requires x to be restricted to a finite interval, which can always be rescaled to $x \in [-1, 1]$ without any loss of generality. If γ and D are chosen such that $p_F(x)$ is restricted well away from the boundaries $|x| = 1$, then this finite interval in x is an excellent match to the analytically more convenient infinite extent. Taking $\gamma \leq 0.7$ and $D \leq 10^{-3}$ ensures that the bimodal peaks at $x = \pm\sqrt{\gamma}$ are still sufficiently far from the boundaries, and sufficiently narrow, that $p = 0$ can simply be imposed as the boundary condition at $x = \pm 1$. The details of the numerical implementation then involve second-order finite-differencing in x and t, using up to $O(10^6)$ grid-points in space, and time-steps as small as $O(10^{-7})$. This spatial resolution allows D to be reduced down to 10^{-7} while still fully resolving the bimodal peaks of width $O(D^{1/2})$.

A. Overall comparison of FP and BP

One of the most significant differences between FP and BP is the timescale on which the process evolves and settles in. There are many diagnostic quantities that could be used to quantify this, but a useful one is simply the ratio $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle$. This can be evaluated analytically for the two endstates, yielding 1.48 for $p_B(x)$ and 1 for $p_F(x)$ (taking $D \ll 1$). The evolution in time must therefore be that FP is $1.48 \to 1$ and BP is $1 \to 1.48$, and the question is on what timescales this happens.

As shown in Fig. 1(a), for FP the dependence on D is such that every reduction in D by a factor of 100 shifts the curves by a constant amount in t. That is, the ratio does not deviate significantly from 1.48 until a time $c \ln D^{-1}$ has elapsed, but except for this shift the three curves are essentially identical. The numerically determined value of c is 0.355 and is very close to the factor of $\frac{1}{4\gamma} = 0.357$ in t_2 in Eq. (23), discussed later. It is interesting that this value of c is exactly half of that in [41], but otherwise the scaling with D is the same. The reason for the faster adjustment in this case is because the initial condition already starts out much broader, with a width $O(D^{1/4})$ here as opposed to $O(D^{1/2})$ in [41]. We recall that $x = 0$ is an unstable equilibrium point when $\xi = 0$, and the instability slowly builds up due to ξ (e.g. see [46, 47]) and a finite width of the initial PDF until $t \propto O(|\ln D|)$. If the initial condition is already broader, then it is not surprising that the instability can develop sooner. However, this factor of half in the settling time does not mean that the evolution of PDFs in on- and off-quenching processes has any similarity. In fact, we will show that they are quite different and that the off-quenching cannot simply be made up of the two phases where effectively $\lambda = \gamma \to 0$ and $\lambda = 0 \to -\gamma$ (as for the on-quenching case here).

Fig. 1(b) shows that for BP the dependence on D is very different, with every reduction in D by a factor of 100 shifting the curves by a factor of 10; that is, time scales as $D^{-1/2}$. The backward process is initially driven by the movement of the two peaks towards $x = 0$ before diffusion becomes crucial in forming a single peak at $x = 0$; it is this final diffusive adjustment process that requires an $O(D^{-1/2})$ time to happen. This is in sharp contrast with the results in [41], where FP and BP both exhibited the same $c \ln D^{-1}$ scaling (and even with the same value of c). In comparison with the off-quenching in [41], the

on-quenching considered here thus has a forward process that is faster by a factor of 2, but a backward process that is much slower, with a completely different $D^{-1/2}$ as opposed to $\ln D^{-1}$ scaling.
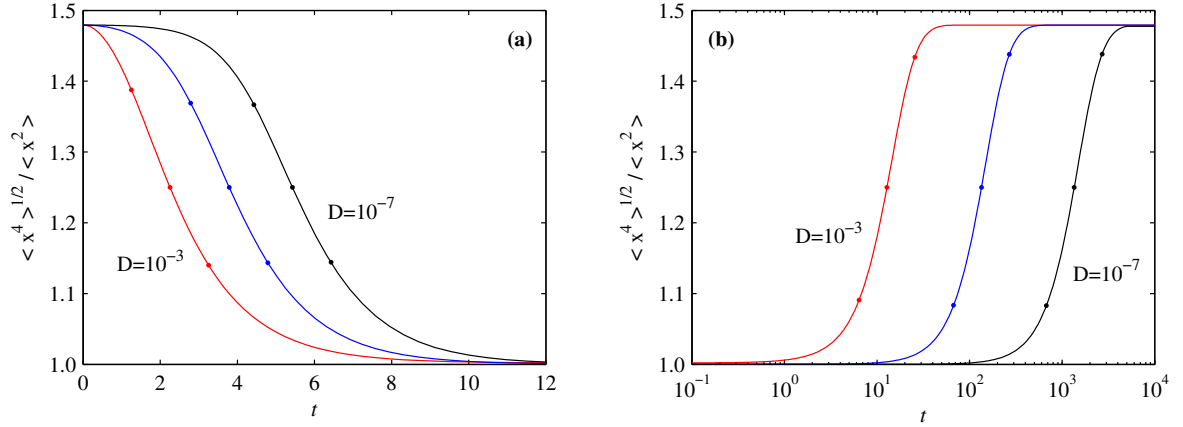

FIG. 1: The ratio $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle$ as a function of t, for the three values $D = 10^{-3}$ (red), 10^{-5} (blue) and 10^{-7} (black). Panel (a) shows FP with t on a linear scale, (b) shows BP with t on a logarithmic scale. All six curves are for $\gamma = 0.7$. The three dots on each curve correspond to the solutions shown in Figs. 2 and 3. The central dot is always when $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle = 1.25$. In (a) the other two dots are at $t \pm 1$ relative to the central one; in (b) they are at $t/2$ and $2t$ relative to the central one.

B. PDF of Forward Process

Fig. 2 shows the structure of $p(x,t)$ for FP. The particular times are chosen to take the shift $c \ln D^{-1}$ into account; that is, different values of D are shown at the times where they have the same ratios $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle$. The results are seen to be identical for the three different values of D. The initial condition obviously does depend on D (as indicated also by the red curves), but once a certain broadening has occurred, in this time frame $c \ln D^{-1}$, the subsequent evolution is independent of D, relying only on the instability process (as measured by γ). It is only in the last stages of the evolution (not shown in Fig. 2), when the solution settles in to the final bimodal structure, that diffusion plays a role again and determines the $O(D^{1/2})$ width of the peaks. Comparing with results in [41], it is interesting

to note also that here a double-peak emerges essentially immediately, whereas in [41] a finite time had to elapse before the central peak split into two separate peaks. The reason for this is that here the initial condition is at a critical state with a much broader profile than the Gaussian profile at the subcritical state considered in [41].
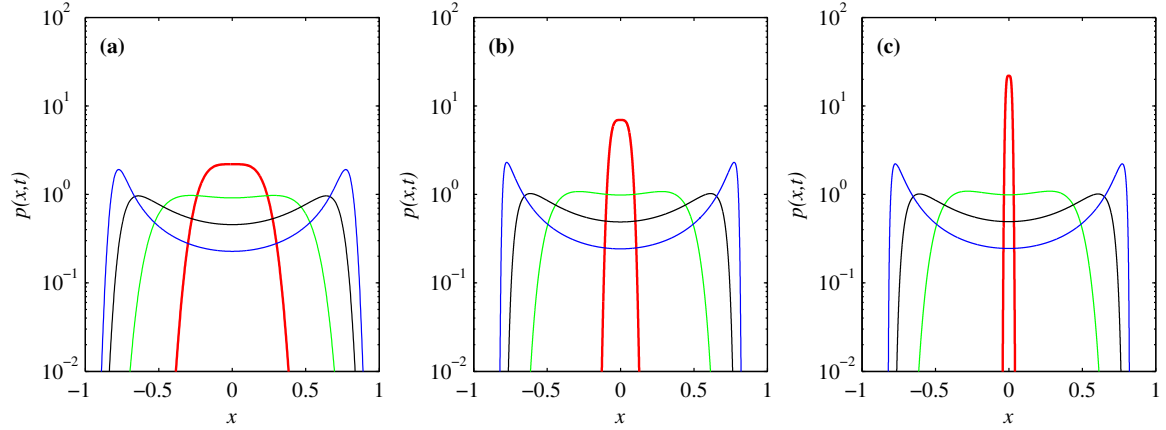

FIG. 2: The PDFs for FP, for (a) $D = 10^{-3}$, (b) $D = 10^{-5}$, (c) $D = 10^{-7}$. The initial condition is indicated by the red central peak. Subsequent times are indicated in green, black and blue, taking the $c \ln D^{-1}$ shift into account. For (a) these are $t = 1.25$, 2.25 and 3.25; for (b) $t = 2.79$, 3.79 and 4.79; for (c) $t = 4.43$, 5.43 and 6.43. The middle time (the black line) is always when the ratio $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle = 1.25$; the green line is always $t - 1$ relative to that, and the black line is $t + 1$. See also the three dots on each curve in Fig. 1(a).

To understand these results better, it is of value to perform analytical analysis in the limiting cases. To this end, we transform the nonlinear term in Eq. (3) into a linear (anti)damping term [45] by seeking a variable y such that Eq. (3) becomes $dy/dt = \gamma y + \xi F(y)$ where $F(y)$ is a function of y. We then easily show that $dy/dx = \gamma y/(\gamma x - \mu x^3)$ which has the solution $x = y/\sqrt{1 + \alpha y^2}$ ($\alpha = \frac{\mu}{\gamma}$). Specifically, y satisfies

$$\frac{dy}{dt} = \gamma y + \xi (1 + \alpha y^2)^{\frac{3}{2}}. \tag{9}$$

Eq. (9) provides a convenient way of computing the PDF of x through y by approximating $\xi(1 + \alpha y^2)^{\frac{3}{2}} \sim \xi$ for small y [46]. Thus, to leading order y is a Gaussian process, simply given

9

by the Ornstein–Uhlenbeck process [44] with a negative damping. The transition probability is thus Gaussian:

$$p(y, t; y_0, 0) = \sqrt{\frac{\beta_1}{\pi}} e^{-\beta_1 (y - y_0)^2}, \tag{10}$$

where $y_0 = \frac{x_0}{\sqrt{1 - \alpha x_0^2}}$ and $\frac{1}{\beta_1(t)} = \frac{D}{\gamma}(e^{2\gamma t} - 1)$.

By using the conservation of the probability $p(x, t)dx = p(y, t)dy$ and Eq. (10), we obtain the transition probability of x as follows:

$$p(x, t; x_0, 0) = \frac{1}{(1 - \alpha x^2)^{\frac{3}{2}}} \sqrt{\frac{\beta_1}{\pi}} \exp\left[-\beta_1 \left(\frac{x}{\sqrt{1 - \alpha x^2}} - \frac{x_0}{\sqrt{1 - \alpha x_0^2}} e^{\gamma t}\right)^2\right], \tag{11}$$

which recovers the previous results [46, 47] when $\alpha = 1$ and $x_0 = 0$.

To obtain a (marginal) PDF of x, we recall that FP has an initial PDF given by a quartic exponential:

$$p(x_0, 0) = M \beta_0^{\frac{1}{4}} e^{-\beta_0 x_0^4}. \tag{12}$$

Here, $\beta_0 = \frac{\mu}{4D}$; $M = 2\Gamma\left(\frac{1}{4}\right)^{-1}$ is a normalisation constant where $\Gamma(x)$ is the Gamma function. Eqs. (10) and (12) then give:

$$p(x, t) = M \beta_0^{\frac{1}{4}} \sqrt{\frac{\beta_1}{\pi}} \int_{-\infty}^{\infty} dx_0 \frac{1}{(1 - \alpha x^2)^{\frac{3}{2}}} \exp\left[-\beta_1 \left(\frac{x}{\sqrt{1 - \alpha x^2}} - \frac{x_0}{\sqrt{1 - \alpha x_0^2}} e^{\gamma t}\right)^2\right] e^{-\beta_0 x_0^4}. \tag{13}$$

Here, $\frac{1}{\beta_1(t)} = \frac{D}{\gamma}(e^{2\gamma t} - 1)$.

We now show that the initial quartic exponential PDF (12) undergoes roughly two stages of the time evolution: stage (i) driven by diffusion/advection with the continuous movement of the PDF peak from $x = 0$ towards $\sqrt{\frac{\gamma}{\mu}}$, and then stage (ii) of settling into an equilibrium PDF with the adjustment of the PDF shape. To this end, we examine the behaviour of $p(x, t)$ in Eq. (13) for a sufficiently small β_1 such that

$$\beta_1 = \frac{\gamma}{D(e^{2\gamma t} - 1)} \gg 1. \tag{14}$$

Eq. (14) will later be shown to be valid in stage (i) (e.g. for $t < t_2$ where t_2 is defined in Eq. (23)). By using Eq. (14), we can approximate the exponential function in Eq. (13):

$$\sqrt{\frac{\beta_1}{\pi}} \exp\left[-\beta_1\left(\frac{x}{\sqrt{1-\alpha x^2}} - \frac{x_0}{\sqrt{1-\alpha x_0^2}}e^{\gamma t}\right)^2\right] \sim \delta\left(\frac{x}{\sqrt{1-\alpha x^2}} - \frac{x_0}{\sqrt{1-\alpha x_0^2}}e^{\gamma t}\right)$$

$$= e^{-\gamma t}(1-\alpha x^2)^{\frac{3}{2}}(1-\overline{\alpha}x^2)^{-\frac{3}{2}}\,\delta\left(x_0 - \frac{xe^{-\gamma t}}{\sqrt{1-\overline{\alpha}x^2}}\right), \tag{15}$$

where

$$\overline{\alpha} \equiv \alpha(1-e^{-2\gamma t}). \tag{16}$$

Then, using Eq. (15) in Eq. (13) gives us

$$p(x,t) \sim M\beta_0^{\frac{1}{4}}\frac{e^{-\gamma t}}{(1-\overline{\alpha}x^2)^{\frac{3}{2}}}e^{-\beta_0 e^{-4\gamma t}\left(\frac{x^2}{1-\overline{\alpha}x^2}\right)^2} \equiv M\beta_0^{\frac{1}{4}}e^{-\gamma t}e^{-\psi}, \tag{17}$$

where ψ is defined by

$$\psi = \beta_0 e^{-4\gamma t}\left(\frac{x^2}{1-\overline{\alpha}x^2}\right)^2 + \frac{3}{2}\ln(1-\overline{\alpha}x^2). \tag{18}$$

By using $\overline{\alpha} = 0$ at $t = 0$, we can easily show that Eq. (17) matches the initial PDF $p(x,0)$ at $t = 0$ in Eq. (12). From $\frac{\partial\psi}{\partial x}|_{x=x_{1,2}} = 0$, we find the value of $x = x_{1,2}$ where the PDF takes its local maximum or minimum:

$$x_1 = 0, \quad 4\beta_0 e^{-4\gamma t}x_2^2 = 3\overline{\alpha}(1-\overline{\alpha}x_2^2)^2. \tag{19}$$

Since $\partial_{xx}\ln p(x,t) = -\partial_{xx}\psi > 0$ at $x = 0$, $x = x_1 = 0$ is a local minimum of $p(x,t)$ for all $t > 0$. On the other hand, two values of x_2 (note $\overline{\alpha}x_2^2 < 1$) represent the location of the local maximum in $x > 0$ and $x < 0$, respectively. Thus, $x = x_1 = 0$, the (global) maximum at $t = 0$, becomes a local minimum for any infinitesimal time $t > 0$, two peaks forming at x_2. For instance, for $\overline{\alpha}x_2^2 \ll 1$, we can easily show that

$$x_2^2 \sim \frac{3D}{\gamma}\left(e^{4\gamma t} - e^{2\gamma t}\right). \tag{20}$$

This reveals the diffusive nature of the peak movement from $x = 0$ due to instability γ towards the equilibrium value $\pm\sqrt{\frac{\gamma}{\mu}}$. In the limit of a very small time $t \ll \frac{1}{4\gamma}$, Eq. (20) gives $x_2 \sim \pm\sqrt{6Dt}$ by using $e^{-2\gamma t} = 1 - 2\gamma t + \dots$ and $\beta_0 = \frac{\mu}{4D}$, showing that the initial movement of the two peaks is via random walk. Our numerical solutions confirmed the predicted scaling of $x_2 \propto D^{1/2}$ in Eq. (20) as well as $x_2 \propto \sqrt{Dt}$ for small time, followed by

11

almost exponential increase (no figure is shown here).

To examine the evolution in more detail, we consider the characteristic time t_2 where the width of the PDF in Eq. (17) becomes comparable to the peak position x_2 in Eq. (20). To estimate the PDF width R in Eq. (17), we find $x = R$ when the exponent becomes unity:

$$1 \sim \beta_0 e^{-4\gamma t} \left(\frac{x^2}{1 - \overline{\alpha} x^2} \right)^2 \sim \beta_0 e^{-4\gamma t} x^4, \tag{21}$$

for $\overline{\alpha} x^2 \ll 1$. We then find $x = R$ from Eq. (21) as

$$R^2 \sim e^{2\gamma t} \left(\frac{4D}{\mu} \right)^{1/2}. \tag{22}$$

By equating R in Eq. (22) and x_2 in Eq. (20) and using $e^{4\gamma t} - e^{2\gamma t} \sim e^{4\gamma t}$ for $\gamma t \gg 1$, we find the characteristic time t_2 as follows:

$$t_2 \sim \frac{1}{4\gamma} \ln \left(\frac{4\gamma^2}{3\mu D} \right), \tag{23}$$

where $\beta_0 = \frac{\mu}{4D}$ was used. For $\gamma = 0.7$ and $D = 10^{-3}$, 10^{-5}, 10^{-7}, $t_2 = 2.3$, 3.9, 5.6. Notably, the value of t_2 will later be shown to be very close to the other time scales t_m signifying order formation. t_2 in Eq. (23) marks the time when the peak position becomes comparable to the rms (Gaussian) fluctuation. For $t \gg t_2$, PDF in settling into the final equilibrium PDF is approximated by the Gaussian [46]. We have confirmed this prediction from our numerical solutions (as discussed in more detail later). Finally, we have checked that Eq. (14) is valid for t_2 given in Eq. (23).

C.   PDF of Backward Process

We recall that BP starts with a bimodal PDF p_F which has two peaks at $\pm\sqrt{\gamma/\mu}$, which is the final equilibrium PDF of FP. For sufficiently small D, the distance between these two peaks is much larger than the width of the PDFs and are thus well separated so that PDF is approximated as the sum of the two Gaussian (double Gaussian) PDFs given by Eq. (8). The latter evolve almost independently in $x > 0$ and $x < 0$, respectively, until $t \sim O(1)$ when PDFs undergo significant change in the shape with large fluctuation. Since the Gaussian evolution is completely determined by mean value and variance, we now compute mean

value and variance in $x > 0$ or $x < 0$ separately by taking advantage of small fluctuations compared to mean value. Specifically, we let $x = z + \delta x$ where $z = \langle x \rangle$ is the mean component averaged over ξ and the initial PDF in $x > 0$ (or $x < 0$) while δx is the fluctuation $\langle \delta x \rangle = 0$. This gives us

$$\frac{dz}{dt} = -\mu z^3 - 3\mu \langle (\delta x)^2 \rangle \sim -\mu z^3, \tag{24}$$

$$\frac{d\sigma}{dt} = -6\mu z^2 \sigma + 2D. \tag{25}$$

The solutions to Eqs. (24) and (25) are as follows (c.f. [48]):

$$z(t) = \frac{z_0}{(1 + 2\mu z_0^2 t)^{1/2}},$$

$$\sigma(t) = \frac{\sigma_0}{G'(t)} + 2D\frac{G(t)}{G'(t)}, \tag{26}$$

where $G'(t) = (1 + 2\mu z_0^2 t)^3$ and $G = \int_0^t G'(t)\, dt$, $z_0 = \sqrt{\frac{\gamma}{\mu}}$ and $\sigma_0 = \frac{D}{2\gamma}$ (see Eq. (8) and Table 1). Eqs. (25) and (26) will be used for computing $\mathcal{L}$ in the next section.

We show $p(x,t)$ for BP in Fig. 3. We choose the particular times again to take the $D^{-1/2}$ scaling into account, and show results at the times where they have the same ratios $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle$. The initial evolution (not shown in Fig. 3) consists simply of a motion toward the origin, with the width of the peaks remaining $O(D^{1/2})$. The scaling of this movement is found to be consistent with the prediction in Eq. (26).

Once the peaks get within a distance $D^{1/4}$ of the origin they start to sense the presence of the potential well, and diffusion starts to collapse them to a single peak. As seen in Fig. 3, if x is rescaled as $D^{1/4}$, and p correspondingly rescaled as $D^{-1/4}$, then the results again look the same for all three values of D. This final diffusive adjustment to the single central peak is very slow though, resulting in the $D^{-1/2}$ scaling in time.

D.   Energy Diagnostics

We now elucidate the role of the linear growth term (positive feedback) and cubic damping (negative feedback) in FP in energy balance and geodesic. To this end, we multiply Eq. (3)
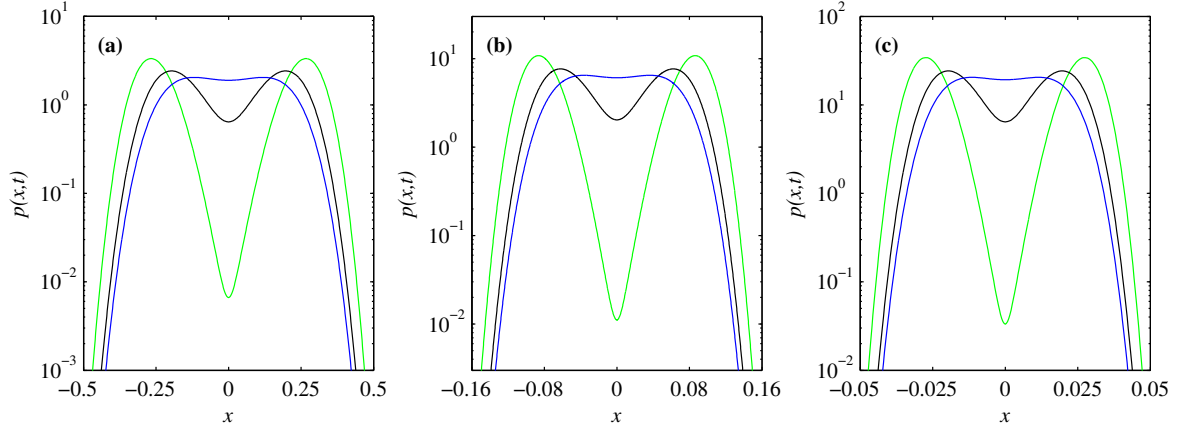
FIG. 3: The PDFs for BP, for (a) $D = 10^{-3}$, (b) $D = 10^{-5}$, (c) $D = 10^{-7}$. The times are indicated in green, black and blue, taking the $D^{-1/2}$ shift into account. For (a) these are $t = 6.4$, 12.8 and 25.6; for (b) $t = 67$, 134 and 268; for (c) $t = 675$, 1350 and 2700. The middle time (the black line) is always when the ratio $\sqrt{\langle x^4 \rangle}/\langle x^2 \rangle = 1.25$; the green line is always $t/2$ relative to that, and the black line is $2t$. See also the three dots on each curve in Fig. 1(b). Note finally how x and p have been rescaled according to $D^{\pm 1/4}$.
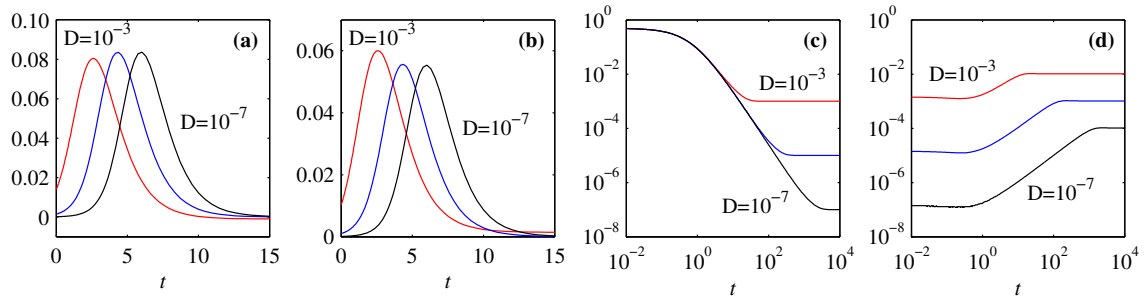

FIG. 4: (a) $H = 0.7\langle x^2 \rangle - \langle x^4 \rangle$ as a function of time for FP. (b) $\Delta = \sqrt{\langle x^4 \rangle} - \langle x^2 \rangle$ as a function of time for FP. (c) $-H = \langle x^4 \rangle$ as a function of time for BP. (d) $\Delta = \sqrt{\langle x^4 \rangle} - \langle x^2 \rangle$ as a function of time for BP.

by x and take the average over ξ and initial condition to obtain the following equation

$$\frac{1}{2}\frac{d\langle x^2 \rangle}{dt} = \gamma\langle x^2 \rangle - \mu\langle x^4 \rangle + D. \tag{27}$$

Here, the last term D, representing the rate of energy injection by ξ, was calculated as $\langle \xi(t)x(t) \rangle = \langle \xi(t) \int_0^t dt_1 [\gamma x(t_1) - \mu x(t_1)^3 + \xi(t_1)] \rangle = D$ (also confirmed by the numerical calculations). The middle term $\gamma\langle x^2 \rangle - \mu\langle x^4 \rangle \equiv H$ represents the energy into the system

14

or environment, depending on the sign. When $H > 0$, the energy goes into the system, contributing to the increase in $\langle x^2 \rangle$; when $H < 0$, the energy is dissipated in the system, increasing heat in the environment.

Fig. 4(a) shows $H = \gamma \langle x^2 \rangle - \mu \langle x^4 \rangle$ for $D = 10^{-3}$, 10^{-5}, 10^{-7}. Unlike $\langle x^2 \rangle$ and $\langle x^4 \rangle$, which each increase monotonically in time, H reaches a peak at some time $t = t_m$, and then decreases to the negative value $-D$ in settling in to the equilibrium PDF. The maximum H signifies when the positive feedback by the linear growth rate most dominates over the negative feedback by the nonlinear damping. It is notable that the times $t_m = 2.6$, 4.3, 5.9 in Fig. 4(a) for the maximum H are similar to the times $t_2 = 2.3$, 3.9, 5.6 given in Eq. (23), with both exhibiting the same $c \ln D^{-1}$ ($c = 1/4\gamma$) scaling. t_m will also be shown to be very close to the time for the maximum entropy in Section V.

Physically, $t_m \sim t_2$ signifies the start of order formation. Another diagnostic for the latter is $\Delta = \sqrt{\langle x^4 \rangle} - \langle x^2 \rangle$, also shown in Fig. 4(b), where similar non-monotonic behaviour is prominent, with Δ peaking at the same times as H. This large fluctuation Δ signifies the phase transition from disordered to ordered states due to the development of the two peaks, which occurs on timescales increasing with $c \ln D^{-1}$ as discussed above.

For BP, $H = -\langle x^4 \rangle$ and Δ in Fig. 4(c)-(d) are monotonic during the return to the disordered state. The monotonic evolution of H and Δ for BP is also reflected in the evolution of the differential entropy in Section V.

IV. INFORMATION LENGTH

We calculate information length in Eq. (1) and explore geometric structure during phase transition. Figs. 5 and 6, for FP and BP respectively, show how $\mathcal{E}$ and $\mathcal{L}$ evolve in time, as well as how the total $\mathcal{L}(t \to \infty) = \mathcal{L}_\infty$ depends on γ. Since FP and BP switch between $\lambda = 0$ and $-\gamma \neq 0$, γ in Figs. 5-6 always refers to the non-zero value. We are especially interested also in comparing the on-quenching results computed here with the previous off-quenching results from [41], shown as the dashed lines in Figs. 5-6.

A.   Forward Process

For FP, it is useful to consider times less than or greater than t_2 in Eq. (23) separately, by approximating the time-dependent PDF as a quartic exponential and Gaussian in $t < t_2$ and $t > t_2$, respectively. First, for $t < t_2$, by ignoring the contribution from the mean value $\langle y \rangle = z$ compared with that from the variance, we obtain $\tau(t)$ in Eq. (2) (see Appendix B):

$$\frac{1}{[\tau(t)]^2} \; \sim \; \frac{1}{4\beta(t)^2} \left(\frac{d\beta}{dt} \right)^2 . \tag{28}$$

Eqs. (2), (23) and (28) then give $\mathcal{L}$ between the time $t = 0$ and $t \sim t_2$:

$$\mathcal{L}(t_2) \sim \frac{1}{2} \left| \ln \left(\frac{\beta_0 e^{-4\gamma t_2}}{\beta_0} \right) \right| \sim \frac{1}{2} \left| \ln \left(\frac{3\mu D}{\gamma^2} \right) \right| = \frac{1}{2} \ln \left(\frac{\gamma^2}{3\mu D} \right). \tag{29}$$

On the other hand, during the time between $t \sim t_2$ and $t \to \infty$, the PDF settles into the double Gaussians so that we can estimate the total $\mathcal{L}$ between $t \sim t_2$ and $t \to \infty$ by using Eq. (C1) (see Appendix C) as

$$\mathcal{L}_\infty - \mathcal{L}(t \sim t_2) \; \sim \; \frac{1}{\sqrt{2}} \left| \ln \left(\frac{\sigma(t_2)}{\sigma(t \to \infty)} \right) \right| \sim \frac{1}{\sqrt{2}} \ln \left[\frac{2\gamma^2 \Gamma\left(\frac{3}{4}\right)}{\sqrt{3}\mu D \, \Gamma\left(\frac{1}{4}\right)} \right], \tag{30}$$

where $\sigma(t_2) = \frac{\Gamma\left(\frac{3}{4}\right)}{\Gamma\left(\frac{1}{4}\right)} \frac{1}{\sqrt{2\beta_0 e^{-4\gamma t_2}}} = \frac{\Gamma\left(\frac{3}{4}\right)}{\Gamma\left(\frac{1}{4}\right)} \sqrt{\frac{4\gamma^2}{3\mu^2}}$ for FP (see above), $\alpha = \frac{\mu}{\gamma}$ and $\sigma(t \to \infty) = \frac{D}{2\gamma}$ were used. Eqs. (29) and (30) have the same dependence on D, μ and γ. The sum of Eqs. (29) and (30) gives the total

$$\mathcal{L}_\infty \sim -1.2 + \frac{\sqrt{2}+1}{2} \ln \left(\frac{\gamma^2}{D} \right), \tag{31}$$

when $\mu = 1$ and numerical values for the Γ functions are inserted.

Fig. 5 shows the numerically computed $\mathcal{E}$ and $\mathcal{L}$ for FP. We see how $\mathcal{E}$ starts out essentially constant, corresponding to a geodesic solution [27]. This constant plateau continues until the $O(|\ln D|)$ equilibration timescale previously also seen in Fig. 1. After this time $\mathcal{E}$ decreases exponentially. Comparing $\mathcal{E}$ here with the previous off-quenching results, we notice three differences: (a) the previous initial adjustment before the plateau regime is absent here, and the curves are essentially flat from the initial condition onward;

(b) the plateau here is *higher* than before; (c) the equilibration, and hence the exponential decrease in $\mathcal{E}$, happens sooner.

Turning to $\mathcal{L}$ next, the combination that the plateau is higher, but ends sooner, has the interesting consequence that initially $\mathcal{L}$ is greater than in the off-quenching case, but the final values $\mathcal{L}_\infty$ are always lower. Fig. 5(c) shows the variation of $\mathcal{L}_\infty$ with γ, and the same pattern persists throughout; $\mathcal{L}_\infty$ is consistently ~ 1 less than before, with the resulting best-fit formula

$$\mathcal{L}_\infty \approx 2.1 \ln \gamma - 1.05 \ln D. \tag{32}$$

The coefficients of $\ln \gamma$ and $\ln D$ are both in generally good agreement with the analytic predictions from Eq. (31), which has 2.4 and -1.2. The constant terms, 0 versus -1.2, match less well, but this term is also strongly affected by the best fit to the $\ln D$ term, since e.g. $|\ln 10^{-7}| = 16$ is already as large as the largest $\mathcal{L}_\infty$ values. (Note finally that the deviation from straight lines for large D and small γ has the same origin as before in [41]: the 'initial' and 'final' states are then so broad (large D) and so close to each other (small γ) that they overlap, causing the dynamics to be different, but also not very interesting in this regime.)

B. Backward Process

Fig. 6 shows corresponding results for BP. $\mathcal{E}$ now starts off lower than in the off-quenching case, but the final equilibration is much slower, again as seen previously in Fig. 1. The result of the initially smaller $\mathcal{E}$ is that for small times $\mathcal{L}$ is a factor of 2 less than in the off-quenching case. See also Eq. (33) below, which confirms this analytically. Because the equilibration is so slow though, there is an additional contribution to $\mathcal{L}_\infty$ that is not present before. Curiously, this seems to result in the final $\mathcal{L}_\infty$ values always being a factor of 1.5 less than in the off-quenching case. The precise origin of this particular factor, or indeed why it is always the same, independent of D, is not fully understood. As seen in Fig. 6(c), the results are summarized by the formula $\mathcal{L}_\infty \approx 0.9 \gamma D^{-1/2}$.
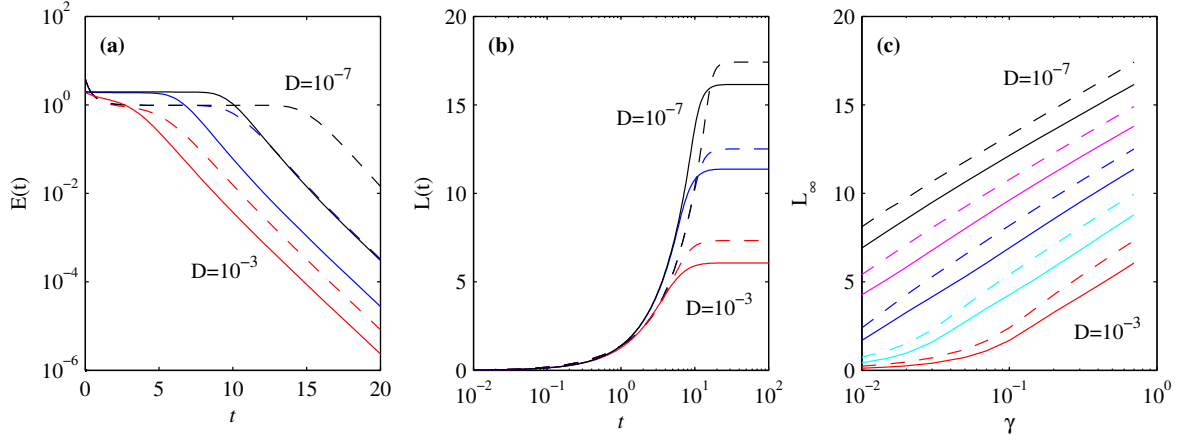
FIG. 5: (a) and (b) show $\mathcal{E}$ and $\mathcal{L}$, respectively, as functions of time, for $\gamma = 0.7$. (c) shows $\mathcal{L}_\infty$ as a function of γ. All three panels are for FP only. The solid lines are the on-quenching process considered here; the dashed lines are for the off-quenching process considered in [41]. $D = 10^{-3}$ to 10^{-7} as indicated. Note the different combinations of linear and logarithmic scales to emphasize different features in different quantities.

To quantify this scaling, we use Eq. (C1) with Eq. (C12) and $z \sim 0$:

$$
\begin{aligned}
\mathcal{L}_\infty &\sim \int_0^\infty dt \frac{1}{\sigma} \frac{dz}{dt} \\
&\sim c_1(z_0) \int_0^\infty \frac{dt}{\sqrt{\sigma_0 + 2DG}} \\
&\sim \mu z_0^3 \int_0^{\frac{1}{2\gamma}} \frac{dt}{\sqrt{\sigma_0 + 2Dt}} \sim \frac{\sqrt{3}-1}{\sqrt{2}} \frac{\gamma}{\sqrt{\mu D}},
\end{aligned}
\tag{33}
$$

where $\sigma_0 = \frac{D}{2\gamma}$ for BP (see Table 1) was used. The variation with γ and D is exactly as in the numerical results, whereas the constant factor is an under-estimate, 0.5 versus 0.9. Given that Eq. (33) only represents the early-time contribution to $\mathcal{L}$ though, we would expect the true $\mathcal{L}_\infty$ to be larger.

V. DIFFERENTIAL ENTROPY AND FISHER INFORMATION

Entropy is most commonly used to describe complexity. In a continuous system, it is given by the (Gibbs) differential entropy (e.g. see [49]) defined by
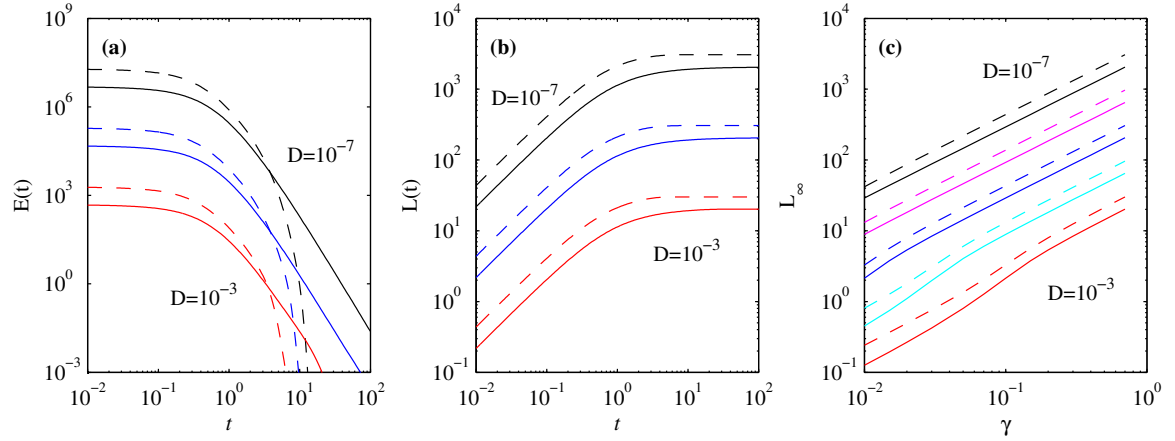
$$
S(t) = -\int dx\, p(x,t) \ln p(x,t).
\tag{34}
$$

18

FIG. 6: (a) and (b) show $\mathcal{E}$ and $\mathcal{L}$, respectively, as functions of time, for $\gamma = 0.7$. (c) shows $\mathcal{L}_\infty$ as a function of γ. All three panels are for BP only. The solid lines are the on-quenching process considered here; the dashed lines are for the off-quenching process considered in [41]. $D = 10^{-3}$ to 10^{-7} as indicated.

Here we use units in which the Boltzmann constant $K_B = 1$. Unlike the usual entropy, the absolute value of the differential entropy does not have a physical meaning, only the difference between two values of the differential entropy being meaningful.

To elucidate the difference in S between the critical and subcritical states, we use equilibrium PDFs of FP and BP (p_F and p_B in Eqs. (8) and (7), respectively) and quantify the difference between $S(t = 0)$ and $S(t \to \infty)$ in FP and BP. For the equilibrium of FP p_F in Eq. (8), we can show that the entropy Eq. (34) takes the following form [49]

$$S_F = \frac{1}{2}\left[1 + \ln\frac{\pi}{\beta_F}\right] + 2\beta_F x_0^2\left[1 - erf(\sqrt{\beta_F}x_0)\right] - \sqrt{\frac{\beta_F}{\pi}}2x_0 e^{-\beta_F x_0^2} + \Delta. \qquad (35)$$

Here, $erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x du \, \exp(-u^2)$ is the error function; $\beta_F = \gamma/D$; Δ is a function of β_F and x_0, taking the value $0 \leq \Delta \leq \ln 2$. For a sufficiently narrow PDF with $\beta_F x_0^2 \gg 1$, Δ takes the maximum value $\ln 2$ (see [49]). Since in this limit $\beta_F x_0^2 \gg 1$, $erf(\sqrt{\beta_F}x_0) \to 1$, Eq. (35) is simplified as

$$S_F \sim \frac{1}{2}\left[1 + \ln\frac{\pi}{\beta_F}\right] + \ln 2 = \frac{1}{2}\left[1 + \ln\frac{\pi D}{\gamma}\right] + \ln 2. \qquad (36)$$

For small values of D as used in our numerical computations, S_F is negative, signifying a strongly localised PDF.

19

For BF, for simplicity we use the final equilibrium p_B in Eq. (7) or (12) and $\beta_B = \frac{\gamma}{2D}$ to obtain the differential entropy S_B

$$S_B = -\int_{-\infty}^{\infty} dx\, p_B \ln p_B = \frac{1}{4}\left[\ln\frac{\Gamma(\frac{1}{4})}{2} + 1 + \ln\frac{4D}{\mu}\right]. \tag{37}$$

For small D, $S_F \sim \frac{1}{2}\ln\frac{D\pi}{\gamma}$ while $S_B \sim \ln\frac{4D}{\mu}$. Thus, the difference in differential entropy between p_F and p_B is

$$\Delta S = S_F - S_B = \frac{1}{4}\ln\frac{D\mu\pi}{\gamma^2}, \tag{38}$$

which is negative for small D. That is, the quartic exponential PDF at the critical state has much larger entropy than the bimodal PDF at the subcritical state.
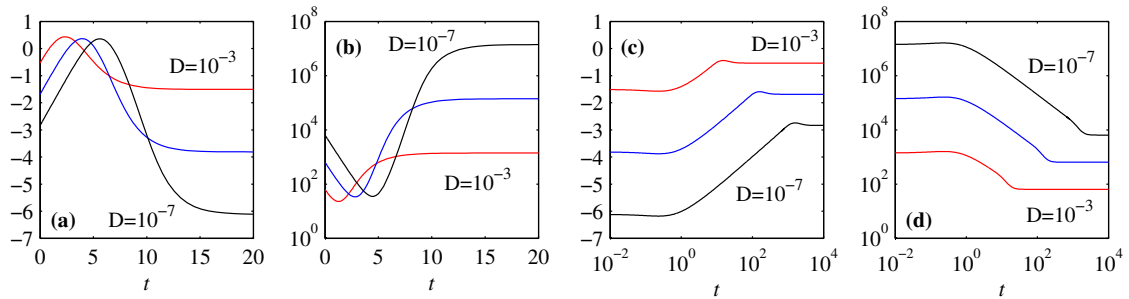

FIG. 7: (a) Entropy $S(t)$ for FP; (b) Fisher information $I(t)$ for FP; (c) Entropy $S(t)$ for BP; (d) Fisher information $I(t)$ for BP. Note how S and I are essentially opposites of each other. Again also note how the equilibration timescale is $O(\ln D^{-1})$ for FP, and $O(D^{-1/2})$ for BP.

Fig. 7 shows the time evolution of (Gibbs) differential entropy defined by Eq. (34). As theoretically predicted above, we see much larger difference between the initial and final states compared with the off-quenching case [41]. It is interesting to observe that for the forward process, S takes its maximum values at times 2.4, 4.0, and 5.5, very close to where H took its maximum values, and both broadly following the t_2 scaling from Eq. (23).

To complement $S(t)$, we also show in Fig. 7 the Fisher information defined by

$$I(t) = \int \frac{1}{p}\left[\frac{\partial p(x,t)}{\partial x}\right]^2 dx. \tag{39}$$

As the Fisher information measures the degree of 'order', increasing as the PDF develops large gradients, it shows the opposite tendency to $S(t)$, which increases with the degree of 'disorder'. In particular, the Fisher information $I(t)$ in Fig. 7(b) for FP takes the minimum value around t_m where the entropy $S(t)$ is maximum and starts increasing beyond $t > t_m$. These results thus confirm that t_m marks the start of the formation of order, as noted previously.

VI. CONCLUSION

We investigated information geometry associated with order-to-disorder and disorder-to-order transitions in a 0D Ginzburg-Landau model where the formation (disappearance) of an ordered state is modelled by the transition from a unimodal (bimodal) to bimodal (unimodal) PDF of a stochastic variable x. Our 0D model permitted us to perform a detailed statistical analysis. We considered on-critical quenching with a pair of forward and backward processes FP and BP for disorder-to-order (critical to subcritical) and order-to-disorder (subcritical to critical) transitions, respectively by selecting the initial PDF of FP/BP the same as the final equilibrium PDF of BP/FP. A pair of disorder-to-order and order-to-disorder transitions models a burst, for example, in the gene expression consisting of a pair of induction and repression (e.g. see [50]). In such bistable systems, a continuous switching between ordered and disordered states is often observed, the transition occurring in bursts interspersed by a quiescent period (e.g. see [50]). For our cyclic order-disorder transition, an initial condition represents the 'resting' state between the two bursts. We thus paid particular attention to the effect of initial conditions on information change by comparing on-quenching and off-quenching cases.

We showed that FP and BP exhibit strikingly different evolution of time-dependent PDFs during transient relaxation due to non-equilibrium initial PDFs. In particular, FP driven by instability undergoes the broadening of the PDF with large increase in (anomalous) fluctuations before the transition to the ordered state accompanied by narrowing the PDF width/decreased fluctuation. This large fluctuation essentially facilitates the existence of a geodesic solution in FP. This geodesic solution is a result of the self-regulation between

the positive feedback (γx) and the negative feedback $(-\mu x^3)$ which regulate each other, minimising the information change. In a biological context, this minimal geodesic path could be understood in terms of 'fitness' in the growth phase (e.g. gene expression). This suggests that the predator-prey type self-regulation with a nonlinear interaction facilitates a geodesic. In comparison, BP is mainly driven by the macroscopic motion due to the movement of the PDF peak, with much less prominent appearance of a geodesic solution. Specifically, the information length $\mathcal{L}$ was found to be much larger in BP than in FP, scaling as $0.9\gamma D^{-1/2}$ for BP, but only $2.1\ln(\gamma D^{-1/2})$ for FP, where D is the strength of an additive stochastic noise with a short correlation time. These results demonstrate a great advantage of $\mathcal{L}$ in revealing different physical processes (diffusion/advection) and different role of diffusion D in transition.

To elucidate the importance of the initial condition between two bursts in cyclic transition, we summarise the striking differences between on-quenching and off-quenching as follows: i) for FP, double-peaks emerge essentially immediately in on-quenching compared to their appearance only after a finite time in off-quenching; ii) for FP, the on-quenching has the equilibration time shorter by a factor of 2 and information length $\mathcal{L}_\infty$ slightly less than in off-quenching; iii) for BP, the equilibration time is much longer in on-quenching than in off-quenching, because the final state is at critical; iv) for BP, the information length $\mathcal{L}_\infty$ is nevertheless reduced by a factor of 1.5 than in off-quenching. It is worth noting that from the perspective of a system's 'fitness', the result ii) could be advantageous when adjusting to a changing environment is costly, and thus the minimum total change (measured by $\mathcal{L}_\infty$) and the minimum equilibration time are beneficial (see below). We highlight that $\mathcal{L}_\infty$ is a 'Lagrangian' measure that quantifies the total change in information content in the system over time. We discuss this further in the following.

We note that our control parameter models the effect of environment (e.g. the temperature of the heat bath, etc) and thus a sudden change in the control parameter represents a sudden change in environment. The time-evolution of PDFs occurs in order for the system to reach a new equilibrium state as the equilibrium state is optimal for the given new parameter (for the new environment). On the other hand, the smaller information length represents the smaller number of different states that a system undergoes to reach this

22

new equilibrium state. Intriguingly, these seem to be closely related to the novel concept in microbial metabolism that states evolve under the trade-off between two principles: optimality under one given condition and minimal adjustment between conditions [51]. That is, when an environment changes, the initial state (optimal for the old environment) should change to the new optimal state (the final equilibrium) by undergoing time-evolution. And the smaller the information length, the less change in the system in adjustment. Thus our results suggest that the initial "critical" state would be more advantageous for the system in changing environment.

Finally, in future work it is planned to extend this model to more realistic cases (e.g. 1D or 2D models, a system of coupled equations, etc.).

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

We first show the relation between $\tau(t)$ in Eq. (2) and the second derivative of the relative entropy (or Kullback-Leibler divergence) $D(p_1, p_2) = \int dx\, p_2 \ln\left(p_2/p_1\right)$ where $p_1 = p(x, t_1)$ and $p_2 = p(x, t_2)$ as follows:

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

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

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

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

By taking the limit where $t_2 \to t_1 = t$ ($p_2 \to p_1 = p$) and by using the total probability conservation (e.g. $\int dx \partial_t p = 0$), Eqs. (A1) and (A3) above lead to

$$\lim_{t_2 \to t_1 = t} \frac{\partial}{\partial t_1} D(p_1, p_2) = \lim_{t_2 \to t_1 = t} \frac{\partial}{\partial t_2} D(p_1, p_2) = \int dx \partial_t p = 0,$$

while Eqs. (A2) and (A4) give

$$\lim_{t_2 \to t_1 = t} \frac{\partial^2}{\partial t_1^2} D(p_1, p_2) = \lim_{t_2 \to t_1 = t} \frac{\partial^2}{\partial t_2^2} D(p_1, p_2) = \int dx \frac{(\partial_t p)^2}{p}.$$

See also [37] for similar derivation.

To link this to information length $\mathcal{L}$, we then express $D(p_1, p_2)$ for small $dt = t_2 - t_1$ as

$$D(p_1, p_2) = \left[\int dx \frac{(\partial_{t_1} p(x, t_1))^2}{p} \right] (dt)^2 + O((dt)^3), \qquad (A5)$$

where $O((dt)^3)$ is higher order term in dt. We define the infinitesimal distance (information length) $dl(t_1)$ between t_1 and $t_1 + dt$ by

$$dl(t_1) = \sqrt{D(p_1, p_2)} = \sqrt{\int dx \frac{(\partial_t p)^2}{p} dt} + O((dt)^{3/2}). \qquad (A6)$$

The total change in information between time 0 and t is then obtained by summing over $dt(t_1)$ and then taking the limit of $dt \to 0$ as

$$
\begin{aligned}
\mathcal{L}(t) &= \lim_{dt \to 0} [dl(0) + dl(dt) + dl(2dt) + dl(3dt) + \cdots dl(t - dt)] \\
&= \lim_{dt \to 0} \left[\sqrt{D(p(x, 0), p(x, dt))} + \sqrt{D(p(x, dt), p(x, 2dt))} + \cdots \sqrt{D(p(x, t - dt), p(x, t))} \right] \\
&\propto \int_0^t dt_1 \sqrt{\int dx \frac{(\partial_{t_1} p)^2}{p}}. \qquad (A7)
\end{aligned}
$$

Appendix B: Derivation of Eq. (28)

For small $\overline{\alpha} x^2 < 1$, we approximate $p(x, t)$ in Eq. (17)

$$
\begin{aligned}
p(x, t) &\sim M \beta_0^{\frac{1}{4}} \frac{e^{-\gamma t}}{(1 - \overline{\alpha} x^2)^{\frac{3}{2}}} e^{-\beta_0 e^{-4\gamma t} \left(\frac{x^2}{1 - \overline{\alpha} x^2} \right)^2} \\
&\sim M \beta(t)^{\frac{1}{4}} e^{-\beta(t) x^4}, \qquad (B1)
\end{aligned}
$$

where the normalisation factor M and $\beta(t)$ are

$$
\begin{aligned}
M &= 2 \left[\Gamma\left(\frac{1}{4}\right) \right]^{-1}, \\
\beta(t) &= \beta_0 e^{-\gamma t}, \\
\beta_0 &= \frac{\mu}{4D}. \qquad (B2)
\end{aligned}
$$

Then,

$$\frac{\partial p}{\partial t} = \frac{d\beta(t)}{dt} \left(\frac{1}{4\beta} - x^4 \right) p(x, t). \qquad (B3)$$

24

Thus, Eq. (2) becomes

$$
\begin{aligned}
\frac{1}{[\tau(t)]^2} &= \int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t} \right]^2 \\
&= \dot{\beta}^2 \left[\frac{1}{16\beta^2} - \frac{1}{2\beta} \langle x^4 \rangle + \langle x^8 \rangle \right] \\
&= \frac{\dot{\beta}^2}{4\beta^2}.
\end{aligned}
\tag{B4}
$$

Here, we used $\langle x^4 \rangle = \frac{1}{4\beta}$ and $\langle x^8 \rangle = \frac{5}{16\beta^2}$; the dot denotes the time derivative. Thus, using Eqs. (B3) and (B4) in $\mathcal{L}(t)$ in Eq. (1) gives us

$$
\mathcal{L}(t) = \int_0^t dt_1 \frac{1}{2\beta} \frac{d\beta(t)}{dt} = \frac{1}{2} \left| \ln \left(\frac{\beta(t)}{\beta(t=0)} \right) \right|.
\tag{B5}
$$

Appendix C: Properties of the sum of two Gaussian PDFs

We recall that for a single Gaussian PDF with mean value $z = \langle x \rangle$ and variance $\sigma = \langle (\delta x)^2 \rangle$, τ in Eq. (2) is given by (e.g. [26, 27])

$$
\begin{aligned}
\frac{1}{[\tau(t)]^2} &= \frac{1}{2\beta(t)^2} \left(\frac{d\sigma}{dt} \right)^2 + 2\beta \left(\frac{dz}{dt} \right)^2 \\
&= \frac{1}{2\sigma(t)^2} \left(\frac{d\sigma}{dt} \right)^2 + \frac{1}{\sigma} \left(\frac{dz}{dt} \right)^2.
\end{aligned}
\tag{C1}
$$

Here, $\sigma = 1/2\beta$.

We now show the information length for double Gaussian PDFs which are well-separated is approximately the same as that for a single Gaussian PDF. To this end, for a double Gaussian, we let

$$
\begin{aligned}
p &= p_1 + p_2 = N(t)[\tilde{p}_1 + \tilde{p}_2], \\
N(t) &= \frac{\sqrt{\beta(t)}}{2\sqrt{\pi}}, \\
\tilde{p}_1 &= e^{-\beta(t)(x+x_0)^2} = e^{-\beta(t)x_1^2}, \\
\tilde{p}_2 &= e^{-\beta(t)(x-x_0)^2} = e^{-\beta(t)x_2^2}.
\end{aligned}
\tag{C2}
$$

Here, N is the normalisation constant (e.g. $N^{-1} = \int dx(\tilde{p}_1 + \tilde{p}_2)$) and $x_1 = x + x_0$ and $x_2 = x - x_0$.

To show Eq. (28), we assume x_0 is constant given by the peak location $x_0 = \sqrt{\frac{\gamma}{\mu}}$ in $x > 0$ while $\beta = \beta(t)$ depending on time. Then, we can show

$$\frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2 = \frac{\dot{N}^2}{N}(\tilde{p}_1 + \tilde{p}_2) + 2\dot{N}(\dot{\tilde{p}}_1 + \dot{\tilde{p}}_2) + N\frac{(\dot{\tilde{p}}_1 + \dot{\tilde{p}}_2)^2}{\tilde{p}_1 + \tilde{p}_2}. \tag{C3}$$

Now, we compute the various quantities in Eq. (C3) as follows:

$$\dot{\tilde{p}}_1 = -\dot{\beta}x_1^2\tilde{p}_1 = \dot{\beta}\partial_\beta\tilde{p}_1,$$
$$(\dot{\tilde{p}}_1)^2 = \dot{\beta}^2\tilde{p}_1\partial_{\beta\beta}\tilde{p}_1. \tag{C4}$$

Similarly,

$$\dot{\tilde{p}}_2 = -\dot{\beta}x_2^2\tilde{p}_2 = \dot{\beta}\partial_\beta\tilde{p}_2,$$
$$(\dot{\tilde{p}}_2)^2 = \dot{\beta}^2\tilde{p}_2\partial_{\beta\beta}\tilde{p}_2. \tag{C5}$$

Thus, by using Eqs. (C4) and (C5), we calculate the last term in Eq. (C3) as follows:

$$(\dot{\tilde{p}}_1 + \dot{\tilde{p}}_2)^2 = \dot{\beta}^2\left[\tilde{p}_1\partial_{\beta\beta}\tilde{p}_1 + \tilde{p}_2\partial_{\beta\beta}\tilde{p}_2 + 2\partial_\beta\tilde{p}_1\partial_\beta\tilde{p}_2\right] \tag{C6}$$
$$= \dot{\beta}^2\left[(\tilde{p}_1 + \tilde{p}_2)\partial_{\beta\beta}\tilde{p}_1 + (\tilde{p}_1 + \tilde{p}_2)\partial_{\beta\beta}\tilde{p}_2 + G_1\right] \tag{C7}$$
$$= \dot{\beta}^2\left[(\tilde{p}_1 + \tilde{p}_2)\partial_{\beta\beta}(\tilde{p}_1 + \tilde{p}_2) + G_2\right], \tag{C8}$$

where G_1 and G_2 are terms involving the product of $\tilde{p}_1$ and $\tilde{p}_2$. For the PDF peaks that are well-separated and thus independent, there is no overlap between $\tilde{p}_1$ and $\tilde{p}_2$ in x, leading to $\int dx\tilde{p}_1(x)\tilde{p}_2(x) = 0$. That is, in this case, $\int dx\, G_1 = \int dx\, G_2 = 0$. Thus, these terms G_1 and G_2 do not contribute to Eq. (2). By using these results in Eq. (2), we obtain

$$\int dx\frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2 = \frac{\dot{N}^2}{N^2} + 2\dot{\beta}\dot{N}\partial_\beta\frac{1}{N} + N\dot{\beta}^2\partial_{\beta\beta}\frac{1}{N}. \tag{C9}$$

By using $N = \frac{1}{2}\sqrt{\frac{\beta}{\pi}}$, we simplify Eq. (C9) as

$$\int dx\frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2 = \frac{\dot{\beta}^2}{2\beta^2} = \frac{\dot{\sigma}^2}{2\sigma^2}. \tag{C10}$$

Thus, Eq. (C10) is the same as Eq. (C1) in the limit $z = 0$. We note that Eq. (28) is obtained by the time integral of Eq. (C10) by using the results in Appendix B.

26

Next to show Eq. (30), we need to consider the case where β is constant in Eq. (C2) while $x_0 = x_0(t)$ depends on time. In this case, we have

$$
\begin{aligned}
(\dot{\tilde{p}}_1 + \dot{\tilde{p}}_2)^2 &= 4\beta^2 \dot{x_0}^2 N^2 \left[x^2(\tilde{p}_1 + \tilde{p}_2)^2 + 2xx_0(\tilde{p}_1^2 - \tilde{p}_2^2) + x_0^2(\tilde{p}_1 - \tilde{p}_2)^2 \right] \\
&= 4\beta^2 \dot{x_0}^2 N^2 \left[x^2(\tilde{p}_1 + \tilde{p}_2)^2 + 2xx_0(\tilde{p}_1^2 - \tilde{p}_2^2) + x_0^2(\tilde{p}_1 + \tilde{p}_2)^2 + G_3 \right], \quad \text{(C11)}
\end{aligned}
$$

where G_3 is a function depending on the product of $\tilde{p}_1$ and $\tilde{p}_2$, which vanishes upon integral over x when $\tilde{p}_1$ and $\tilde{p}_2$ are well-separated with negligible overlap. In this case,

$$
\begin{aligned}
\int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t} \right]^2 &= 4\beta^2 \dot{x_0}^2 N \int dx \left[(x+x_0)^2 \tilde{p}_1 + (x-x_0)^2 \tilde{p}_2 \right] \\
&= -4\beta^2 \dot{x_0}^2 N \partial_\beta \int dx \, (\tilde{p}_1 + \tilde{p}_2) \\
&= -4\beta^2 \dot{x_0}^2 N \partial_\beta \frac{1}{N} \\
&= 2\beta \dot{x_0}^2, \quad \text{(C12)}
\end{aligned}
$$

where we used $N = \frac{1}{2}\sqrt{\frac{\beta}{\pi}}$ and thus $\partial_\beta \frac{1}{N} = -\frac{1}{2\beta N}$. Eq. (C12) is the same as Eq. (C1) in the opposite limit where $z = x_0$ and $\dot{\beta} = 0$.

---

[1] T.W.B. Kibble, Phys. Rep. **67**, 183 (1980).

[2] Y. Nagashima and Y. Nambu, Elementary Particle Physics: Quantum Field Theory and Particles (Wiley-VCH, Weinheim, 2010).

[3] G. Mazenko, Physica A **204** 437 (1994).

[4] G. Longo and M. Montévil, Progress in Biophysics and Molecular Biology. Systems Biology and Cancer **106**, 340 (2011)

[5] T. Bossomaier, R. Barnett and M. Harré, Complex Adaptive Systems Modeling **1**, 9 (2013).

[6] H. Haken, Information and Self-Organization: A Macroscopic Approach to Complex Systems (Springer, Berlin, 2006).

[7] E. Kim and P.H. Diamond, Phys. Rev. Lett. **90**,185006 (2003).

[8] E. Kim, Phys. Rev. Lett. **96**, 084504 (2006).

[9] K. Srinivasan and W.R. Young, J. Atmos. Sci. **69**, 1633 (2012).

[10] K.M. Sayanagi, A.P. Showman and T.E. Dowling, J. Atmos. Sci. **65**, 12 (2008).

[11] E. Kim, H. Liu and J. Anderson, Phys. Plasmas **16**. 052304 (2009).

[12] A.P.L. Newton, E. Kim and H.-L. Liu, Phys. Plasmas **20**, 092306 (2013)

[13] M. Tsuchiya, A. Giuliani, M. Hashimoto, J. Erenpreisa and K. Yoshikawa, PLoS One **10**, e0128565 (2015).

[14] C. Tang and P. Bak, J. Stat. Phys. **51**, 797 (1988).

[15] H.J. Jensen, Self-organized Criticality: Emergent Complex Behavior in Physical and Biological Systems (Cambridge Univ. Press, Cambridge, 1998).

[16] G. Pruessner, Self-organised Criticality (Cambridge Univ. Press, Cambridge, 2012).

[17] S. Fauve and F. Heslot, Phys. Lett. A **97**, 5 (1983).

[18] D. Angeli, J.E. Ferrell, and E. D. Sontag, Proc. Nat. Acad. Sci. **101**, 1822 (2004).

[19] D. Holcman and M. Tsodyks, PLoS Comp Biol **2**, e23 (2006).

[20] J.F. Mejias, H.J. Kappen, and J.J. Torres, PLoS One **5**, e13651 (2010).

[21] J. Hidalgo, L. F. Seoane, J. M. Cortés, and M. A. Munoz, PloS One **7**, e40710 (2012).

[22] S. Tyagi, Mol. Syst. Biol. **11**, 805 (2015).

[23] S. di Santo, R. Burioni, A. Vezzani, and M.A. Muñoz, Phys. Rev. Lett. **116**, 240601 (2016).

[24] S.B. Nicholson and E. Kim, Phys. Lett. A. **379**, 8388 (2015).

[25] S.B. Nicholson and E. Kim, Entropy **18**, 258, e18070258 (2016).

[26] J. Heseltine and E. Kim, J. Phys. A **49**, 175002 (2016).

[27] E. Kim, U. Lee, J. Heseltine and R. Hollerbach, Phys. Rev. E **93**, 062127 (2016).

[28] E. Kim and R. Hollerbach, Phys. Rev. E **95**, 022137 (2017).

[29] A.L. Gibbs and F.E. Su, Int. Stat. Rev. **70**, 419 (2002).

[30] W.K. Wootters, Phys. Rev. D **23**, 357 (1981).

[31] G. Ruppeiner, Phys. Rev. A. **20**, 1608 (1979).

[32] F. Schlögl, Z. Phys. B – Cond. Matt. **59**, 449 (1985).

[33] E.H. Feng and G.E. Crooks, Phys. Rev. E. **79**, 012104 (2009).

[34] S.L. Braunstein and C.M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[35] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D.B. Hume, L. Pezzé, A. Smerzi, M.K. Oberthaler, Science **345**, 424 (2014).

[36] J. Nulton, P. Salamon, B. Andresen and Qi. Anmin, J Chem. Phys. **83**, 334 (1985).

[37] G.E. Crooks, Phys. Rev. Lett. **99**, 100602 (2007).

[38] D.A. Sivak and G.E. Crooks Phys. Rev. Lett. **8**, 190602 (2012).

[39] P. Salamon. J.D. Nulton, G. Siragusa, A. Limon, D. Bedeaus and S. Kjelstrup, J. Non-Equilib. Thermodyn. **27**, 45 (2002).

[40] J.K. Bhattacharjee, P. Meakin and D.J. Scalapino, Phys. Rev. A. **30**, 1026 (1984).

[41] E. Kim and R. Hollerbach, Phys. Rev. E, in press (2017).

[42] E. Kim and R. Hollerbach, Phys. Rev. E **94**, 052118 (2016).

[43] F. Klebaner, Introduction to Stochastic Calculus with Applications (Imperial College Press, London, 2012).

[44] H. Risken, The Fokker-Planck Equation: Methods of Solution and Applications (Springer, Berlin, 1996).

[45] M. Suzuki, Phys. Lett. A **67**, 339, (1978).

[46] M. Suzuki, Prog. Theo. Phys. **56**, 77 (1976); J. Stat. Phys. **16**, 477 (1977).

[47] B. Caroli, C. Caroli and B. Roulet, J. Stat. Phys. **21**, 415 (1979).

[48] R. Kubo, K. Matsuo and K. Kitahara, J. Stat. Phys. **9**, 51 (1973).

[49] J.V. Michalowicz, J.M. Nichols and F. Bucholtz, Entropy **10**, 200 (2008).

[50] M.L. Ferguson, D. Le Coq, M. Jules, S. Aymerich, O. Radulescue, N. Declerck and C.A. Royer, PNAS, **109**, 155 (2012).

[51] R. Schuetz,N. Zamboni, M. Zampieri, M. Heinemann and U. Sauer, Science **336**, 601 (2012).