Geometric structure and geodesic in a solvable model of nonequilibrium process

Eun-jin Kim,[1] UnJin Lee,[2] James Heseltine,[1] and Rainer Hollerbach[3]
[1]*School of Mathematics and Statistics, University of Sheffield, Sheffield, S3 7RH, United Kingdom*
[2]*Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA*
[3]*Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom*
(Received 19 December 2015; revised manuscript received 25 April 2016; published 20 June 2016)

We investigate the geometric structure of a nonequilibrium process and its geodesic solutions. By employing an exactly solvable model of a driven dissipative system (generalized nonautonomous Ornstein-Uhlenbeck process), we compute the time-dependent probability density functions (PDFs) and investigate the evolution of this system in a statistical metric space where the distance between two points (the so-called information length) quantifies the change in information along a trajectory of the PDFs. In this metric space, we find a geodesic for which the information propagates at constant speed, and demonstrate its utility as an optimal path to reduce the total time and total dissipated energy. In particular, through examples of physical realizations of such geodesic solutions satisfying boundary conditions, we present a resonance phenomenon in the geodesic solution and the discretization into cyclic geodesic solutions. Implications for controlling population growth are further discussed in a stochastic logistic model, where a periodic modulation of the diffusion coefficient and the deterministic force by a small amount is shown to have a significant controlling effect.

DOI: 10.1103/PhysRevE.93.062127

I. INTRODUCTION

A probabilistic description is essential for understanding the dynamics of stochastic systems far from equilibrium, given the uncertainty inherent in such systems. To compare different probability density functions (PDFs), it is extremely useful to quantify the differences among PDFs by assigning an appropriate metric to probability. This metric structure then provides a key link between stochastic systems and geometry. Depending on the question of interest, different metrics have been proposed (e.g., Refs. [1–10] and further references therein). For instance, the Wasserstein metric has been studied extensively by many authors in the optimal transport problem [9] in which the key problem is to minimize transport cost, which is typically taken to increase quadratically with the distance between two locations. For Gaussian measures, the Wasserstein metric is defined in the product space consisting of Euclidean and positive symmetric matrices for the mean and variance, respectively (e.g., see Ref. [4]). Compared with the Wasserstein metric, whose application has established itself as a branch of applied mathematics, the geometric structure associated with the information change in the Fisher (or Fisher-Rao) metric seems to be explored much less. Unlike the Wasserstein distance, the Fisher metric provides a hyperbolic geometry in the upper half plane (e.g., Refs. [2,7]) where the distance is measured in units of the width of the PDF. That is, the distance in the Fisher metric is dimensionless and represents the number of different states in the statistical space. Such a notion was proposed in the seminal work [11] where statistical distance was introduced as the number of distinguishable states between two PDFs. The purpose of our paper is to generalize this concept to nonequilibrium systems and to quantify the rate of information flow by computing the change in the number of indistinguishable states within these processes. This generalization will endow nonequilibrium processes with geometric structure, providing an alternative perspective on the link between stochastic processes and geometry.

The Fisher metric for Gaussian measures is related to the covariance, thereby also relating to fluctuations in the systems. Specifically, Ref. [12] related the second moment of fluctuations to the inverse of a metric tensor since strong correlations between any point and its neighbors may emerge from large fluctuations, resulting in shorter distances around that point. In other words, the distance between different thermodynamic states is normalized by the resolution (the unit of distance) set by the strength of fluctuations. Such fluctuation-based metrics in thermodynamic states have been studied near equilibrium [11–14], for instance, in the comparison of two equilibrium states via a statistical distance, or the interpretation of the interaction in a system via the curvature of the metric tensor (e.g., near phase transitions). A similar metric structure was also utilized in quantum systems [11,15,16]. Generalization of this concept to nonequilibrium systems was attempted by different authors, although they tend to be limited to the analysis of systems in near equilibrium [14,17–19]. Recent efforts include the application of this concept to minimize entropy production within a controlled system, or even the experimental measurement of statistical distance as a tool to validate theory [14,20–22]. As many systems in nature are not near equilibrium due to intrinsic variability, heterogeneity, or uncertainty in a system [23,34], our recent work [23–25] focused on physical implications of the metric for the structure of an attractor and the information flow in a strongly out-of-equilibrium system (e.g., music). In particular, Ref. [25] presents a mapping between the nonequilibrium state and the distance to an attractor by information length; Ref. [24] analyzed classical music by constructing time-dependent PDFs from the music data stored in MIDI files.

The purpose of this paper is to investigate the information change associated with nonequilibrium stochastic processes by using the Fisher information metric and to provide a link between geometric structure and a nonequilibrium process within a strongly out-of-equilibrium system using an exactly solvable model. We then examine implications of a geodesic for which the information propagates at a constant speed.

In particular, we show how our results can be utilized in controlling a system. It is the aim of our work to inform the potential utility of information length, which can provide a powerful tool to unify different nonequilibrium processes. The remainder of the paper is organized as follows. Section II provides an information interpretation of out-of-equilibrium processes. Section III introduces our model (the generalized nonautonomous Ornstein-Uhlenbeck process) and provides some important statistical relations. A general geodesic solution is presented in Sec. IV, and specific solutions with prescribed boundary conditions are investigated in Sec. V. Section VI expands on physical realizability of a geodesic solution. An example of controlling a system by utilizing a geodesic motion is presented in Sec. VII. Conclusions are provided in Sec. VIII. Appendixes A–E contain detailed steps in deriving equations used in the text, as well as the derivation of fluctuating Hamiltonian in relation to information velocity and an equation of motion in the curved metric space, the Christoffel and curvature tensors. Some of the included derivations are quite basic and are similar to related analyses by other researchers but are nevertheless included here to make this paper self-contained.

II. INFORMATION CHANGE AND FLOW

As noted in Sec. I, the fluctuation-based Fisher metric provides the number of states measured in units of the resolution, which is set by the strength of fluctuations. To elucidate the meaning of the resolution, it is worth recalling that in equilibrium thermodynamics, the fluctuation of random variables is determined by the properties of the heat bath, which is assumed to be fixed with infinite capacity. For instance, for the Maxwell-Boltzmann distribution, the probability distribution function (PDF) of the state energy E is given by

$$p(E) = \beta e^{-\beta E}, \tag{1}$$

where $\beta = 1/k_B T$ (k_B is the Boltzmann constant, T the temperature of the heat bath) is the inverse temperature. When $E \propto x^2$ (where x is the velocity of a particle), fluctuations in velocity x in the system are proportional to the thermal energy $k_B T$ of the heat bath, which determines the variance (the width) of the PDF in Eq. (1), and consequently the resolution on which the state E is differentiated. The finer the resolution, the more distinguishable different states are, and therefore the amount of accessible information in the system increases. This can alternatively be interpreted that the thermal energy of the heat bath provides a unit of energy for the probability, setting the unit of the information. This is consistent with the view that the information increases with the increase in the gradient of the PDF.

Many systems in nature are, however, far from equilibrium and there is no fixed environment that can serve as a heat bath for these systems [26–34]. In fact, one of the important characteristics of these systems is that they are open and continuously interact with their environment. The resolution of the PDFs and thus the unit of information evolve dynamically at the same time as the PDFs change with time. It is thus important to model a system without advocating the presence of an unphysical, artificial heat bath. One such method is to use a stochastic forcing, for instance, via the following Langevin equation:

$$\frac{dx}{dt} = F(x) + \xi. \tag{2}$$

Here, x is a random variable, and F is a deterministic force; ξ is a stochastic forcing, which can for simplicity be taken as a short correlated random forcing as follows:

$$\langle \xi(t)\xi(t')\rangle = 2D(t)\delta(t-t'). \tag{3}$$

In Eq. (3), the angular brackets represent the average over ξ, $\langle \xi \rangle = 0$, and $D(t)$ is the strength of the forcing, which can be prescribed as a function of time t. In this model, the stochastic forcing ξ plays the role of heat reservoir in terms of the maintenance of the fluctuations in the system, and in equilibrium the energy provided by the stochastic forcing is balanced by the energy dissipation (the so-called fluctuation-dissipation theorem). This model permits us to investigate the time evolution of a strongly out-of-equilibrium system and the associated change in information.

As a system evolves out of equilibrium, the PDF of the state evolves in time, and subsequent information change in the system is quantified by comparing the PDFs at different times. In order to quantify the difference in PDFs, which are changing with time, we use the rate at which fluctuations change in time as the (time-dependent) resolution of the PDFs. Specifically, the rate of change in PDFs defines the following information velocity $v(t)$:

$$v^2(t) = \int dx \frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2. \tag{4}$$

The velocity v in Eq. (4) has units of inverse time, and quantifies the rate at which the (dimensionless) information changes. As shown in Appendix A, the information velocity is the rms of the fluctuating Hamiltonian in a stochastic system, and $\tau = 1/v$ thus provides a dynamic time unit as far as information is concerned. As shall be discussed shortly, a geodesic is a special path that has a constant v where the metric is locally flat with no net force acting on it. This is reminiscent of the constant speed of light as a photon travels along a geodesic in curved space time.

The total accumulated change in information is then obtained by computing the total time between the initial and final times $t = 0$ and $t = t_F$ in units of τ as:

$$\mathcal{L}(t_F) = \int_0^{t_F} dt \frac{1}{\tau(t)} = \int_0^{t_F} dt \sqrt{\int dx \frac{1}{p(x,t)}\left[\frac{\partial p(x,t)}{\partial t}\right]^2}. \tag{5}$$

The accumulated change in information given by Eq. (5) provides the total change in the information, and is the total distance between the initial and final PDFs in the statistical space. We call $\mathcal{L}(t)$ in Eq. (5) the information length. The utility of a geodesic as an optimal path that minimizes the dissipated energy (or entropy production) has been previously invoked through the inequality relation between $\mathcal{L}$ and $\mathcal{J}$ as $\mathcal{J}(t_F)t_F \geqslant [\mathcal{L}(t_F)]^2$ where $\mathcal{J}(t_F) = \int_0^{t_F} dt\ [v(t)]^2 = \int_0^{t_F} dt \int dx \frac{1}{p(x,t)}[\frac{\partial p(x,t)}{\partial t}]^2$ is the time integral of v^2. Note that this inequality follows from the Cauchy-Schwartz inequality

$\int v^2 dt \int u^2 dt \geqslant (\int vu dt)^2$ with $u = 1$. The equality holds for the minimum path where v is constant (e.g., Refs. [22,35]), and the deviation from this equality quantifies the amount of disorder in an irreversible process [35], or deviation from a geodesic. Given initial and final points in the parameter space, a geodesic is an extreme path that minimizes $\mathcal{L}$; this is discussed in Sec. III in an exactly solvable model.

A clearer geometric interpretation of the information velocity and length is possible when control parameters λ^i ($i = 1,2,3,\ldots$) of a system are known, in which case Eqs. (4)–(5) can be expressed in terms of the metric g_{ij} based on the Fisher information (see, e.g., Refs. [14,36,37]) as follows:

$$v^2(t) = \mathcal{E} = \int dx \frac{d\lambda^i}{dt} g_{ij} \frac{d\lambda^j}{dt}, \tag{6}$$

where

$$g_{ij} = \int dx\ p(x,t) \frac{\partial \ln p(x,t)}{\partial \lambda^i} \frac{\partial \ln p(x,t)}{\partial \lambda^j}. \tag{7}$$

In Eq. (6), the velocity is defined in the control parameter space λ^i, where the metric tensor g_{ij} in Eq. (7) gives the Riemannian metric [38]. For the Gaussian process that we will consider later, λ^i represents the mean value and variance [see Eqs. (27) and (28)]. That is, the evolution of a nonequilibrium system can be viewed as the motion of a particle with unit mass traveling in the parameter space with the velocity $v(t)$. Here, the distance the particle travels represents the information change. This dimensionless distance represents the number of indistinguishable states that a system undergoes during the time evolution. As shown in Appendix A, the information velocity is a measure of the rms value of fluctuating energy, and the square of the information velocity is related to the second derivative of the relative entropy (or Kullback-Leibler divergence) (see Appendix in Ref. [25]).

While v^2 is given either by Eq. (5) or Eq. (6), Eq. (5) has the advantage of enabling the computation of information velocity and length directly from experimental and observational data as long as the time-dependent PDFs can be constructed, even when control parameters or governing equations of the system are not available. For instance, Ref. [24] has analyzed the information flow and length in classical music by computing time-dependent PDFs from the music MIDI files while Ref. [23] investigated an attractor structure in a logistic map by using numerically computed time-dependent PDFs.

III. A SOLVABLE MODEL

The numerical computation of time-dependent PDFs is often extremely demanding. In order to gain a key insight into the implication of information length and geodesics, it is thus invaluable to utilize an exactly solvable model. To this end, we consider a linear, driven-dissipative system for a stochastic variable x, which damps due to a friction γ while driven by an external stochastic forcing ξ as follows:

$$\frac{dx}{dt} = -\gamma(t)[x - f(t)] + \xi. \tag{8}$$

Here, $\gamma(t)$ is a non-negative friction constant; $f(t)$ is a deterministic force, which controls the location of the equilibrium position. $\gamma(t)$ or $f(t)$ will be prescribed as a time-dependent function for our purpose of finding a geodesic motion later. For simplicity, we take the stochastic forcing ξ to have a short correlation time with the correlation function given in Eq. (2) with the amplitude $D = D(t)$, which can depend on time in general. When $f = 0$ and γ and D are constant, Eq. (8) is the Ornstein-Uhlenbeck process, which is a prototypical model for a noisy relaxation system and has been utilized and extended in many areas of physical science and financial mathematics (e.g., Refs. [39,40]).

Given an initial condition $x = x_0$ at $t = 0$, the solution to the stochastic differential equation (8) is simply

$$x(t) = x_0 e^{-G(t)} + \int_0^t dt_1 e^{-[G(t)-G(t_1)]} [\gamma(t_1) f(t_1) + \xi(t_1)], \tag{9}$$

where $G(t) = \int_0^t dt' \gamma(t')$ and $G(t_1) = \int_0^{t_1} dt' \gamma(t')$.

For the assumed Gaussian process ξ, the transition probability between the position x_0 at $t = 0$ and the final position x at time t is given by

$$P(x,t;x_0,0) = \sqrt{\frac{\beta_1(t)}{\pi}} \exp\{-\beta_1(t)[x - y_1(t)]^2\}. \tag{10}$$

Letting the angular brackets denote the average over ξ, we then have for $y_1(t)$ and β_1 as the mean values of x and the inverse temperature, respectively:

$$y_1(t) = \langle x \rangle = x_0 e^{-G(t)} + F(t), \tag{11}$$

$$\frac{1}{2\beta_1(t)} = \langle [x(t) - y_1(t)]^2 \rangle = \int_0^t dt_1 e^{-2[G(t)-G(t_1)]} 2D(t_1), \tag{12}$$

$$F(t) = \int_0^t dt_1 e^{-[G(t)-G(t_1)]} \gamma(t_1) f(t_1). \tag{13}$$

To facilitate the analysis in the general case where the initial position x_0 is random with the mean value μ, we assume the initial distribution of x_0 to be the Gaussian distribution with the inverse temperature β_0

$$P(x_0,0) = \sqrt{\frac{\beta_0}{\pi}} \exp[-\beta_0(x_0 - \mu)^2], \tag{14}$$

which has a peak at $x_0 = \mu$. The PDF at a later time is obtained by integrating the product of the transition probability (10) and the initial PDF (14) over x_0 as:

$$\begin{aligned} P(x,t) &= \int dx_0 P(x,t;x_0,0) P(x_0,0) \\ &= \sqrt{\frac{\beta(t)}{\pi}} \exp[-\beta(t)\{x - y(t))^2\}. \end{aligned} \tag{15}$$

In Eq. (15), $y(t)$ and $\beta(t)$ are the mean values averaged over x_0 and ξ as:

$$y(t) = \langle \langle x \rangle \rangle = \mu e^{-G(t)} + F(t), \tag{16}$$

$$\frac{1}{2\beta(t)} = \langle \langle [x(t) - y(t)]^2 \rangle \rangle = \langle \langle (\delta x)^2 \rangle \rangle = \frac{e^{-2G(t)}}{2\beta_0} + \frac{1}{2\beta_1}. \tag{17}$$

Here, β_1 is given in Eq. (12); the double angular brackets $\langle\langle \ldots \rangle\rangle$ now denote the average over both x_0 and ξ; $\mu = \langle\langle x_0 \rangle\rangle$ and $\delta x = x - \langle\langle x \rangle\rangle$. It is useful to note that β in Eq. (17) satisfies the following relation:

$$\beta = \frac{\beta_0 \beta_1}{\beta_1 e^{-2G(t)} + \beta_0} = \frac{\beta_0}{e^{-2G(t)} + q(t)}. \tag{18}$$

$$q(t) = \frac{\beta_0}{\beta_1} = -e^{-2G(t)} + \frac{\beta_0}{\beta}. \tag{19}$$

For simplicity, the average over both ξ and the initial x_0 will now be denoted by angular brackets (instead of double angular brackets) unless there is ambiguity in their meaning.

A. Energy budget

In order to understand the role of $D(t)$, $f(t)$, and $\gamma(t)$ in relation to information length, it is useful to examine energy relations involving the second moment of x for macroscopic energy $y^2 = \langle x \rangle^2$ and fluctuating energy $\langle (\delta x)^2 \rangle$, where $\delta x = x - y$. From Eq. (8), we obtain by using the Stratonovich calculus [39–41]:

$$\frac{d}{dt}\left(\frac{x^2}{2}\right) = -\gamma(t) x [x - f(t)] + \xi x. \tag{20}$$

The total work W_ξ by the external forcing between the initial time $t = 0$ and final time t is obtained by the time integral of the last term in Eq. (20). This is computed by using Eqs. (9) and (3):

$$\langle \xi(t) x(t) \rangle = \langle \xi(t) \delta x(t) \rangle = \int_0^t dt_1 e^{-[G(t)-G(t_1)]} \langle \xi(t) \xi(t_1) \rangle = D(t), \tag{21}$$

$$W_\xi = \int_0^t dt_1 \langle \xi(t_1) x(t_1) \rangle = \int_0^t dt_1 D(t_1). \tag{22}$$

The average of the first term on the right-hand side of Eq. (20) gives us

$$\langle \gamma (x - f) x \rangle = \gamma y^2 + \gamma \langle (\delta x)^2 \rangle - \gamma f y = \gamma y (y - f) + \frac{\gamma}{2\beta}. \tag{23}$$

Noting that the first and second terms on the right-hand side of Eq. (23) are due to the mean and fluctuations, we can separate the time integral of the average of Eq. (20) as

$$\frac{1}{2}(y^2 - \mu^2) = \int_0^t dt_1 \gamma(t_1) y(t_1) [y(t_1) - f(t_1)], \tag{24}$$

$$\frac{1}{2}\left(\frac{1}{\beta} - \frac{1}{\beta_0}\right) = -W_\gamma + W_\xi, \tag{25}$$

where W_γ is the frictional energy loss from fluctuations to the environment:

$$W_\gamma = \int_0^t dt_1 \frac{\gamma(t_1)}{2\beta(t_1)}, \tag{26}$$

and $\mu = y(t = 0) = \langle x(t = 0) \rangle$ and $\beta_0 = \beta(t = 0)$. When $\beta(t) = \beta_0$ at some time t, the left-hand side of Eq. (25) vanishes, therefore $W_\xi = W_\gamma$. That is, when the temperature is equal at the initial and final times, the work W_ξ is balanced by the total energy dissipation W_γ. Alternatively, if W_ξ and W_γ are not equal, then $(W_\xi - W_\gamma)$ and $(\beta_0 - \beta)$ having the same sign implies that the temperature, and hence the PDF width, will increase or decrease according to whether W_ξ is greater or less, respectively, than W_γ.

IV. GEODESIC MOTION

For the PDF in Eq. (15), a lengthy but straightforward algebra yields the information velocity in Eq. (4) in the following form [25]:

$$v^2 = \mathcal{E} = \frac{1}{2\beta^2}\dot{\beta}^2 + 2\beta \dot{y}^2, \tag{27}$$

where $\dot{\beta} = \frac{d\beta}{dt}$ and $\dot{y} = \frac{dy}{dt}$. By comparing Eq. (27) with Eq. (6), we can easily read off the metric tensor and control parameters as follows:

$$g_{ij} = \begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix}, \quad \lambda^i = \begin{pmatrix} \beta \\ y \end{pmatrix}. \tag{28}$$

By using the Euler-Lagrange equations

$$\frac{d\mathcal{E}}{d\beta} - \frac{d}{dt}\frac{d\mathcal{E}}{d\dot{\beta}} = 0, \tag{29}$$

$$\frac{d\mathcal{E}}{dy} - \frac{d}{dt}\frac{d\mathcal{E}}{d\dot{y}} = 0, \tag{30}$$

we obtain the coupled equations for the geodesic motion

$$\ddot{\beta} - \frac{\dot{\beta}^2}{\beta} - 2\beta^2 \dot{y}^2 = 0, \tag{31}$$

$$\frac{d}{dt}[\beta \dot{y}] = 0. \tag{32}$$

Here, Eq. (32) can be written as

$$\beta \dot{y} = c, \tag{33}$$

where c is constant. When $c = 0$, the geodesic becomes $y = \text{const}$ and $\beta \propto \ln t$. By using Eq. (33) in Eq. (31) and after some straightforward manipulation (see Appendix B), we obtain

$$\dot{\beta}^2 = -4c^2 \beta + \alpha \beta^2, \tag{34}$$

where α is another constant. To understand its physical meaning, we use Eq. (33) and Eq. (34) in Eq. (31) to obtain:

$$v^2 = \frac{1}{2\beta^2}\dot{\beta}^2 + 2c^2\beta = \frac{\alpha}{2}. \tag{35}$$

Thus, Eq. (35) implies that α is related to the information velocity as

$$v = \sqrt{\frac{\alpha}{2}}. \tag{36}$$

A solution to Eq. (34) is found after some lengthy algebra (see Appendix B) as

$$\beta(t) = \frac{2c^2}{\alpha}[\cosh \sqrt{\alpha}(t - A) + 1] = \frac{4c^2}{\alpha} \cosh^2\left[\frac{1}{2}\sqrt{\alpha}(t - A)\right], \tag{37}$$

where A is constant. By using Eqs. (37) in (33), we then find the solution for y (see also Appendix B):

$$y(t) = -\frac{\sqrt{\alpha}}{c}\frac{1}{1+e^{\sqrt{\alpha}(t-A)}} + B$$
$$= \frac{\sqrt{\alpha}}{2c}\tanh\left[\frac{1}{2}\sqrt{\alpha}(t-A)\right] - \frac{\sqrt{\alpha}}{2c} + B, \quad (38)$$

where B is another constant.

The identity $\mathrm{sech}^2\theta + \tanh^2\theta = 1$ permits us to derive a useful relationship between $\beta(t)$ and $y(t)$ from Eqs. (37) and (38) as follows:

$$\left(y + \frac{s}{\sqrt{\beta_*}} - B\right)^2 + \frac{1}{\beta} = \frac{1}{\beta_*}. \quad (39)$$

Here

$$\beta_* = \frac{4c^2}{\alpha}, \quad s = \frac{c}{|c|}, \quad (40)$$

where s represents the sign of c. That is, if we think of y and $\frac{1}{\sqrt{\beta}}$ being the variables, they are related via a circle, with radius $\frac{1}{\beta_*}$ and centered at $(0, B - \frac{s}{\sqrt{\beta_*}})$. Geodesic motions are then along portions of this circle. This is a reflection of a hyperbolic geometry (the upper half Poincaré model) formed by y and the square root of the temperature $\frac{1}{\sqrt{\beta}}$ where the center of the circle Eq. (39) is at the boundary (i.e., on the axis where $1/\sqrt{\beta} = 0$) of the upper half plane. The location of the center and the radius of the circle depend on the particular problem of interest (see the next section).

To summarize, Eqs. (37)–(38) are general solutions for the geodesic, and the values of the four constants c, α, A, and B are to be fixed by the boundary conditions at the initial $t = 0$ and final time t_F, depending on the problem of interest. A few specific examples are shown in the following sections.

V. GEODESIC EXAMPLES AND SIGNIFICANCE

We now consider specific cases of the geodesic Eqs. (37)–(38) and examine the implications for the total time required for the system to reach the final state, and the total amount of work required. As an illustration, we consider the time evolution of a nonequilibrium state shown in Fig. 1(a), where the mean position starts with the initial value $y = y_0 = \mu$ and approaches another nonequilibrium state $y = y_F$, which is closer to the equilibrium. As boundary conditions, we consider the case where the initial and final temperatures are equal. In terms of the PDFs, the width of the initial and final PDFs is thus the same, as shown in Fig. 1(a), while the mean position of the PDF moves to the final point $y = y_F$. The key question of interest would be to find a path connecting the initial and final states, which minimizes the total information change, and to examine whether this path also minimizes the time in addition to total energy dissipation. For instance, such minimization will be particularly useful when the initial state is very harmful, causing a lot of damage (e.g., a large population of bacteria, etc., causing illness).

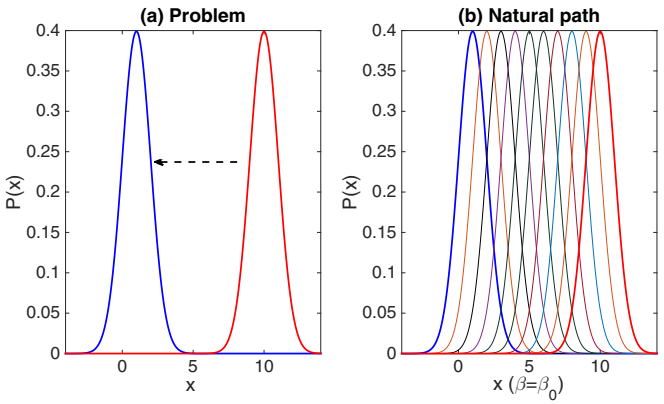

FIG. 1. (a) A sketch illustrating the problem of moving a PDF from a larger (y_0) to smaller mean position ($y_F < y_0$), where the temperature (width) of the PDF is the same at the initial and final times $t = 0$ and $t = t_F$. (b) A natural path with the same temperature $\beta = \beta_0$ for all time between $t = 0$ and $t = t_F$. The units of x are arbitrary.

A. Nongeodesic: $\beta(t) = \beta_0$, $\gamma(t) = \gamma_0$, and $f = 0$

As illustrated in Fig. 1(b), one possible and perhaps natural path connecting the initial and final points would be to decrease the mean position while keeping the same temperature as $\beta(t) = \beta_0$ for all time. This is achieved by using a constant $D = \gamma_0/2\beta_0$ and $f = 0$ while decreasing y exponentially in time through the constant frictional force as $y = y_0\exp(-\gamma t)$. For a specific example, we consider the situation where $y = y_0 = \mu$ at the initial time $t = 0$ and $y = y_F$ at the final time $t = t_F$. As β is the same at the initial and final times, Eq. (25) gives a simple relation that the work done by ξ is dissipated by the frictional force, that is $W_\gamma = W_\xi$. These enable us to compute the total time t_F and energy dissipation W_ξ in Eq. (22) simply as

$$t_F = \frac{1}{\gamma}\ln\frac{y_0}{y_F}, \quad (41)$$

$$W_\xi = W_\gamma = \int_0^{t_F} dt\, D = Dt_F = \frac{1}{2\beta_0}\ln\frac{y_0}{y_F}. \quad (42)$$

We can also compute the total information length using the result in Ref. [25] as follows:

$$\mathcal{L} = \sqrt{2\beta_0}(y_0 - y_F). \quad (43)$$

Solid black curves in Figs. 2(a), 2(b), and 2(c), respectively, show the total time t_F, $\mathcal{L}$ and W_ξ in Eqs. (24)–(43) against β_0. These will be compared with results obtained for the geodesic path.

B. Geodesics

The natural path discussed above is not a geodesic as it does not satisfy Eq. (39). For a circular geodesic motion, the change in y in time should be compensated by the change in β. Specifically, since the initial and final β are equal, β should decrease in time initially and then eventually increase back to the initial value β_0 at the final time. That is, temperature increases from $1/\beta_0$ to $1/\beta_* > 1/\beta_0$ initially, and then at some point should decrease back to recover the value $1/\beta_0$ at the

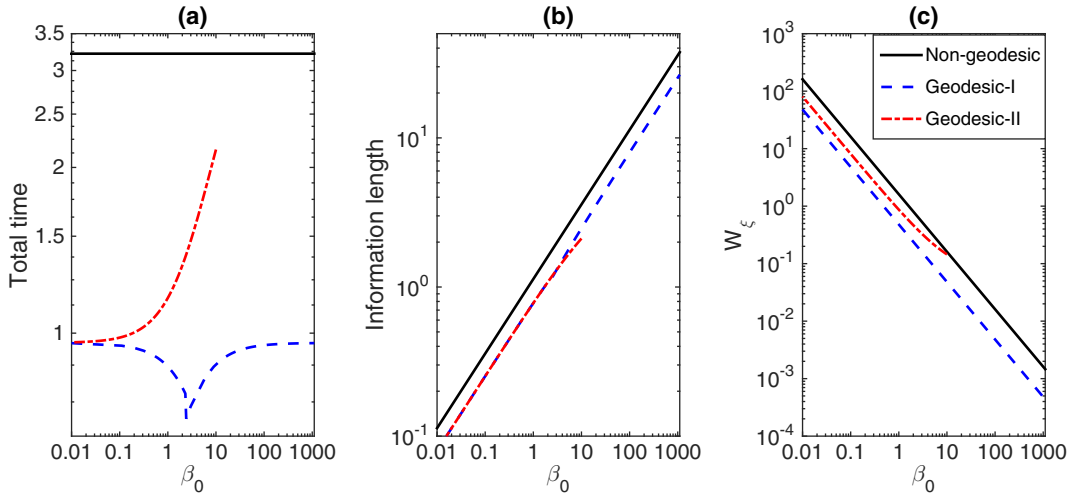

FIG. 2. (a) Total time; (b) information length; (c) W_ξ against β_0 in the nongeodesic case (in solid black) and geodesic I (in dashed blue) and II (in dash-dotted red); $y(t=0)=y_0=5/6$ and $y(t=t_F)=y_F=1/30$. Geodesic II (dash-dotted red) is shown for the value of β_0 where the diffusion (D_{II}) is non-negative. A distinct minimum in the total time is observed in geodesic I caused by the resonance (the matching of $\Delta=\Delta_m$).

final time, and furthermore the system may need to go through several cycles of periodic increase and decrease in temperature to satisfy a physical realizability [e.g., see Fig. 5(d)].

It is useful to start with the simplest case of one cycle. To be specific, we take the total time along the path to be $t_F=2A$, and the values of the temperature and y at the midpoint $t=A$ to be $\beta(t=A)=\beta_*=4c^2/\alpha<\beta_0$ and $y(t=A)=(y_0+y_F)/2\equiv y_M$, respectively. Then, from Eqs. (37) and (38), we obtain

$$\beta(t)=\beta_*\cosh^2\left[\frac{1}{2}\sqrt{\alpha}(t-A)\right], \tag{44}$$

$$y(t)=-\frac{1}{\sqrt{\beta_*}}\tanh\left[\frac{1}{2}\sqrt{\alpha}(t-A)\right]+y_M, \tag{45}$$

where we used $c<0$ (as $y_0>y_F$) and $\sqrt{\alpha}/2c=-1/\sqrt{\beta_*}$. Therefore, the conditions $y(t=0)=y_0$, $y(t=t_F)=y_F$, $\beta(t=0)=\beta(t=t_F)=\beta_0$ give us the following relations:

$$y_0-y_F\equiv\Delta=\frac{2}{\sqrt{\beta_*}}\tanh\left[\frac{1}{2}\sqrt{\alpha}A\right], \tag{46}$$

$$\sqrt{\frac{\beta_0}{\beta_*}}=\cosh\left[\frac{1}{2}\sqrt{\alpha}A\right]. \tag{47}$$

Typical behavior of y and $\beta^{-1/2}$ against time are shown in Fig. 3(a) and Fig. 3(c) for $\beta_0=0.3$ and 3, respectively, where we use $y(t=0)=y_0=5/6$ and $y(t=t_F)=y_F=1/30$ and the value of α obtained in Sec. IV B 1. Note that throughout the paper, we will use these same values $y_0=5/6$ and $y_F=1/30$ to facilitate comparison among different cases.

We recast Eq. (46) by using Eq. (47) to eliminate β_*:

$$\Delta=\frac{2}{\sqrt{\beta_0}}\sinh\left[\frac{1}{2}\sqrt{\alpha}A\right]. \tag{48}$$

Equations (47)–(48) then give us

$$A=\frac{2}{\sqrt{\alpha}}\sinh^{-1}\left[\frac{\Delta\sqrt{\beta_0}}{2}\right]=\frac{2}{\sqrt{\alpha}}\cosh^{-1}\sqrt{\frac{\beta_0}{\beta_*}}, \tag{49}$$

$$\sqrt{\frac{\beta_0}{\beta_*}}=\cosh\left[\sinh^{-1}\left(\frac{\Delta\sqrt{\beta_0}}{2}\right)\right]. \tag{50}$$

In order to determine the total time $2A$ and associated energy dissipation, we will shortly find the value of α and choose $\gamma(t)$, $D(t)$, and $f(t)$ in Eq. (8) to satisfy our derived equations above. Before doing this through specific examples, it is useful to visualize the time evolution of general geodesic solutions. To this end, we note that $\beta(t)$ and $y(t)$ in Eqs. (44) and (45) satisfy Eq. (39) where $B-s/\sqrt{\beta_*}$ is replaced by $(y_0+y_F)/2$ as

$$(y-y_M)^2+\frac{1}{\beta}=\frac{1}{\beta_*}, \tag{51}$$

where $y_M=\frac{1}{2}(y_0+y_F)$. As Eq. (51) only depends on y_0, y_F, $\Delta=y_0-y_F$, β_0 through Eq. (50), Eq. (51) is independent of the information velocity ($\sqrt{\alpha/2}$). Without specifying the value of α, we can plot the geodesic motion from Eq. (51) in Figs. 3(b) and 3(d) by using our fixed parameter values $y(t=0)=y_0=5/6$ and $y(t=t_F)=y_F=1/30$ for the two different values of $\beta_0=0.3$ and 3, respectively. They clearly show the part of a circular motion in the upper half plane y and $\beta^{-1/2}$. Although α does not change the shape of the geodesic circular motion, it affects the speed at which a trajectory travels along it. That is, the time scale on which y and β in Fig. 3(a) and 3(c) evolve depends on α (i.e., larger α, faster evolution).

For completeness, the corresponding time evolution of the PDFs is shown in Figs. 4(a) and 4(b) where the initial and final PDFs are plotted in red on the right and blue on the left. The increase followed by decrease in the width of the PDFs ($\propto\beta^{-1/2}$) with time is clearly seen. Figs. 4(c) and 4(d) are shown for larger value of β_0 to highlight the effect of β_0.

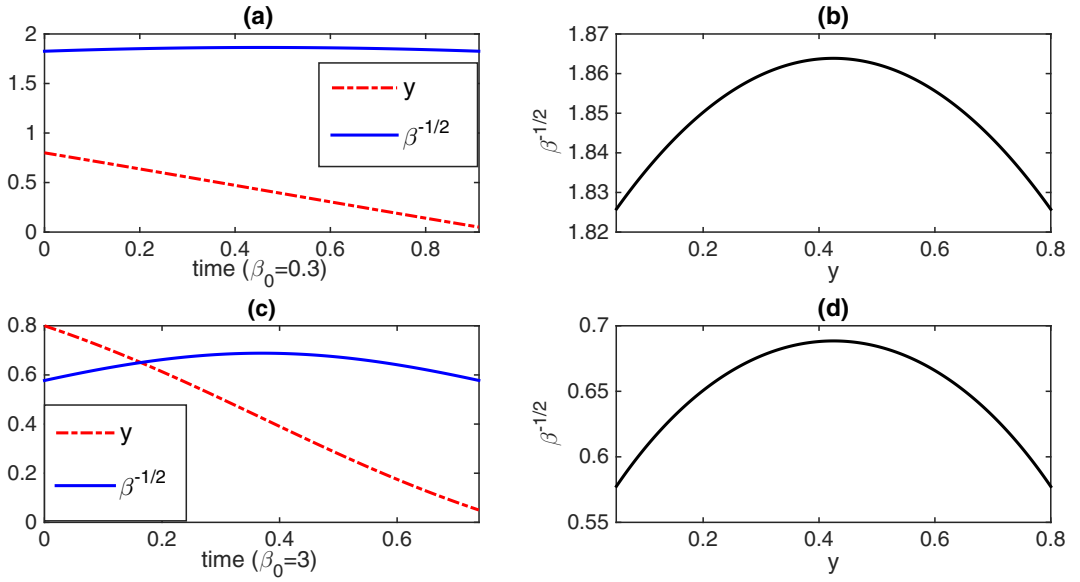

FIG. 3. y and $\beta^{-1/2}$ against time for $\beta_0 = 0.3$ and 3 in (a) and (c), respectively; the corresponding geodesic circular segments in the $(y,\beta^{-1/2})$ upper half plane in (b) and (d), respectively. In both cases, $y_0 = 5/6$ and $y_F = 1/30$.

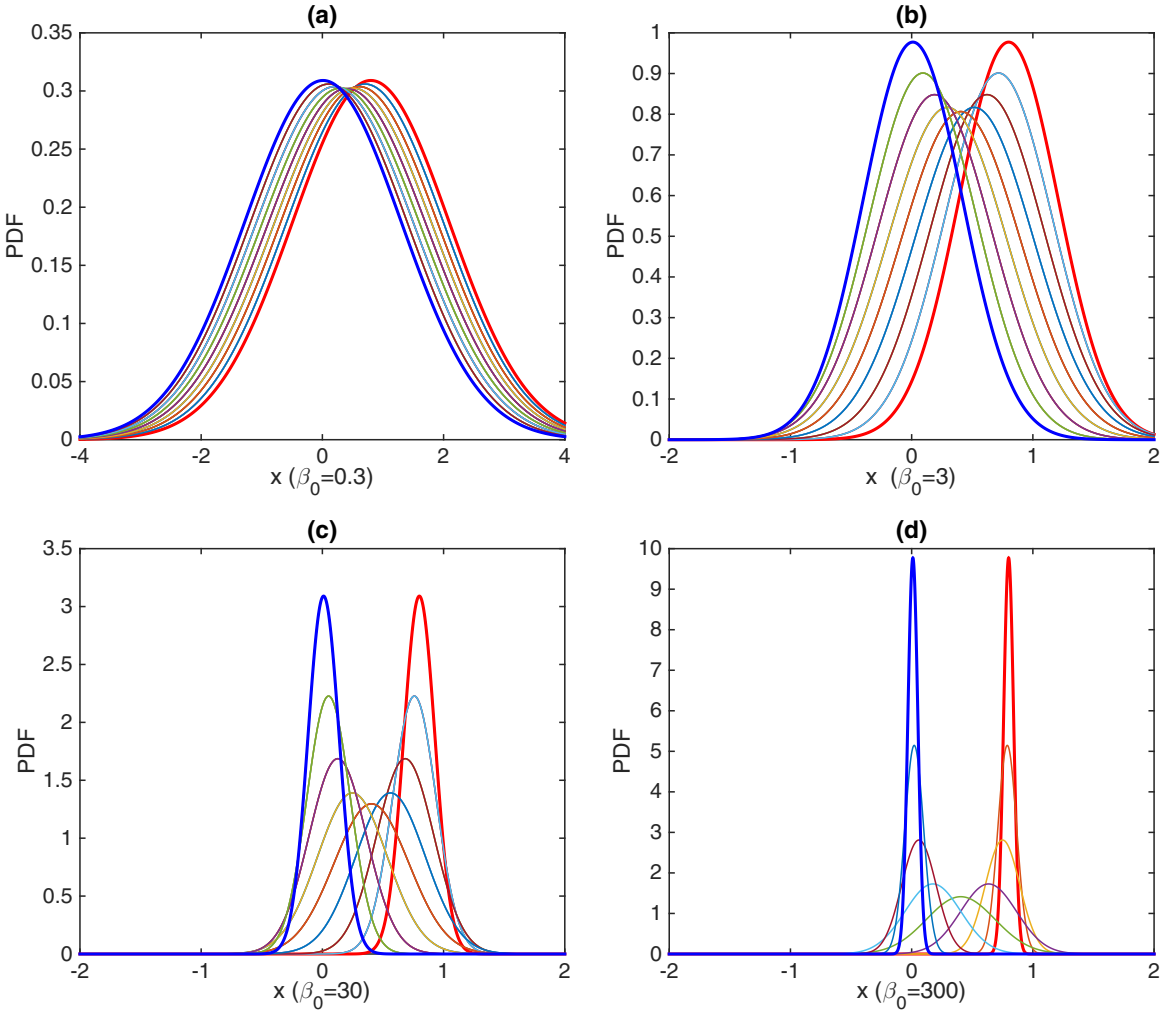

FIG. 4. Time evolution of PDFs against x for $\beta_0 = 0.3,3,30$, and 300 in (a)–(d), respectively. $y_0 = \langle x(t=0)\rangle = 5/6$ and $y_F = \langle x(t = t_F)\rangle = 1/30$. The initial and final PDFs are shown by thick red lines on the right and blue lines on the left, respectively.

1. Geodesic I: Time varying $D(t)$ and $f(t) \neq 0$ with constant friction $\gamma(t) = \gamma_0$

The total time $t_F = 2A$ depends on the value of α (or c), which is in turn determined by the condition on f. Specifically, we require $f(t=0) = 0$ at $t = 0$ in the following.

Since $\dot{y} = -\gamma_0(y - f)$ and $\beta\dot{y} = c$, we recast f as

$$f = y + \frac{1}{\gamma_0}\dot{y} = y + \frac{c}{\gamma_0\beta}. \tag{52}$$

Thus, using $f(t=0) = 0$ in Eq. (52) fixes the value of c as

$$c = -\gamma_0 y_0 \beta_0. \tag{53}$$

By using Eq. (53) in Eq. (40), we obtain the value of α as

$$\sqrt{\alpha} = \frac{2\gamma_0\beta_0 y_0}{\sqrt{\beta_*}}. \tag{54}$$

Thus, from Eqs. (46), (49), (50), and (54), we obtain the total time $t_F = 2A$ for the geodesic

$$2A = \frac{2}{\gamma_0\sqrt{\beta_0}y_0}\frac{Q}{\cosh Q}, \tag{55}$$

where

$$\phi = \sqrt{\beta_0}(y_0 - y_F), \quad Q = \sinh^{-1}\left[\frac{\phi}{2}\right]. \tag{56}$$

To compute the total work done by ξ, we observe that when $\gamma(t) = \gamma_0$ is constant, $G(t) = \gamma_0 t$ and $G(t) - G(t_1) = \gamma_0(t - t_1)$. By using them in Eqs. (12) and (17) and letting $D = D_I(t)$ for geodesic I, we obtain

$$2e^{2\gamma_0 t}D_I(t) = \frac{d}{dt}\left(\frac{e^{2\gamma_0 t}}{2\beta}\right), \tag{57}$$

and thus

$$\begin{aligned} 2D_I(t) &= \frac{\gamma_0}{\beta} + \frac{1}{2}\frac{d}{dt}\left(\frac{1}{\beta}\right) \\ &= \frac{1}{\beta}\left[\gamma_0 - \frac{\sqrt{\alpha}}{2}\tanh\left[\frac{1}{2}\sqrt{\alpha}(t - A)\right]\right]. \end{aligned} \tag{58}$$

There is no contribution from the second term on the right-hand side of Eq. (58) to W_ξ for our prescribed boundary condition $\beta = \beta_0$ at $t = 0$ and t_F; the contribution from the first term is found by using $\frac{1}{\beta} = \frac{1}{c}\frac{dy}{dt}$ [Eq. (33)] in Eq. (58):

$$W_\xi = \int_0^{t_F} dt\, D_I(t) = \int_0^{2A} dt \frac{\gamma_0}{2c}\frac{dy}{dt} = \frac{1}{2\beta_0}\left(1 - \frac{y_F}{y_0}\right). \tag{59}$$

While Eq. (59) is correct, a careful examination of the value of $D_I(t)$ in Eq. (58) reveals an interesting aspect about the information velocity $\sqrt{\alpha/2}$. That is, when α in Eq. (54) is used in Eq. (58), D_I can be shown to be negative for approximately the second half of the time interval when the initial β_0 is sufficiently large, specifically, when $\beta_0 > 3$ for $y_0 = 5/6$ and $y_F = 1/30$ and for the fixed value $\gamma_0 = 1$. The detailed discussion regarding the origin of a negative D_I is provided in Sec. VII. In order to satisfy a physically realistic condition that D_I is non-negative, we need to impose the constraint that the maximum value that $\tanh\theta$ [$\theta = \frac{\sqrt{\alpha}}{2}(t - A)$] can take as

$$(\tanh\theta)_{max} = \frac{\gamma_0}{\sqrt{\alpha}/2} = \frac{\sqrt{\beta_*}}{\beta_0 y_0}. \tag{60}$$

On the other hand, since Eq. (45) implies that the maximum value of $\tanh\theta$ is $\sqrt{\beta_*}(y_M - y_F)$, we have

$$(\tanh\theta)_{max} = \frac{1}{2}\sqrt{\beta_*}\Delta, \tag{61}$$

where we used $y_M = \frac{1}{2}(y_0 + y_F)$ and $\Delta = y_0 - y_F$.

By equating Eqs. (60) and (61), we obtain the maximum value, say Δ_m, of Δ as

$$\Delta_m = \frac{2}{\beta_0 y_0} = \frac{4\sigma_0^2}{y_0}, \tag{62}$$

where $\sigma_0 = 1/\sqrt{2\beta_0}$ is the standard deviation of the initial and final PDF. Δ_m is the largest displacement in y that can be made before bringing the temperature back to the initial value, and is referred to as the length of one cycle. The physical meaning of Δ_m as the maximum variation in y for a geodesic subject to the boundary conditions of the equal temperature $\beta = \beta_0$ at the initial and final time is provided in Sec, VII. When Δ_m is smaller than $\Delta = y_0 - y_F$, we will shortly show how to construct a geodesic solution that satisfies boundary conditions. In a very special case where Δ_m exactly matches Δ—the so-called resonance between two length scales—we obtain an interesting relation

$$\Delta = \frac{2}{\beta_0 y_0} = \frac{4\sigma_0^2}{y_0}. \tag{63}$$

At this resonant point, t_F takes the minimum value [see Fig. 2(a)], as discussed later.

We now present some detailed analysis on how to construct a geodesic solution when $\Delta_m < \Delta$. Leaving the most general analysis for future work, for the purpose of this paper it suffices to consider a simple quantized case where there are an integer number of cycles of length Δ_m in Δ:

$$N = \frac{\Delta}{\Delta_m} = \frac{\Delta\beta_0 y_0}{2}, \tag{64}$$

and divide the path between y_0 and y_F into N small cycles of length Δ_m. For example, the geodesic for $N = 10$ is shown in Figs. 5(c)–5(d) together with the case where $N = 1$ in Figs. 5(a)–5(b) for comparison. Since all the cycles have the same time evolution of β while the mean position of the ith cycle changes as

$$y_M^{(i)} = \left[(N - i) + \frac{1}{2}\right]\Delta_m, \tag{65}$$

where $i = 1,2,3,\ldots,N$, we can write down the geodesic equation for the ith cycle by using Eqs. (44), (45), and Eq. (65):

$$\beta^{(i)}(t) = \beta_* \cosh^2\left[\frac{1}{2}\sqrt{\alpha}(t^{(i)} - A_m)\right], \tag{66}$$

$$y^{(i)}(t) = -\frac{1}{\sqrt{\beta_*}}\tanh\left[\frac{1}{2}\sqrt{\alpha}(t^{(i)} - A_m)\right] + y_M^{(i)}, \tag{67}$$

where $t^{(i)} = [0, 2A_m]$; $2A_m$ in Eqs. (66) and (67) is the time duration of the ith cycle, which can easily be found from

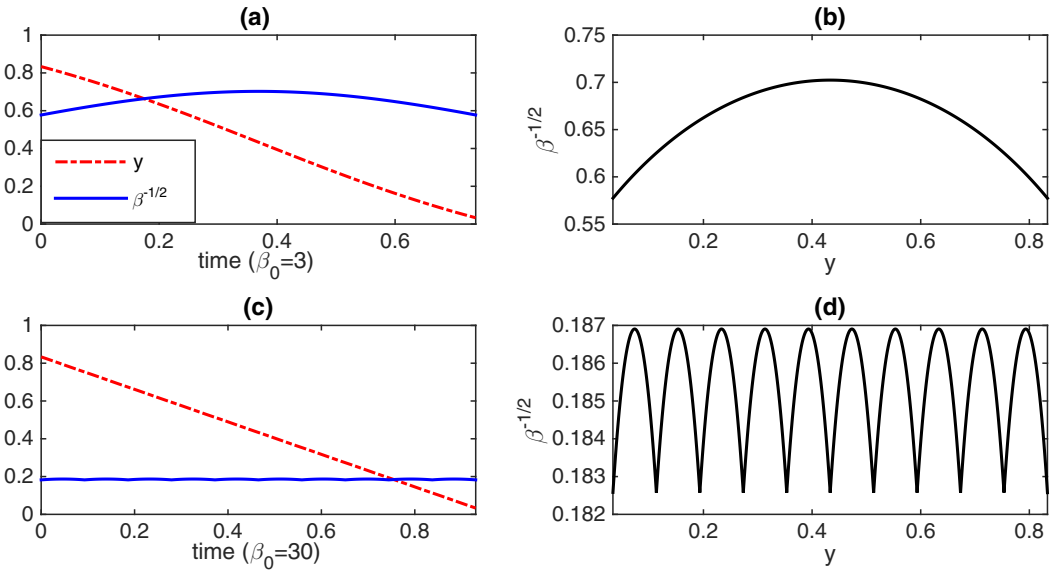

FIG. 5. y and $\beta^{-1/2}$ against time in (a) and (c); The circular geodesics in the upper half plane y and $\beta^{-1/2}$ in (b) and (d) for $\beta_0 = 3$ and 30, respectively. In both cases, $y_0 = 5/6$ and $y_F = 1/30$.

Eq. (46) by replacing Δ by Δ_m and by using Eq. (54) as

$$A_m = \frac{2\sqrt{\beta_*}}{\beta_0 y_0 \gamma_0} \sinh^{-1}\left(\frac{\Delta_m\sqrt{\beta_0}}{2}\right). \tag{68}$$

Note that the time t in $\beta(t)$ and $y(t)$ in Eqs. (66) and (67) is the cumulative time over all the cycles, computed as:

$$t = t^{(i)} + (i-1)(2A_m), \tag{69}$$

for $i = 1,2,3,\ldots,N$.

Figure 5(c) shows the time history of β and y, which undergo ten small-amplitude periodic modulations; this modulation is more visible in Fig. 5(d). We note that the parameter values in this figure were chosen to ensure an integer number of (specifically, ten) cycles.

By using these results, we now compute the total information length and total time by adding the contributions from all i paths ($i = 1,2,3,\ldots,N$) as follows:

$$t_F = 2A_m N = \frac{\sqrt{\beta_*}\Delta}{\gamma_0} \sinh^{-1}\left(\frac{1}{\sqrt{\beta_0}y_0}\right), \tag{70}$$

$$\mathcal{L} = t_F\sqrt{\frac{\alpha}{2}} = \sqrt{2}\Delta\beta_0 y_0 \sinh^{-1}\left(\frac{1}{\sqrt{\beta_0}y_0}\right). \tag{71}$$

To present results, we compute t_F, $\mathcal{L}$, and W_ξ by varying the value of β_0 for $y_0 = 5/6$ and $y_F = 1/30$, noting that for these values, $\Delta = 0.8$ and resonance $\Delta = \Delta_m$ occurs when $\beta_0 = 3$ for geodesic I. Thus, when $\beta_0 < 3$, $N = 1$ and we use results obtained for one cycle where $\Delta = 0.8$ [e.g., Eqs. (44)–(51)]. When $\beta_0 > 3$, we use the integer N number of cycles for geodesic I, by using Eqs. (70)–(71), and Eq. (59), together with Eqs. (64)–(69). Results in Fig. 2 reveal a very interesting utility of geodesic I. First, we observe that geodesic I results in much smaller values not only for $\mathcal{L}$ and W_ξ but also for t_F, compared with the nongeodesic case. Furthermore, a distinct minimum in the total time is observed in geodesic I around $\beta_0 = 3$ due to the aforementioned resonance ($\Delta_m = \Delta$). A corresponding time evolution of the PDFs for this resonant case is shown in Fig. 4(b). This implies that for the given initial y_0 and final y_F mean position, there exists an optimal initial temperature (β_0), which moves the PDF from y_0 to y_F in the least time. These results imply the interesting possibility of utilizing a geodesic to optimize total time in addition to total dissipated energy. Recalling that these results are obtained for a particular realization of a geodesic consisting of a number of cycles with shorter length, further investigation into other realizations would clearly also be worthwhile.

Finally, the time evolution of $D_I(t)$ and $f(t)$ in Eqs. (58) and (52) are shown in Fig. 6(a) for $\beta_0 = 30$, respectively. We see ten cycles ($N = 10$) of periodic modulation in D_I and $f(t)$. The sign of f remains negative, the significance of which will be discussed in a specific problem in Sec. VI.

2. *Geodesic II: Time-varying friction $\gamma = \gamma(t)$ and $f(t) = 0$*

To determine the value of $\gamma(t)$ that is consistent with Eqs. (44) and (45), we utilize Eq. (8) and $\beta\frac{dy}{dt} = c$ in Eq. (33):

$$\gamma = -\frac{1}{y}\frac{dy}{dt} = -\frac{c}{\beta y}. \tag{72}$$

Then, the use of the condition $\gamma(t=0) = \gamma_0$ in Eq. (72) gives us the value of c as

$$c = -\gamma_0 y_0 \beta_0. \tag{73}$$

By using Eq. (73) in Eq. (40), we obtain the value of α as

$$\sqrt{\alpha} = \frac{2\gamma_0\beta_0 y_0}{\sqrt{\beta_*}}, \tag{74}$$

which is the same as Eq. (54). As in the case of geodesic I in Sec. V B1, the diffusion $D(t)$ can also become negative for sufficiently large β_0. For the purpose of formulating a theoretical framework in this paper, in the following, we limit our study to the one-cycle case for small β_0 where D_{II} is positive between $t = 0$ and t_F. In this case, from Eqs. (46), (49), (50), and (74), we obtain the total time $t_F = 2A$

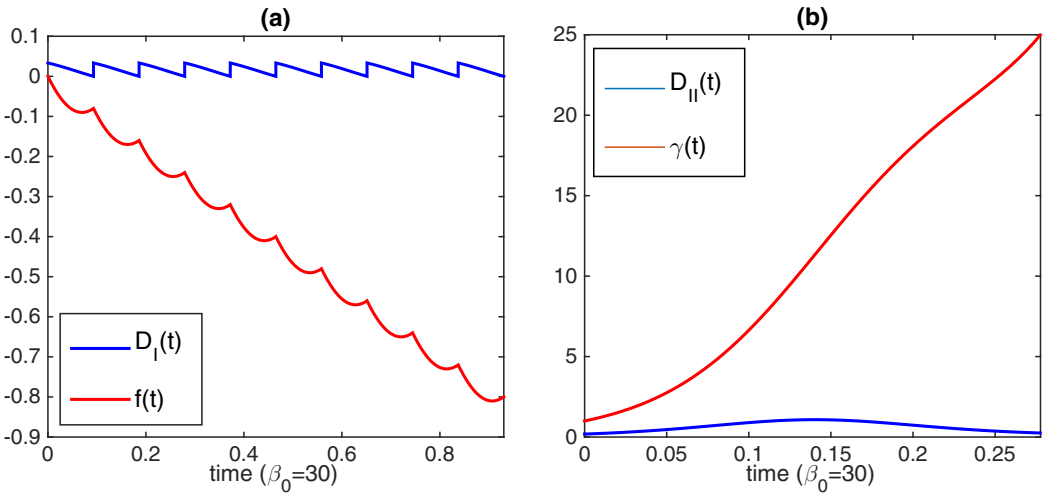

FIG. 6. (a) The blue upper curve shows $D_I(t)$, and the red lower curve shows $f(t)$. (b) The blue lower curve shows $D_{II}(t)$, and the red upper curve shows $\gamma(t)$. For both panels $\beta_0 = 30$, $y(t=0) = y_0 = 0.08$, and $y(t=t_F) = y_F = 0.05$.

for the geodesic

$$2A = \frac{2}{\gamma_0\sqrt{\beta_0}y_0}\frac{Q}{\cosh Q}, \tag{75}$$

where Q is defined as

$$\phi = \sqrt{\beta_0}(y_0 - y_F), \quad Q = \sinh^{-1}\left[\frac{\phi}{2}\right]. \tag{76}$$

The information length for the geodesic motion then simply follows from Eqs. (36), (74), and (55) as:

$$\mathcal{L} = \int_0^{2A} dt\, v = \sqrt{2\alpha}A. \tag{77}$$

The computation of the total dissipated energy W_ξ requires lengthier algebra. We refer D as D_{II} for geodesic II and obtain from Eqs. (12) and (17) the following:

$$2e^{2G(t)}D_{II}(t) = \frac{d}{dt}\left(\frac{e^{2G(t)}}{2\beta}\right), \tag{78}$$

which essentially leads to the same equation (58) as,

$$\begin{aligned} 2D_{II}(t) &= \frac{\gamma}{\beta} + \frac{1}{2}\frac{d}{dt}\left(\frac{1}{\beta}\right) \\ &= \frac{1}{\beta}\left[\gamma - \frac{\sqrt{\alpha}}{2}\tanh\left[\frac{1}{2}\sqrt{\alpha}(t-A)\right]\right]. \end{aligned} \tag{79}$$

We again note that there is no contribution from the second term on the right-hand side of Eq. (79) to W_ξ for the same $\beta = \beta_0$ at $t = 0$ and t_F, and the contribution from the first term depends on $\gamma(t)$. By using Eq. (72) in Eq. (79), we can rewrite W_ξ as:

$$W_\xi = \int_0^{t_F} dt\, D_{II}(t) = -\int_0^{t_F} dt\frac{c}{\beta^2 y}. \tag{80}$$

By using Eqs. (44), (45), and (73) in Eq. (80), and after further lengthy algebra (see Appendix E), we obtain

$$W_\xi = \frac{1}{2}\left[\frac{1}{\beta_0} - y_0 y_F\right]\ln\frac{y_0}{y_F} + \frac{1}{4}\left(y_0^2 - y_F^2\right). \tag{81}$$

Time evolution of $D_{II}(t)$ and $\gamma(t)$ for geodesic II is shown in Fig. 6(b) for $\beta_0 = 30$, by using the same values of y_0 and y_F as previously. We observe that the increase in the frictional energy loss by larger γ is now responsible for reducing the temperature back to the smaller value (larger β) during the second half of the time evolution. Figure 2 shows $\mathcal{L}$, t_F, and W_ξ against β_0, where they are seen to take small values compared to the nongeodesic case. Information length for geodesic I and II is observed to be small for sufficiently small β_0. Note that in this figure, geodesic II (in dash-dotted red) is shown for sufficiently small value of β_0 where the diffusion (D_{II}) is non-negative. These results again point to the interesting possibility of its application in optimization.

VI. PHYSICAL REALIZABILITY OF A GEODESIC SOLUTION

In Sec. V B1, we noted that the diffusion coefficient can become negative as the initial inverse temperature β_0 becomes too large for fixed parameter values γ_0, y_0, and y_F. In this section, we expand on its physical meaning and realizability of a geodesic solution.

We begin by looking at the meaning of $\Delta_m = 2/\beta_0 y_0 = 4\sigma_0^2/y_0$ in Eq. (62) in relation to the information transfer by expressing the information velocity in Eq. (36) in terms of Δ_m as

$$v = \sqrt{\frac{\alpha}{2}} = \frac{\gamma_0 y_0 \sigma_*}{\sigma_0^2} = \frac{\gamma_0(2\sigma_*)}{\Delta_m/2}. \tag{82}$$

Here, $\sigma_0 = \sqrt{1/2\beta_0}$ is the standard deviation of the initial and final PDF and $\sigma_* = \sqrt{1/2\beta_*}$ is the largest standard deviation (width) of the PDF when $t = A_m$, where the inverse temperature takes the smallest value β_*. Recall that the radius of the geodesic circle is $1/\sqrt{\beta_*} = \sqrt{2}\sigma_*$. Equation (82) illustrates that the rate of information propagation across $\Delta_m/2$ is balanced by the frictional dissipation rate across the width of the PDF ($2\sigma_*$) at the middle point. Alternatively, $\Delta_m/2$ is the largest distance over which the information can be transferred physically for a given γ_0.

In order to highlight the effect of β_0 on cyclic solutions, it is useful to find an approximate expression for σ_* from Eq. (51) evaluated for the first cycle at $\beta = \beta_0$, $y = y_0$, and $\Delta_m/2 = y_0 - y_M$:

$$\left(\frac{\Delta_m}{2}\right)^2 + 2\sigma_0^2 = 2\sigma_*^2, \tag{83}$$

which gives

$$\sigma_* = \sigma_0\left[1 + 2\frac{\sigma_0^2}{y_0^2}\right]^{\frac{1}{2}}, \tag{84}$$

where Eq. (62) was used. We define the change in σ_0 and y_0 for the first half cyclic motion as

$$D\sigma_0 \equiv \sigma_* - \sigma_0, \quad Dy_0 = -\frac{\Delta_m}{2} = -\frac{2\sigma_0^2}{y_0}, \tag{85}$$

and examine how they are related to each other in the two cases depending on the relative ratio of the width σ_0 of the initial PDF to y_0.

In the first case where $\sigma_0 > y_0/\sqrt{2}$, Eq. (84) is approximated as

$$\sigma_* \sim \sqrt{2}\frac{\sigma_0^2}{y_0} \gg \sigma_0, \tag{86}$$

leading to $D\sigma_0 \sim \sigma_*$. Thus, Eqs. (85)–(86) give us

$$\frac{D\sigma_0}{Dy_0} \sim -\frac{1}{\sqrt{2}}. \tag{87}$$

Note that this limit supports a geodesic solution with $N = 1$. In the opposite limit of $\sigma_0 < y_0/\sqrt{2}$, Eq. (84) is approximated as

$$\sigma_* \sim \sigma_0 + \frac{\sigma_0^3}{y_0^2}, \tag{88}$$

which leads to $D\sigma_0 \sim \frac{\sigma_0^3}{y_0^2}$, and thus

$$\frac{D\sigma_0}{Dy_0} \sim -\frac{\sigma_0}{2y_0} = \frac{Dy_0}{4\sigma_0}, \tag{89}$$

where Eq. (85) is used (e.g., to eliminate y_0 in place of Dy_0). It is intriguing that Eqs. (87) and (89) suggest very different scaling relations between $D\sigma_0$ and Dy_0 depending on whether the initial PDF has a width much narrower or wider than y_0. Specifically, Eqs. (87) gives a simple linear relation as

$$\frac{D\sigma_0}{\sigma_0} \sim -\frac{1}{\sqrt{2}}\frac{Dy_0}{\sigma_0}, \tag{90}$$

where the normalization of $D\sigma_0$ and Dy_0 was made by the resolution σ_0. In comparison, Eq. (89) gives an interesting power-law relation, which can be expressed as follows:

$$\left|\frac{Dy_0}{2\sigma_0}\right| \sim \left(\frac{D\sigma_0}{\sigma_0}\right)^{1/2}, \tag{91}$$

where the normalization by the resolution σ_0 was again made. In comparison with Eq. (90), Eq. (91) implies a much smaller change in σ_0 than y_0 as a power law. This is suggestive of a fractal structure for small σ_0 (near $1/\beta = 0$ axis, which is the lower boundary of the Poincaré half plane).

To examine the physical realizability of a geodesic solution, we rewrite Eqs. (91) and (90) in terms of $Dy_0 = -\Delta_m/2$ as

$$|D\sigma_0| \propto \begin{cases} \Delta_m & \text{if } \sigma_0 \gg y_0 \\ \Delta_m\left(\frac{\Delta_m}{\sigma_0}\right) & \text{if } \sigma_0 \ll y_0 \end{cases}, \tag{92}$$

where Δ_m/σ_0 is factored out to highlight that for small σ_0, $|D\sigma_0|$ is larger than Δ_m by this factor $\Delta_m/\sigma_0 \gg 1$. If Δ_m were to be the whole interval $\Delta = y_0 - y_F$, $|D\sigma_0| \propto \Delta^2/\sigma_0$, which becomes very large for small σ_0. Although a geodesic solution is permitted for any value of $|D\sigma_0|$, too large $|D\sigma_0|$ can be problematic in its physical realization in a particular model. To see this, we recall that from Sec. III, the change in the PDF width is due to the competition between W_ξ and W_γ. According to Eq. (92), when the total distance ($\Delta = y_0 - y_F$) that the PDF needs to move is too large compared to the narrow width of the initial PDF, the required change in σ becomes large; the PDF needs to become much wider than the initial one along the geodesic (e.g., at the midpoint) and then become narrow to recover β_0. In order for the PDF to become narrower, the fluctuating energy (which is large for a broad PDF at the middle point) needs to be removed by W_γ via frictional damping γ_0, which transfers the energy to the environment. For a fixed γ_0 and y_0 and y_F, there is a critical value of β_0, above which the frictional damping is insufficient to accomplish this task, causing a negative diffusion D. Alternatively, for the given initial β_0 and y_0, there is upper bound Δ_m on Δ for a physically realizable geodesic solution.

VII. APPLICATION TO POPULATION GROWTH

A logistic-type equation is a popular model for population growth, which has been widely used to understand nonlinear equilibration in many different systems. The merit of this model is the simplicity in incorporating two conflicting effects of the positive feedback (promoting the growth) and of the negative feedback (inhibiting the growth) via nonlinear damping. In a stochastic internal environment, the logistic model can be written in the following form [32]:

$$\frac{du}{dt} = \gamma u - (\epsilon - \xi)u^2 - g(t)u^2. \tag{93}$$

Here $u \geqslant 0$ is a non-negative random variable for the population; the terms involving $\gamma > 0$ and $\epsilon > 0$ represent linear positive and nonlinear negative feedbacks, responsible for the linear growth and the nonlinear saturation through competition, respectively. ξ is the stochastic random part of the negative feedback, which accounts for a stochastic component of competition. For simplicity, ξ is assumed to be a short-memory noise given by Eq. (3). The nonlinear term $-g(t)u^2$ represents the reduction of the population by a prescribed deterministic force, which preferentially decreases larger populations, specifically, as a quadratic power of u. Note that $-\epsilon u^2$ represents the internal damping (negative feedback) mechanism while $-g(t)u^2$ is damping by an external force, which can be controlled for a geodesic solution. Since time and u can always be normalized by γ and ϵ, respectively, we fix the value of $\gamma = 1$ and $\epsilon = 1$ while varying other parameters for our study in this paper.

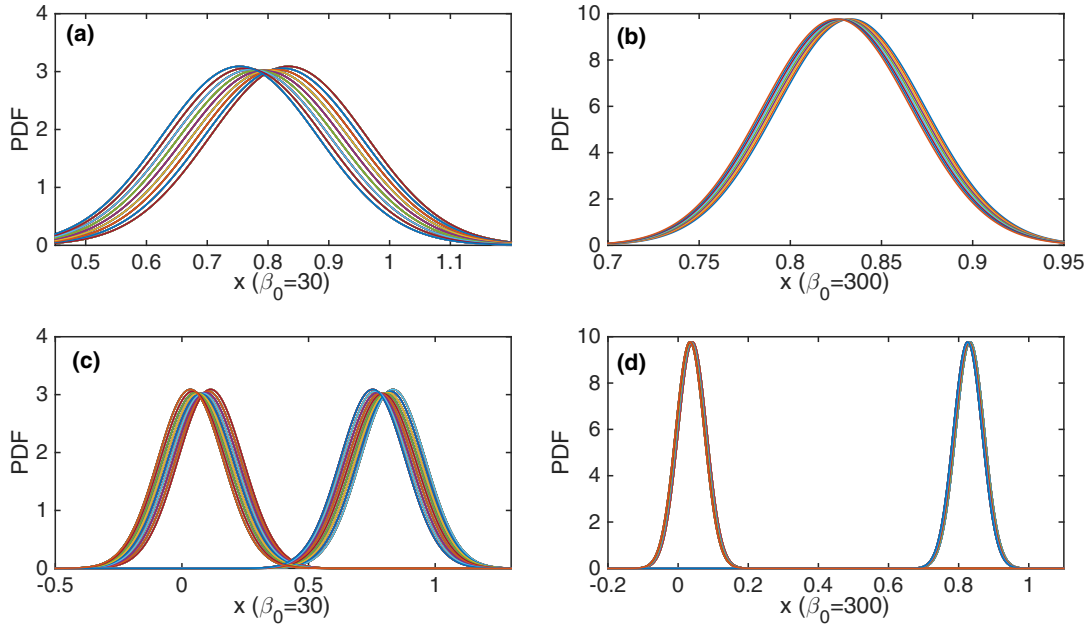

FIG. 7. Time evolution of PDFs of x over the first cycle for $\beta_0 = 30$ and 300 in (a) and (b), respectively. (c)–(d) are the evolution of PDFs over the first and the last cycles [tenth and hundredth cycle for (c) and (d), respectively] shown at the same time. $y_0 = 5/6$ and $y_F = 1/30$.

We envision the situation where we can control the time dependence of the prescribed forcing $g(t)$ and the strength of the stochastic noise ξ between the initial and the final states and are interested in finding a best treatment protocol, which reduces the population size in the least time. This could potentially be very beneficial when a fast reduction of the population (e.g., treatment of disease) is desired. This optimal protocol is provided by a geodesic found in Sec. V.

In order to utilize the results obtained in previous sections, we transform the nonlinear equation (93) into the form of Eq. (8) by using the change of variable $x = -1/u + \epsilon/\gamma$ as

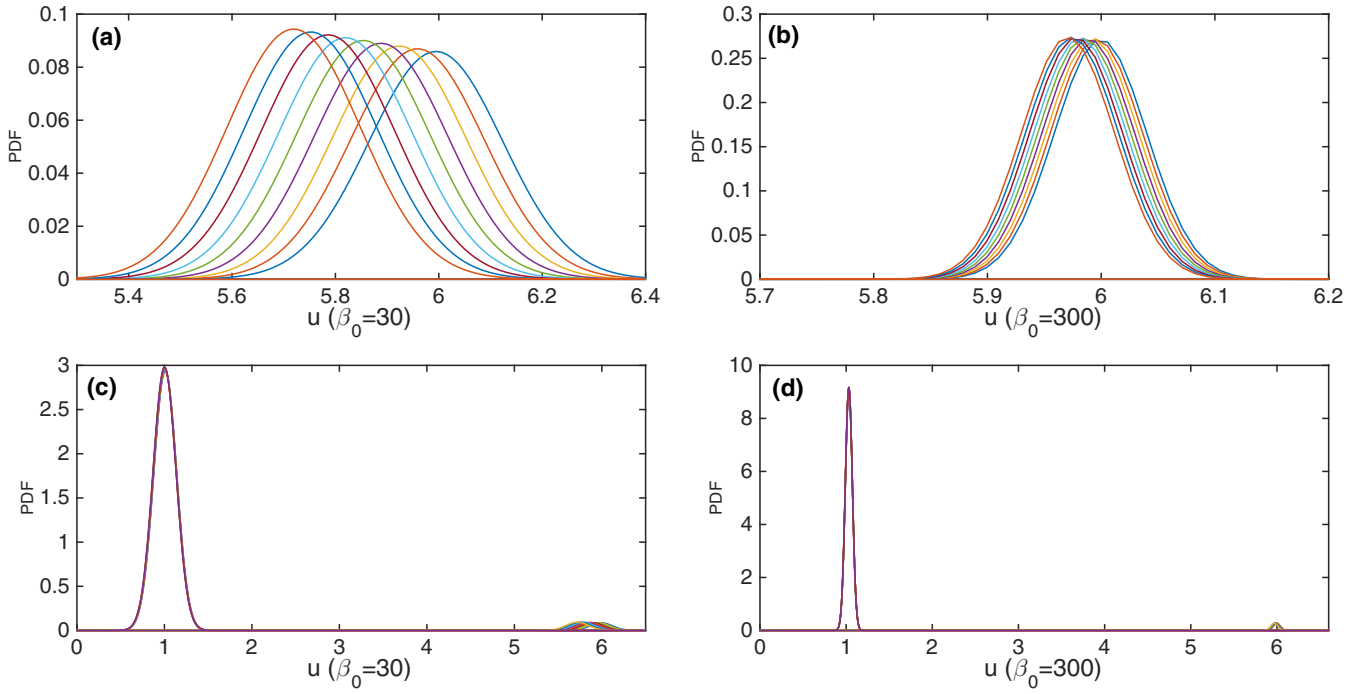

FIG. 8. Time evolution of PDFs of the population u for $\beta_0 = 30$ and 300, corresponding to Fig. 7. $\langle u(t=0)\rangle = 6xs\langle u(t=t_F)\rangle = 1.0345$, which is close to the carrying capacity $u_\infty = 1$ ($\gamma = 1$, $\epsilon = 1$). (a)–(b) show the PDF during the first cycle as in Figs. 7(a)–7(b), where the PDF at $t = 0$ can be identified with the peak at $u = 6$. (c)–(d) are the evolution of PDFs over the first and the last cycles [tenth and hundredth cycle for (c) and (d), respectively] shown at the same time.

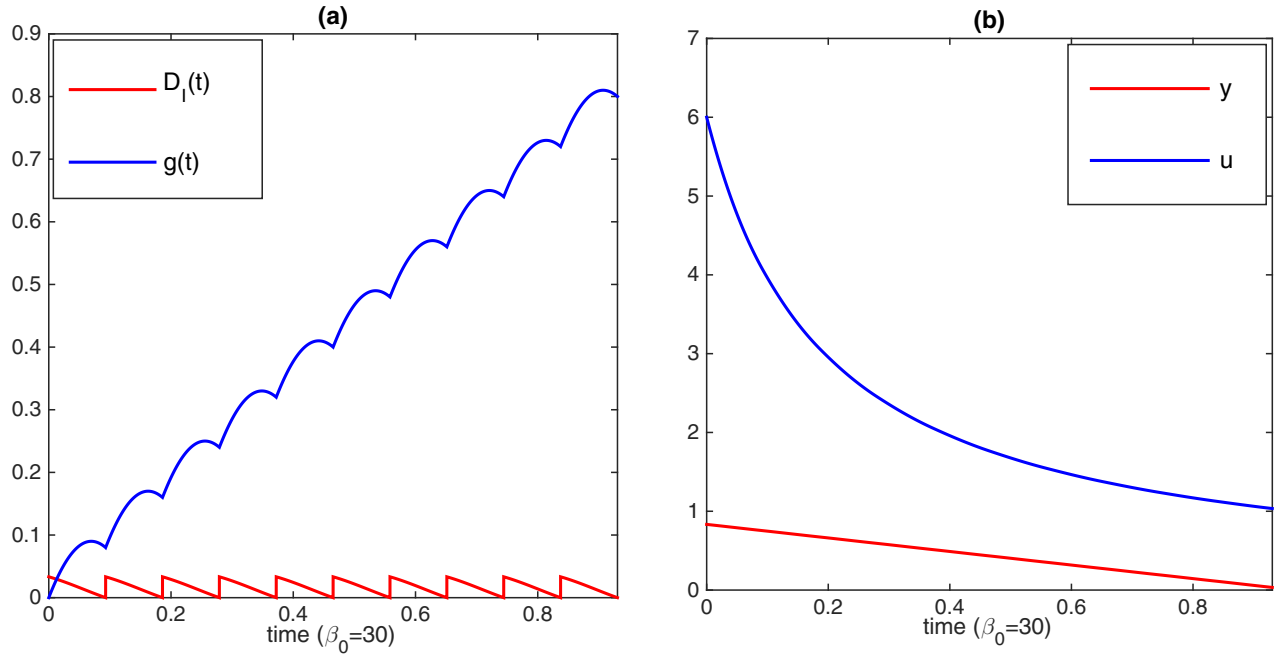

FIG. 9. (a) The red lower curve shows $D_I(t)$, and the blue upper curve shows $g(t)$. (b) The red lower curve shows y (the mean value of x), and the blue upper curve shows u. For both panels $\beta_0 = 30$, the carrying capacity $\gamma/\epsilon = u_\infty = 1$, $y(t=0) = y_0 = \langle x(t=0)\rangle = 5/6$, and $y(t=t_F) = y_F = \langle x(t=t_F)\rangle = 5/6$.

follows:

$$\frac{dx}{dt} = -\gamma\left(x + \frac{g}{\gamma}\right) + \xi. \tag{94}$$

By comparing with Eqs. (8) and (94), we identify that f is replaced by $-g(t)/\gamma$. Therefore, very conveniently, if we assume that $x = -1/u + \epsilon/\gamma$ has the initial Gaussian distribution with inverse temperature β_0 and that the mean position $y = \langle x\rangle$ takes the value of y_0 and y_F at the initial and final times $t = 0$ and t_F, respectively, the inverse temperature $\beta(t)$ and the mean position $y(t)$ satisfy the same equations as in Eqs. (64)–(68). We recall that the mean value denoted by the angular brackets is obtained by the average over both ξ and initial position x_0. Furthermore, by considering the situation where the objective is to reduce the average population by keeping the same inverse temperature $\beta = \beta_0$ at the initial and final times and the constant growth rate $\gamma = \gamma_0$, we can find the best treatment protocol, which minimizes the time and the associated dissipated energy by using a geodesic solution (geodesic I) for N cycle. We consider $\beta_0 = 30$ and $\beta_0 = 300$ for the fixed values of $y_0 = 5/6$ and $y = 1/30$, as in previous sections. Therefore, $N = 10$ and 100 for $\beta_0 = 30$ and 300, respectively. (Modest values of β_0 are used for this model to ensure a negligible escape rate of the population to $+\infty$.) The optimal treatment schedule $g(t)$ is obtained from the geodesic solution via Eq. (52):

$$g(t) = -\gamma_0 f(t) = -\gamma_0\left(y + \frac{c}{\gamma_0\beta}\right). \tag{95}$$

Results are shown in Figs. 7–9 for the case where the carrying capacity $u_\infty = 1$ and $y_0 = \langle x(t=0)\rangle = 5/6$ and $y_F = \langle x(t=t_F)\rangle = 1/30$, the same values used in all other figures. Figure 7 shows time-dependent PDFs of $x = 1 - 1/u$ for $\beta_0 = 30$ and 300 where a geodesic consists of 10 ($N = 10$) and 100 ($N = 100$) cycles, respectively. Specifically, the time evolution of PDFs of x over the first cycle for $\beta_0 = 30$ and 300 are shown in Figs. 7(a) and 7(b), respectively, while the evolution of PDFs over the first and the last cycles (tenth and hundredth cycle for $\beta_0 = 30$ and 300) are shown in Figs. 7(c) and 7(d), respectively. The corresponding time evolution of PDFs of the population u is shown in Fig. 8. We note that the PDF of u at $t = 0$ has a peak at $\langle u(t=0)\rangle = 6$, corresponding to $y_0 = 5/6$; $\langle u(t=t_F)\rangle = 1.0345$, which is close to the carrying capacity $u_\infty = 1$ ($\gamma = 1, \epsilon = 1$). Finally, Fig. 9 shows $D_I(t)$, $g(t)$, mean values of x and u against time for $\beta_0 = 30$, where the sign of $g(t)$ is seen to be positive. Interestingly, the observation that the total change in $g(t)$ is comparable to $\Delta = y_0 - y_F$ in Fig. 9 reveals how the geodesic solution is established by slowly moving the PDF peak by the deterministic force.

VIII. CONCLUSION

Far from equilibrium, the level of fluctuations in a system changes with time and becomes a dynamical variable itself, and the importance of a full knowledge of the evolution of PDFs cannot be overemphasized. As the computation of time-dependent PDFs is highly demanding and expensive numerically, we utilized one analytically solvable model of a driven dissipative system (a generalized nonautonomous Ornstein-Uhlenbeck process) by including the time-dependent deterministic forcing and time-dependent strength of the stochastic noise (diffusion). By generalizing our familiar concept of distance by using a dynamical ruler whose

resolution is set by time-dependent fluctuations, we mapped the time evolution of our system onto the trajectory in the statistical metric space given by the Poincáre upper half plane consisting of the mean position and the standard deviation $1/\sqrt{\beta}$. We computed the information velocity and length and found geodesic solutions for which the information propagates at constant speed to be either in the form of a line of a constant mean position or a circle. We then demonstrated how to construct a particular realization of a geodesic, which satisfies boundary conditions at the initial and final times in the two specific cases of geodesic I and II and showed that in both cases, our realization of a geodesic provided a path that ensures not only small information length and dissipated energy but also smaller total time along the path in comparison with a nongeodesic path. A resonance phenomenon in the geodesic solution due to the matching of $\Delta = \Delta_m$ was reported. Application of our results to a stochastic logistic model demonstrated a significant improvement in controlling population growth by a periodic modulation of diffusion coefficient and deterministic force by a small amount. This optimization can have a significant implication for damage control where the prolongation of a system near an initial condition (e.g., a large population of harmful bacteria, a strong tornado, etc.) is harmful.

Although we utilized an exact solution in this paper, our methodology is general and does not rely on the existence of exact PDFs, nor even the existence of basic equations, which govern the evolution of systems. This is because the information velocity and length can be computed directly from Eqs. (4) and (5) by constructing time-dependent PDFs from experimental, observational, and numerical data. For instance, Ref. [23] numerically computed PDFs by simulating a logistic map and the information velocity and length for the purpose of investigating the attractor structure (e.g., stable and unstable points); Ref. [24] studied the information velocity and length in classical music by computing time-dependent PDFs from the music MIDI files, elucidating different classical music in terms of the information flow and the role of geodesics in classical music. Application of our methodology to other data (e.g., heart rhythm) is under progress. A geodesic solution can also be implemented numerically, for instance, as has been done in Ref. [10]. In addition to this very applicability to a variety of systems whose evolution is far too complex to be modeled by a system of equations, our methodology will provide a unifying framework for understanding seemingly different phenomena by using system-independent variables (information length and geodesics).

In summary, this paper provides a key theoretical framework for understanding nonequilibrium processes in terms of information change and a scope for investigation and application of a geodesic to different nonequilibrium systems, particularly for the purpose of optimization. Given our discovery of a resonance phenomenon, further investigation into different realizations of a geodesic solution would be of great interest. Detailed analysis on different realizations and different situations with appropriate boundary conditions is a subject for future work. Further application and extension of this work to different systems is also a subject for future research.

APPENDIX A: FLUCTUATING HAMILTONIAN $\mathcal{E}$

To appreciate the relation between information velocity or energy $\mathcal{E}$ and fluctuating energy, we express the PDF $p(x,t)$ as

$$p(x,t) = \sqrt{\frac{\beta}{\pi}} e^{-S_A} \equiv e^{-S_A+\mathcal{F}}. \tag{A1}$$

Here, $\mathcal{F} = \frac{1}{2}\ln\frac{\beta}{\pi}$ is the free energy; S_A is the effective action, which can be related to the Hamiltonian H of the stochastic system (see Ref. [42]) as

$$H = -\frac{\partial S_A}{\partial t}, \tag{A2}$$

which is a stochastic analogy to the Hamilton-Jacobi relation [42,43]. Specifically, it was shown in Ref. [42] by a path integral formulation that H is given in terms of

$$H(t) = -\frac{\partial S_A}{\partial t} = \frac{D}{2}\Pi^2 - \mu\Pi x,$$

where Π is the conjugate momentum. Note that Π stems from the stochastic noise. Taking the time derivative of Eq. (A1) gives us

$$\frac{\partial p(x,t)}{\partial t} = (\dot{\mathcal{F}} + H)p(x,t), \tag{A3}$$

where $\dot{\mathcal{F}} = \frac{d\mathcal{F}}{dt}$. First, we integrate both sides of Eq. (A3) over x and use the conservation of the total probability as follows:

$$0 = \int dx \frac{\partial p}{\partial t} = \int dx(\dot{\mathcal{F}} + H)p(x,t) = \dot{\mathcal{F}} + \langle H\rangle, \tag{A4}$$

where $\langle H\rangle$ is the mean (average) value of the Hamiltonian. Therefore,

$$\dot{\mathcal{F}} = -\langle H\rangle. \tag{A5}$$

That is, the mean value of the Hamiltonian compensates for the change in free energy to conserve the total probability. We now compute the second moment, which is related to $\mathcal{E}$ in Eq. (6) as

$$\begin{aligned}\mathcal{E} &= \int dx \frac{1}{p}\left(\frac{\partial p}{\partial t}\right)^2 = \int dx(H+\dot{\mathcal{F}})^2 p(x,t)\\ &= \langle (H+\dot{\mathcal{F}})^2\rangle = \langle (\delta H)^2\rangle,\end{aligned} \tag{A6}$$

where $\delta H = H - \langle H\rangle = H + \dot{\mathcal{F}}$ is the fluctuating Hamiltonian. By using Eq. (A5), it is interesting to observe that

$$\langle (\delta H)^2\rangle = \langle H^2\rangle + 2\langle H\rangle\dot{\mathcal{F}} + \dot{\mathcal{F}}^2 = \langle H^2\rangle - \langle H\rangle^2.$$

APPENDIX B: DERIVATION OF EQS. (34), (37), AND (38)

By using Eq. (33) in (31), we obtain

$$\ddot{\beta} - \frac{1}{\beta}\dot{\beta}^2 = 2c^2, \tag{B1}$$

which can be written as

$$\beta^2 \frac{\partial}{\partial\beta}\left[\frac{\dot{\beta}^2}{\beta^2}\right] = 4c^2. \tag{B2}$$

Dividing Eq. (B2) by β^2 and integrating over β gives us

$$\frac{\dot{\beta}^2}{\beta^2} = -\frac{4c^2}{\beta} + \alpha, \tag{B3}$$

where α is an integration constant. The rearrangement of Eq. (B2) gives Eq. (34) in the text.

From Eq. (34), we obtain

$$\frac{d\beta}{dt} = \sqrt{\alpha}\sqrt{\beta^2 - \frac{4c^2}{\alpha}\beta}, \tag{B4}$$

which can be integrated as

$$\sqrt{\alpha}\int dt = \int \frac{d\beta}{\sqrt{\beta^2 - \frac{4c^2}{\alpha}\beta}}, \tag{B5}$$

to obtain

$$\sqrt{\alpha}t = A + \cosh^{-1}\left(\frac{\alpha\beta}{2c^2} - 1\right), \tag{B6}$$

where A is constant. Solving Eq. (B6) for β gives us Eq. (37) in the text.

To find the solution for y, we solve Eq. (33) for y by using Eq. (33) as follows:

$$\begin{aligned}\frac{1}{c}y(t) &= \int dt\frac{1}{\beta} = \frac{\alpha}{2c^2}\int_0^t \frac{dt}{\cosh\theta + 1} \\ &= \frac{\alpha}{c^2}\int \frac{2e^\theta dt}{(e^\theta+1)^2} \\ &= -\frac{\sqrt{\alpha}}{c^2}\frac{1}{e^\theta+1} + B,\end{aligned} \tag{B7}$$

where $\theta = \sqrt{\alpha}t - A$ and B is constant.

APPENDIX C: CHRISTOFFEL AND RICCI-CURVATURE TENSORS

This Appendix is included for completeness of our paper. From Eq. (27), the metric tensor g_{ij} and its inverse g^{ij} can be found as:

$$g_{ij} = \begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix}, \quad g^{ij} = \begin{pmatrix} 2\beta^2 & 0 \\ 0 & \frac{2}{2\beta} \end{pmatrix}. \tag{C1}$$

The connection tensor $\Gamma_{ijk} = \frac{1}{2}[\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}]$ ($\Gamma^i_{jk} = g^{im}\Gamma_{jkm}$) can be found to have the following nonzero components

$$\Gamma^1_{11} = -\frac{1}{\beta}, \Gamma^1_{22} = -2\beta^2, \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{2\beta}. \tag{C2}$$

The Riemann curvature tensor $R^i_{kmn} = \partial_m\Gamma^i_{nk} + \Gamma^i_{mp}\Gamma^p_{nk} - \partial_n\Gamma^i_{mk} - \Gamma^i_{np}\Gamma^p_{mk}$ can then be shown to have the following nonvanishing components:

$$R^1_{212} = -R^1_{221} = -\beta, \;\; R^2_{112} = -R^2_{121} = \frac{1}{4\beta^2}. \tag{C3}$$

As the curvature tensors do not vanish for certain components, the metric space is not flat but curved. The Ricci tensor is then computed by contracting the curvature tensor as $R_{ij} = R^k_{ikj}$:

$$R_{11} = -\frac{1}{4\beta^2}, \quad R_{22} = -\beta, \quad R_{12} = R_{21} = 0, \tag{C4}$$

leading to the Ricci scalar

$$R = g^{ij}R_{ij} = -1. \tag{C5}$$

We now make an analogy to the Einstein field equation where $G_{ij} = 8\pi T_{ij}$ (where T_{ij} is the stress-energy tensor). By using $R = -1$ for G_{ij}

$$\begin{aligned}G_{ij} &= R_{ij} - \frac{1}{2}R\,g_{ij} \\ &= \begin{pmatrix} -\frac{1}{4\beta^2} & 0 \\ 0 & -\beta \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix} = 0.\end{aligned} \tag{C6}$$

Therefore, the stress-energy tensor $T_{ij} = 0$.

APPENDIX D: GEODESIC EQUATION

It is worth noting that the Euler-Lagrange equations (29)–(30) can also be derived from the following geodesic motion for $\mathcal{E}$ in Eq. (27) by using the Christoffel tensors in Eq. (C3):

$$\frac{d^2\lambda^i}{dt^2} + \Gamma^i_{mk}\frac{d\lambda^m}{dt}\frac{d\lambda^k}{dt}, \tag{D1}$$

where $\lambda^i = (\beta, y)$. Specifically, Eq. (D1) becomes

$$0 = \ddot{\beta} + \Gamma^1_{11}\dot{\beta}^2 + \Gamma^1_{22}\dot{y}^2, \tag{D2}$$

$$0 = \ddot{y} + \Gamma^2_{12}\dot{\beta}\dot{y} + \Gamma^2_{21}\dot{\beta}\dot{y}. \tag{D3}$$

Using Eq. (C2) in Eqs. (D2)–(D3) gives us Eqs. (31)–(33).

APPENDIX E: DERIVATION OF EQ. (81)

By using Eqs. (44), (45), and (73) in Eq. (80) and by letting $\theta = \frac{1}{2}\sqrt{\alpha}(t - A)$, we can derive

$$\begin{aligned}2\beta_* W_\xi &= \int d\theta \frac{1}{\cosh^4\theta(b - \tanh\theta)} \\ &= -\int d\theta[\ln(b - \tanh\theta)]\mathrm{sech}^2\theta \\ &= \left[-\frac{\ln(b - \tanh\theta)}{\cosh^2\theta}\right]_{\theta_0}^{\theta_F} - 2J \\ &\equiv I(t = 2A) - I(t = A).\end{aligned} \tag{E1}$$

Here, $b = \sqrt{\beta_*}y_M$ and $y_M = \frac{1}{2}(y_0 + y_F)$; θ_0 and θ_F are the values of θ at $t = 0$ and $t = 2A$, respectively; $I(t = 0)$ and $I(t = 2A)$ are the value of integral evaluated at $t = 0$ and $t = 2A$, respectively. J is defined as follows:

$$\begin{aligned}J &= \int d\theta \ln(b - \tanh\theta)\tanh\theta\,\mathrm{sech}^2\theta \\ &= -\int dw(\ln w)(b - w) \\ &= \left[-b[w\ln w - w] + \frac{1}{2}w^2\ln w - \frac{1}{4}w^2\right]_{w_0}^{w_F} \\ &= \left[-\frac{1}{2}(b^2 - \tanh^2\theta)\ln(b - \tanh\theta)\right. \\ &\quad \left. + \frac{1}{4}(b - \tanh\theta)(3b + \tanh\theta)\right]_{\theta_0}^{\theta_F},\end{aligned} \tag{E2}$$

where $w = (b - \tanh\theta)$ was used; w_0 and w_F are evaluated at $t = 0$ and $t = 2A$, respectively. In order to compute W_ξ in

Eq. (E1), we need to evaluate the various terms at $t = 0$ and $t = 2A$. For $t = 0$, we can show that

$$b - \tanh\theta = y_0\sqrt{\beta_*},\ b + \tanh\theta = y_F\sqrt{\beta_*},$$
$$3b + \tanh\theta = (2y_F + y_0)\sqrt{\beta_*},\ \cosh^2\theta = \frac{\beta_0}{\beta_*},$$
$$b^2 - \tanh^2\theta = y_0 y_F \beta_*,$$
$$(b - \tanh\theta)(3b + \tanh\theta) = \beta_* y_0(2y_F + y_0), \tag{E3}$$

leading to

$$I(t=0) = -\frac{\ln(y_0\sqrt{\beta_*})}{\beta_0/\beta_*} + y_0 y_F \beta_* \ln(y_0\sqrt{\beta_*}) - \frac{1}{2}\beta_* y_0(2y_F + y_0). \tag{E4}$$

Similarly, at $t = 2A$, we can show that

$$b - \tanh\theta = y_F\sqrt{\beta_*},\ b + \tanh\theta = y_0\sqrt{\beta_*},$$
$$3b + \tanh\theta = (y_F + 2y_0)\sqrt{\beta_*},\ \cosh^2\theta = \frac{\beta_0}{\beta_*},$$
$$b^2 - \tanh^2\theta = y_0 y_F \beta_*,$$
$$(b - \tanh\theta)(3b + \tanh\theta) = \beta_* y_F(y_F + 2y_0), \tag{E5}$$

leading to

$$I(t=2A) = -\frac{\ln(y_F\sqrt{\beta_*})}{\beta_0/\beta_*} + y_0 y_F \beta_* \ln(y_F\sqrt{\beta_*}) - \frac{1}{2}\beta_* y_F(y_F + 2y_0). \tag{E6}$$

Therefore, by using Eqs. (E4) and (E6) in Eq. (E1), we obtain Eq. (81) in the main text.

[1] A. L. Gibbs and F. E. Su, Int. Stat. Rev. **70**, 419 (2002).
[2] E. Chevallier, E. Kalunga, and J. Angulo, hal.archives-ouvertes.fr/hal-01245712.
[3] R. Jordan, D. Kinderlehrer, and F. Otto, SIAM J. Math. Anal. **29**, 1 (1998).
[4] A. Takatsu, Osaka J. Math. **48**, 1005 (2011).
[5] J. Lott, Commun. Math. Phys. **277**, 423 (2008).
[6] F. Otto, Comm. Part. Diff. Eqs. **26**, 101 (2001).
[7] S. I. R. Costa, S. A. Santos, and J. E. Strapasson, Discrete Appl. Math. **197**, 59 (2015).
[8] B. R. Frieden, *Science from Fisher Information*, Vol. 2 (Cambridge University Press, Cambridge, 2004).
[9] W. Gangbo and R. J. McCann, C. R. Acad. Sci. Paris **321**, 1653 (1995).
[10] S. Ferradans, G.-S. Xia, G. Peyré, and J.-F. Aujol, Lect. Notes Comput. Sci. **7893**, 137 (2013).
[11] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[12] G. Ruppeiner, Phys. Rev. A **20**, 1608 (1979).
[13] F. Schlögl, Z. Phys. B **59**, 449 (1985).
[14] E. H. Feng and G. E. Crooks, Phys. Rev. E **79**, 012104 (2009).
[15] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[16] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzé, A. Smerzi, and M. K. Oberthaler, Science **345**, 424 (2014).
[17] P. Salamon, J. D. Nulton, and R. S. Berry, J. Chem. Phys. **82**, 2433 (1985).
[18] J. Nulton, P. Salamon, B. Andresen, and Qi, Anmin, J. Chem. Phys. **83**, 334 (1985).
[19] G. E. Crooks, Phys. Rev. Lett. **99**, 100602 (2007).
[20] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. **108**, 190602 (2012).
[21] L. Diósi, K. Kulacsy, B. Lukács, and A. Rácz, J. Chem. Phys. **105**, 11220 (1996).
[22] P. Salamon, J. D. Nulton, G. Siragusa, A. Limon, D. Bedeaus, and S. Kjelstrup, J. Non-Equilib. Thermodyn. **27**, 45 (2002).
[23] S. B. Nicholson and E. Kim, Phys. Lett. A **379**, 83 (2015).
[24] S. B. Nicholson and E. Kim (unpublished).
[25] J. Heseltine and E. Kim, J. Phys. A **49**, 175002 (2016).
[26] A. P. L. Newton, E. Kim, and H.-L. Liu, Phys. Plasmas **20**, 092306 (2013).
[27] E. Kim, H. Liu, and J. Anderson, Phys. Plasmas **16**, 0552304 (2009).
[28] E. Kim and P. H. Diamond, Phys. Rev. Lett. **90**, 185006 (2003).
[29] A. P. Feinberg and R. A. Irizarry, Proc. Natl. Acad. Sci. USA **107**, 1757 (2010).
[30] N. X. Wang, X. M. Zhang, and X. B. Han, Braz. J. Phys. **42**, 253 (2012).
[31] J. Lee, K. S. Farquhar, J. Yun, C. Frankenberger, E. Bevilacqua, K. Yeung, E. Kim, G. Balázsi and M. R. Rosner, Proc. Natl. Acad. Sci. USA **111**, E364 (2014).
[32] U. Lee, J. J. Skinner, J. Reinitz, M. R. Rosner, and E. Kim, PLoS One **10**, e0132397 (2015).
[33] A. d'Onofrio, Chaos Solitons Fractals **41**, 875 (2009).
[34] H. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, 3rd ed. (Springer, Berlin, 2006), pp. 63-64.
[35] S. W. Flynn, H. C. Zhao, and J. R. Green, J. Chem. Phys. **141**, 104107 (2014); J. W. Nichols, S. W. Flynn, and J. R. Green, *ibid.* **142**, 064113 (2015).
[36] M. Polettini and M. Esposito, Phys. Rev. E **88**, 012112 (2013).
[37] S. Bordel, J. Stat. Mech. (2011) P05013.
[38] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[39] C. Klebaner, *Introduction to Stochastic Calculus with Applications* (Imperial College Press, London, 2012), Chap. 5.5.
[40] C. Gardiner, *Stochastic Methods*, 4th ed. (Springer, Berlin, 2008), Chap. 4.4.
[41] E. Wong and M. Zakai, Ann. Math. Stat. **36**, 1560 (1960).
[42] E. Kim and S. Nicholson, Phys. Lett. A **379**, 1613 (2015).
[43] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integral* (McGraw-Hill, New York, 1965), Chap. 2; R. P. Feynman, *Statistical Mechanics* (W. A. Benjamin, San Francisco, 1972), Chap. 3.