PHYSICAL REVIEW E **91**, 032120 (2015)

Pathwise thermodynamic structure in population dynamics

Yuki Sughiyama,[1] Tetsuya J. Kobayashi,[2] Koji Tsumura,[3] and Kazuyuki Aihara[2]
[1]*Department of Basic Science, School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan*
[2]*Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan*
[3]*Department of Information Physics and Computing, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

(Received 5 August 2014; revised manuscript received 21 January 2015; published 12 March 2015)

We reveal thermodynamic structure in population dynamics with phenotype switching. Mean fitness for a population of organisms is determined by a thermodynamic variational principle described by the large deviation of phenotype-switching dynamics. Owing to this variational principle, a response relation of the mean fitness with respect to changes of environments and phenotype-switching dynamics is represented as a thermodynamic differential form. Furthermore, we discuss the strength of the selection by using the difference between time-forward and time-backward (retrospective) processes.

DOI: 10.1103/PhysRevE.91.032120 PACS number(s): 05.70.Ce, 87.23.Kg, 05.10.Gg, 87.10.Mn

I. INTRODUCTION

Equilibrium statistical physics [1] enables us to reconstruct thermodynamics on the basis of microscopic ensembles (e.g., microcanonical and canonical ensembles) by employing the large deviation principle (LDP) [2–5]. Recently, such a framework (so-called thermodynamic structure) has been extensively applied to analysis of statistics on paths generated by Markov processes [6] in the field of nonequilibrium physics [7–13]. Whereas calculations of the sample mean on Markov paths are essential in an equilibrium situation, the main focus shifts to evaluation of statistics concerned with jumps (e.g., current and entropy production) in nonequilibrium situations [7,8,10,11,13]. The mathematics of LDP on Markov processes was first established by Donsker, Varadhan, and Ellis [14,15]. Based on such mathematics, Nemoto, and Sasa [10] proposed a method for observing the rate function [5] characterizing nonequilibrium behavior in actual experiments. Chetrite and Touchette [12] also elucidated an equivalence between microcanonical and canonical ensembles on paths.

On the other hand, path-ensemble representations have been introduced in the field of biology (especially in population dynamics) to analyze interplay between growth and phenotype switching in a growing population [16–21]. In these studies, the difference between time-forward and time-backward (retrospective) processes [16,20,21] plays an essential role. By using this difference, the linear response relation for mean fitness was constructed [16], and the new approach for measuring the strength of selection was also reported [18]. However, the mathematics connecting time-forward and retrospective processes is still primitive. In this study, we reveal the hidden structure of population dynamics that shares the same mathematical basis as those of thermodynamics and statistical physics. By importing *the cost of the phenotype switching* (which is defined in Sec. IV) to ordinary population dynamics, we obtain a transparent and unified structure. Owing to this structure, the mean fitness is determined by a thermodynamic variational principle. Furthermore, it is found that response coefficients are given by the large deviation of phenotype-switching dynamics through a variational form.

This paper is organized as follows (see Fig. 1).

In the next section, we introduce the thermodynamic structure of Markov processes. Here, we firstly review LDP for the pair empirical measure and secondly discuss LDP for statistics on paths in terms of the contraction principle. In Sec. III, we show several applications in nonequilibrium physics which employ the thermodynamic structure of Markov processes. After Sec. III, we are going to derive the main result of this paper, which is the thermodynamic structure in population dynamics. In Sec. IV, we give a concrete setup for population dynamics with phenotype switching. Secondly, for preparations of the following sections, we review the linear response theory for the mean fitness which has been discussed in Ref. [16]. Here, we show that response coefficients are represented by the pair ancestral distribution [i.e., Eq. (18)]. In Sec. V, with the thermodynamic structure of Markov processes, we analyze the population dynamics with phenotype switching. Here, we find that the mean fitness is determined by the Legendre transform of the rate function for the pair empirical measure [i.e., Eq. (23)]. Furthermore, by using a thermodynamic differential form Eq. (26), we elucidate that responses of the mean fitness to changes of environments and phenotype switching are given by the large deviation for the pair empirical measure of phenotype-switching dynamics through the variational form Eq. (27). In Sec. VI, to clarify the consistency between our response theory and that discussed in Ref. [16], we introduce retrospective processes, which are obtained by solving the variational form. Here, we also show that the retrospective process coincides with Doob's h-transformed tilted process [12]. In Sec. VII, we discuss the strength of selection by using the Kullback-Leibler divergence between time-forward and retrospective processes. Finally, we conclude this paper in Sec. VIII.

II. THERMODYNAMIC STRUCTURE OF MARKOV PROCESSES

Before working on the problem of population dynamics, we reformulate the thermodynamic structure of Markov processes in terms of "Level-3" LDP; similar calculations have been done in Refs. [10,12,13]. Let $X_\tau = \{x_\tau, \ldots, x_1, x_0\}$ be a path of a finite-state discrete-time ergodic Markov process evolving over a time interval $t \in [0,\tau]$, which is generated by a transition matrix $T(\cdot|\cdot')$. Consider the frequency of a jump from x' to x in a Markov path X_τ. This quantity is represented by the pair empirical measure defined by

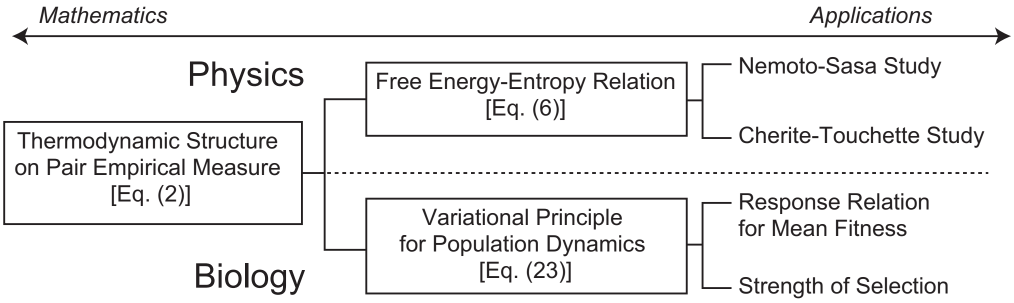

FIG. 1. Overview of this paper.

$L_{2,\tau}(x,x') \equiv 1/\tau \Sigma_{t=1}^{\tau} \delta_{x,x_t}\delta_{x',x_{t-1}}$. Although the pair empirical measure is a random quantity, it converges to the typical measure $J^{\rm typical}(x,x')$ in the limit $\tau \to \infty$, due to the law of large numbers. Now, evaluate the large deviation around this typical measure, which is given by a rate function defined as $I_T[J] \equiv \lim_{\tau\to\infty} 1/\tau \log {\rm Prob}[L_{2,\tau} = J]$. Here, the typical measure is characterized as $I_T[J^{\rm typical}] = 0$. Owing to Sanov's theorem [2,5], we can obtain the explicit form of the rate function as

$$I_T[J] = \begin{cases} D_{KL}[J|T \times q] & \text{if } J \text{ is shift invariant,} \\ \infty & \text{otherwise,} \end{cases} \tag{1}$$

where $D_{KL}[J|T \times q]$ denotes the Kullback-Leibler divergence as $D_{KL}[J|T \times q] = \Sigma_{x,x'} J(x,x') \log\{J(x,x')/T(x|x')q(x')\}$, and q represents the marginal distribution of J. Here, "shift invariant" means that J satisfies the following equation, $\Sigma_{x'} J(x',x) = \Sigma_{x'} J(x,x') = q(x)$. On the other hand, we consider the scaled cumulant generating function of the pair empirical measure $L_{2,\tau}(x,x')$ as

$$\begin{aligned} \lambda[k,T] &\equiv \lim_{\tau\to\infty} \frac{1}{\tau} \log\langle e^{\tau \sum_{x,x'} k(x,x') L_{2,\tau}(x,x')}\rangle_{\mathbb{P}^T} \\ &= \lim_{\tau\to\infty} \frac{1}{\tau} \log\langle e^{\sum_{i=1}^{\tau} k(x_i,x_{i-1})}\rangle_{\mathbb{P}^T}, \end{aligned} \tag{2}$$

where $\langle\cdot\rangle_{\mathbb{P}^T}$ denotes expectation with respect to the path probability $\mathbb{P}^T(X_\tau) = \Pi_{i=1}^{\tau} T(x_i|x_{i-1}) P_0(x_0)$ and P_0 is an initial condition. From Varadhan's theorem [22], the rate function $I_T[J]$ connects to the cumulant generating function $\lambda[k,T]$ by the Legendre transform,

$$\lambda[k,T] = \max_J \left\{ \sum_{x,x'} k(x,x') J(x,x') - I_T[J] \right\}. \tag{3}$$

An alternative derivation of Eqs. (1) and (3) without employing the knowledge of LDP is shown in Appendix A. The variational form (3) gives the mathematical basis of the thermodynamic structure of Markov Processes.

In many situations of statistical physics, major interest is statistics on paths (e.g., the sample mean and the time-averaged current) rather than the pair empirical measure. To clarify the link between them, we consider the following statistic on paths:

$$A_\tau[X_\tau] = \frac{1}{\tau} \sum_{i=1}^{\tau} \{f(x_i) + g(x_i,x_{i-1})\}, \tag{4}$$

where the first and the second terms represent the sample mean of f and a time-averaged current-like quantity (e.g., heat and entropy flow), respectively. Here, we defined $g(x,x) = 0$. The scaled cumulant generating function for A_τ is evaluated as

$$\begin{aligned} \hat{\lambda}[h,T] &\equiv \lim_{\tau\to\infty} \frac{1}{\tau} \log\langle e^{\tau h A_\tau[X_\tau]}\rangle_{\mathbb{P}^T} \\ &= \max_J \left\{ h \sum_{x,x'} \{f(x) + g(x,x')\} J(x,x') - I_T[J] \right\}, \end{aligned} \tag{5}$$

where we use $\hat{\lambda}[h,T] = \lambda[h\{f+g\},T]$ and Eq. (3). By defining $a \equiv \Sigma_{x,x'}\{f(x) + g(x,x')\}J(x,x')$, we obtain

$$\hat{\lambda}[h,T] = \max_a \{ha - \hat{I}_T[a]\}, \tag{6}$$

where $\hat{I}_T$ represents the rate function of the statistic A_τ such that

$$\hat{I}_T[a] = \min_{J : \Sigma_{x,x'}\{f(x)+g(x,x')\}J(x,x')=a} I_T[J]. \tag{7}$$

Equation (7) expresses the contraction principle in LDP [2,5]. The variational form (6) implies the relationship between free energy and entropy in ordinary thermodynamics, where $\hat{\lambda}$ and $\hat{I}_T$ correspond to free energy and entropy, respectively. h and a also represent inverse temperature and internal energy, respectively. Finally, by differentiating both sides of Eq. (6) with respect to h, we obtain the Gibbs-Helmholtz equation, which represents a relationship between free energy and internal energy as

$$\frac{\partial \hat{\lambda}[h,T]}{\partial h} = a^* \equiv \arg\max_a \{ha - \hat{I}_T[a]\}. \tag{8}$$

By solving the maximization in Eq. (8), we also obtain the temperature-entropy relation,

$$\frac{d\hat{I}_T[a^*]}{da^*} = h. \tag{9}$$

III. APPLICATIONS OF THE VARIATIONAL FORM IN PHYSICS

Applications of the thermodynamic structure discussed in the previous section yield many important results. We begin with rewriting Eq. (5) into a useful form for applications. Note that the shift invariant pair empirical measure J can be decomposed as $J(x,x') = R(x|x')P_R^{\rm st}(x')$, in which R is a stochastic-irreducible matrix and $P_R^{\rm st}$ represents its stationary distribution [i.e., $\Sigma_{x'} R(x|x')P_R^{\rm st}(x') = P_R^{\rm st}(x)$], and the marginal distribution q can be written as $q(x) = P_R^{\mathrm{st}}(x)$. Then, by substituting these into Eq. (5), we obtain

$$\hat{\lambda}[h,T] \equiv \max_{R\in\mathcal{R}} \left\{ h \sum_{x,x'} \{f(x) + g(x,x')\} R(x|x') P_R^{\mathrm{st}}(x') - \sum_{x,x'} R(x|x') P_R^{\mathrm{st}}(x') \log \frac{R(x|x')}{T(x|x')} \right\}, \tag{10}$$

where $\mathcal{R}$ denotes a set of stochastic-irreducible matrices. By differentiating both sides of Eq. (10) with respect to h, we have

$$\frac{\partial \hat{\lambda}[h,T]}{\partial h} = \sum_{x,x'} \{f(x) + g(x,x')\} R^*(x|x') P_{R^*}^{\mathrm{st}}(x'), \tag{11}$$

where R^* attains the maximum of Eq. (10) as

$$R^*(\cdot|\cdot') \equiv \arg\max_{R\in\mathcal{R}} \left\{ h \sum_{x,x'} \{f(x) + g(x,x')\} R(x|x') P_R^{\mathrm{st}}(x') - \sum_{x,x'} R(x|x') P_R^{\mathrm{st}}(x') \log \frac{R(x|x')}{T(x|x')} \right\}. \tag{12}$$

In Eq. (11), the differentiation of the cumulant generating function for the Markov process T gives the typical (expectation) value of statistic A_τ on the modified process R^* which is determined by Eq. (12). With this framework, we can estimate, in principle, the rate function $\hat{I}_T$ experimentally as follows [23]: (i) We construct the modified processes following Eq. (12) for all h; (ii) for each modified process, we observe the typical value of A_τ; (iii) $\hat{\lambda}$ is evaluated by integrating both sides of Eq. (11), and $\hat{I}_T$ is finally obtained by the Legendre transform [24].

This framework can also be used as a method for constructing processes R_{target} which give the typical value of A_τ as a_{target} [25]. We attempt to make such processes by modifying the original process T; that is, we make R_{target} by using the modified process R^*. Here, we assume that we know the form of the rate function for the original process, $\hat{I}_T[a]$ and that it is convex. Solving Eq. (12), we obtain

$$R^*(x|x') = u(x) e^{h\{f(x)+g(x,x')\}-\hat{\lambda}} T(x|x')/u(x'), \tag{13}$$

where u denotes the largest left eigenvector of the matrix $e^{h\{f(\cdot)+g(\cdot,\cdot')\}}T(\cdot|\cdot')$ whose eigenvalue is $e^{\hat{\lambda}[h,T]}$, i.e., $\Sigma_x u(x) e^{h\{f(x)+g(x,x')\}} T(x|x') = u(x') e^{\hat{\lambda}[h,T]}$. [The derivation of Eq. (13) is shown in Appendix B.] This process is called Doob's h-transformed tilted process [12]. By comparing Eq. (11) with Eq. (8), the typical value of A_τ on the modified process R^* is represented as

$$\sum_{x,x'} \{f(x) + g(x,x')\} R^*(x|x') P_{R^*}^{\mathrm{st}}(x') = a^*(h). \tag{14}$$

Since the parameter h in Eq. (13) is free, we can obtain a_{target} by tuning $a^*(h)$ (by tuning h). By employing the fact that $a^*(h)$ is also given by solving the temperature-entropy relation Eq. (9), we can fix the free parameter h at $h_{\mathrm{target}} = d\hat{I}_T[a_{\mathrm{target}}]/da_{\mathrm{target}}$. Thus, by substituting h_{target} into Eq. (13), we find the target process as $R_{\mathrm{target}} = u(x) e^{h_{\mathrm{target}}\{f(x)+g(x,x')\}-\hat{\lambda}} T(x|x')/u(x')$.

IV. POPULATION DYNAMICS AND LINEAR RESPONSE THEORY

We are going to investigate population dynamics with phenotype switching, which describes growing populations of organisms and cells. We first give a concrete setup for population dynamics. In this paper, it is assumed that the doubling rate of an organism depends on its phenotype and environment. Let x be a phenotype of an organism in the considered system, and its time evolution is given by a discrete-time ergodic Markov jump process generated by a transition matrix $T(\cdot|\cdot')$. The doubling rate of an organism with phenotype x is denoted by $\mu(x)$, whose functional form depends on the environment. For convenience, we define the reproduction rate as $\alpha(x) \equiv \log\{1 + \mu(x)\}$. That is, an organism with phenotype x reproduces $e^{\alpha(x)}$ daughters, in the sense of expectation for each time step. The daughters share the same phenotype as their mother. The above setup is common in the field of population dynamics [16,20]. In addition to this setup, we take *the cost of the phenotype switching* into account in this paper. This cost is defined as a decrease in the reproduction rate during the phenotype switching by $\beta(x,x')$. Organisms pay this cost, e.g., due to dying with the probability $1 - e^{\beta(x,x')}$ when the phenotype switches from x' to x [i.e., $\beta(x,x') < 0, \beta(x,x') = 0$ for $x = x'$]. While β is not so common in the modeling of population dynamics, it is biologically relevant because phenotype switching usually accompanies large rearrangement of intercellular states. The functional form of β again depends on the environment. Under this extended setup, time evolution of the total population size N_t^{tot} is given by

$$N_{t+1}^{\mathrm{tot}} = \sum_{x,x'} e^{k(x,x')} T(x|x') P_t(x') N_t^{\mathrm{tot}}, \tag{15}$$

where $P_t(x)$ denotes the phenotype distribution at time t, and we define the growth rate as $k(x,x') \equiv \alpha(x) + \beta(x,x')$ for simplicity. In the field of population dynamics, it is important to analyze the stationary mean fitness, which represents the growth rate of the total population size in a stationary situation. Here, Let us review the linear response theory for this mean fitness, following the method of reference [16]. The mean fitness is defined as $\lambda[k,T] \equiv \log N_{t+1}^{\mathrm{tot}}/N_t^{\mathrm{tot}}$. In the stationary situation, it is calculated as

$$\lambda[k,T] = \log \sum_{x,x'} e^{k(x,x')} T(x|x') v(x'), \tag{16}$$

where $v(x)$ represents the phenotype distribution in the stationary situation. By taking into account the fact that the time evolution of P_t is represented by $P_{t+1}(x) = \Sigma_{x'} e^{k(x,x')-\lambda[k,T]} T(x|x') P_t(x')$ and the matrix $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$ is irreducible, $v(x)$ coincides with the largest right eigenvector of $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$ and the mean fitness $\lambda[k,T]$ is evaluated by its eigenvalue as $\Sigma_{x'} e^{k(x,x')} T(x|x') v(x') = e^{\lambda[k,T]} v(x)$. Variation of the mean fitness Eq. (16) leads to the response relation,

$$\delta\lambda[k,T] = \sum_{x,x'} J^*(x,x') \delta k(x,x') + \sum_{x,x'} \frac{J^*(x,x')}{T(x|x')} \delta T(x|x'), \tag{17}$$

where the response coefficient $J^*(x,x')$ is given by

$$J^*(x,x') = u(x)e^{k(x,x')-\lambda[k,T]}T(x|x')v(x'). \tag{18}$$

Here $u(x)$ represents the largest left eigenvector of the time evolution matrix $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$, which is normalized as $\Sigma_x u(x)v(x) = 1$; the derivation of Eqs. (17) and (18) is shown in Appendix C. The left-hand side of Eq. (18) is called the pair ancestral distribution [16]. Its biological meaning is as follows: Let us pick an arbitrary organism from the stationary population and consider its ancestral parent at an infinitely earlier time $t = -\infty$ and its daughter at $t = -\infty + 1$. Then, the pair ancestral distribution $J^*(x,x')$ represents the probability that phenotypes of the ancestral parent and its daughter are x' and x. More precise interpretation of the pair ancestral distribution is in discussion of retrospective processes in Sec. VI. Accordingly, we find that the response coefficients of the mean fitness to changes of environments and phenotype switching are evaluated by the pair ancestral distribution.

Before closing this section, we show that the response relation Eq. (17) reduces to that derived in Ref. [16] if we neglect the cost of the phenotype switching [i.e., $\beta(x,x') = 0$ and $k(x,x') = \alpha(x)$]. By using the property of stochastic matrices, $\delta T(x|x) = -\Sigma_{x':x'\neq x}\delta T(x'|x)$, we can rewrite the response relation Eq. (17) as

$$\begin{aligned}\delta\lambda[k,T] &= \sum_{x,x'} J^*(x,x')\delta k(x,x') \\ &\quad + \sum_{x,x':x'\neq x}\left(\frac{J^*(x,x')}{T(x|x')} - \frac{J^*(x,x')}{T(x'|x')}\right)\delta T(x|x'). \end{aligned}\tag{19}$$

By substituting $\delta k(x,x') = \delta\alpha(x)$ and Eq. (18) into Eq. (19), we obtain the response relation,

$$\begin{aligned}\delta\lambda[k,T] &= \sum_x u(x)v(x)\delta\alpha(x) \\ &\quad + \sum_{x,x':x'\neq x} e^{-\lambda}(e^{\alpha(x)}u(x) - e^{\alpha(x')}u(x')) \\ &\quad \times v(x')\delta T(x|x'). \end{aligned}\tag{20}$$

Here, $u(x)v(x)$ represents the ancestral distribution, which implies the probability that the phenotype of an ancestor at $t = -\infty$ is x. Equation (20) is the same as the discrete-time version of the response relation in Ref. [16].

V. VARIATIONAL PRINCIPLE IN POPULATION DYNAMICS

With the above preparation, now we are in a position to discuss the thermodynamic structure in population dynamics with phenotype switching. In the following analysis, we will show that the mean fitness is determined by a variational principle described by the large deviation for phenotype switching dynamics. Owing to this variational principle, we can reformulate the response theory for the mean fitness in terms of a variational principle. As a result, response coefficients of the mean fitness are evaluated by the large deviation for the pair empirical measure of phenotype-switching dynamics through the variational form.

Let us evaluate the stationary mean fitness by employing a path-integral approach. Suppose that a system initially has N_0^{tot} organisms whose phenotype distribution is $P_0(x_0)$. The population of organisms at time τ who have undergone a phenotype-switching path (history) $X_\tau = \{x_\tau,\ldots,x_1,x_0\}$ is represented as

$$N_\tau(X_\tau) = \prod_{i=1}^{\tau} e^{k(x_i,x_{i-1})}T(x_i|x_{i-1})P_0(x_0)N_0^{\mathrm{tot}}, \tag{21}$$

where we again denote the growth rate as $k(x,x') \equiv \alpha(x) + \beta(x,x')$ for simplicity. Since the total population at time τ is given by integrating both sides of Eq. (21) with all paths, we obtain $N_\tau^{\mathrm{tot}} = \Sigma_{X_\tau}\Pi_{i=1}^{\tau}e^{k(x_i,x_{i-1})}T(x_i|x_{i-1})P_0(x_0)N_0^{\mathrm{tot}}$. Accordingly, the mean fitness is given by

$$\begin{aligned}\lambda[k,T] &= \lim_{\tau\to\infty}\frac{1}{\tau}\log\frac{N_\tau^{\mathrm{tot}}}{N_0^{\mathrm{tot}}} \\ &= \lim_{\tau\to\infty}\frac{1}{\tau}\log\langle e^{\sum_{i=1}^{\tau}k(x_i,x_{i-1})}\rangle_{\mathbb{P}^T}, \end{aligned}\tag{22}$$

where $\langle\cdot\rangle_{\mathbb{P}^T}$ denotes the expectation with respect to the path probability characterizing the phenotype-switching dynamics, $\mathbb{P}^T(X_\tau) = \Pi_{i=1}^{\tau}T(x_i|x_{i-1})P_0(x_0)$. Since the form of the last line in Eq. (22) is completely equivalent to the scaled cumulant generating function (2), we immediately obtain the variational form as follows:

$$\lambda[k,T] = \max_J\left\{\sum_{x,x'}k(x,x')J(x,x') - I_T[J]\right\}, \tag{23}$$

where I_T is the rate function of the pair empirical measure on the phenotype-switching dynamics T, which is expressed as Eq. (1). This variational principle can easily be extended to continuous-time cases (see Appendix D). The variational form (23) constitutes the foundation of biological discussions in this paper, which expresses that the mean fitness and the rate function of the phenotype-switching dynamics T are the Legendre dual [26]. This fact is crucial, because the variational form (23) enables us to estimate the large deviation of the phenotype switching (internal behavior) by measuring the mean fitness (external quantity). This duality contributes not only to population dynamics but also to the evaluation of the rate function of Markov processes [8]. The detailed discussion of the latter is shown in Appendix E.

Let us again investigate cases where we can neglect the cost of the phenotype switching, $\beta(x,x') = 0$ and $k(x,x') = \alpha(x)$. Consequently, we show that our variational principle Eq. (23) reduces to the study by Baake and Georgii [20]. In such cases, we can employ the contraction principle from "Level-3" LDP to "Level-2" LDP [2,5]. Thus, we obtain

$$\lambda[\alpha,T] = \max_q\left\{\sum_x \alpha(x)q(x) - \check{I}_T[q]\right\}, \tag{24}$$

where $\check{I}_T[q]$ denotes the rate function of the empirical measure $q(x)$. From Ellis's theorem [15], $\check{I}_T[q]$ is concretely expressed as

$$\check{I}_T[q] = \max_{v>0}\left\{\sum_x q(x)\log\frac{v(x)}{\sum_{x'}T(x|x')v(x')}\right\}. \tag{25}$$

Equations (24) and (25) coincide with the variational principle by Baake and Georgii.

In contrast to the Baake-Georgii variational principle Eqs. (24) and (25), our variational principle with Eqs. (23) and (1) is written in "Level-3" LDP. Therefore, we can easily differentiate both sides of Eq. (23); on the other hand, the differentiation of Eq. (24) is very complicated since the rate function Eq. (25) is expressed by a maximization problem. By differentiating Eq. (23) with respect to k and T, we obtain the thermodynamic differential form (the response relation) as

$$\delta\lambda[k,T] = \sum_{x,x'} J^*(x,x')\delta k(x,x') + \sum_{x,x'} \frac{J^*(x,x')}{T(x|x')}\delta T(x|x'), \tag{26}$$

where J^* is given by

$$J^*(\cdot,\cdot') = \arg\max_J \left\{ \sum_{x,x'} k(x,x')J(x,x') - I_T[J] \right\}. \tag{27}$$

Here, we used the explicit form of the rate function, Eq. (1). Coefficients of the first and the second terms in Eq. (26) represent responses to changes of environments and the phenotype-switching dynamics, respectively. In addition, Eq. (27) indicates that these response coefficients are connected with the large deviation of the phenotype-switching dynamics T through the variational form. To be more precise, the response coefficients in the environmental condition $k(\cdot,\cdot')$ are evaluated by solving $k(\cdot,\cdot') = \delta I_T[J^*]/\delta J^*(\cdot,\cdot')$.

VI. RETROSPECTIVE PROCESSES

Here, we will clarify the consistency between our response theory discussed above and that discussed in Ref. [16]. By using the decomposition discussed in Sec. III, $J(x,x') = R(x|x')P_R^{\mathrm{st}}(x')$, we can rewrite Eq. (23) as

$$\lambda[k,T] \equiv \max_{R\in\mathcal{R}} \left\{ \sum_{x,x'} k(x,x')R(x|x')P_R^{\mathrm{st}}(x') - \sum_{x,x'} R(x|x')P_R^{\mathrm{st}}(x') \log\frac{R(x|x')}{T(x|x')} \right\}, \tag{28}$$

where $\mathcal{R}$ denotes a set of stochastic-irreducible matrices. We also define

$$R^*(\cdot|\cdot') \equiv \arg\max_{R\in\mathcal{R}} \left\{ \sum_{x,x'} k(x,x')R(x|x')P_R^{\mathrm{st}}(x') - \sum_{x,x'} R(x|x')P_R^{\mathrm{st}}(x') \log\frac{R(x|x')}{T(x|x')} \right\}. \tag{29}$$

By solving the maximization in Eq. (29), we obtain the corresponding Doob's h-transformed tilted process as

$$R^*(x|x') = u(x)e^{k(x,x')-\lambda[k,T]}T(x|x')/u(x'), \tag{30}$$

where u again denotes the largest left eigenvector of the matrix $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$ whose eigenvalue is $e^{\lambda[k,T]}$. The Markov process generated by the matrix (30) is called the retrospective process [16,20,21] in the field of population dynamics. Its biological meaning is as follows: Let us pick an arbitrary organism from the stationary population and backwardly trace its phenotype-switching history (i.e., lineage), $\{y_0,y_{-1},\ldots,y_t,\ldots,y_{-\infty+1},y_{-\infty}\}$. Here, we consider the partial history during the time interval from $t=-\infty$ to $t=-\infty+\tau$, $\{y_{-\infty+\tau},\ldots,y_{-\infty+1},y_{-\infty}\}$. Then, the stationary Markov process generated by the transition matrix R^* represents the following path probability, $R^*(x_\tau|x_{\tau-1})\times\cdots\times R^*(x_1|x_0)P_{R^*}^{\mathrm{st}}(x_0) = \mathrm{Prob}[\{y_{\tau-\infty},\ldots,y_{1-\infty},y_{-\infty}\} = \{x_\tau,\ldots,x_1,x_0\}]$. By using Eq. (30), we can evaluate J^* as

$$\begin{aligned} J^*(x,x') &= R^*(x|x')P_{R^*}^{\mathrm{st}}(x') \\ &= u(x)e^{k(x,x')-\lambda[k,T]}T(x|x')v(x'), \end{aligned} \tag{31}$$

where we use $P_{R^*}^{\mathrm{st}}(x) = u(x)v(x)$, and v represents the largest right eigenvector of the matrix $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$. By taking into account the above discussion for retrospective processes and comparing Eq. (31) with Eq. (18), we find that the response coefficient $J^*(x,x')$ expresses the pair ancestral distribution, which implies the probability that the switch of the phenotype during the time interval from $t=-\infty$ to $t=-\infty+1$ is the jump from x' to x. In addition, we find that $u(x)v(x)[=P_{R^*}^{\mathrm{st}}(x)]$ represents the ancestral distribution.

VII. STRENGTH OF SELECTION

In the previous section, the retrospective process is given by the tradeoff between effects of the growth k and the phenotype switching T through the variational form (29). Owing to this variational form, we can discuss strength of the selection as follows. Suppose the case where there is no selection; i.e., the function $k(x,x')$ depends neither on x nor x'. Then, we can neglect the first term in Eq. (29) and obtain $R^* = T$. In contrast, if the selection is very strong, the deviation of R^* from T becomes large because the first term in Eq. (29) strongly contributes to the maximization. Accordingly, the distance between the phenotype-switching dynamics T and the retrospective process R^*, $\Sigma_{x,x'}R^*(x|x')P_{R^*}^{\mathrm{st}}(x')\log\{R^*(x|x')/T(x|x')\}$, implies the strength of the selection.

Before completing this paper, we attempt to express this strength of the selection by employing the historical mean fitness introduced in Ref. [18]. Consider the case, $k(x,x') = h\gamma(x,x')$ (constant change of the growth rate). Then, the historical mean fitness is defined as $\Gamma(\gamma,T) \equiv d\lambda[h\gamma,T]/dh = \Sigma_{x,x'}\gamma(x,x')R^*(x|x')P_{R^*}^{\mathrm{st}}(x')$. By substituting $k(x,x') = h\gamma(x,x')$ into Eq. (28), we obtain

$$\begin{aligned} \lambda[h\gamma,T] &= \max_{R\in\mathcal{R}} \left\{ h\sum_{x,x'} \gamma(x,x')R(x|x')P_R^{\mathrm{st}}(x') - \sum_{x,x'} R(x|x')P_R^{\mathrm{st}}(x')\log\frac{R(x|x')}{T(x|x')} \right\} \\ &= h\sum_{x,x'}\gamma(x,x')R^*(x|x')P_{R^*}^{\mathrm{st}}(x') - \sum_{x,x'} R^*(x|x')P_{R^*}^{\mathrm{st}}(x')\log\frac{R^*(x|x')}{T(x|x')}, \end{aligned} \tag{32}$$

where R^* represents the retrospective process in the case with $k(x,x') = h\gamma(x,x')$. Thus, we find

$$\sum_{x,x'} R^*(x|x')P^{\mathrm{st}}_{R^*}(x')\log\frac{R^*(x|x')}{T(x|x')} = h\Gamma(\gamma,T) - \lambda[h\gamma,T]. \tag{33}$$

This relation implies that the strength of the selection is given by the difference between the mean fitness and the historical mean fitness; a similar discussion is implicitly employed in Ref. [18].

VIII. CONCLUSION

We have shown the thermodynamic structure in population dynamics with phenotype switching. The mean fitness for a population of organisms has been determined by the variational principle (the Legendre transform) for the rate function of the phenotype-switching dynamics. We have also found that the response coefficients of the mean fitness are given by using this variational principle. Finally, we have discussed the strength of the selection by employing the difference between the time-forward and time-backward (retrospective) processes.

However, we have just dealt with the situation in which environments can be regarded as constant for the time scale of the growing dynamics. That is, our thermodynamic structure in population dynamics is equivalent to that of the equilibrium thermodynamics. In a future paper [27], we will discuss cases where environments change in population growth.

ACKNOWLEDGMENTS

Y.S. would like to thank Prof. Yoichiro Takahashi for fruitful discussions on the large deviation principle. This research is partially supported by Platform for Dynamic Approaches to Living System from the Ministry of Education, Culture, Sports, Science and Technology, Japan, the Aihara Innovative Mathematical Modelling Project, JSPS through the FIRST Program, CSTP, Japan, and the Specially Promoted Project of the Toyota Physical and Chemical Research Institute.

APPENDIX A

Here, we derive Eqs. (1) and (3) from Eq. (2). The scaled cumulant generating function in the finite time τ is calculated as

$$\begin{aligned}\lambda_\tau[k,T] &= \frac{1}{\tau}\log\sum_{\{x_i\}}\prod_{i=1}^{\tau} e^{k(x_i,x_{i-1})}T(x_i|x_{i-1})P_0(x_0)\\ &= \frac{1}{\tau}\log\sum_{\{x_i\}}\mathbb{P}(x_\tau,\ldots,x_1,x_0)\\ &\quad\times\frac{\prod_{i=1}^{\tau} e^{k(x_i,x_{i-1})}T(x_i|x_{i-1})P_0(x_0)}{\mathbb{P}(x_\tau,\ldots,x_1,x_0)},\end{aligned} \tag{A1}$$

where $\mathbb{P}(x_\tau,\ldots,x_1,x_0)$ denotes an arbitrary probability measure on the path $\{x_\tau,\ldots,x_1,x_0\}$. By using Jensen's inequality, we obtain

$$\begin{aligned}\lambda_\tau[k,T] &\geqslant \frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\\ &\quad\times\log\frac{e^{\sum_{i=1}^{\tau}k(x_i,x_{i-1})}\prod_{i=1}^{\tau}T(x_i|x_{i-1})P_0(x_0)}{\mathbb{P}(\{x_i\})}\\ &= \frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\sum_{i=1}^{\tau}k(x_i,x_{i-1})\\ &\quad-\frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\log\frac{\mathbb{P}(\{x_i\})}{\prod_{i=1}^{\tau}T(x_i|x_{i-1})P_0(x_0)}.\end{aligned} \tag{A2}$$

Thus, we have

$$\begin{aligned}\lambda_\tau[k,T] = \max_{\mathbb{P}}\Bigg\{&\frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\sum_{i=1}^{\tau}k(x_i,x_{i-1})\\ &-\frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\log\frac{\mathbb{P}(\{x_i\})}{\prod_{i=1}^{\tau}T(x_i|x_{i-1})P_0(x_0)}\Bigg\}.\end{aligned} \tag{A3}$$

We also define a probability measure, which attains the maximum in Eq. (A3), as

$$\begin{aligned}\mathbb{P}^*(\{x_i\}) \equiv \arg\max_{\mathbb{P}}\Bigg\{&\frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\sum_{i=1}^{\tau}k(x_i,x_{i-1})\\ &-\frac{1}{\tau}\sum_{\{x_i\}}\mathbb{P}(\{x_i\})\log\frac{\mathbb{P}(\{x_i\})}{\prod_{i=1}^{\tau}T(x_i|x_{i-1})P_0(x_0)}\Bigg\}.\end{aligned} \tag{A4}$$

By solving Eq. (A4), we reach

$$\begin{aligned}\mathbb{P}^*(\{x_i\}) &= \prod_{i=1}^{\tau}e^{k(x_i,x_{i-1})}T(x_i|x_{i-1})P_0(x_0)\\ &\quad\times\left\{\sum_{\{x_i\}}\prod_{i=1}^{\tau}e^{k(x_i,x_{i-1})}T(x_i|x_{i-1})P_0(x_0)\right\}^{-1}\\ &= \prod_{i=1}^{\tau}e^{k(x_i,x_{i-1})-\lambda_\tau[k,T]}T(x_i|x_{i-1})P_0(x_0),\end{aligned} \tag{A5}$$

where we again use the definition of the scaled cumulant generating function, Eq. (A1). Now, assume a stochastic matrix

$$R^*(x|x') = u(x)e^{k(x,x')-\lambda}T(x|x')/u(x'), \tag{A6}$$

where u denotes the largest left eigenvector of the matrix $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$ and e^{λ} is its eigenvalue, i.e., $\Sigma_x u(x)e^{k(x,x')}T(x|x') = u(x')e^{\lambda}$. By using this matrix (A6), we write a Markov process as

$$\begin{aligned}\prod_{i=1}^{\tau}R^*(x_i|x_{i-1})P_0'(x_0) &= u(x_\tau)\prod_{i=1}^{\tau}e^{k(x_i,x_{i-1})-\lambda}T(x_i|x_{i-1})\\ &\quad\times\frac{1}{u(x_0)}P_0'(x_0),\end{aligned} \tag{A7}$$

where P_0' denotes an arbitrary initial distribution. Let us consider that the scaled cumulant generating function $\lambda_\tau[k,T]$ approaches the eigenvalue λ in the limit $\tau \to \infty$, i.e., $\lambda[k,T] \equiv \lim_{\tau\to\infty} \lambda_\tau[k,T] = \lambda$. By neglecting boundary effects, we can approximate $\mathbb{P}^*$ by the Markov process generated by $R^*(x_i|x_{i-1})$ as

$$\mathbb{P}^*(\{x_i\}) \approx \prod_{i=1}^{\tau} R^*(x_i|x_{i-1})P_0'(x_0). \tag{A8}$$

Since the maximization with respect to $\mathbb{P}$ in Eq. (A3) can be reduced to that with respect to the stochastic matrix R, we obtain

$$\begin{aligned}\lambda_\tau[k,T] \approx \max_{R\in\mathcal{R}} \Biggl\{ &\frac{1}{\tau}\sum_{\{x_i\}}\prod_{i=1}^{\tau} R(x_i|x_{i-1})P_0'(x_0)\sum_{i=1}^{\tau} k(x_i,x_{i-1}) \\ &-\frac{1}{\tau}\sum_{\{x_i\}}\prod_{i=1}^{\tau} R(x_i|x_{i-1})P_0'(x_0) \\ &\times \log\frac{\prod_{i=1}^{\tau} R(x_i|x_{i-1})P_0'(x_0)}{\prod_{i=1}^{\tau} T(x_i|x_{i-1})P_0(x_0)}\Biggr\}.\end{aligned} \tag{A9}$$

In the limit $\tau \to \infty$, the first and the second terms are evaluated as follows. The first term is

$$\begin{aligned}&\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\{x_i\}}\prod_{i=1}^{\tau} R(x_i|x_{i-1})P_0'(x_0)\sum_{i=1}^{\tau} k(x_i,x_{i-1}) \\ &= \sum_{x,x'} R(x|x')P_R^{\rm st}(x')k(x,x'),\end{aligned} \tag{A10}$$

where $P_R^{\rm st}$ represents stationary distribution of R. The second term is

$$\begin{aligned}&\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\{x_i\}}\prod_{i=1}^{\tau} R(x_i|x_{i-1})P_0'(x_0)\log\frac{\prod_{i=1}^{\tau} R(x_i|x_{i-1})P_0'(x_0)}{\prod_{i=1}^{\tau} T(x_i|x_{i-1})P_0(x_0)} \\ &= \sum_{x,x'} R(x|x')P_R^{\rm st}(x')\log\frac{R(x|x')}{T(x|x')}.\end{aligned} \tag{A11}$$

By substituting Eqs. (A10) and (A11) into Eq. (A9) and taking the limit $\tau \to \infty$, we reach

$$\begin{aligned}\lambda[k,T] = \max_{R\in\mathcal{R}} \Biggl\{ &\sum_{x,x'} k(x,x')R(x|x')P_R^{\rm st}(x') \\ &- \sum_{x,x'} R(x|x')P_R^{\rm st}(x')\log\frac{R(x|x')}{T(x|x')}\Biggr\},\end{aligned} \tag{A12}$$

where $\mathcal{R}$ denotes a set of the stochastic matrices. This equation (A12) is equivalent to the variational form Eq. (10) in the main text. By considering $J(x,x') = R(x|x')P_R^{\rm st}(x')$ discussed in the main text, we finally get Eqs. (1) and (3).

APPENDIX B

Here, we briefly prove that the solution of Eq. (12) is given by Eq. (13). Before going to the main proof, we show the relationship between the differential of the stochastic matrix, δR, and the differential of its stationary distribution, $\delta P_R^{\rm st}$. By taking into account the fact that $P_R^{\rm st}$ represents the right eigenvector of R with the unit eigenvalue, we obtain

$$\begin{aligned}&\sum_{x'}(R(x|x') + \delta R(x|x'))\left(P_R^{\rm st}(x') + \delta P_R^{\rm st}(x')\right) \\ &= P_R^{\rm st}(x) + \delta P_R^{\rm st}(x).\end{aligned} \tag{B1}$$

By neglecting the second order of δ, we attain

$$\sum_{x'}\left(\delta R(x|x')P_R^{\rm st}(x') + R(x|x')\delta P_R^{\rm st}(x')\right) = \delta P_R^{\rm st}(x). \tag{B2}$$

Now, let us start the main proof. We will prove the following identity:

$$\begin{aligned}&\delta\Biggl\{\sum_{x,x'} h\{f(x)+g(x,x')\}R(x|x')P_R^{\rm st}(x') \\ &\quad - \sum_{x,x'} R(x|x')P_R^{\rm st}(x')\log\frac{R(x|x')}{T(x|x')}\Biggr\}\Biggr|_{R=R^*} \\ &= 0.\end{aligned} \tag{B3}$$

The calculation of the variation in the left-hand side of Eq. (B3) gives us

$$\begin{aligned}&\delta x\Biggl\{\sum_{x,x'}\{f(x)+g(x,x')\}R(x|x')P_R^{\rm st}(x') \\ &\quad - \sum_{x,x'} R(x|x')P_R^{\rm st}(x')\log\frac{R(x|x')}{T(x|x')}\Biggr\}\Biggr|_{R=R^*} \\ &= \sum_{x,x'}\left(h\{f(x)+g(x,x')\} - \log\frac{R^*(x|x')}{T(x|x')}\right) \\ &\quad \times (R^*(x|x')\delta P_{R^*}^{\rm st}(x') + \delta R^*(x|x')P_{R^*}^{\rm st}(x')),\end{aligned} \tag{B4}$$

where we use the property of the stochastic matrix, $\Sigma_x \delta R^*(x|x') = 0$. By substituting Eq. (13) into Eq. (B4), we reach

$$\begin{aligned}&\sum_{x,x'}\left(\hat{\lambda} - \log\frac{u(x)}{u(x')}\right) \\ &\quad\times \left(R^*(x|x')\delta P_{R^*}^{\rm st}(x') + \delta R^*(x|x')P_{R^*}^{\rm st}(x')\right) \\ &= \sum_{x,x'}(\log u(x') - \log u(x)) \\ &\quad\times \left(R^*(x|x')\delta P_{R^*}^{\rm st}(x') + \delta R^*(x|x')P_{R^*}^{\rm st}(x')\right) \\ &= \sum_{x'}\log u(x')\delta P_{R^*}^{\rm st}(x') - \sum_{x,x'}\log u(x) \\ &\quad\times \left(R^*(x|x')\delta P_{R^*}^{\rm st}(x') + \delta R^*(x|x')P_{R^*}^{\rm st}(x')\right),\end{aligned} \tag{B5}$$

where we again use the properties of the stochastic matrix: $\Sigma_x R^*(x|x') = 1$, $\Sigma_x \delta R^*(x|x') = 0$, and $\Sigma_x \delta P_{R^*}^{\rm st}(x) = 0$. By employing Eq. (B2), we finally obtain

$$\begin{aligned}
&\sum_{x'} \log u(x')\delta P^{\mathrm{st}}_{R^*}(x') - \sum_{x,x'} \log u\,(x) \\
&\quad \times \left(R^*(x|x')\delta P^{\mathrm{st}}_{R^*}(x') + \delta R^*(x|x')P^{\mathrm{st}}_{R^*}(x')\right) \\
&= \sum_{x} \log u\,(x) \Bigg\{ \delta P^{\mathrm{st}}_{R^*}\,(x) - \sum_{x'} R^*(x|x')\delta P^{\mathrm{st}}_{R^*}(x') \\
&\quad - \sum_{x'} \delta R^*(x|x')P^{\mathrm{st}}_{R^*}(x') \Bigg\} \\
&= 0. \qquad \text{(B6)}
\end{aligned}$$

Accordingly, we have found the identity (B3).

APPENDIX C

Here, we derive Eqs. (17) and (18). We denote the largest right and left eigenvectors of the time evolution matrix $e^{k(\cdot,\cdot')}T(\cdot|\cdot')$ as v and u, respectively. These eigenvectors are normalized as $\Sigma_x v\,(x) = 1$ and $\Sigma_x u\,(x)\,v\,(x) = 1$. By using the notations above, the mean fitness Eq. (16) can be expressed as

$$\lambda[k,T] = \log \sum_{x,x'} u(x)e^{k(x,x')}T(x|x')v(x'). \qquad \text{(C1)}$$

Variation of both sides in Eq. (C1) leads us to

$$\begin{aligned}
\delta\lambda[k,T] &= e^{-\lambda[k,T]} \\
&\quad \times \Bigg\{ \sum_{x,x'} u(x)e^{k(x,x')}T(x|x')v(x')\delta k(x,x') \\
&\quad + \sum_{x,x'} u(x)e^{k(x,x')}v(x')\delta T(x|x') \\
&\quad + \sum_{x,x'} \delta u(x)e^{k(x,x')}T(x|x')v(x') \\
&\quad + \sum_{x,x'} u(x)e^{k(x,x')}T(x|x')\delta v(x') \Bigg\}. \qquad \text{(C2)}
\end{aligned}$$

The third and fourth terms can be evaluated as

$$\begin{aligned}
&\sum_{x,x'} \delta u(x)e^{k(x,x')}T(x|x')v(x') + \sum_{x,x'} u(x)e^{k(x,x')}T(x|x')\delta v(x') \\
&= e^{\lambda[k,T]} \sum_{x} \{\delta u(x)v(x) + u(x)\delta v(x)\} \\
&= e^{\lambda[k,T]}\delta \sum_{x} u(x)v(x) = 0, \qquad \text{(C3)}
\end{aligned}$$

where we use the above normalization condition. Thus, we obtain Eqs. (17) and (18).

APPENDIX D

Here, we discuss continuous-time cases, which are more common in biological models. We can move to continuous-time cases by using the following correspondences:

$$\begin{aligned}
T(x|x') &\to \delta_{x,x'} + \tilde{T}(x|x')\Delta t, \\
R(x|x') &\to \delta_{x,x'} + \tilde{R}(x|x')\Delta t, \\
k(x,x') &\to \tilde{\mu}(x)\Delta t + \beta(x,x'), \\
\lambda\,[k,T] &\to \tilde{\lambda}\,[\tilde{\mu},\beta,T]\,\Delta t,
\end{aligned} \qquad \text{(D1)}$$

where $\tilde{T}$ and $\tilde{R}$ denote transition rate matrices satisfying $\Sigma_x \tilde{T}(x|x') = 0$ and $\Sigma_x \tilde{R}(x|x') = 0$, respectively. Here, we used the fact that the reproduction rate $\tilde{\alpha}$ coincides with the doubling rate $\tilde{\mu}$ in continuous-time cases. The reproduction rate is defined as $\tilde{\alpha}(x)\Delta t \equiv \log\{1 + \tilde{\mu}(x)\Delta t\}$ in continuous-time cases. Thus, $\tilde{\alpha}(x) = \tilde{\mu}(x)$ in the limit $\Delta t \to 0$. By substituting these correspondences, Eq. (D1), into the variational form Eq. (28), we obtain

$$\begin{aligned}
\tilde{\lambda}\,[\tilde{\mu},\beta,T]\,\Delta t &= \max_{\tilde{R}\in\tilde{\mathcal{R}}} \Bigg\{ \sum_{x,x'} (\tilde{\mu}(x)\Delta t + \beta(x,x')) \\
&\quad \times (\delta_{x,x'} + \tilde{R}(x|x')\Delta t)P^{\mathrm{st}}_{\tilde{R}}(x') \\
&\quad - \sum_{x,x'} (\delta_{x,x'} + \tilde{R}(x|x')\Delta t)P^{\mathrm{st}}_{\tilde{R}}(x') \\
&\quad \times \log \frac{\delta_{x,x'} + \tilde{R}(x|x')\Delta t}{\delta_{x,x'} + \tilde{T}(x|x')\Delta t} \Bigg\}, \qquad \text{(D2)}
\end{aligned}$$

where $\tilde{\mathcal{R}}$ represents a set of transition rate matrices. Here, we used the fact that R and $\tilde{R}$ have the same stationary distribution, $P^{\mathrm{st}}_{R} = P^{\mathrm{st}}_{\tilde{R}}$. Taking the limit $\Delta t \to 0$, we attain the variational form for continuous-time cases,

$$\begin{aligned}
\tilde{\lambda}\,[\tilde{\mu},\beta,T] &= \max_{\tilde{R}\in\tilde{\mathcal{R}}} \Bigg\{ \sum_{x} \tilde{\mu}(x)P^{\mathrm{st}}_{\tilde{R}}(x) \\
&\quad + \sum_{x,x':x\neq x'} \beta(x,x')\tilde{R}(x|x')P^{\mathrm{st}}_{\tilde{R}}(x') \\
&\quad - \sum_{x} P^{\mathrm{st}}_{\tilde{R}}(x)(\tilde{R}\,(x|x) - \tilde{T}\,(x|x)) \\
&\quad - \sum_{x,x':x\neq x'} \tilde{R}(x|x')P^{\mathrm{st}}_{\tilde{R}}(x') \log \frac{\tilde{R}(x|x')}{\tilde{T}(x|x')} \Bigg\}. \qquad \text{(D3)}
\end{aligned}$$

In addition, this variational principle Eq. (D3) leads to the response relation,

$$\begin{aligned}
\delta\tilde{\lambda}\,[\tilde{\mu},\beta,T] &= \sum_{s} P^{\mathrm{st}}_{\tilde{R}^*}(x)\delta\tilde{\mu}(x) \\
&\quad + \sum_{x,x':x\neq x'} \tilde{R}^*(x|x')P^{\mathrm{st}}_{\tilde{R}^*}(x')\delta\beta(x,x') \\
&\quad + \sum_{x,x':x\neq x'} \left(\frac{\tilde{R}^*(x|x')P^{\mathrm{st}}_{\tilde{R}^*}(x')}{\tilde{T}(x|x')} - P^{\mathrm{st}}_{\tilde{R}^*}(x') \right) \\
&\quad \times \delta\tilde{T}(x|x'), \qquad \text{(D4)}
\end{aligned}$$

where we use the property of the transition rate matrix, $\delta\tilde{T}(x|x) = -\Sigma_{x':x'\neq x}\delta\tilde{T}(x'|x)$.

Finally, we calculate the maximization in Eq. (D3). Then, we obtain

$$\tilde{R}^*(x|x') = u(x)(H(x|x') - \tilde{\lambda}\delta_{x.x'})/u(x'), \tag{D5}$$

where H represents the time-evolution operator of the population distribution, $H(x|x') = \delta_{x,x'}\tilde{\mu}(x) + e^{\beta(x,x')}\tilde{T}(x|x')$. Equation (D5) expresses the continuous-time version of the retrospective process.

APPENDIX E

Authors of Ref. [8] propose a method for estimating the rate function of Markov processes by employing the "copying procedure." To be more precise, by artificially constructing a growing system in simulation, they evaluate the rate function. Here, we will review this method in terms of population dynamics. First, suppose the situation where we want to evaluate the rate function of the pair empirical measure on a given Markov process, $T(x|x')$. Then, if we use Eq. (23), we can find the rate function by observing the mean fitness for populations of organisms whose phenotype switching is described by $T(x|x')$; to be more precise, we obtain the rate function by the Legendre transform of the mean fitness. Second, consider the case where we want to compute the rate function of a statistic on Markov path generated by $T(x|x')$:

$$A_\tau\,[X_\tau] = \frac{1}{\tau}\sum_{i=1}^{\tau}\{f\,(x_i) + g\,(x_i,x_{i-1})\}\,. \tag{E1}$$

The mean fitness in the environmental condition $k(x,x') = h\{f(x) + g(x,x')\}$ is calculated as

$$\lambda[h\{f(x)+g(x,x')\},T] \equiv \lim_{\tau\to\infty}\frac{1}{\tau}\log\langle e^{\tau h A_\tau[X_\tau]}\rangle_{\mathbb{P}^T}. \tag{E2}$$

By using discussions regarding Eq. (6), the Legendre transform of this mean fitness gives the rate function of A_τ. Thus, we can obtain the rate function by measuring mean fitnesses for all environmental conditions given by $h \in \mathbb{R}$.

---

[1] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Pergamon, Oxford, 1980).

[2] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. (Springer, Berlin, 1998).

[3] R. S. Ellis, Entropy, *Large Deviations, and Statistical mechanics* (Springer, Berlin, 1985).

[4] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, Berlin, 1984).

[5] H. Touchette, Phys. Rep. **478**, 1 (2009).

[6] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer, Berlin, 1985)

[7] R. M. L. Evans, Phys. Rev. Lett. **92**, 150601 (2004).

[8] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).

[9] T. Speck and J. P. Garrahan, Eur. Phys. J. B **79**, 1 (2011).

[10] T. Nemoto and S.-I. Sasa, Phys. Rev. E **83**, 030105(R) (2011); **84**, 061113 (2011).

[11] T. Speck, A. Engel, and U. Seifet, J. Stat. Mech. (2012) P12001.

[12] R. Chetrite and H. Touchette, Phys. Rev. Lett. **111**, 120601 (2013); arXiv:1405.5157.

[13] Y. Sughiyama and M. Ohzeki, Interdiscip. Inform. Sci. **19**, 93 (2013); arXiv:1307.6111.

[14] M. D. Donsker and S. R. S. Varadhan, Commum. Pure Appl. Math. **28**, 1 (1975).

[15] R. S. Ellis, Ann. Prob. **12**, 1 (1984).

[16] J. Hermisson, O. Redner, H. Wagner, and E. Baake, Theor. Pop. Biol. **62**, 9 (2002).

[17] V. Mustonen and M. Lässing, Proce. Natl. Acad. Sci. **107**, 4248 (2010).

[18] S. Leibler and E. Kussell, Proc. Natl. Acad Sci. **107**, 13183 (2010).

[19] Y. Wakamoto, A. Y. Grosberg, and E. Kussell, Evolution **66**, 115 (2011).

[20] E. Baake and H.-O. Georgii, J. Math. Biol. **54**, 257 (2007).

[21] H.-O. Georgii and E. Baake, Adv. Appl. Prob. **35**, 1090 (2003).

[22] S. R. S. Varadhan, Commum. Pure Appl. Math. **19**, 261 (1966).

[23] This work is first done by Nemoto and Sasa [10] on diffusion processes. Here, the modification is given by an additional force.

[24] If $\hat{I}_T\,[a]$ is nonconvex, we just obtain the envelope of $\hat{I}_T\,[a]$. Therefore, we cannot use this method in such a case.

[25] This method is employed by Chetrite and Touchette in Refs. [12].

[26] Because of the convexity of $I_T\,[J]$, we can obtain $I_T\,[J] = \max_k\{\Sigma_{x,x'}k\,(x,x')\,J\,(x,x') - \lambda\,[k,T]\}$.

[27] T. J. Kobayashi and Y. Sughiyama (unpublished).