Stochastic and information-thermodynamic structures of population dynamics in a fluctuating environment

Tetsuya J. Kobayashi and Yuki Sughiyama

Stochastic and Information-thermodynamic Structures of Population Dynamics in Fluctuating Environment

Tetsuya J. Kobayashi[1, 2, a)] and Yuki Sughiyama[1]

[1)] *Institute of Industrial Science, the University of Tokyo.*
[2)] *PRESTO, JST.*

Adaptation in a fluctuating environment is a process of fueling environmental information to gain fitness. Living systems have gradually developed strategies for adaptation from random and passive diversification of the phenotype to more proactive decision making, in which environmental information is sensed and exploited more actively and effectively. Understanding the fundamental relation between fitness and information is therefore crucial to clarify the limits and universal properties of adaptation. In this work, we elucidate the underlying stochastic and information-thermodynamic structure in this process, by deriving causal fluctuation relations (FRs) of fitness and information. Combined with a duality between phenotypic and environmental dynamics, the FRs reveal the limit of fitness gain, the relation of time reversibility with the achievability of the limit, and the possibility and condition for gaining excess fitness due to environmental fluctuation. The loss of fitness due to causal constraints and the limited capacity of real organisms is shown to be the difference between time-forward and time-backward path probabilities of phenotypic and environmental dynamics. Furthermore, the FRs generalize the concept of evolutionary stable state (ESS) for fluctuating environment by giving the probability that the optimal strategy on average can be invaded by a suboptimal one owing to rare environmental fluctuation. These results clarify the information thermodynamic structures in adaptation and evolution.

I. INTRODUCTION

A. Adaptation in fluctuating environment

Adaptation is fundamental to all organisms for their survival and evolutionary success in a changing environment. In the course of evolution, living systems have gradually attained and developed more active and efficient strategies for adaptation, which generally accompany more effective use of environmental information. Understanding how the use of information is linked to the efficiency of adaptation is crucial to clarify the fundamental limits and universal properties of biological adaptations[1,2].

The most primitive strategy for adaptation is to randomly generate genetic and phenotypic heterogeneity in a population[3–6]. Provided that a sufficiently large heterogeneity is constantly generated in the population, a fraction of organisms can, by chance, have the types adaptive to the upcoming environmental state and circumvent extinction of the population at the cost of others with non-adaptive types[7,8]. Such a strategy is known as bet-hedging or phenotypic diversification and works even if the organisms are completely blind to the environment, without any *a priori* knowledge of its dynamics. The bet-hedging is a passive and *a posteriori*

adaptation in the sense that the adaptation is achieved extrinsically by and after the impact of environmental selection[9]. The evolutionary advantage of the bet-hedging strategy is demonstrated by the persistence of bacteria, pathogens, and cancer cells to antibiotic or anticancer drug treatments[5,6,10–12]. The gain of fitness by bet-hedging can be optimized if the population evolves to generate an appropriate pattern of heterogeneity by learning the environmental statistics[13]. Nevertheless, the gain of fitness by bet-hedging is fundamentally limited because of the passive and *a posteriori* nature of the strategy, in which the individual organisms have no access to the microscopic information of which environmental states will actually be realized.

With any access to such information, the loss can be avoided further by decision making: directly sensing the current environmental state, predicting the upcoming state, and switching into the phenotypic state that is adaptive to that state[14–17]. The strategy of adaptation via sensing is active and *a priori* in a sense that adaptation is intrinsically achieved by the predictive actions of the organisms[9]. In biologically relevant situations, both passive and active aspects of adaptation are intermingled because perfect sensing and prediction of environment are impossible with the limited capacities of biological systems.

[a)]Electronic mail: tetsuya@mail.crmind.net; http://research.crmind.net

B. Notions of information and analogy with physics in biological adaptation

At an analogical level, the problem of the fundamental law and the limits of adaptation and evolution shares several aspects with physics, especially with thermodynamics, which drove the long-lasting attempts to establish the thermodynamics of biological adaptation and evolution[18–24]. Among other areas, the fundamental limit of fitness in a changing environment and the value of environmental information have been a major focus in evolutionary biology[1,2,7,25–27]. Haccou and Iwasa may be the first who linked, albeit implicitly, environmental information with the gain of fitness in a stochastic environment[28]. Bergstrom and Lachmann pursued the fitness value of information by directly incorporating mutual information[29–31]. Others also pointed out some quantitative relations between fitness and information measures such as relative entropy and Jeffreys' divergence[32–34]. More recently, Rivoire and Liebler conducted a comprehensive analysis by employing an analogy between bet-hedging of organisms and horse race gambling[36], the link of which to information theory was revealed in the seminal work by Kelly in 1956[37]. However, all previous works either imposed certain restrictions on their models to derive the information-theoretic measures of fitness value[32,36,38] or had to introduce phenomenological measures for the value of information to accommodate more general situations[36,38], because they lacked an appropriate method to handle the mixture of the passive and active aspects in adaptation.

We recently resolved this problem[9] by combining a path integral formulation of population dynamics[39–42], a retrospective characterization of the selected population[41,43,44], and a variational structure in population dynamics[42,45]. The results we obtained generalized the limits of fitness gain by sensing and revealed that the gain satisfies fluctuation relations (FRs) that fundamentally constrain not only its average but also its fluctuation. These relations, alongside a previous work in the line of neutral theory[24], imply that fitness in the fluctuation environment shares, at least mathematically, similar structures to those of stochastic and information thermodynamics[46,47]. In our FRs, the fluctuation of fitness of a given population is evaluated by the difference from the fitness that achieves the maximum average fitness over all possible phenotypic histories of organisms. Conceptually, this means that we postulate a Darwinian demon, an imaginary organism, that can exhibit any type of behavior without imposing any constraint not only on biological capacity but also on the causality of dynamics. The FRs characterize the loss of fitness of a realistic organism from such an idealized organism. Thus, understanding the properties of the Darwinian demon and the deviation from it by a realistic organism are central to a deeper understanding of the behavior of populations in a changing environment. However, the implicit definition of the demon

as the maximizer of the average fitness hampers the explicit characterization of the demon and obscures the formal link to stochastic thermodynamics, in which a variational characterization is not common[46,47]. More practically, without an explicit characterization, we are unable to simulate possible behaviors of the demon even numerically.

C. Outline of main results

In this paper, we resolve these problems by deriving FRs of fitness without using the variational approach. To this end, we first formulate and generalize the problem of adaptation in a changing environment so that individual organisms can change not only their strategy of switching phenotypic states but also the strategy of allocating metabolic resources to each phenotypic state (Sec. II). Combined with the path integral formulation of population dynamics, this generalization enables us to obtain a decomposition of fitness with a combination of time-forward (chronological) and time-backward (retrospective) path probabilities (Sec. III). The decomposition naturally spells out an explicit representation of the upper bound of the average fitness, which was implicitly defined in our previous work[9].

For the bet-hedging problem without a sensing environment (Sec. IV), the decomposition directly leads to FRs of the fitness loss, which has a very similar form as the FRs of entropy production in stochastic thermodynamics[46]. After numerically verifying the derived FRs (Sec. IV B), we investigate the biological meanings and achievability of the FRs (Sec. V). The average FR is related to the evolutionary stable state (ESS), under which the strategy with maximal average fitness cannot be invaded by any other strategies. The detailed and integral FRs generalize the ESS by giving the probability that a suboptimal strategy outperforms the optimal one within a finite time interval owing to rare fluctuation of the environment[48]. By using a dualistic relation between phenotypic and environmental dynamics, the detailed FR is shown to be represented as the ratio of the path probability of the actual environment and that of the conjugate environment under which a given strategy of the organisms becomes optimal. The duality also clarifies that the average loss of fitness is directly related to the imperfectness of the adaptive behavior of the organisms, originating both from physical constraints and from the suboptimality of the behaviors.

The introduction of a sensing signal extends the FRs to accommodate the mutual information between the environment and the signal as the gain of fitness by sensing, in the same manner as mutual information bounds the negative gain of entropy production in information thermodynamics[47] (Sec. VI). Although the extended FRs cover very general situations, the FRs are not tight and therefore the mutual information overestimate the value of fitness by sensing. By explicitly assuming a

causal relation between the environment and the signal, the FRs are further modified to involve the directed information as a tighter bound of fitness gain (Sec. VII). This modification clarifies how the loss of fitness from the upper bound is related to the causality, the inaccessibility to perfect information of the environment, and the imperfect implementation of information processing. Finally, three quantities are introduced to account for the fitness loss of inappropriate sensing and the imperfectness of metabolic allocation and phenotypic switching strategies in general situations (Sec. VII C). The summary and future directions are described in Discussion (Sec. VIII).

II. MODELING ADAPTATION OF POPULATION IN CHANGING ENVIRONMENT

Let $x_t \in \mathfrak{S}_x$, $y_t \in \mathfrak{S}_y$, and $z_t \in \mathfrak{S}_z$ be the phenotype of a living organism, the state of the environment, and the state of the sensing signal at time t, respectively (Fig. 1 (A)). For simplicity, possible phenotypic, environmental, and signal states are assume to be discrete as in references[9,30,34,36]. The paths (histories) of the states up to time t are defined as $\mathcal{X}_t := \{x_\tau | \tau \in [0, t]\} \in \mathfrak{S}_{\mathcal{X}} := \mathfrak{S}_x^{\times(t+1)}$, $\mathcal{Y}_t := \{y_\tau | \tau \in [0, t]\} \in \mathfrak{S}_{\mathcal{Y}} := \mathfrak{S}_y^{\times(t+1)}$, and $\mathcal{Z}_t := \{z_\tau | \tau \in [0, t]\} \in \mathfrak{S}_{\mathcal{Z}} := \mathfrak{S}_z^{\times(t+1)}$, respectively. Time is also treated as discrete in this work.

A. Modeling phenotype switching

The phenotype of an organism, in general, switches stochastically over time, depending on its past phenotypic state and the sensed signal (Fig. 1 (A) and (B)). The switching dynamics is modeled, for example, by a Markov transition probability $\mathbb{T}_F(x_{t+1} | x_t, z_{t+1})$, which satisfies $\sum_{x_{t+1}} \mathbb{T}_F(x_{t+1} | x_t, z_{t+1}) = 1$ for all x_t and z_{t+1}. Although we mainly focus on the Markov switching, our result can be extended for causal switching $\mathbb{T}_F(x_{t+1} | \mathcal{X}_t, \mathcal{Z}_{t+1})$ in which the next phenotypic state x_t depends on both past phenotypic and signal histories $\mathcal{X}_t$ and $\mathcal{Z}_{t+1}$.

B. Modeling metabolic resource allocation

Next, we model the strategy of metabolic resource allocation of the organisms. Each organism is assumed to duplicate asexually to produce e^k daughter organisms on average. For each state of environment y, the organisms have a maximum replication rate, $e^{k_{max}(y)}$, that can be achieved only when the organism allocates its all metabolic resources for adapting only to that environmental state. Because the environment changes, however, the organism usually distributes its resources for different environmental conditions[49].

To represent this situation, we introduce a conditional probability $\mathbb{T}_K(y | x)$ that quantitatively represents the fraction of resources allocated to the environmental state y in a phenotype x. An instantaneous replication rate $k(x, y)$ of the phenotype x under the environmental state y is then assumed to be represented as

$$e^{k(x,y)} = e^{k_{max}(y)} \mathbb{T}_K(y | x), \qquad (1)$$

(Fig. 1 (C)). Note that such a decomposition of $e^{k(x,y)}$ can date back to Haccou and Iwasa[28], at least. This relation between the resource allocation strategy and the replication rate may appear to be restrictive because of the linear relation between the allocation strategy $\mathbb{T}_K(y | x)$ and the actual replication rate $e^{k(x,y)}$. Nevertheless, this decomposition of $k(x, y)$ is general enough because, for a given $k(x, y)$, we can find a pair of $e^{k_{max}(y)}$ and $\mathbb{T}_K(y | x)$ as long as the possible phenotypic states are fewer than those of the environment (refer to Appendix A). Such a situation is biologically plausible because the environment is usually more complex than the phenotype of an organism. Moreover, our setting includes the special situation that has been intensively investigated in previous works[32,36]. For example, when the numbers of phenotypic and environmental states are equal as $\#\mathfrak{S}_x = \#\mathfrak{S}_y$, $\mathbb{T}_K(y | x) = \delta_{x,y}$ corresponds to Kelly's horse race gambling[36,37], in which each phenotypic state allocates all metabolic resources to a certain environmental state and, as a result, can survive and grow only when the phenotypic state matches the realized environmental state.

C. Modeling sensing processes

We finally model the sensing process. We consider the case that organisms can obtain a sensing signal $z(t)$, which correlates with the environment $y(t)$. The sensed signal $z(t)$ is assumed to be common to all the organisms in the population (Fig. 1 (A)). A biologically relevant situation is that $z(t)$ is a vector of concentrations of extracellular signaling molecules that cannot be consumed as metabolites but correlate with the available metabolites. Another situation is that $z(t)$ is a subset of $y(t)$ to which the organisms have sensors. In either situation, the sensing noise should be negligibly small because all the organisms receive the same sensing signal $z(t)$. Even though sensing of a common signal cannot cover all biologically realistic situations such as individual sensing with noisy receptors[14–17,50,51], the common signal has been investigated in various works[9,28,36]. In this work, we mainly focus on the common sensing problem and touch on the individual sensing in Discussion.

D. Population dynamics of organisms

By combining the phenotype switching, the metabolic allocation, and the sensing strategies, we can explicitly
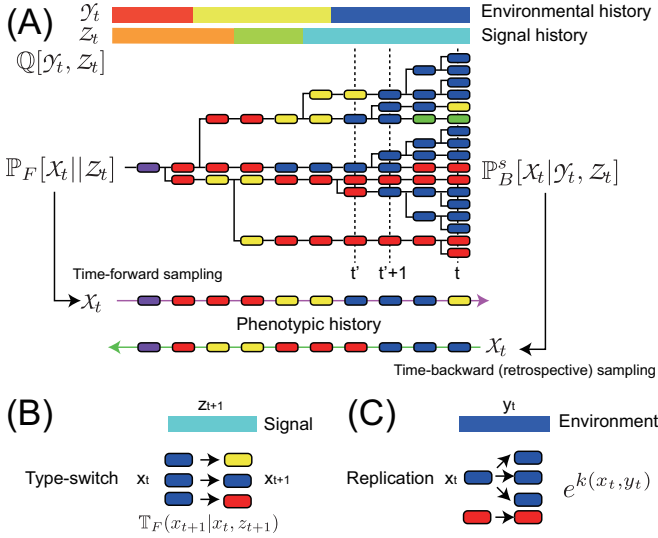
FIG. 1. Schematic diagrams for population dynamics of organisms with multiple types and sensing in a changing environment. (A) A tree representation of a growing population in a changing environment. Colors indicate different phenotypic states, environmental states, and signal over time. By tracking the phenotypic history from $t = 0$ to t in a time-forward manner, we can effectively obtain a sample $\mathcal{X}_t$ with probability $\mathbb{P}_F[\mathcal{X}_t || \mathcal{Z}_t]$. By tracking the phenotypic history retrospectively from t to $t = 0$ in a time-backward manner, we can obtain a sample $\mathcal{X}_t$ with probability $\mathbb{P}_B^s[\mathcal{X}_t | \mathcal{Y}_t, \mathcal{Z}_t]$. Note that we assume that the size of the population is sufficiently large when we introduce the population dynamics of the organisms eq. (2). (B) Phenotypic switching from x_t to x_{t+1} in response to environmental signal z_{t+1}. The switching probability is represented by $\mathbb{T}_F(x_{t+1} | x_t, z_{t+1})$. (C) Replication of organisms with phenotypic state x_t under environmental state y_t. Each organism generates $e^{k(x_t, y_t)}$ descendants on average.

derive the dynamics of the population of the organisms. Because both environmental and sensing histories are external factors of the organisms, the population dynamics of the organisms is described for a given pair of environmental and signaling histories, $\mathcal{Y}_t$ and $\mathcal{Z}_t$. Let $\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t) \in \mathbb{R}_{\geq 0}$ be the number of organisms whose phenotypic state is x_t at time t under the realization of environmental and signal histories $\mathcal{Y}_t$ and $\mathcal{Z}_t$. The population size of the organisms is assumed to be sufficiently large so that $\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t)$ can be well approximated as a continuous variable. Because the large population size enable us to effectively ignore the demographic fluctuation due to finite number of the organisms, we can obtain the update rule[9] of $\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t)$ as

$$\mathcal{N}_{t+1}^{\mathcal{Y}, \mathcal{Z}}(x_{t+1}) = e^{k(x_{t+1}, y_{t+1})} \sum_{x_t \in \mathbb{S}_x} \mathbb{T}_F(x_{t+1} | x_t, z_{t+1}) \mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t).$$
(2)

If we need to work on a population, the size of which is not sufficiently large, modeling of the population e.g., by a branching process is required. The statistical properties of the environmental and signal histories are generally

characterized by a joint path probability $\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]$.

III. PATH-WISE FORMULATION AND FITNESS DECOMPOSITION

By using $\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t)$ in the previous section, we define the fitness of the population and derive its path integral formulation. The cumulative fitness $\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]$ of the population at t under environmental and signal histories $\mathcal{Y}_t$ and $\mathcal{Z}_t$ is defined by the exponential expansion of the total population size as follows:

$$\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] := \ln \frac{\sum_{x_t} \mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t)}{\sum_{x_0} \mathcal{N}_0^{\mathcal{Y}, \mathcal{Z}}(x_0)} = \ln \frac{\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}}{\mathcal{N}_0^{\mathcal{Y}, \mathcal{Z}}}.$$
(3)

We here define $\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}} := \sum_{x_t} \mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}(x_t)$. The ensemble average of the cumulative fitness for different realizations of the environmental and signal histories is represented as

$$\langle \Psi_t^s \rangle_{\mathbb{Q}} := \langle \Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] \rangle_{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}.$$
(4)

A. Path-wise and retrospective formulation

As derived in[9,39,42], the cumulative fitness at time t can be represented with a path-wise (path integral) formulation. Let us first define the time-forward path probability of the phenotype switching as

$$\mathbb{P}_F[\mathcal{X}_t || \mathcal{Z}_t] := \left[\prod_{\tau=0}^{t-1} \mathbb{T}_F(x_{\tau+1} | x_\tau, z_{\tau+1}) \right] p(x_0),$$
(5)

where $p(x_0) := \mathcal{N}_0^{\mathcal{Y}, \mathcal{Z}}(x_0) / \sum_{x_0} \mathcal{N}_0^{\mathcal{Y}, \mathcal{Z}}(x_0)$. We here use Kramer's causal conditioning $||$ rather than the usual conditioning $|$ in order to indicate that the path probability $\mathbb{P}_F[\mathcal{X}_t || \mathcal{Z}_t]$ is causally generated by the Markov transition matrix $\mathbb{T}_F(x_{t+1} | x_t, z_{t+1})$, which depends only on the past phenotypic state x_t and the signal z_{t+1}[52,53]. The single bar $|$ is also used for the normal conditioning of a path probability that does not necessarily satisfy the causal relation between conditioning and conditioned histories. We also define the path-wise (historical) fitness of a phenotypic history $\mathcal{X}_t$ under an environmental history $\mathcal{Y}_t$ as

$$K[\mathcal{X}_t, \mathcal{Y}_t] := \sum_{\tau=0}^{t-1} k(x_{\tau+1}, y_{\tau+1}).$$
(6)

where $K[\mathcal{X}_t, \mathcal{Y}_t]$ is defined over all $\{\mathcal{X}_t, \mathcal{Y}_t\} \in \mathbb{S}_{\mathcal{X}} \times \mathbb{S}_{\mathcal{Y}}$. With these path-wise quantities[9], we obtain the population size of the organisms at time t, the past phenotypic history of which is $\mathcal{X}_t$ as

$$\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}[\mathcal{X}_t] = e^{K[\mathcal{X}_t, \mathcal{Y}_t]} \mathbb{P}_F[\mathcal{X}_t || \mathcal{Z}_t] \mathcal{N}_0^{\mathcal{Y}, \mathcal{Z}}.$$

Because $\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}} = \sum_{\mathcal{X}_t} \mathcal{N}_t^{\mathcal{Y}_t,\mathcal{Z}_t}[\mathcal{X}_t]$, the cumulative fitness with sensing $\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]$ is explicitly described as

$$\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] = \ln \frac{\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}}}{\mathcal{N}_0^{\mathcal{Y},\mathcal{Z}}} = \ln \left\langle e^{K[\mathcal{X}_t,\mathcal{Y}_t]} \right\rangle_{\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]}. \quad (7)$$

From this representation, the fitness can be described variationally[9,42] as

$$\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] = \max_{\mathbb{P}[\mathcal{X}_t]} \left[\langle K[\mathcal{X}_t, \mathcal{Y}_t] \rangle_{\mathbb{P}[\mathcal{X}_t]} - \mathcal{D}\Big[\mathbb{P}[\mathcal{X}_t]\Big|\Big|\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\Big] \right], \quad (8)$$

where $\mathcal{D}[\mathbb{P}||\mathbb{P}'] := \sum_{\mathcal{X}_t} \mathbb{P}[\mathcal{X}_t] \ln \mathbb{P}[\mathcal{X}_t]/\mathbb{P}'[\mathcal{X}_t]$ is the Kullback-Leibler divergence (relative entropy)[31,54] between two path measures $\mathbb{P}$ and $\mathbb{P}'$. It should be noted that both KL divergence and Kramer's causal conditioning use the double bar $||$ but their meanings are different. The maximization is achieved with the time-backward retrospective path probability defined by

$$\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t] := e^{K[\mathcal{X}_t,\mathcal{Y}_t]-\Psi^s[\mathcal{Y}_t,\mathcal{Z}_t]}\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t] = \frac{\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}}[\mathcal{X}_t]}{\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}}}. \quad (9)$$

If the phenotypic switching does not depend on the sensing signal as $\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t] = \mathbb{P}_F[\mathcal{X}_t]$, which corresponds to the bet-hedging by random phenotypic switching, $\mathbb{P}_B^s$ is reduced to

$$\mathbb{P}_B^b[\mathcal{X}_t|\mathcal{Y}_t] := e^{K[\mathcal{X}_t,\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t]}\mathbb{P}_F[\mathcal{X}_t] = \frac{\mathcal{N}_t^{\mathcal{Y}}[\mathcal{X}_t]}{\mathcal{N}_t^{\mathcal{Y}}}, \quad (10)$$

where the superscript b denotes bet-hedging and $\Psi^b[\mathcal{Y}_t] := \ln \left\langle e^{K[\mathcal{X}_t,\mathcal{Y}_t]} \right\rangle_{\mathbb{P}_F[\mathcal{X}_t]}$. Because $\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}}[\mathcal{X}_t]$ is the number of organisms with phenotypic history $\mathcal{X}_t$ in the population at time t, the second equality in eq. (9) indicates that $\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]$ is the fraction of the organisms that has phenotpyic history $\mathcal{X}_t$. This property of $\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]$ leads to an interpretation of $\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]$ as the probability of observing a certain phenotypic history $\mathcal{X}_t$ under a realization of the environmental and signal histories $\mathcal{Y}_t$ and $\mathcal{Z}_t$ when we sample an organism in the population at time t and track its phenotypic history in a time-backward manner, retrospectively[9,41–44]. Because the organisms grow more if their phenotypic histories are more adaptive than others for the given environmental realization $\mathcal{Y}_t$, the chance to observe a certain phenotypic history $\mathcal{X}_t$ after selection under $\mathcal{Y}_t$ is biased to $\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]$ from the probability $\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]$ with which the same phenotypic history is intrinsically generated. Because the selected phenotypic histories strongly depend on the actual realization of the environmental history, $\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]$ is conditional on $\mathcal{Y}_t$. It should be noted that $\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]$ is not necessarily causal, because of which we use the normal conditioning $|$[9,42].

B. Decomposition of fitness

In order to understand the relation between fitness and information obtained by sensing, we decompose the cumulative fitnesses into biologically relevant components. To obtain the decompositions, we first define a constant ϕ_0 and a probability distribution $q_0(y)$ by using $k_{\max}(y)$ as $\phi_0 := -\ln \sum_y e^{-k_{\max}(y)}$ and $q_0(y) := e^{\phi_0} e^{-k_{\max}(y)}$. From these definitions, $k_{\max}(y)$ can be described as $e^{k_{max}(y)} = e^{\phi_0}/q_0(y)$. By defining $\mathbb{Q}_0[\mathcal{Y}_t] := \prod_{\tau=0}^{t-1} q_0(y_{\tau+1})$, $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t] := \prod_{\tau=0}^{t-1} \mathbb{T}_K(y_{\tau+1}|x_{\tau+1})$, and $\Phi_0 := t\phi_0$, we obtain the following decomposition of K:

$$K[\mathcal{X}_t, \mathcal{Y}_t] = \Phi_0 + \ln \frac{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]}{\mathbb{Q}_0[\mathcal{Y}_t]}, \quad (11)$$

where we use eq. (1). By defining $K_{\max}[\mathcal{Y}_t] := \sum_{\tau=0}^{t-1} k_{max}(y_{\tau+1}) = \Phi_0 - \ln \mathbb{Q}_0[\mathcal{Y}_t]$, $K[\mathcal{X}_t, \mathcal{Y}_t]$ can also be described as

$$K[\mathcal{X}_t, \mathcal{Y}_t] = K_{\max}[\mathcal{Y}_t] + \ln \mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]. \quad (12)$$

With eqs. (10) and (11), for the bet-hedging problem, we obtain a decomposition of the fitness

$$\Psi^b[\mathcal{Y}_t] = \Phi_0 - \ln \mathbb{Q}_0[\mathcal{Y}_t] - \ln \frac{\mathbb{P}_B^b[\mathcal{X}_t|\mathcal{Y}_t]}{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]}. \quad (13)$$

For a given environmental statistics $\mathbb{Q}[\mathcal{Y}_t]$, the fitness is represented by the ratio of the time-forward and time-backward path probabilities as

$$\Psi^b[\mathcal{Y}_t] = \Psi_0[\mathcal{Y}_t] - \ln \frac{\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t]}{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]}, \quad (14)$$

where $\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t] := \mathbb{P}_B^b[\mathcal{X}_t|\mathcal{Y}_t]\mathbb{Q}[\mathcal{Y}_t]$ is the time-backward joint probability of the phenotypic and environmental histories $\mathcal{X}_t$ and $\mathcal{Y}_t$. Here we also define

$$\Psi_0[\mathcal{Y}_t] := \Phi_0 + \ln \frac{\mathbb{Q}[\mathcal{Y}_t]}{\mathbb{Q}_0[\mathcal{Y}_t]} = K_{\max}[\mathcal{Y}_t] + \ln \mathbb{Q}[\mathcal{Y}_t]. \quad (15)$$

If the organisms can perfectly foresee that the environmental state at time τ becomes y_τ and if they can choose the phenotype that allocates all metabolic resource to the environmental state y_τ, the maximum replication rate at time τ reaches $e^{k_{\max}(y_\tau)}$. Therefore, $K_{\max}[\mathcal{Y}_t]$ is interpreted as the maximum replication over an environmental path $\mathcal{Y}_t$ that can be achieved only when the organisms perfectly foresee what kind of environmental history will be realized in advance. In contrast, $\ln \mathbb{Q}[\mathcal{Y}_t]$ is the entropic loss of fitness due to the lack of knowledge of which environmental history will be realized[32,36]. Therefore, $\Psi_0[\mathcal{Y}_t]$ is the maximum replication when the organisms cannot know which environmental state will be realized but know the statistics of the future environmental state. The relevance of this interpretation and the biological meaning of some quantities such as Φ_0 and $\mathbb{Q}_0[\mathcal{Y}_t]$ are explicitly shown by using the FRs derived in the following sections.

For the case with the sensing signal, we can similarly obtain a decomposition of $\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]$ as

$$\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] = \Psi_0[\mathcal{Y}_t] - \ln \frac{\mathbb{P}_B^s[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t]}{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t|\mathcal{Y}_t]} \tag{16}$$

where $\mathbb{P}_B^s[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t] := \mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t]\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]$ is the time-backward joint path probability among $\mathcal{X}_t$, $\mathcal{Y}_t$, and $\mathcal{Z}_t$. It should be noted that $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t|\mathcal{Y}_t]$ is not a joint path probability because of the circular noncausal dependency among $\mathcal{X}_t$, $\mathcal{Y}_t$, and $\mathcal{Z}_t$. Finally, by using the decomposition in eq. (11), the time-backward conditional path probabilities, eq. (10) and eq. (9), are reduced to

$$\mathbb{P}_B^b[\mathcal{X}_t|\mathcal{Y}_t] = \frac{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]]\mathbb{P}_F[\mathcal{X}_t]}{\mathbb{P}_{K,F}[\mathcal{Y}_t]}, \tag{17}$$

$$\mathbb{P}_B^s[\mathcal{X}_t|\mathcal{Y}_t, \mathcal{Z}_t] = \frac{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]}{\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]}, \tag{18}$$

where the normalization factors are

$$\mathbb{P}_{K,F}[\mathcal{Y}_t] := \sum_{\mathcal{X}_t} \mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]]\mathbb{P}_F[\mathcal{X}_t],$$

$$\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t] := \sum_{\mathcal{X}_t} \mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t].$$

Because $\mathbb{P}_F[\mathcal{X}_t]$ is the probability to intrinsically generate the repertoire of phenotypic histories in the population and $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]$ is the probability that an organism with a phenotypic history $\mathcal{X}_t$ allocates resources to each history of environment $\mathcal{Y}_t$, the normalization factor $\mathbb{P}_{K,F}[\mathcal{Y}_t]$ can be interpreted as the marginal resource allocation to the environmental history $\mathcal{Y}_t$ at the population level. $\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]$ can similarly be interpreted as the population-level resource allocation to $\mathcal{Y}_t$ when signal history $\mathcal{Z}_t$ is received. By using FRs in the next section, we clarify that $\mathbb{P}_F[\mathcal{X}_t]$ and $\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]$ also have meaning as the conjugate environment under which the given strategy $\{\mathbb{T}_F, \mathbb{T}_K\}$ becomes optimal.

IV. CAUSAL FRS FOR BET-HEDGING STRATEGY

By rearranging the decomposition of $\Psi^b[\mathcal{Y}_t]$ in eq. (14), we can immediately obtain a detailed causal FR for fitness difference $\Psi_0[\mathcal{Y}_t] - \Psi^b[\mathcal{Y}_t]$ as

$$e^{-(\Psi_0[\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t])} = \frac{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]}{\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t]} = \frac{\mathbb{P}_{K,F}[\mathcal{Y}_t]}{\mathbb{Q}[\mathcal{Y}_t]}, \quad (19)$$

where we use eq. (17) to obtain the last equality. The first equality means that the fitness difference is the log ratio of time-forward and time-backward path probabilities $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]$ and $\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t]$. $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]$ is the time-forward probability of observing an organism that takes phenotypic history $\mathcal{X}_t$ and then allocates $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]$ of metabolic resources to $\mathcal{Y}_t$ a priori to selection by conducting time-forward tracking of the histories.

$\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t]$ is the time-backward probability of observing the realization of environmental history $\mathcal{Y}_t$ and the time-backward phenotypic history $\mathcal{X}_t$ a posteriori to selection by conducting time-backward tracking of the histories. The second equality also indicates that the fitness difference is the log ratio of the percentage of resource allocated to environmental history $\mathcal{Y}_t$ at the population level and the probability of observing environmental history $\mathcal{Y}_t$.

By averaging eq. (19) with respect to $\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t]$ or $\mathbb{Q}[\mathcal{Y}_t]$, we can derive an integral FR as

$$\left\langle e^{-(\Psi_0[\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t])} \right\rangle_{\mathbb{Q}[\mathcal{Y}_t]} = 1. \tag{20}$$

If we average eq. (19) after taking the logarithm of both sides, we obtain an average FR as

$$\left\langle \Psi^b \right\rangle_{\mathbb{Q}} = \left\langle \Psi_0 \right\rangle_{\mathbb{Q}} - \mathcal{D}_{\text{loss}}^b, \tag{21}$$

where $\left\langle \Psi^b \right\rangle_{\mathbb{Q}} := \left\langle \Psi^b[\mathcal{Y}_t] \right\rangle_{\mathbb{Q}[\mathcal{Y}_t]}$, $\left\langle \Psi_0 \right\rangle_{\mathbb{Q}} := \left\langle \Psi_0[\mathcal{Y}_t] \right\rangle_{\mathbb{Q}[\mathcal{Y}_t]}$, and

$$\mathcal{D}_{\text{loss}}^b := \mathcal{D}\left[\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t] \middle|\middle| \mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]\right] \tag{22}$$

$$= \mathcal{D}\left[\mathbb{Q}[\mathcal{Y}_t] \middle|\middle| \mathbb{P}_{K,F}[\mathcal{Y}_t]\right] \tag{23}$$

Because of the non-negativity of the relative entropy $\mathcal{D}_{\text{loss}}^b$, we can easily see that $\left\langle \Psi_0 \right\rangle_{\mathbb{Q}}$ is an upper bound of the average fitness $\left\langle \Psi^b \right\rangle_{\mathbb{Q}}$ of a bet-hedging strategy:

$$\left\langle \Psi_0 \right\rangle_{\mathbb{Q}} \geq \max_{\{\mathbb{T}_F, \mathbb{T}_K\}} \left\langle \Psi^b \right\rangle_{\mathbb{Q}}, \tag{24}$$

where we use the fact that $\Psi_0[\mathcal{Y}_t]$ is dependent neither on the phenotype switching strategy $\mathbb{T}_F$ nor on the metabolic allocation strategy $\mathbb{T}_K$. These FRs are basically the same as those we derived in our previous work by using a variational approach. It should be also noted that the FRs derived by Mustonen and Lässig[24] are different from ours because their relations are based on a model describing the dynamics of an ensemble of populations whereas ours is one describing the dynamics of a population.

A. Biological meaning of Ψ_0 and Φ_0

From eq. (15), the upper bound of the average fitness, $\left\langle \Psi_0 \right\rangle_{\mathbb{Q}}$ admits two different representations:

$$\left\langle \Psi_0 \right\rangle_{\mathbb{Q}} = \left\langle K_{\max} \right\rangle_{\mathbb{Q}} - \mathcal{S}[\mathbb{Q}] = \Phi_0 + \mathcal{D}\left[\mathbb{Q} \middle|\middle| \mathbb{Q}_0\right], \tag{25}$$

where $\mathcal{S}[\mathbb{Q}] := -\left\langle \ln \mathbb{Q}[\mathcal{Y}_t] \right\rangle_{\mathbb{Q}[\mathcal{Y}_t]}$ is the entropy of $\mathbb{Q}[\mathcal{Y}_t]$ and $\left\langle K_{\max} \right\rangle_{\mathbb{Q}}$ is the average fitness under environmental statistics $\mathbb{Q}[\mathcal{Y}_t]$ that is attained only when organisms have perfect knowledge of the future environment. Therefore, the first equality means that the randomness of the environment quantified by the entropy $\mathcal{S}[\mathbb{Q}]$ works as the
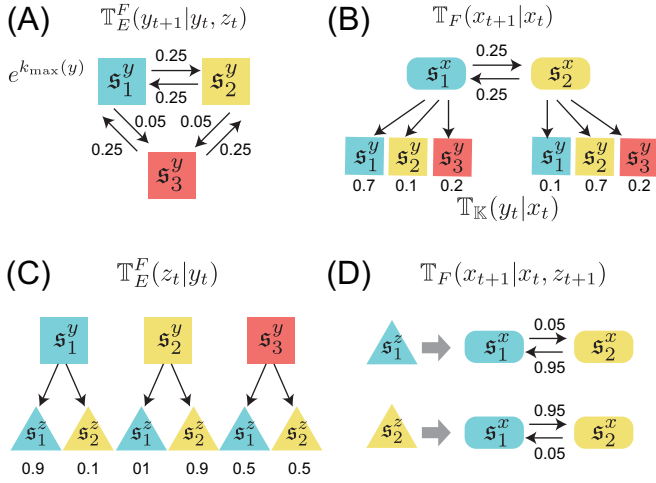
FIG. 2. Schematic representations of environmental dynamics, phenotypic switching, and sensing used for numerical verification of FRs in Figs. 3 and 5. (A) Characteristics of the environment. The environment has three states: $y_t \in \mathfrak{S}_y = \{\mathfrak{s}_1^y, \mathfrak{s}_2^y, \mathfrak{s}_3^y\}$. $\mathfrak{s}_1^y$ and $\mathfrak{s}_2^y$ are nutrient rich states that have different sources of nutrients. $\mathfrak{s}_3^y$ is a nutrient-poor state under which replication of the organisms is severely restricted. The maximum growth under each environment is $e^{k_{max}(\mathfrak{s}_1^y)} = 3.2$, $e^{k_{max}(\mathfrak{s}_2^y)} = 3.2$, and $e^{k_{max}(\mathfrak{s}_3^y)} = 0.4$. The transition rates $\mathbb{T}_E^F(y_{t+1}|y_t)$ between the states are shown on the arrows. The environmental states usually fluctuate between $\mathfrak{s}_1^y$ and $\mathfrak{s}_2^y$ but occasionally flip to $\mathfrak{s}_3^y$ with a 5% chance. (B) Characteristics of phenotypic states without sensing. Organisms have two phenotypic states: $x_t \in \mathfrak{S}_x = \{\mathfrak{s}_1^x, \mathfrak{s}_2^x\}$. The transition rates $\mathbb{T}_F(x_{t+1}|x_t)$ between the states are shown on the arrows. $\mathfrak{s}_1^x$ is the phenotypic state allocating more resources to the environmental state $\mathfrak{s}_1^y$ (70%) than $\mathfrak{s}_2^y$ (10%) whereas $\mathfrak{s}_2^x$ allocates more to $\mathfrak{s}_2^x$ (70%) than $\mathfrak{s}_1^x$ (10%). Both states allocate 20% of the resources to the starving environmental state $\mathfrak{s}_3^y$. (C) Characteristic of sensing signal. The sensing signal has two states as $z_t \in \mathfrak{S}_z = \{\mathfrak{s}_1^z, \mathfrak{s}_2^z\}$. When the environmental state is $\mathfrak{s}_1^y$ ($\mathfrak{s}_2^y$), the organisms obtain $\mathfrak{s}_1^z$ ($\mathfrak{s}_2^z$) as the sensing signal with a 90% accuracy. If the environment is in $\mathfrak{s}_3^y$, the sensing signal produces $\mathfrak{s}_1^z$ or $\mathfrak{s}_2^z$ with equal probability. (D) Characteristic of phenotypic switching with sensing. When the organism obtains the sensing signal $\mathfrak{s}_1^z$ ($\mathfrak{s}_2^z$), it switches its phenotypic state into $\mathfrak{s}_1^x$ ($\mathfrak{s}_2^x$) 95% of the time.

inevitable loss of fitness due to the lack of knowledge of which environmental history will be realized in the future. If the environment fluctuates more unpredictably, we have a higher $\mathcal{S}[\mathbb{Q}]$ and a lower upper bound of the average fitness. Note that these properties of $\langle \Psi_0 \rangle_{\mathbb{Q}}$ have been pointed out previously and repeatedly[28,32,36].

The meaning of Φ_0 and $\mathbb{Q}_0[\mathcal{Y}_t]$ in the second equality becomes explicit by considering the minimization of $\langle \Psi_0 \rangle_{\mathbb{Q}}$ with respect to $\mathbb{Q}[\mathcal{Y}_t]$ as follows:

$$\Phi_0 = \min_{\mathbb{Q}} \langle \Psi_0[\mathcal{Y}_t] \rangle_{\mathbb{Q}[\mathcal{Y}_t]} \geq \min_{\mathbb{Q}} \max_{\{\mathbb{T}_F, \mathbb{T}_K\}} \langle \Psi^b[\mathcal{Y}_t] \rangle_{\mathbb{Q}[\mathcal{Y}_t]},$$
(26)

where we use

$$\min_{\mathbb{Q}} \langle \Psi_0[\mathcal{Y}_t] \rangle_{\mathbb{Q}[\mathcal{Y}_t]} = \Phi_0 + \min_{\mathbb{Q}} \mathcal{D}[\mathbb{Q}||\mathbb{Q}_0] = \Phi_0.$$
(27)

This relation indicates that Φ_0 is the minimum of the maximum average fitness and that $\mathbb{Q}_0[\mathcal{Y}_t]$ is the worst environment for the organisms under which the maximum average fitness is minimized. This min-max characterization of Φ_0 and $\mathbb{Q}_0[\mathcal{Y}_t]$ has been clarified in the context of game theory with a matrix formulation[34].

B. Verification of FRs for fitness

Equations (19–21) indicate that, under quite general situations, the fitness difference $\Psi_0[\mathcal{Y}_t] - \Psi^b[\mathcal{Y}_t]$ satisfies the FRs[9] as the entropy production does in stochastic thermodynamics[46,47]. To demonstrate the relations, we consider an organism with two phenotypic states growing in a Markovian environment with three states as depicted in Fig. 2 (A) and (B). The three environmental states, $\mathfrak{s}_1^y$, $\mathfrak{s}_2^y$, and $\mathfrak{s}_3^y$, describe nutrient A rich, nutrient B rich, and nutrient-poor conditions, respectively (Fig 2(A)). The two phenotypic states, $\mathfrak{s}_1^x$ and $\mathfrak{s}_2^x$, employ strategies to allocate 70% of the metabolic resources to $\mathfrak{s}_1^y$ and $\mathfrak{s}_2^y$, respectively. Both states allocate 10% resources to the rest of two states (see Appendix B for more details). This setting abstractly and simply represents the fact that organisms generally have much less phenotypic and sensing states than possible environmental states because of the limited physical complexity of the organisms[35].

Figures 3 (A) and (B) show the population dynamics of the organisms with the phenotypic states $\mathfrak{s}_1^x$ and $\mathfrak{s}_2^x$ under two different realizations of the environmental history alongside $\Psi^b[\mathcal{Y}_t]$ and $\Psi_0[\mathcal{Y}_t]$. Depending on the actual realization of the environment, the relative relations among $\mathcal{N}_t^{\mathcal{Y}}(\mathfrak{s}_1^x)$, $\mathcal{N}_t^{\mathcal{Y}}(\mathfrak{s}_2^x)$, $e^{\Psi^b[\mathcal{Y}_t]} = \mathcal{N}_t^{\mathcal{Y}}$, and $e^{\Psi_0[\mathcal{Y}_t]}$ change over time stochastically. In Fig 3 (A), $e^{\Psi_0[\mathcal{Y}_t]}$ is mostly greater than $\mathcal{N}_t^{\mathcal{Y}} = e^{\Psi_t^b[\mathcal{Y}_t]}$, which reflects the average relation $\langle \Psi_0 \rangle_{\mathbb{Q}} \geq \langle \Psi_t^b \rangle_{\mathbb{Q}}$. On the contrary, $\mathcal{N}_t^{\mathcal{Y}}$ frequently becomes greater than $e^{\Psi_0[\mathcal{Y}_t]}$ in Fig 3 (B). Figures 3 (C) and (D) show that the environmental fluctuation induces a large fluctuation in both $\Psi^b[\mathcal{Y}_t]$ and $\Psi_0[\mathcal{Y}_t]$. As shown in Fig 3 (E), although most environmental fluctuations result in positive fitness differences, that is, $\Psi_0[\mathcal{Y}_t] - \Psi^b[\mathcal{Y}_t] > 0$, rare environmental fluctuations lead to a negative fitness difference in a finite time interval, meaning that the fitness of the suboptimal strategy $\Psi^b[\mathcal{Y}_t]$ outperforms the average upper bound $\Psi_0[\mathcal{Y}_t]$. This is analogue to the reversed heat flow in a small thermal system[55]. Such rare events are balanced to satisfy the integral FR in eq. (20) as verified numerically in Fig 3 (F).
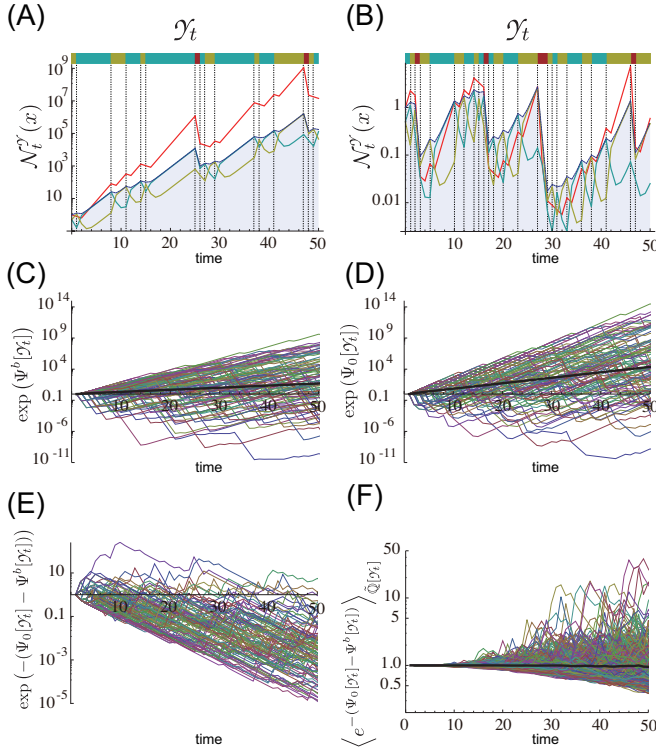
FIG. 3. Numerical simulation of the population dynamics without sensing defined in Fig. 2. (A, B) Two sample histories of environment $\mathcal{Y}_t$ and the dynamics of $\mathcal{N}_t^{\mathcal{Y}}(\mathfrak{s}_1^x)$, $\mathcal{N}_t^{\mathcal{Y}}(\mathfrak{s}_2^x)$, $\mathcal{N}_t^{\mathcal{Y}} = e^{\Psi^b[\mathcal{Y}_t]}$, and $e^{\Psi_0[\mathcal{Y}_t]}$ under $\mathcal{Y}_t$ obtained by solving eq. (2). We set $\mathcal{N}_0^{\mathcal{Y}} = 1$. The color bars on the graphs represent the state of the environment at each time point. The correspondence between the colors and environmental states is as shown in Fig. 2 (A). Cyan and yellow lines represent $\mathcal{N}_t^{\mathcal{Y}}(\mathfrak{s}_1^x)$ and $\mathcal{N}_t^{\mathcal{Y}}(\mathfrak{s}_2^x)$, respectively. The dashed blue line with filled grey style is $\mathcal{N}_t^{\mathcal{Y}}$. The red line is $\Psi_0[\mathcal{Y}]$. (C) Stochastic behaviors of population fitness $\Psi^b[\mathcal{Y}_t]$ for 100 independent samples of the environmental histories. Each colored thin line represents $e^{\Psi^b[\mathcal{Y}_t]} = \mathcal{N}_t^{\mathcal{Y}}/\mathcal{N}_0^{\mathcal{Y}}$ for each realization of environmental history. The thick black line is $e^{\langle \Psi_b[\mathcal{Y}_t]\rangle_{\mathbb{Q}[\mathcal{Y}_t]}}$. (D) Stochastic behaviors of $\Psi_0[\mathcal{Y}_t]$ for the same 100 samples of environmental histories as those in (C). Each colored thin line represents $e^{\Psi_0[\mathcal{Y}_t]}$ for each realization of the same environmental history as in (C). (E) Stochastic behaviors of $e^{-(\Psi_0[\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t])}$ for the same 100 samples of environmental histories as those in (C). Each colored thin line represents $e^{-(\Psi_0[\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t])}$ for each realization of the same environmental history as in (C). (F) $\left\langle e^{-(\Psi_0[\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t])}\right\rangle_{\mathbb{Q}[\mathcal{Y}_t]}$ calculated empirically by the numerical simulations of sample paths of $e^{-(\Psi_0[\mathcal{Y}_t]-\Psi^b[\mathcal{Y}_t])}$. The thick line is the average of 10^8 independent samples of environmental histories. To illustrate the fluctuation, the 10^8 histories are dissected into 100 groups of histories, each contains 10^5 histories. Each thin line is obtained as the average of 10^5 histories in each group.

V. BIOLOGICAL MEANING OF THE FITNESS FRS

In the original detailed FR over paths[47], the entropy production is the log ratio of the path probability of a system's trajectory and its time reversal. The average entropy production attains its minimum 0 only when the time reversibility of the system holds in the sense that the probabilities of observing the time-forward and the time-reversed trajectories are equal. Thus, the FRs are related to the extent of the time reversibility of the system. In contrast, the detailed FRs for the fitness difference (eq. (19)) is the log ratio between the probability $\mathbb{Q}[\mathcal{Y}_t]$ of observing the environmental history and the percentage $\mathbb{P}_{K,F}[\mathcal{Y}_t]$ of the marginal resource allocation to $\mathcal{Y}_t$, or that between the time-backward path probability $\mathbb{P}_B^b[\mathcal{X}_t, \mathcal{Y}_t]$ and time-forward path probability $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t]$. By investigating the FRs, we clarify a dualistic structure, a conjugacy of these quantities, and the time-reversal condition for the equality attained in eq. (24).

A. Dualistic relation between strategy and environment

The average FR in eq. (21) implies that the maximization of $\left\langle \Psi^b\right\rangle_{\mathbb{Q}}$ with respect to the strategies is dual to the minimization of the relative entropy $\mathcal{D}_{\text{loss}}^b$ as follows:

$$\{\mathbb{T}_F^\dagger, \mathbb{T}_K^\dagger\} := \arg \max_{\{\mathbb{T}_F, \mathbb{T}_K\}} \left\langle \Psi^b\right\rangle_{\mathbb{Q}} = \arg \min_{\{\mathbb{T}_F, \mathbb{T}_K\}} \mathcal{D}_{\text{loss}}^b, \tag{28}$$

because $\Psi_0[\mathcal{Y}_t]$ is independent of $\mathbb{P}_F$ and $\mathbb{P}_K$. This duality indicates that maximizing the average fitness by choosing the best strategy is equivalent to the organisms to implicitly learn and prepare for the environmental statistics $\mathbb{Q}[\mathcal{Y}_t]$ so that the marginal resource allocation $\mathbb{P}_{K,F}[\mathcal{Y}_t]$ to $\mathcal{Y}_t$ becomes close to $\mathbb{Q}[\mathcal{Y}_t]$ because $\mathcal{D}_{\text{loss}}^b = \mathcal{D}[\mathbb{Q}[\mathcal{Y}_t]||\mathbb{P}_{K,F}[\mathcal{Y}_t]]$. The upper bound of the average fitness in eq. (24) is achieved as $\Psi^b[\mathcal{Y}_t] = \Psi_0[\mathcal{Y}_t]$ if and only if $\{\mathbb{T}_F^\dagger, \mathbb{T}_K^\dagger\}$ satisfies $\mathbb{Q}[\mathcal{Y}_t] = \mathbb{P}_{K,F}^\dagger[\mathcal{Y}_t]$ where $\mathbb{P}_{K,F}^\dagger[\mathcal{Y}_t] := \sum_{\mathcal{X}_t} \mathbb{T}_K^\dagger[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{T}_F^\dagger[\mathcal{X}_t]$, meaning that the environmental statistics and the marginal resource allocation match perfectly.

B. Meaning of $\mathbb{P}_{F,K}[\mathcal{Y}_t]$ as conjugate environment

For a given environment, the strategy that achieves the bound may not always exist. In contrast, for a given pair of strategies $\{\mathbb{T}_F, \mathbb{T}_K\}$, there always exists the environment $\mathbb{Q}^\dagger[\mathcal{Y}_t]$ under which the pair achieves the bound

$$\left\langle \Psi_0\right\rangle_{\mathbb{Q}^\dagger} = \left\langle \Psi^b[\mathcal{Y}_t]\right\rangle_{\mathbb{Q}^\dagger[\mathcal{Y}_t]} \geq \max_{\{\mathbb{T}_F', \mathbb{T}_K'\}} \left\langle \Psi^{b'}[\mathcal{Y}_t]\right\rangle_{\mathbb{Q}^\dagger[\mathcal{Y}_t]}, \tag{29}$$

and

$$\{\mathbb{T}_F, \mathbb{T}_K\} = \arg \max_{\{\mathbb{T}_F', \mathbb{T}_K'\}} \left\langle \Psi^{b'}[\mathcal{Y}_t]\right\rangle_{\mathbb{Q}^\dagger[\mathcal{Y}_t]}, \tag{30}$$

is satisfied. Because of the duality shown in eq. (28), $\mathbb{Q}^\dagger[\mathcal{Y}_t]$ is explicitly obtained as $\mathbb{Q}^\dagger[\mathcal{Y}_t] = \mathbb{P}_{F,K}[\mathcal{Y}_t]$. Therefore, $\mathbb{P}_{K,F}[\mathcal{Y}_t]$ can also be regarded as the conjugate environment to the strategy $\{\mathbb{T}_F, \mathbb{T}_K\}$ or $\{\mathbb{P}_F, \mathbb{P}_K\}$ under which they are optimal. Therefore, the fitness difference in eq. (19) is the log ratio of the actual environmental statistics $\mathbb{Q}[\mathcal{Y}_t]$ and that of the conjugate environment $\mathbb{Q}^\dagger[\mathcal{Y}_t] = \mathbb{P}_{K,F}[\mathcal{Y}_t]$ of the given strategy.

$\Psi^b[\mathcal{Y}_t]$ is bounded by $\Psi_0[\mathcal{Y}_t]$ on average, and therefore, the optimal strategy that attains $\Psi_0[\mathcal{Y}_t]$ cannot be invaded by any other strategy if we consider an infinitely large population and the asymptotic dynamics. Thus, the optimal strategy is a version of ESS in a fluctuating environment. Within a finite time interval, however, $\Psi^b[\mathcal{Y}_t]$ becomes greater than $\Psi_0[\mathcal{Y}_t]$ under certain realizations of environment $\mathcal{Y}_t$ that satisfy $\mathbb{Q}[\mathcal{Y}_t] < \mathbb{Q}^\dagger[\mathcal{Y}_t]$. In stochastic thermodynamics, such realizations correspond to the temporal reversed heat flow in a small thermal system[55]. In biology, they temporally violate ESS because a suboptimal strategy $\{\mathbb{T}_F, \mathbb{P}_K\}$ under $\mathbb{Q}$ can outperform the optimal strategy or near optimal strategies with the aid of the environmental fluctuation within a finite time interval. The integral FR in eq. (20) tells us that the violation of ESS can always occur with a small but finite probability in a finite time interval. If the population size of the optimal strategy is finite, such violation can leads to extinction of the optimal population with a finite probability[48]. Moreover, the detailed FR in eq. (19) implies that a greater violation can occur under a realization of environmental history $\mathcal{Y}_t$ if $\mathbb{Q}[\mathcal{Y}_t]$ is small but $\mathbb{Q}^\dagger[\mathcal{Y}_t]$ is large. This fact can be intuitively understood as follows: the greater violation is induced by the environmental history $\mathcal{Y}_t$ that rarely occurs in the actual environmental statistics $\mathbb{Q}[\mathcal{Y}_t]$ but is adaptive and advantageous for the given strategy. A crucial fact is that this intuitive understanding is supported by a quantitative relation as in eq. (19). Furthermore, the detailed FR suggests that greater violation can occur for specific suboptimal strategies than for others if $\mathbb{Q}[\mathcal{Y}_t]$ contains many rare environmental histories. In contrast, if $\mathbb{Q}[\mathcal{Y}_t]$ is perfectly random as $\mathbb{Q}[\mathcal{Y}_t] = $ const., the extent of the violation is limited and all the suboptimal strategies have only an even chance of violation. This implies that structured environmental fluctuation promotes impactful violation by some specific suboptimal strategies.

C. Time reversibility and optimality

If $\Psi^b[\mathcal{Y}_t] = \Psi_0[\mathcal{Y}_t]$ holds, moreover, the time-backward probability $\mathbb{P}_B^{b,\dagger}[\mathcal{X}_t, \mathcal{Y}_t]$ and the time-forward path probability $\mathbb{P}_K^\dagger[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F^\dagger[\mathcal{X}_t]$ become equal:

$$\mathbb{P}_B^{b,\dagger}[\mathcal{X}_t, \mathcal{Y}_t] = \mathbb{P}_B^{b,\dagger}[\mathcal{X}_t|\mathcal{Y}_t]\mathbb{Q}[\mathcal{Y}_t] = \mathbb{P}_K^\dagger[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F^\dagger[\mathcal{X}_t] \quad (31)$$

where we use the first equality in eq. (23). Marginalization of this equality with respect to $\mathcal{Y}_t$ indicates that the marginalized time-backward path probability $\mathbb{P}_B^b[\mathcal{X}_t]$ sat-

isfies the consistency condition $\mathbb{P}_B^{b,\dagger}[\mathcal{X}_t] = \mathbb{P}_F^\dagger[\mathcal{X}_t]$ as shown previously[9]. Thus, the optimal strategies to achieve the bound have time reversibility in the sense that the ensembles of the time-forward and time-backward phenotypic histories are indistinguishable without knowing the actual environmental history that the population experienced. While the definition of time reversibility is different, this result is closely related to the fact that the average entropy production attains 0 when the time reversibility is satisfied. Biologically, this result is quite important because we can evaluate the optimality of an organism in a changing environment by just observing its phenotypic dynamics without directly measuring the environment that the organisms experience. It should be noted, however, that $\mathbb{P}_B^{b,\dagger}[\mathcal{X}_t] = \mathbb{P}_F^\dagger[\mathcal{X}_t]$ is a necessary but not sufficient condition for $\Psi^b[\mathcal{Y}_t] = \Psi_0[\mathcal{Y}_t]$.

D. Achievability of the fitness upper bound

The duality eq. (28) indicates that the achievability of the bound in eq. (24) depends on the actual property of the environmental statistics and biological constraints on the selectable strategies. If the environment is a time-homogeneous Markov process as $\mathbb{Q}[\mathcal{Y}_t] = \prod_{\tau=0}^{t-1} \mathbb{T}_E(y_{\tau+1}|y_\tau)q(y_0)$ and if the number of possible phenotypic states is the same as that of the environmental states, that is, $\#\mathfrak{S}_\mathcal{X} = \#\mathfrak{S}_\mathcal{Y}$, the bound can be achieved by a pair of strategies:

$$\{\mathbb{T}_F^\dagger(x'|x), \mathbb{T}_K^\dagger(y|x)\} = \{\mathbb{T}_E(x'|x), \delta_{x,y}\} \quad (32)$$

where $\mathbb{T}_E(x'|x) := \mathbb{T}_E(y_{\tau+1}|y_\tau)|_{y_{\tau+1}=x',y_\tau=x}$ and $\delta_{x,y}$ is the Kronecker delta. This pair is equivalent to the optimal bet-hedging strategy in Kelly's horse race gambling because $\mathbb{T}_K^\dagger(y|x) = \delta_{x,y}$ means that organisms can survive only when their phenotypic state matches the current environmental state; they die out otherwise. To be specific, we call $\mathbb{T}_K(y|x) = \delta_{x,y}$ Kelly's strategy of metabolic allocation. Under biologically realistic constraints, however, Kelly's strategy cannot be the optimal strategy because the possible phenotypic states are usually much fewer than those of the environment as in Fig 2 and 3. Allocating all resources to a specific environment easily leads to extinction if the phenotypic states cannot cover all the possible environmental states. With a limited capacity in possible phenotypic states, $\mathcal{D}_{loss}^b = 0$ can be attained only if the environmental fluctuation has a hidden structure the dimensionality of which is sufficiently low. This is manifested by the fact that the conjugate environment $\mathbb{Q}^\dagger$ is the environment with a hidden dynamics $\mathbb{P}_F[\mathcal{X}_t']$ that generates the actual environmental history as $\mathbb{Q}^\dagger[\mathcal{Y}_t] = \sum_{\mathcal{X}_t'} \mathbb{P}_K[\mathcal{Y}_t|\mathcal{X}_t']\mathbb{P}_F[\mathcal{X}_t']$. Because such a low dimensional structure may not always exist, however,

$$\mathcal{D}_{loss}^{b\dagger} := \min_{\{\mathbb{T}_F, \mathbb{T}_K\}} \mathcal{D}_{loss}^b$$

is generally not zero but finite and positive under a biological constraint that the possible phenotypic states are

fewer than environmental ones. Therefore, $\langle\Psi_0\rangle_{\mathbb{Q}}$ is generally attained only by a Darwinian demon that cannot perfectly foresee the future environment but has sufficient capacity in its phenotypic properties to perfectly learn and prepare for the environmental fluctuation $\mathbb{Q}[\mathcal{Y}_t]$.

Even when the bound is not achieved, $\mathcal{D}_{\mathrm{loss}}^{b\dagger}$ has explicit meaning as the component in the environmental fluctuation that cannot be learned or represented by the dynamics of the cell's strategy. For example, when $\mathbb{T}_F$ is memoryless and $\mathbb{Q}[\mathcal{Y}_t]$ is a stationary Markov process with the stationary probability $q(y)$, that is, $\mathbb{Q}[\mathcal{Y}_t] = \prod_{\tau=0}^{t-1}\mathbb{T}_E^F(y_{\tau+1}|y_\tau)q(y_0)$, then

$$
\begin{aligned}
\mathcal{D}_{\mathrm{loss}}^{b\dagger} &= \min_{\{\mathbb{T}_F,\mathbb{T}_K\}}\mathcal{D}_{\mathrm{loss}}^b \\
&= \min_{\{\mathbb{T}_F,\mathbb{T}_K\}}\left\langle \ln\frac{\mathbb{T}_E^F(y'|y)q(y)}{\langle\mathbb{T}_K(y'|x)\rangle_{\mathbb{T}_F(x)}}\right\rangle_{\mathbb{T}_E^F(y'|y)q(y)} \\
&= \min_{\{\mathbb{T}_F,\mathbb{T}_K\}}\left[\mathcal{I}^{x_t;x_{t+1}} + \mathcal{D}\left[q(y)\Big|\Big|\langle\mathbb{T}_K(y|x)\rangle_{\mathbb{T}_F(x)}\right]\right] \\
&= \mathcal{I}^{x_t;x_{t+1}},
\end{aligned}
$$

where $\mathcal{I}^{x_t;x_{t+1}} := \mathcal{D}\left[\mathbb{T}_E^F(y'|y)q(y))\Big|\Big|q(y)q(y')\right]$ is the mutual information that measures the correlation of environmental states between consecutive time points, which cannot be learned or mimicd by the memoryless phenotypic switching.

VI. FRS WITH SIGNAL SENSING

Next, we consider the case in which the organisms can exploit the information obtained from the sensing signal. The fitness decomposition in eq. (16) can be rearranged as

$$
\Psi^s[\mathcal{Y}_t,\mathcal{Z}_t] = \Psi_0[\mathcal{Y}_t] + i[\mathcal{Y}_t;\mathcal{Z}_t] - \ln\frac{\mathbb{P}_B^s[\mathcal{X}_t,\mathcal{Y}_t,\mathcal{Z}_t]}{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]}
\tag{33}
$$

$$
= \Psi_0[\mathcal{Y}_t] + i[\mathcal{Y}_t;\mathcal{Z}_t] - \ln\frac{\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]}{\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]},
\tag{34}
$$

where $i[\mathcal{Y}_t;\mathcal{Z}_t] := \ln\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]/\mathbb{Q}[\mathcal{Y}_t]\mathbb{Q}[\mathcal{Z}_t]$ is the bare mutual information between $\mathcal{Y}_t$ and $\mathcal{Z}_t$ and we use eq. (18) to derive the last equality. From this, we can similarly obtain detailed, integral, and average FRs with information as follows:

$$
e^{-(\Psi_0[\mathcal{Y}_t]+i[\mathcal{Y}_t;\mathcal{Z}_t]-\Psi^s[\mathcal{Y}_t,\mathcal{Z}_t])} = \frac{\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]}{\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]},
\tag{35}
$$

$$
\left\langle e^{-(\Psi_0[\mathcal{Y}_t]+i[\mathcal{Y}_t;\mathcal{Z}_t]-\Psi^s[\mathcal{Y}_t,\mathcal{Z}_t])}\right\rangle_{\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]} = 1,
\tag{36}
$$

and

$$
\langle\Psi^s\rangle_{\mathbb{Q}} = \langle\Psi_0\rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Y};\mathcal{Z}} - \mathcal{D}_{\mathrm{loss}}^s,
\tag{37}
$$

where $\mathcal{I}^{\mathcal{Y};\mathcal{Z}} := \langle i[\mathcal{Y}_t;\mathcal{Z}_t]\rangle_{\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]}$ is the path-wise mutual information between $\mathcal{Y}_t$ and $\mathcal{Z}_t$, and

$$
\begin{aligned}
\mathcal{D}_{\mathrm{loss}}^s &:= \mathcal{D}\left[\mathbb{P}_B[\mathcal{X}_t,\mathcal{Y}_t,\mathcal{Z}_t]\Big|\Big|\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]\right] \\
&= \mathcal{D}\left[\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]\Big|\Big|\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]\right].
\end{aligned}
\tag{38}
$$

The way information terms involved in Eqs. (35)-(37) is the same as the way those appearing in the Sagawa-Ueda relations, where the Maxwell demon and feedback regulation are involved[47]. Because of the non-negativity of $\mathcal{D}_{\mathrm{loss}}^s$, $\langle\Psi^s\rangle_{\mathbb{Q}}$ is upper bounded by $\langle\Psi_0\rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Y};\mathcal{Z}}$:

$$
\langle\Psi^s\rangle_{\mathbb{Q}} \le \langle\Psi_0\rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Y};\mathcal{Z}} = \langle K_{\max}\rangle_{\mathbb{Q}} - \mathcal{S}_{\mathcal{Y}|\mathcal{Z}}[\mathbb{Q}],
\tag{39}
$$

where $\mathcal{S}_{\mathcal{Y}|\mathcal{Z}}[\mathbb{Q}] := -\langle\ln\mathbb{Q}[\mathcal{Y}_t|\mathcal{Z}_t]\rangle_{\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]}$ is the conditional entropy of $\mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]$. If the history of signal $\mathcal{Z}_t$ has perfect information on the history of $\mathcal{Y}_t$, the upper bound reaches $\langle K_{\max}\rangle_{\mathbb{Q}}$. As in the bet-hediging situation, maximization of the average fitness $\langle\Psi^s\rangle_{\mathbb{Q}}$ with sensing is also dual to the minimization of the relative entropy $\mathcal{D}_{\mathrm{loss}}^s$:

$$
\{\mathbb{T}_F^*,\mathbb{T}_K^*\} := \arg\max_{\{\mathbb{T}_F,\mathbb{T}_K\}}\langle\Psi^s\rangle_{\mathbb{Q}} = \arg\min_{\{\mathbb{T}_F,\mathbb{T}_K\}}\mathcal{D}_{\mathrm{loss}}^s,
\tag{40}
$$

where we use $*$ to denote the optimal $\mathbb{T}_F$ and $\mathbb{T}_K$ with sensing. The duality indicates that $\langle\Psi^s\rangle_{\mathbb{Q}}$ achieves the bound $\langle\Psi_0\rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Y};\mathcal{Z}}$ only when $\mathbb{P}_{K,F}^*[\mathcal{Y}_t|\mathcal{Z}_t] = \mathbb{Q}[\mathcal{Y}_t|\mathcal{Z}_t]$ holds. As in the bet-heding, if $\mathbb{P}_{K,F}^*[\mathcal{Y}_t|\mathcal{Z}_t] = \mathbb{Q}[\mathcal{Y}_t|\mathcal{Z}_t]$ holds, the time backward path probability $\mathbb{P}_B^{s,*}[\mathcal{X}_t,\mathcal{Y}_t,\mathcal{Z}_t]$ equals the time-forward path probability:

$$
\mathbb{P}_B^{s,*}[\mathcal{X}_t,\mathcal{Y}_t,\mathcal{Z}_t] = \mathbb{P}_K^*[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F^*[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t].
\tag{41}
$$

Marginalization of this equation leads to the consistency condition $\mathbb{P}_F^*[\mathcal{X}_t||\mathcal{Z}_t] = \mathbb{P}_B^{s,*}[\mathcal{X}_t|\mathcal{Z}_t](:= \sum_{\mathcal{Y}_t}\mathbb{P}_B^{s,*}[\mathcal{X}_t|,\mathcal{Y}_t,\mathcal{Z}_t]\mathbb{Q}[\mathcal{Y}_t|\mathcal{Z}_t])$ derived in reference[9]. Moreover, $\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]$ is the conjugate environment and signal of a given pair of strategies $\{\mathbb{T}_F,\mathbb{T}_K\}$ under which it achieves the bound. From the detailed FR in eq. (35), we also see that the fitness of a given strategy can exceed the bound by chance as $\Psi^s[\mathcal{Y}_t,\mathcal{Z}_t] > \Psi_0[\mathcal{Y}_t] + i[\mathcal{Y}_t;\mathcal{Z}_t]$ when the realized pair of environmental and signal histories $\{\mathcal{Y}_t,\mathcal{Z}_t\}$ appears more frequently in the conjugate environment than in the actual environment as $\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t] > \mathbb{Q}[\mathcal{Y}_t,\mathcal{Z}_t]$.

A. Achievability of the bound and causality

The necessary and sufficient condition $\mathbb{P}_{K,F}^*[\mathcal{Y}_t|\mathcal{Z}_t] = \mathbb{Q}[\mathcal{Y}_t|\mathcal{Z}_t]$ for achieving the bound means that the optimal metabolic allocation and phenotype switching strategy together implement the Bayesian computation of the posterior distribution $\mathbb{Q}[\mathcal{Y}_t|\mathcal{Z}_t]$ of $\mathcal{Y}_t$ given the history of the sensed signal $\mathcal{Z}_t$. Under the constraint that $\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]$ satisfies a causal relation as $\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t] =$

$\sum_{X_t} \mathbb{P}_K[\mathcal{Y}_t||X_t]\mathbb{P}_F[X_t||Z_t]$, $\mathbb{P}^*_{K,F}[\mathcal{Y}_t|Z_t] = \mathbb{Q}[\mathcal{Y}_t|Z_t]$ cannot be realized in general because $\mathbb{Q}[\mathcal{Y}_t|Z_t]$ does not necessarily satisfy the causality relation between $\mathcal{Y}_t$ and Z_t. If, for example, the metabolic allocation strategy is of Kelly's type as $\mathbb{T}_K(y|x) = \delta_{y,x}$, the phenotypic switching strategy must satisfy $\mathbb{P}^*_F[X_t||Z_t] = \mathbb{Q}[X_t|Z_t]$ to achieve the bound where $\mathbb{Q}[X_t|Z_t] := \mathbb{Q}[\mathcal{Y}_t|Z_t]|_{\mathcal{Y}_t=X_t}$. By definition, $\mathbb{P}_F[X_t||Z_t]$ should satisfy the causal relation that $x(t)$ depends only on the past and/or current states of $z(t)$. However, $\mathbb{Q}[\mathcal{Y}_t|Z_t]$ may not necessarily be causal because the conditioning by histories biases the past states of the conditioned history. For example, consider the case that the signal Z_t is causally generated from the environment $\mathcal{Y}_t$ and there is no feedback from Z_t to $\mathcal{Y}_t$ as in Fig.2(A) and (C). If we observe $Z_t = \{\cdots, z(t-2), z(t-1), z(t)\} = \{\cdots, \mathfrak{s}^z_1, \mathfrak{s}^z_1, \mathfrak{s}^z_2\}$, we may infer that the environmental state changes from $\mathfrak{s}^y_1$ to $\mathfrak{s}^y_2$ at time t. However, if we further observe $\{z(t+1), z(t+2)\} = \{\mathfrak{s}^z_1, \mathfrak{s}^z_1\}$, then we may change our prediction such that $y(t) = \mathfrak{s}^y_1$ and $z(t) = \mathfrak{s}^z_2$ was simply an error of the signal. This intuitive observation illustrates that the inferred state of $y(t)$ can be affected by the future observation of the signal, e.g., $z(t+1)$ and $z(t+2)$. For exactly the same reason, in $\mathbb{Q}[\mathcal{Y}_t|Z_t]$, the past state in the environmental history $\mathcal{Y}_t$ is affected by the future observations of the signal in Z_t. Because of the dependency on future state of the signal, $\mathbb{Q}[\mathcal{Y}_t|Z_t]$ cannot be represented causally in general. An exception is that the environmental history is causally and memorylessly generated from the signal as $\mathbb{Q}[\mathcal{Y}_t||Z_t] = \prod_{\tau=0}^{t} \mathbb{T}_E(y_\tau|z_\tau)$, which is not realistic because we usually expect the signal to be generated from environment and not *vice versa*. Thus,

$$\mathcal{D}^{s^*}_{\mathrm{loss}} := \min_{\{\mathbb{T}_F, \mathbb{T}_K\}} \mathcal{D}^s_{\mathrm{loss}}, \qquad (42)$$

is not generally zero, and $\mathcal{D}^s_{\mathrm{loss}}$ contains not only the loss due to suboptimality of strategies, but also the loss from the causal constraints in exploiting the information of Z_t for phenotype switching.

VII. DIRECTED INFORMATION AND CAUSAL FRS WITH SENSING

By taking account of the causality relation between $\mathcal{Y}_t$ and Z_t in more detail, we can obtain tighter FRs, which illustrate the problem of the causality more explicitly. Let us additionally assume that $\mathbb{Q}[\mathcal{Y}_t, Z_t]$ has a causal and Markov relation as depicted in Fig. 4 (A). The relation is represented as

$$\mathbb{Q}[\mathcal{Y}_t, Z_t] = \left[\prod_{\tau=0}^{t-1} \mathbb{T}^F_E(z_{\tau+1}|y_{\tau+1}, z_\tau)\mathbb{T}^F_E(y_{\tau+1}|y_\tau, z_\tau)\right] q(y_0, z_0),$$
$$:= \mathbb{Q}_F[Z_t||\mathcal{Y}_t]\mathbb{Q}_F[\mathcal{Y}_t||Z_{t-1}], \qquad (43)$$

where $\mathbb{Q}_F[Z_t||\mathcal{Y}_t] := \prod_{\tau=0}^{t-1} \mathbb{T}^F_E(z_{\tau+1}|y_{\tau+1}, z_\tau)q(z_0|y_0)$ and $\mathbb{Q}_F[\mathcal{Y}_t||Z_{t-1}] := \prod_{\tau=0}^{t-1} \mathbb{T}^F_E(y_{\tau+1}|y_\tau, z_\tau)q(y_0)$. Here, the
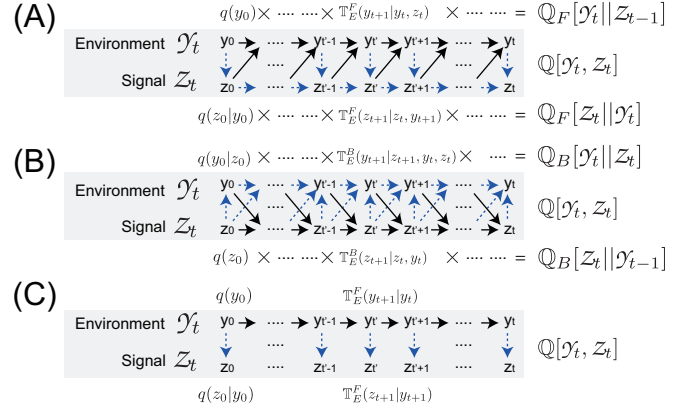

FIG. 4. Causal dependency of environmental and signal histories. (A) A causal relation between environment and signal, in which the environment causally drives the signal. Black solid arrows indicate that y_{t+1} is affected by its one-step past state y_t and that of the signal z_t. Blue dashed arrows represent that z_{t+1} is affected by the current environment y_{t+1} and the one-step past signal z_t. (B) The causal decomposition of $\mathbb{Q}[\mathcal{Y}_t, Z_t]$ with the reversed causal relation between the environment and the signal, in which the signal causally drives the environment. Black solid arrows indicate that z_{t+1} is affected by its one-step past environment y_t and that of the signal z_t. Blue dashed arrows represent that y_{t+1} is affected by the current and one-step past signal z_{t+1} and z_t and the one-step past environment y_t. (C) The causal relation between the environment and the signal described in Figs. 2 (A) and (C), and used in the numerical simulation (Figs. 3 and 5).

signal $z(t)$ and environment $y(t)$ are Markovean and dependent on the past states of the environment and the signal. By applying Bayes' theorem, $\mathbb{Q}[\mathcal{Y}_t, Z_t]$ admits another causal decomposition (Fig. 4 (B)) with reversed causality between $\mathcal{Y}_t$ and Z_t:

$$\mathbb{Q}[\mathcal{Y}_t, Z_t] = \left[\prod_{t=0}^{t-1} \mathbb{T}^B_E(y_{t+1}|z_{t+1}, y_t, z_t)\right]\left[\prod_{t=0}^{t-1} \mathbb{T}^B_E(z_{t+1}|y_t, z_t)\right],$$
$$\times q(y_0, z_0)$$
$$:= \mathbb{Q}_B[\mathcal{Y}_t||Z_t]\mathbb{Q}_B[Z_t||\mathcal{Y}_{t-1}], \qquad (44)$$

where

$$\mathbb{T}^B_E(z_{t+1}|y_t, z_t) := \sum_{y_{t+1}} \mathbb{T}^F_E(z_{t+1}|y_{t+1}, z_t)\mathbb{T}^F_E(y_{t+1}|y_t, z_t), \qquad (45)$$

$$\mathbb{T}^B_E(y_{t+1}|z_{t+1}, y_t, z_t) := \frac{\mathbb{T}^F_E(z_{t+1}|y_{t+1}, z_t)\mathbb{T}^F_E(y_{t+1}|y_t, z_t)}{\mathbb{T}^B_E(z_{t+1}|y_t, z_t)}, \qquad (46)$$

$\mathbb{Q}_B[\mathcal{Y}_t||Z_t] := \prod_{t=0}^{t-1} \mathbb{T}^B_E(y_{t+1}|z_{t+1}, y_t, z_t)q(y_0|z_0)$, and $\mathbb{Q}_B[Z_t||\mathcal{Y}_{t-1}] := \prod_{t=0}^{t-1} \mathbb{T}^B_E(z_{t+1}|y_t, z_t)q(z_0)$. Because of the meaning of Bayes' theorem, $\mathbb{T}^B_E(y_{t+1}|z_{t+1}, y_t, z_t)$ is interpreted as the posterior distribution of y_{t+1} inferred from observations, z_{t+1}, y_t, and z_t. Thus, $\mathbb{Q}_B[\mathcal{Y}_t||Z_t]$

is equivalent to the posterior path probability of the environment inferred by the sequential Bayesian inference of y_{t+1} given the observation of the signal z_t and the past state of the environment y_t. Note that $\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t)$ is slightly different from the usual sequential Bayesean inference to estimate the hidden state $y(t)$ only from observation $\mathcal{Z}_t$ in which we cannot use the past environmental state y_t for inference. Moreover, $\mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t]$ is also different from $\mathbb{Q}_B[\mathcal{Y}_t|\mathcal{Z}_t]$ because the posterior path probability $\mathbb{Q}_B[\mathcal{Y}_t|\mathcal{Z}_t]$ is computed after observing the whole history of $\mathcal{Z}_t$ rather than by conducting Bayesian inference sequentially as in $\mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t]$.

Equation (44) yields a tighter decomposition of the fitness:

$$\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] = \Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]$$
$$- \ln \frac{\mathbb{P}_B^s[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t]}{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]}. \quad (47)$$

where $i[\mathcal{Z}_t \to \mathcal{Y}_t] := \ln \mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]/\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]\mathbb{Q}[\mathcal{Y}_t]$ is the bare directed information from $\mathcal{Z}_t$ to $\mathcal{Y}_t$[53]. This decomposition leads to the following detailed FR:

$$e^{-(\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t] - \Psi^s[\mathcal{Y}_t, \mathcal{Z}_t])}$$
$$= \frac{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]}{\mathbb{P}_B^s[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t]},$$
$$= \frac{\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]}{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}. \quad (48)$$

Because $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]$ forms a joint path probability among $\mathcal{X}_t$, $\mathcal{Y}_t$, and $\mathcal{Z}_t$, which satisfies $\sum_{\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t} \mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}] = 1$, we obtain the integral and average FRs as

$$\left\langle e^{-(\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t] - \Psi^s[\mathcal{Y}_t, \mathcal{Z}_t])} \right\rangle_{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]} = 1, \quad (49)$$

and

$$\langle \Psi^s \rangle_{\mathbb{Q}} = \langle \Psi_0 \rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t} - \mathcal{D}_{\mathrm{loss}}^{s,d} \le \langle \Psi_0 \rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t}, \quad (50)$$

where

$$\mathcal{D}_{\mathrm{loss}}^{s,d} := \mathcal{D}\Big[\mathbb{P}_B[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t]\Big|\Big|\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]\Big],$$
$$= \mathcal{D}\Big[\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]\Big|\Big|\mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]\Big] \quad (51)$$

and $\mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t} := \mathcal{D}\Big[\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]\Big|\Big|\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]\mathbb{Q}[\mathcal{Y}_t]\Big]$. $\mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t}$ is the directed information that quantifies the amount of information $\mathcal{Z}_t$ has for inferring $\mathcal{Y}_t$[53]. If $\mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t] = \mathbb{Q}[\mathcal{Y}_t]$, $\mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t} = 0$ as well as $i[\mathcal{Z}_t \to \mathcal{Y}_t] = 0$ hold, which means that $\mathcal{Z}_t$ is useless for inferring $\mathcal{Y}_t$ causally. $\langle \Psi_0 \rangle_{\mathbb{Q}} + \mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t}$ in eq. (50) is a tighter bound for the average fitness with sensing because

$$\mathcal{I}^{\mathcal{Y};\mathcal{Z}} = \left\langle \ln \frac{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}{\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]\mathbb{Q}[\mathcal{Y}_t]} + \ln \frac{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}{\mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]} \right\rangle_{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}$$
$$= \mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t} + \mathcal{I}^{\mathcal{Y}_t \to \mathcal{Z}_t} \ge \mathcal{I}^{\mathcal{Z}_t \to \mathcal{Y}_t} \quad (52)$$

holds, where $\mathcal{I}^{\mathcal{Y}_t \to \mathcal{Z}_t} := \mathcal{D}\Big[\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]\Big|\Big|\mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]\Big]$ is the directed information from $\mathcal{Y}_t$ to $\mathcal{Z}_t$. Moreover, $\mathcal{I}^{\mathcal{Y}_t \to \mathcal{Z}_t}$ is the residual information in $\mathcal{I}^{\mathcal{Y};\mathcal{Z}}$ that has no fitness value under the causal constraints.

A. Numerical verification of FRs with sensing

To demonstrate the validity of FRs in Eqs. (48)-(50), we conduct a modified numerical simulation of the population dynamics to incorporate sensing (see also Appendix B). The stochastic laws of environmental dynamics $\mathbb{T}_E^F(y_{t+1}|y_t, z_t)$ (Fig.2 (A)) and the metabolic allocation strategy $\mathbb{T}_K(y_t|x_t)$ (Fig.2 (B)) are the same as those of the simulations in Fig. 3. The sensing signal is assumed to have two states, $\mathfrak{s}_1^z$ and $\mathfrak{s}_2^z$ (Fig. 2 (C)). If the environment is either in $\mathfrak{s}_1^y$ or in $\mathfrak{s}_2^y$, the signal gives the corresponding $\mathfrak{s}_1^z$ and $\mathfrak{s}_2^z$ with 90% accuracy. If the environment is in $\mathfrak{s}_3^y$, the signal produces either $\mathfrak{s}_1^z$ or $\mathfrak{s}_2^z$ with eqaul probability (Fig. 2 (C)). Because the signal is memorylessly generated from the environment at every time point, the environment and signal together form a Markov relation as depicted in Fig. 4 (C). By employing the sensed signal, the organisms switch to the phenotypic state $\mathfrak{s}_1^x$ with a 95% chance when the sensed signal is $\mathfrak{s}_1^z$ and to $\mathfrak{s}_2^x$ when the signal is $\mathfrak{s}_2^z$ as shown in Fig. 2 (D).

Figures 5 (A) and (B) show two trajectories of the population size under the same realizations of environment as in Figs. 3 (A) and (B), respectively. In Fig. 5 (A), the total population size $\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}} = e^{\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]}$ (dashed line) is less than $e^{\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]}$ (red line) whereas $\mathcal{N}_t^{\mathcal{Y},\mathcal{Z}}$ frequently becomes greater than $e^{\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]}$ in Fig. 5 (B). Even with sensing, both $\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]$ and $\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]$ fluctuate substantially as shown in Figs. 5 (C) and (D) depending on the realization of the environment and signal. Due to this fluctuation, $\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]$ sometimes become greater than $e^{\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]}$ as in Fig. 5 (E) to satisfy the integral FR (49) as demonstrated in Fig. 5 (F).
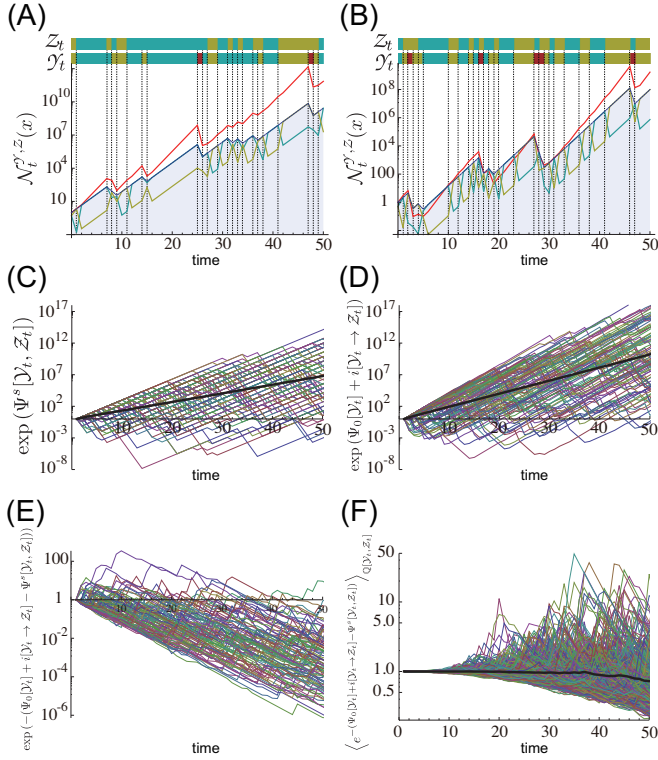
FIG. 5. Numerical simulation of the population dynamics with sensing defined in Fig. 2. (A, B) Two independent sample histories of environment $\mathcal{Y}_t$ and signal $\mathcal{Z}_t$. The dynamic of $\mathcal{N}_t^{\mathcal{Y},z}(x_t)$, $\mathcal{N}_t^{\mathcal{Y},z} = e^{\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]}$, and $e^{\Psi^0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]}$ under the two realizations of the environmental and signal histories obtained by solving eq. (2). The color bars on the graphs represent the states of the environment and the signal at each time point. The correspondence of the colors and environmental and signal states is as shown in Fig. 2 (A) and (C). Cyan and yellow lines represent $\mathcal{N}_t^{\mathcal{Y},z}(\mathfrak{s}_1^x)$ and $\mathcal{N}_t^{\mathcal{Y},z}(\mathfrak{s}_2^x)$, respectively. The dashed blue line with filled grey style is $\mathcal{N}_t^{\mathcal{Y},z}$. The red line is $e^{\Psi_0[\mathcal{Y}] + i[\mathcal{Z}_t \to \mathcal{Y}_t]}$. (C) Stochastic behaviors of population fitness $\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]$ for 100 independent samples of environmental and signal histories. Each colored thin line represents $e^{\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t]} = \mathcal{N}_t^{\mathcal{Y},z}/\mathcal{N}_0^{\mathcal{Y},z}$ for each realization of the environmental and signal histories. The thick black line is $e^{\langle \Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] \rangle_{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}}$. (D) Stochastic behaviors of $\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]$ for the same 100 samples of environmental and signal histories as those in (C). Each colored thin line represents $e^{\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t]}$ for each realization of the same environmental and signal histories as in (C). (E) Stochastic behaviors of $e^{-(\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t] - \Psi^s[\mathcal{Y}_t])}$ for the same 100 samples of environmental and signal histories as those in (C). Each colored thin line represents $e^{-(\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t] - \Psi^s[\mathcal{Y}_t])}$ for each realization of the same environmental and signal histories as in (C). (F) $\left\langle e^{-(\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t] - \Psi^s[\mathcal{Y}_t])} \right\rangle_{\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]}$ calculated empirically by the numerical simulations of sample paths of $e^{-(\Psi_0[\mathcal{Y}_t] + i[\mathcal{Z}_t \to \mathcal{Y}_t] - \Psi^s[\mathcal{Y}_t])}$. The meaning of the thick and thin lines are the same as in Fig. 3.

B. Achievability of the bound

The bound in eq. (50) can be achieved when $\mathbb{P}_{K,F}^*[\mathcal{Y}_t|\mathcal{Z}_t] = \mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t]$ holds, which also means that $\mathbb{Q}^*[\mathcal{Y}_t, \mathcal{Z}_t] := \mathbb{Q}^{*B}[\mathcal{Y}_t||\mathcal{Z}_t]\mathbb{Q}_B[\mathcal{Z}_t||\mathcal{Y}_{t-1}]$ is the conjugate environment for a given $\{\mathbb{T}_F, \mathbb{T}_K\}$ where $\mathbb{Q}^{*B}[\mathcal{Y}_t||\mathcal{Z}_t] = \mathbb{P}_{K,F}[\mathcal{Y}_t|\mathcal{Z}_t]$. Because $\mathbb{Q}_B[\mathcal{Y}_t||\mathcal{Z}_t]$ is the path obtained by a type of sequential Bayesian inference $\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t)$, the bound is attained if phenotype switching with sensing and metabolic allocation, as a whole, implement the same Bayesian computation as $\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t)$. However, to estimate y_{t+1}, the Bayesian inference represented by $\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t)$ uses not only z_{t+1} and z_t but also the error-less information on the past environmental state y_t. If organisms do not have any way to obtain perfect information on the past environment by sensing as they do not for biologically realistic situations, the metabolic allocation strategy should be of Kelly's type as $\mathbb{T}_K^*(y_t|x_t) = \delta_{y_t, x_t}$ to achieve the bound without additional constraints on the environment. With Kelly's strategy, the organisms can effectively obtain information on the past environment because the past phenotypic state $x(t)$ of the survived organisms becomes the same as the past environmental state $y(t)$ under this situation. The optimal phenotype switching with Kelly's strategy for metabolic allocation should be

$$\mathbb{T}_F^*(x_{t+1}|x_t, z_{t+1}, z_t) = \mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t)\big|_{\substack{y_{t+1}=x_{t+1}, \\ y_t = x_t}}. \tag{53}$$

to attain the bound. Another situation in which the bound can be achieved is when the environmental state y_{t+1} depends not on y_t but only on z_t as $\mathbb{T}_E(y_{t+1}|y_t, z_t) = \mathbb{T}_E(y_{t+1}|z_t)$. We then have

$$\mathbb{T}_E^B(z_{t+1}|z_t) = \sum_{y_{t+1}} \mathbb{T}_E^F(z_{t+1}|y_{t+1})\mathbb{T}_E^F(y_{t+1}|z_t), \tag{54}$$

$$\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t) = \frac{\mathbb{T}_E^F(z_{t+1}|y_{t+1})\mathbb{T}_E^F(y_{t+1}|z_t)}{\mathbb{T}_E^F(z_{t+1}|z_t)}, \tag{55}$$

where $\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t)$ is reduced to be independent of y_t as $\mathbb{T}_E^B(y_{t+1}|z_{t+1}, y_t, z_t) = \mathbb{T}_E^B(y_{t+1}|z_{t+1}, z_t)$. Thus, the bound can be attained if the phenotypic switching and metabolic allocation strategies satisfy

$$\sum_{x_t} \mathbb{T}_K^*(y_{t+1}|x_{t+1})\mathbb{T}_F^*(x_{t+1}|z_{t+1}, z_t) = \mathbb{T}_E^B(y_{t+1}|z_{t+1}, z_t).$$

Note that the phenotypic switching need not be dependent on the past phenotypic state in this situation because $z(t+1)$ and $z(t)$ contain all the relevant information on the future state of $y(t+1)$.

For both cases, the achievability of the bound is directly linked to the accessibility of the organisms to the past information that directly drives the current environmental state. In addition, even when the bound is not achieved, $\mathcal{D}_{\text{loss}}^{s,d^*}$ explicitly represents the loss of fitness associated with the inaccessibility.

C. Losses due to limited capacity and suboptimality

For general metabolic allocation strategies other than Kelly's type, the phenotypic states of the survived organisms contain only imperfect information on the past environmental history. In addition, the environment and signal may not always form the circular dependency $\cdots \rightarrow y_t \rightarrow z_t \rightarrow y_{t+1} \rightarrow z_{t+1} \rightarrow \ldots$. In such situations, the loss is most generally represented by $\mathcal{D}^s_{\mathrm{loss}}$, which quantifies the difference between the time-backward path probability $\mathbb{P}_B[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t]$ and the time-forward path probability $\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\mathbb{Q}[\mathcal{Z}_t]$. The loss from different sources can be represented by using the following dissection of the fitness:

$$\Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] = \Psi_0[\mathcal{Y}_t] + i[\mathcal{Y}_t, \mathcal{Z}_t] - i^s_B[\mathcal{Y}_t, \mathcal{Z}_t|\mathcal{X}_t] \quad (56)$$

$$- \ln \frac{\mathbb{P}^s_B[\mathcal{Y}_t|\mathcal{X}_t]}{\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]} - \ln \frac{\mathbb{P}^s_B[\mathcal{X}_t|\mathcal{Z}_t]}{\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]}, \quad (57)$$

where $i^s_B[\mathcal{Y}_t, \mathcal{Z}_t|\mathcal{X}_t] := \ln \mathbb{P}^s_B[\mathcal{Y}_t, \mathcal{Z}_t|\mathcal{X}_t]/\mathbb{P}^s_B[\mathcal{Y}_t|\mathcal{X}_t]\mathbb{P}^s_B[\mathcal{Z}_t|\mathcal{X}_t]$ is the bare conditional backward mutual information. Although this representation does not immediately admit detailed and integral FRs, we can obtain the following decomposition of $\mathcal{D}^s_{\mathrm{loss}}$ into three terms:

$$\mathcal{D}^s_{\mathrm{loss}} = \mathcal{I}^{\mathcal{Y};z|X}_B + \mathcal{D}^{s,K}_{\mathrm{loss}} + \mathcal{D}^{s,F}_{\mathrm{loss}},$$

where

$$\mathcal{I}^{\mathcal{Y};z|X}_B := \langle i^s_B[\mathcal{Y}_t, \mathcal{Z}_t|\mathcal{X}_t] \rangle_{\mathbb{P}_B[\mathcal{X}_t, \mathcal{Y}_t, \mathcal{Z}_t]},$$

$$\mathcal{D}^{s,K}_{\mathrm{loss}} := \left\langle \mathcal{D}\left[\mathbb{P}^s_B[\mathcal{Y}_t|\mathcal{X}_t]\middle\|\mathbb{P}_K[\mathcal{Y}_t||\mathcal{X}_t]\right]\right\rangle_{\mathbb{P}^s_B[\mathcal{X}_t]},$$

$$\mathcal{D}^{s,F}_{\mathrm{loss}} := \left\langle \mathcal{D}\left[\mathbb{P}^s_B[\mathcal{X}_t|\mathcal{Z}_t]\middle\|\mathbb{P}_F[\mathcal{X}_t||\mathcal{Z}_t]\right]\right\rangle_{\mathbb{Q}[\mathcal{Z}_t]}.$$

$\mathcal{I}^{\mathcal{Y};z|X}_B$ is the residual information on the environment that the signal still has even if we know the retrospective history of the phenotype. Because the retrospective correlation between the phenotype and environment is induced by the selection, this residual information represents the information that is not used in the selection. If the signal contains certain information on the environmental fluctuation that is nothing to do with the replication of the organisms, for example, the information on the existence of unmetabolizable artificial molecules, such information cannot be exploited in the phenotypic switching to choose a better phenotype for the current environmental state. Thus, $\mathcal{I}^{\mathcal{Y};z|X}_B$ measures the amount of such useless information for fitness that the sensing signal conveys. $\mathcal{D}^{s,K}_{\mathrm{loss}}$ accounts for the imperfectness of the metabolic allocation strategy, and $\mathcal{D}^{s,F}_{\mathrm{loss}}$ quantifies the imperfectness of phenotype switching due to the causal constraint and suboptimality.

For example, in the case of the causally optimal strategy under Kelly's metabolic allocation in eq. (53) with the causal structure in $\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]$ (eq. (43)), these quantities are reduced to

$$\mathcal{I}^{\mathcal{Y};z|X}_B = 0, \qquad \mathcal{D}^{s,K}_{\mathrm{loss}} = 0, \qquad \mathcal{D}^{s,F}_{\mathrm{loss}} = \mathcal{I}^{\mathcal{Y}_t \rightarrow z_t}.$$

The first equation, $\mathcal{I}^{\mathcal{Y};z|X}_B = 0$, holds because organisms with the metabolic allocation strategy of Kelly's type can survive only when their phenotypic history is identical to the actual environmental history, and the retrospective phenotypic history $\mathcal{X}_t$ contains perfect information on $\mathcal{Y}_t$. Thus, the residual $\mathcal{I}^{\mathcal{Y};z|X}_B$ becomes 0. Similarly, $\mathcal{D}^{s,K}_{\mathrm{loss}} = 0$ because only organisms with their phenotypic history identical to their environmental hisotry can survive, and $\mathbb{P}^s_B[\mathcal{Y}_t|\mathcal{X}_t] = \delta_{\mathcal{X}_t, \mathcal{Y}_t}$ holds for Kelly's strategy. In contrast, $\mathcal{D}^{s,F}_{\mathrm{loss}}$ cannot be zero because the phenotypic switching strategy is causally constrained. The minimum loss due to this constraint is the directed information $\mathcal{I}^{\mathcal{Y}_t \rightarrow z_t}$ that measures the causally useless information in $\mathcal{I}^{\mathcal{Y};z}$.

In a general situation, these three quantities are mutually related. For example, the amount of useless information for fitness $\mathcal{I}^{\mathcal{Y};z|X}_B$ depends on the choice of strategies. If phenotype switching does not use any information obtained from $z(t)$, then $\mathcal{I}^{\mathcal{Y};z|X}_B$ becomes $\mathcal{I}^{\mathcal{Y};z}_B$. They are appropriate information-theoretic quantities that account for the irrelevance of sensing and the imperfectness of metabolic allocation and phenotypic switching for exploiting the sensed information.

VIII. SUMMARY AND DISCUSSION

In this paper, we clarified the stochastic and information thermodynamic structures in population dynamics with and without sensing, to derive the bound of fitness and the fitness gain by sensing, in the form of FRs. The detailed and integral FRs obtained substantially generalized the previous results on the average fitness value of information by showing that not only average but also the fluctuation of fitness is generically constrained as is the entropy production. Such constraints manifest the possibility and the condition that the fitness of organisms with suboptimal strategies can be greater than that of the optimal one by chance due to environmental fluctuations, just as there is a finite probability of observing reversal of heat flow in a small thermal system. Such rare events can be regarded as a violation of ESS because they may induce the takeover of the optimal organism by a suboptimal one in a finite population. Nevertheless, the violation is somehow ruled to follow the integral and detailed FRs. Moreover, the directed information is derived to be the tighter measure of the fitness value of information, in which the causal structure in the problem is explicitly accounted for. The condition for achieving this bound is shown to be related to the ability of the organisms to conduct or implement a type of Bayesian sequential inference from the sensed information. Finally, we derived three quantities that can measure the irrelevance of sensing and the imperfectness of metabolic allocation and phenotypic switching for exploiting the sensed information.

All these results and generalizations are derived by employing the path-wise and the time-backward represen-

tation of population dynamics. Among others, pivotal is the duality between the maximization of the average fitness and the minimization of the difference between the time-forward and time-backward path probabilities (eqs. (28) and (40)). The minimum loss of the average fitness due to causality is clearly described in this dual problem as the deviation between the causal time-forward path probability and the non-causal time-backward path probability (eq. (38)). We believe that the path-wise and retrospective formulation of the population dynamics also play indispensable roles in addressing other biologically relevant problems.

A.   Learning rules

In this paper as well as in most of the previous works, the processes of attaining better strategies by mutation or learning are rarely considered directly and explicitly[11,56], except in a reference[13]. As we have demonstrated, the average optimal strategy may not dominate a population all the time in a fluctuating environment because of a rare event: the takeover by suboptimal strategies. Thus, the organisms and their corresponding strategies are expected to change non-stationarily over time. In such a situation, an advantage may be gained by organisms that acquire an ability to adaptively change and learn the strategies as our brain and immune system do[57]. However, the fitness to be increased by adaptation is a population-level quantity whereas adaptation is conducted at the individual-level. The duality might be used to resolve this problem. As demonstrated in Kelly's strategy in the most extreme fashion, the survived organisms can retrospectively obtain certain information on the environment they experienced. As shown recently[13], such information can be used to adjust time-forward strategies to switch phenotype and to allocate resource so as to reduce the discrepancy between time-forward and backward phenotypic histories rather than directly increasing the population fitness. By employing our path-wise formulation, we may obtain more general results on the adaptive learning of strategies and its evolutionary advantages.

B.   Individual sensing

Another problem that should be addressed is the type of sensing. In this paper, the sensing signal is treated as an extrinsic factor by assuming that all organisms have the same sensing signal. Such an assumption is valid only when the sensing noise of the organisms is negligibly small. A more general and biologically realistic situation is the individual sensing in which each organism receives different sensing signals owing to the sensing noise intrinsic to the organism[14–17]. The fitness value of individual sensing has been rarely addressed, with the exception of a pioneering work[36], which shows that the fitness value

of individual sensing can be greater than the mutual information under restricted situations[36]. A generalization of this result may be achieved by using our pat-hwise and retrospective formulation.

C.   Thermodynamics and evolution

As we have clarified, thermodynamics and adaptation share the same fundamental mathematical structure. Nevertheless, most attempts to bridge thermodynamics and adaptation or evolution, including ours, are just formal in the sense that the similarity shown is at the level of mathematics[18–21,23]. However, an actual thermodynamics underlies the processes of replication and sensing of organisms[58–60]. The thermodynamics must constrain the rate of replication and the efficiency of sensing, and the latter was investigated intensively in the context of stochastic and information thermodynamics, recently[51,61,62]. Our relation between fitness and information must be consistent with the constraints imposed by the thermodynamics of these processes. Integration of the relation between entropy and information with that between fitness and information would be an indispensable step toward establishing the real thermodynamics of adaptation and evolution.

Appendix A: Decomposition of replication rate

Let us consider a given replication rate $k(x,y)$. $e^{k(x,y)}$ determines a $(\#\mathfrak{S}_x \times \#\mathfrak{S}_y)$ matrix, and we define its row vectors by $\{\mathbf{v}_x \in (\mathbb{R}_{\geq 0})^{\#\mathfrak{S}_y} | v_{x,y} = e^{k(x,y)}, x \in \mathfrak{S}_x\}$. When the number of the environmental states is equal to or greater than the phenotypic ones as $\#\mathfrak{S}_y \geq \#\mathfrak{S}_x$, $\{\mathbf{v}_x | x \in \mathfrak{S}_x\}$ can form a hyperplane, and we can fine a vector $\mathbf{u} \in (\mathbb{R}_{\geq 0})^{\#\mathfrak{S}_y}$ that is orthogonal to the plane. Let $\mathbf{q}_0$ be the normalization of $\mathbf{u}$ as $\mathbf{q}_0 = \mathbf{u}/\sum_y u(y)$. By definition, $\mathbf{q}_0$ satisfies $\mathbf{v}_x \cdot \mathbf{q}_0 = \sum_y e^{k(x,y)} q_0(y) = \phi_0$ for all $x \in \mathfrak{S}_x$ where $\phi_0 > 0$ is a constant. This means that we can define a conditional probability $\mathbb{T}_K(y|x)$ as

$$\mathbb{T}_K(y|x) := \frac{e^{k(x,y)} q_0(y)}{\phi_0}. \tag{A1}$$

Thus, we have the decomposition of $e^{k(x,y)}$ in eq. (1) and its variant used in eq. (11) as

$$e^{k(x,y)} = \frac{\phi_0 \mathbb{T}_K(y|x)}{q_0(y)} = e^{k_{\max}(y)} \mathbb{T}_K(y|x), \qquad \text{(A2)}$$

where $e^{k_{\max}(y)} := \phi_0/q_0(y)$. Thus, when the environmental states are more complex than the phenotypic ones as $\#\mathfrak{S}_y \geq \#\mathfrak{S}_x$, the decomposition of the replication rate, eq. (1), is general enough even if we do not explicitly assume the relation between metabolic allocation strategy and replication rate as in eq. (1).

If $\#\mathfrak{S}_y < \#\mathfrak{S}_x$, the decomposition does not necessarily exist. Such a situation may occur when an organism has redundant phenotypic states, a fraction of which can be effectively realized by linear combinations of the others. While the redundant phenotypic states can appear in the process of mutation, $\#\mathfrak{S}_y < \#\mathfrak{S}_x$ is not biologically realistic because the environment is generally more complex than phenotype.

Finally, we note that the decomposition, eq. (A2), is not unique except when $\#\mathfrak{S}_x = \#\mathfrak{S}_y$ and $\{\mathbf{v}_x | x \in \mathfrak{S}_x\}$ are linearly independent each others. Nevertheless, our FRs hold irrespective of the choice of the decomposition. Achievability of the maximum average fitness, eq. (24), is affected by the choice because Φ_0 depends on the decomposition and because $\mathbb{T}(y|x)$ must be changed so that it satisfies eq. (A2) for a given $k(x,y)$. While we apparently have a freedom to choose the metabolic allocation strategy, $\mathbb{T}_K(y|x)$, for maximization of $\langle \Psi^b \rangle_{\mathbb{Q}}$ in eq. (24), eq. (24) is virtually reduced to the problem of the maximization of the average fitness by changing only phenotypic switching strategy $\mathbb{P}_F$ as

$$\max_{\mathbb{T}_F} \langle \Psi^b \rangle_{\mathbb{Q}} \leq \langle \Psi_0 \rangle_{\mathbb{Q}}, \qquad \text{(A3)}$$

because Ψ^b is independent of the choice of the decomposition of $e^{k(x,y)}$. This is the problem addressed in our previous work[9]. Tight upper bound of $\max_{\mathbb{T}_F} \langle \Psi^b \rangle_{\mathbb{Q}}$ can be obtained by minimization of $\langle \Psi_0 \rangle_{\mathbb{Q}}$ as

$$\max_{\mathbb{T}_F} \langle \Psi^b \rangle_{\mathbb{Q}} \leq \min_{\mathbb{T}_K} \langle \Psi_0 \rangle_{\mathbb{Q}}. \qquad \text{(A4)}$$

We again stress that our integral and detailed FRs generally hold for any given $\mathbb{T}_F$ and $\mathbb{T}_K$.

Appendix B: Numerical verification of FRs

For the numerical simulations shown in Figs 3 and 5, we consider the case that the environmental dynamics is Markovean as $\mathbb{T}_E^F(y_{t+1}|y_t)$, and that the signal is memorylessly generated from the environment as $\mathbb{T}_E^F(z_{t+1}|z_{t+1})$.

Realizations of the environmental history $\mathcal{Y}_t$ and the sensing signal $\mathcal{Z}_t$ are obtained by conducting a finite-state Markov transition by following $\mathbb{T}_E^F(y_{t+1}|y_t)$ and $\mathbb{T}_E^F(z_{t+1}|y_{t+1})$. For a given pair of realizations $\{\mathcal{Y}_t, \mathcal{Z}_t\}$,

the fitnesses $\Psi^b[\mathcal{Y}_t]$ (without sensing) and $\Psi^s[\mathcal{Y}_t]$ (with sensing) are calculated recursively as

$$\Psi^b[\mathcal{Y}_{t+1}] = \Psi^b[\mathcal{Y}_t] + \ln \frac{\mathcal{N}_{t+1}^{\mathcal{Y}}}{\mathcal{N}_t^{\mathcal{Y}}},$$

$$\Psi^s[\mathcal{Y}_{t+1}, \mathcal{Z}_{t+1}] = \Psi^s[\mathcal{Y}_t, \mathcal{Z}_t] + \ln \frac{\mathcal{N}_{t+1}^{\mathcal{Y}, \mathcal{Z}}}{\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}},$$

where $\mathcal{N}_t^{\mathcal{Y}}$ and $\mathcal{N}_t^{\mathcal{Y}, \mathcal{Z}}$ are obtained by solving eq. (2) recursively for the given $\{\mathcal{Y}_t, \mathcal{Z}_t\}$. Similarly, $\Psi_0[\mathcal{Y}_t]$ is recursively computed as

$$\Psi_0[\mathcal{Y}_{t+1}] = \Psi_0[\mathcal{Y}_t] + \left(\phi_0 + \ln \frac{\mathbb{T}_E^F(y_{t+1}|y_t)}{q_0(y_{t+1})} \right).$$

The point-wise directed information $i[\mathcal{Z}_t \to \mathcal{Y}_t]$ in eq. (48) and eq. (49) is recursively obtained as

$$i[\mathcal{Z}_{t+1} \to \mathcal{Y}_{t+1}] = i[\mathcal{Z}_t \to \mathcal{Y}_t] + \ln \frac{\mathbb{T}_E^F(z_{t+1}|y_{t+1})}{\mathbb{T}_E^B(z_{t+1}|y_t)},$$

where $\mathbb{T}_E^B(z_{t+1}|y_t) = \sum_{y_{t+1}} \mathbb{T}_E^F(z_{t+1}|y_{t+1})\mathbb{T}_E^F(y_{t+1}|y_t)$.

Note that, when a feedback relation exists between y_t and z_t, the computation of $\Psi_0[\mathcal{Y}_t]$ and $i[\mathcal{Z}_t \to \mathcal{Y}_t]$ becomes more demanding because $\mathbb{Q}[\mathcal{Y}_t]$ has to be obtained by marginalizing $\mathcal{Z}_t$ in $\mathbb{Q}[\mathcal{Y}_t, \mathcal{Z}_t]$.

REFERENCES

[1] L.-M. Chevin, R. Lande, and G. M. Mace, "Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory," PLoS Biol **8**, e1000357–8 (2010).

[2] S. A. Frank, "Natural selection. I. Variable environments and uncertain returns on investment*," J. Evol. Biol. **24**, 2299–2309 (2011).

[3] M. Slatkin, "Hedging one's evolutionary bets.," Nature **250**, 704–705 (1974).

[4] T. Philippi and J. Seger, "Hedging one's evolutionary bets, revisited." Trends in Ecol Evol **4**, 41–44 (1989).

[5] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, "Bacterial persistence as a phenotypic switch." Science **305**, 1622–1625 (2004).

[6] Y. Wakamoto, N. Dhar, R. Chait, K. Schneider, F. Signorino-Gelo, S. Leibler, and J. D. McKinney, "Dynamic persistence of antibiotic-stressed mycobacteria," Science **339**, 91–95 (2013).

[7] D. Cohen, "Optimizing reproduction in a randomly varying environment," J. Theor. Biol. **12**, 119–129 (1966).

[8] D. Cohen, "Maximizing final yield when growth is limited by time or by limiting resources." J. Theor. Biol. **33**, 299–307 (1971).

[9] T. J. Kobayashi and Y. Sughiyama, "Fluctuation relations of fitness and information in population dynamics," Phys. Rev. Lett. **115**, 238102–238105 (2015).

[10] N. Dhar and J. D. McKinney, "Microbial phenotypic heterogeneity and antibiotic tolerance," Curr Opin Microbiol **10**, 30–38 (2007).

[11] H. J. E. Beaumont, J. Gallie, C. Kost, G. C. Ferguson, and P. B. Rainey, "Experimental evolution of bet hedging," Nature **461**, 90–93 (2009).

[12] C. E. Meacham and S. J. Morrison, "Tumour heterogeneity and cancer cell plasticity," Nature **501**, 328–337 (2013).

[13] B. Xue and S. Leibler, "Evolutionary learning of adaptation to varying environments through a transgenerational feedback" Proc. Natl. Acad. Sci. U.S.A. **113**, 11266–11271 (2016).

[14] T. J. Perkins and P. S. Swain, "Strategies for cellular decision-making." Mol Syst Biol **5**, 326 (2009).

[15] E. Ben-Jacob and D. Schultz, "Bacteria determine fate by playing dice with controlled odds." Proc. Natl. Acad. Sci. U.S.A. **107**, 13197–13198 (2010).

[16] T. J. Kobayashi and A. Kamimura, "Theoretical aspects of cellular decision-making and information-processing." Adv. Exp. Med. Biol. **736**, 275–291 (2012).

[17] M. D. Brennan, R. Cheong, and A. Levchenko, "Systems biology. How information theory handles cell signaling and uncertainty." Science **338**, 334–335 (2012).

[18] A. J. Lotka, "Contribution to the energetics of evolution" Proc. Natl. Acad. Sci. U.S.A. **8**, 147–151 (1922).

[19] A. J. Lotka, "Natural selection as a physical principle" Proc. Natl. Acad. Sci. U.S.A. **8**, 151–154 (1922).

[20] Y. Iwasa, "Free fitness that always increases in evolution" J. Theor. Biol. **135**, 265–281 (1988).

[21] H. P. de Vladar and N. H. Barton, "The contribution of statistical physics to evolutionary biology," Trends in Ecol Evol **26**, 424–432 (2011).

[22] S. A. Frank, "Natural selection. IV. The Price equation*," J. Evol. Biol. **25**, 1002–1019 (2012).

[23] H. Qian, "Fitness and entropy production in a cell population dynamics with epigenetic phenotype switching," Quant Biol **2**, 47–53 (2014).

[24] V. Mustonen and M. Lässig, "Fitness flux and ubiquity of adaptive evolution," Proc. Natl. Acad. Sci. U.S.A. **107**, 4248–4253 (2010).

[25] R. Levins, "Theory of fitness in a heterogeneous environment. V. Optimal genetic systems" Genetics **52**, 891–904 (1965).

[26] R. Levins, *Evolution in Changing Environments*, Some Theoretical Explorations (Princeton University Press, 1968).

[27] O. Rivoire, "Informations in models of evolutionary dynamics," J Stat Phys **162**, 1324–1352 (2015).

[28] P. Haccou and Y. Iwasa, "Optimal mixed strategies in stochastic environments," Theor Pop Biol **47**, 212–243 (1995).

[29] C. T. Bergstrom and M. Lachmann, "Shannon information and biological fitness," IEEE Information Theory Workshop , 50–54 (2004).

[30] M. C. Donaldson-Matasci, C. T. Bergstrom, and M. Lachmann, "The fitness value of information," Oikos **119**, 219–230 (2010).

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, 2012).

[32] E. Kussell and S. Leibler, "Phenotypic diversity, population growth, and information in fluctuating environments," Science **309**, 2075–2078 (2005).

[33] S. A. Frank, "Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory," J. Evol. Biol. **25**, 2377–2396 (2012).

[34] R. Pugatch, N. Barkai, and T. Tlusty, "Asymptotic Cellular growth rate as the effective information utilization rate," arXiv , 1308.0623v3 (2013).

[35] While the difference between phenotypic and environmental states is only 1, this does not affect our numerical verifications because our theory is proved for more general situations. However, if we consider more environmental states, we need more samples to verify the fluctuation relations, numerically.

[36] O. Rivoire and S. Leibler, "The value of information for populations in varying environments," J Stat Phys **142**, 1124–1166 (2011).

[37] J. L. Kelly, "A new interpretation of information rate," Bell Sys-Tech J **35**, 917–926 (1956).

[38] O. Rivoire and S. Leibler, "A model for the generation and transmission of variations in evolution." Proc. Natl. Acad. Sci. U.S.A. **111**, E1940–9 (2014).

[39] S. Leibler and E. Kussell, "Individual histories and selection in heterogeneous populations." Proc. Natl. Acad. Sci. U.S.A. **107**, 13183–13188 (2010).

[40] G. Bianconi and C. Rahmede, "Quantum mechanics formalism for biological evolution," Chaos Solitons Fract **45**, 555–560 (2012).

[41] Y. Wakamoto, A. Y. Grosberg, and E. Kussell, "Optimal lineage principle for age-structured populations." Evolution **66**, 115–134 (2012).

[42] Y. Sughiyama, T. J. Kobayashi, K. Tsumura, and K. Aihara, "Pathwise thermodynamic structure in population dynamics," Phys Rev E Stat Nonlin Soft Matter Phys **91**, 032120 (2015).

[43] E. Baake and H.-O. Georgii, "Mutation, selection, and ancestry in branching models: a variational approach," J. Math. Biol. **54**, 257–303 (2006).

[44] G. Lambert and E. Kussell, "Quantifying selective pressures driving bacterial evolution using lineage analysis," Phys. Rev. X **5**, 011016–10 (2015).

[45] L. Demetrius and V. Gundlach, "Directionality theory and the entropic principle of natural selection," Entropy **16**, 5428–5522 (2014).

[46] U. Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," Rep. Prog. Phys. **75**, 126001 (2012).

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

[48] O. D. King and J. Masel, "The evolution of bet-hedging adaptations to rare scenarios," Theoretical Population Biology **72**, 560–575 (2007).

[49] N. Giordano, F. Mairet, J.-L. Gouzé, J. Geiselmann, and H. de Jong, "Dynamical allocation of cellular resources as an optimal control problem: Novel insights into microbial growth strategies," PLoS Comput Biol **12**, e1004802–28 (2016).

[50] T. J. Kobayashi, "Implementation of dynamic Bayesian decision making by intracellular kinetics." Phys. Rev. Lett. **104**, 228104 (2010).

[51] A. C. Barato, D. Hartich, and U. Seifert, "Nonequilibrium sensing and its analogy to kinetic proofreading," New J. Phys. **17**, 055026–19 (2014).

[52] G. Kramer, *Directed information for channels with feedback*, Ph.D. thesis (1998).

[53] H. H. Permuter, Y.-H. Kim, and T. Weissman, "Interpretations of directed information in portfolio theory, data compression, and hypothesis testing," IEEE Trans. Inform. Theory **57**, 3248–3259 (2011).

[54] S. Kullback and R. A. Leibler, "On information and sufficiency," Ann Math Stat **22**, 79–86 (1951).

[55] C. Jarzynski and D. K. Wójcik, "Classical and quantum fluctuation theorems for heat exchange," Phys. Rev. Lett. **92**, 230602–4 (2004).

[56] G. Lambert and E. Kussell, "Memory and fitness optimization of bacteria under fluctuating environments," PLoS Genet **10**, e1004556–10 (2014).

[57] A. Mayer, T. Mora, O. Rivoire, and A. M. Walczak, "Diversity of immune strategies explained by adaptation to pathogen statistics," Proc. Natl. Acad. Sci. U.S.A. **113**, 8630–8635 (2016).

[58] J. L. England, "Statistical physics of self-replication," J Chem Phys **139**, 121923–9 (2013).

[59] J. L. England, "Dissipative adaptation in driven self-assembly," Nature Publishing Group **10**, 919–923 (2015).

[60] R. Pugatch, "Greedy scheduling of cellular self-replication leads to optimal doubling times with a log-Frechet distribution," Proc. Natl. Acad. Sci. U.S.A. **112**, 2611–2616 (2015).

[61] C. C. Govern and P. R. ten Wolde, "Optimal resource allocation in cellular sensing systems," Proc. Natl. Acad. Sci. U.S.A. **111**, 17486–17491 (2014).

[62] P. R. Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, "Fundamental limits to cellular sensing," J Stat Phys **162**, 1395–1424 (2016).