# Fluctuation theorem for partially masked nonequilibrium dynamics

Naoto Shiraishi and Takahiro Sagawa
*Department of Basic Science, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan*
(Received 17 March 2014; revised manuscript received 25 September 2014; published 16 January 2015)

We establish a generalization of the fluctuation theorem for partially masked nonequilibrium dynamics. We introduce a partial entropy production with a subset of all possible transitions, and show that the partial entropy production satisfies the integral fluctuation theorem. Our result reveals the fundamental properties of a broad class of autonomous as well as nonautonomous nanomachines. In particular, our result gives a unified fluctuation theorem for both autonomous and nonautonomous Maxwell's demons, where mutual information plays a crucial role. Furthermore, we derive a fluctuation-dissipation theorem that relates nonequilibrium stationary current to two kinds of equilibrium fluctuations.

DOI: 10.1103/PhysRevE.91.012130 PACS number(s): 05.70.Ln, 05.40.−a, 87.10.Mn, 89.70.−a

## I. INTRODUCTION

In modern nonequilibrium statistical physics, the fluctuation theorem (FT) is significant for characterizing the foundation of thermodynamic irreversibility [1–6]. The FT has revealed that entropy production is directly related to the probability of the observed trajectory and that of its time reversal. The entropy production is measured by observing the microscopic trajectories, which has been experimentally demonstrated in a variety of systems [7–11].

In many nonequilibrium systems, however, we are not necessarily interested in all of the microscopic transitions. A prominent example is Maxwell's demon, which is a composite system of an engine and a memory. The memory measures the state of the engine and performs feedback control on the engine. If we calculate the entropy production with the engine alone, the engine apparently violates the FT and the second law of thermodynamics. Moreover, in many experimental situations with complicated artificial [12–14] and biological [10,15–19] nanomachines, we cannot observe all of the transitions. If we observe only a some of the transitions, we cannot determine the total amount of entropy production. In such situations, is it still possible to obtain a universal nonequilibrium relation like the FT?

In this paper, we reveal the universal property of partially masked nonequilibrium dynamics. Let $G$ be the set of all possible transitions between microscopic states, and $\Omega$ be a subset of $G$. We call transitions in $\Omega$ *observed*, and its complement *masked* [see Fig. 1(a)]. We then introduce a partial entropy production associated with $\Omega$. Surprisingly, we can show that the integral FT holds for the partial entropy production, which is regarded as an interesting generalization of the FT.

The concept of partial entropy production is straightforwardly applicable to quite a broad class of nanomachines in thermal environments, such as autonomous Maxwell's demons (or bipartite sensing systems) [20–31], molecular motors [10,15,16], ion exchangers [32], bacterial chemotaxis [17–19], and single-electron boxes [14]. In order to discuss the power of our result, we show two applications. First, we apply it to autonomous demons [20,22–25], which reveals the crucial role of mutual information at the level of stochastic trajectories. Our approach reproduces the previous results on nonautonomous Maxwell's demons as a special case [33,34]. Moreover, we derive a fluctuation-dissipation theorem (FDT) for a pair of transitions.

## II. TOTAL ENTROPY PRODUCTION

A thermodynamic system obeys a continuous-time Markov jump process for the time interval $0 \leqslant t \leqslant T$. We assume that the number of states of the system is finite. The transition (i.e., jump) from state $w'$ to state $w$ is written as $w' \to w$, to which we assign a transition probability $P(w' \to w;t)$ that depends on time $t$ in general. The dynamics of the system is described by the master equation

$$\frac{\partial P(w,t)}{\partial t} = J(w,t) := \sum_{w'} J(w' \to w;t), \tag{1}$$

where $P(w,t)$ is the probability of $w$ at time $t$, and $J(w' \to w;t) := P(w',t)P(w' \to w;t) - P(w,t)P(w \to w';t)$ is the probability flux from $w'$ to $w$. We assume that the system is attached to a single heat bath at inverse temperature $\beta$. From the local detailed balance condition, the heat absorbed by the system from the bath during the transition $w' \to w$ at time $t$ is given by

$$Q(w' \to w;t) = -\frac{1}{\beta} \ln \frac{P(w' \to w;t)}{P(w \to w';t)}. \tag{2}$$

Let $\Gamma$ be a realized trajectory of the dynamics, in which transitions occur $N$ times at $t = t_1, t_2, \ldots, t_N$. The state during the time interval $t_i \leqslant t < t_{i+1}$ is denoted by $w_i$ with $t_0 := 0$ and $t_{N+1} := T$. In particular, the initial and the final states are denoted by $w_0$ and $w_N$, respectively. The total entropy production along trajectory $\Gamma$ is then given by

$$\sigma_{\rm tot} := -\beta \sum_{i=1}^{N} Q(w_{i-1} \to w_i;t_i) + \Delta s, \tag{3}$$

where the stochastic entropy at time $t$ is given by $s(w,t) := -\ln P(w,t)$, and its change is given by $\Delta s := s(w_N,T) - s(w_0,0)$.

 

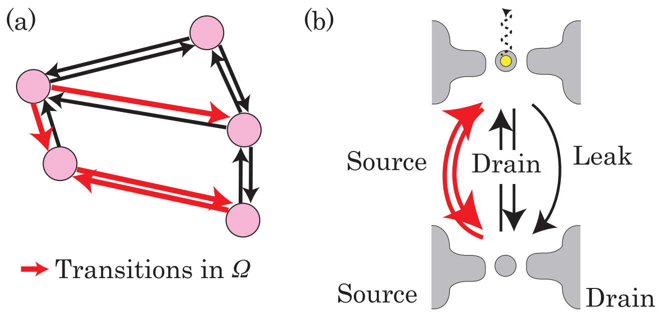


FIG. 1. (Color online) (a) Schematic of a Markov jump process, where the circles indicate the microscopic states and the arrows indicate the paths of possible transitions. The four bold red arrows indicate the observed transitions in Ω and the eight black arrows indicate its complement. (b) Schematic of a quantum dot with at most one electron. Two electron baths, a source and a drain, provide (absorb) electrons to (from) the dot. There is also a leak of an electron to the outer environment. We observe only the transfer of electrons between the source and the dot. Thus, we cannot distinguish the transition associated with the drain from that associated with the leak.

## III. MAIN RESULT

First of all, we define the entropy production associated with a single path of $w' \to w$ (see Fig. 2):

$$\sigma_{w'\to w} := -\beta Q_{w'\to w} + \Delta s_{w'\to w}. \tag{4}$$

The right-hand side (RHS) consists of the following two terms. First, $Q_{w'\to w}$ is the heat absorbed by the system during transitions in $w' \to w$:

$$Q_{w'\to w} := \sum_{i=1}^{N} Q(w_{i-1} \to w_i; t_i)\delta_{w'\to w}(w_{i-1} \to w_i), \tag{5}$$

where $\delta_{w'\to w}(w_{i-1} \to w_i)$ takes the value 1 if $w_{i-1} = w'$ and $w_i = w$, and 0 otherwise. Second, $\Delta s_{w'\to w}$ is the change in the stochastic entropy induced by the transition $w' \to w$:

$$\Delta s_{w'\to w} := s_{w'\to w,\text{jump}} - \int_0^T \frac{J(w' \to w; t)\delta_{w(t),w}}{P(w(t),t)}dt, \tag{6}$$

where $w(t)$ represents the state at time $t$, and $\delta_{w(t),w}$ takes the value 1 if $w(t) = w$ and 0 otherwise. The first term on the RHS in Eq. (6) represents the change in the stochastic entropy due to the realized jumps in $w' \to w$:

$$s_{w'\to w,\text{jump}} := \sum_{i=1}^{N} s(w_i,t_i) - s(w_{i-1},t_i)\delta_{w'\to w}(w_{i-1} \to w_i). \tag{7}$$

The second term on the RHS in Eq. (6) represents the change in the stochastic entropy due to the time evolution of the probability distribution induced by transitions in $w' \to w$. The sum of the second term on the RHS in Eq. (6) for $w'$ equals the time differential of the stochastic entropy:

$$\frac{\partial s(w,t)}{\partial t} = -\sum_{w'} \frac{J(w' \to w; t)}{P(w,t)}. \tag{8}$$

We can then show that the sum of the single-path entropy production for all paths recovers the total entropy production:

$$\sigma_{\text{tot}} = \sum_{w'\to w\in G} \sigma_{w'\to w}, \tag{9}$$

which is a crucial property of the definition (4). By summing up the single-path entropy production over a subset of all paths, we define the partial entropy production with a subset $\Omega \subset G$;

$$\sigma_\Omega := \sum_{w'\to w\in\Omega} \sigma_{w'\to w} = -\beta Q_\Omega + s_{\Omega,\text{jump}} - \int_0^T \frac{J_\Omega(w,t)}{P(w,t)}, \tag{10}$$

where

$$Q_\Omega := \sum_{w'\to w\in\Omega} Q_{w'\to w}, \tag{11}$$

$$s_{\Omega,\text{jump}} := \sum_{w'\to w\in\Omega} s_{w'\to w,\text{jump}}, \tag{12}$$

$$J_\Omega(w,t) := \sum_{\{w'|(w'\to w)\in\Omega\}} J(w' \to w; t). \tag{13}$$

From Eq. (9), we can show that

$$\sigma_{\text{tot}} = \sigma_\Omega + \sigma_{\Omega^c}, \tag{14}$$

where $\Omega^c$ is a complement of Ω. In general, if $G$ is divided into $m$ parts $\Omega_1, \ldots, \Omega_m$, then $\sum_i \sigma_{\Omega_i} = \sigma_{\text{tot}}$ holds. Therefore, our formalism enables additive decompositions of the total entropy production; we call this property *additivity*.

We here discuss a simple example of the choice of Ω. Figure 1(b) shows an experimentally realizable setup of a quantum dot with two electron baths (the source and the drain) [14]. At most one electron is in the dot. Electrons are provided from these two baths. In addition, there is a leak of electrons; an electron sometimes escapes from the dot to the outside environment that is regarded as the third bath. Suppose that we observe transport of the electrons only between the source and the dot. We set transitions associated with the source to Ω, which is denoted by two bold red arrows in Fig. 1(b). Note that we cannot distinguish the transition associated with the drain from that associated with the leak, and thus we cannot calculate the total entropy production from the observed data. Even in such a case, we can calculate the partial entropy production with Ω.

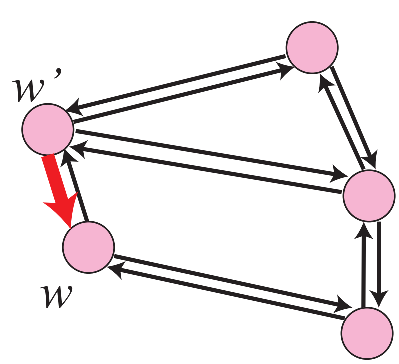


FIG. 2. (Color online) An example of the single path $w' \to w$, which is bold and colored red.

In the above example, there is no backward process of the leak, and therefore the total entropy production and the partial entropy production with $\Omega^c$ are not well defined. However, the partial entropy production with $\Omega$ is still well defined. In general, in order to define the partial entropy production with $\Omega$, we assume only that, for any $w \to w' \in \Omega$ with $P(w \to w';t) \neq 0$, the backward transition probability is also nonzero $P(w' \to w;t) \neq 0$.

We stress that it is highly nontrivial whether $\sigma_\Omega$ satisfies the integral FT. However, we indeed have that for any choice of $\Omega$

$$\langle e^{-\sigma_\Omega} \rangle = 1, \tag{15}$$

which is the main result in this paper.

We prove Eq. (15) as follows. We define another transition rate $P^*$ as

$$P^*(w \to w';t) := \begin{cases} P(w \to w';t), & (w' \to w) \in \Omega, \\ \frac{P(w',t)P(w' \to w;t)}{P(w,t)}, & (w' \to w) \notin \Omega, \end{cases} \tag{16}$$

with the modified escape rate

$$\lambda^*(w,t) := \sum_{w'} P^*(w \to w';t) = \lambda(w,t) + \frac{J_{\Omega^c}(w,t)}{P(w,t)}, \tag{17}$$

where $\lambda$ is the original escape rate of $P$. It is easy to show that

$$P(w \to w';t) e^{\beta Q(w \to w';t) + s(w,t) - s(w',t)\delta_\Omega(w \to w')} = P^*(w' \to w;t)\frac{P(w',t)}{P(w,t)}, \tag{18}$$

where $\delta_\Omega(w \to w')$ takes 1 if $w \to w' \in \Omega$ and 0 otherwise. In addition, $J_\Omega(w,t) + J_{\Omega^c}(w,t) = dP(w,t)/dt$ leads to

$$e^{\int_{t'}^{t''} J_\Omega(w,t)/P(w,t)dt} = \frac{P(w,t'')}{P(w,t')} e^{-\int_{t'}^{t''} J_{\Omega^c}(w,t)/P(w,t)dt}. \tag{19}$$

By using Eqs. (17), (18), and (19), we arrive at our main result

$$\langle e^{-\sigma_\Omega} \rangle = \int d\Gamma P(w_N,T) \prod_{i=1}^{N} P^*(w_i \to w_{i-1};t_i) \prod_{i=0}^{N} e^{-\int_{t_i}^{t_{i+1}} \lambda^*(w_i,t)dt} = 1. \tag{20}$$

Since Eq. (15) is valid for any Markov jump system and any choice of $\Omega$, we obtain many relations in specific situations by applying Eq. (15). In the following, we show two applications. One is to bipartite systems, which clarifies how information is used in autonomous measurement and feedback. The other gives a FDT for a pair of transitions, in which the empirical measure fluctuation plays a role as important as that of the current fluctuation.

## IV. AUTONOMOUS MAXWELL'S DEMONS

We consider a model of autonomous Maxwell's demons, which is a simplification of models discussed in Refs. [20,22,23]. We call the system autonomous when the transition rates are time independent. Suppose that a particle is transported between two particle baths: H with high density and L with low density (see Fig. 3). Between the baths, there is a single site where at most a single particle can come in. Let $x \in \{0,1\}$ be the number of particles in the site. In addition,

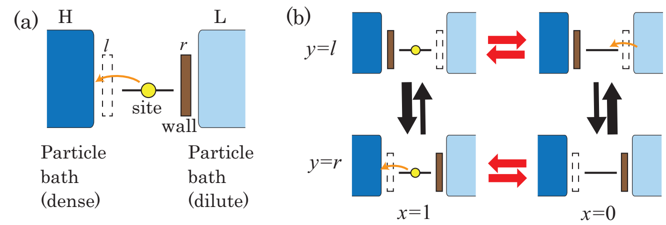


FIG. 3. (Color online) (a) Schematic of the autonomous demon, which consists of two baths, a site for a single particle, and a wall. (b) State space of the model. If a particle is (is not) in the site, the wall tends to go right (left). The red arrows indicate transitions in $\Omega$. With one counterclockwise rotation, one particle is carried from L to H.

we consider a wall that plays the role of the demon. The wall is inserted between the site and one of the baths. Let $y \in \{l,r\}$ be the position of the wall corresponding to left or right. If $y = l$ ($y = r$), the wall prohibits the jump of the particle between the site and the bath H (L). The state of the total system is written as $w := (x,y)$. Correspondingly, we denote $w_i =: (x_i,y_i)$. We assume that the probability of $y = l$ is higher (lower) than $r$ if $x = 0$ ($x = 1$). Intuitively, the wall measures $x$ and then changes its own state depending on the measurement result, which enables the particles to move from L to H against the chemical potential difference. However, since the time intervals for measurement processes and those for feedback processes are not separated from each other, the previous results for nonautonomous demons [33,34], in which the mutual information plays a crucial role, cannot apply to autonomous cases. In more recent works, the role of the mutual information has been clarified for autonomous demons at the level of the ensemble average [24,25,28,29]. Here, by applying our general result, we will show that the mutual information also plays an important role in autonomous demons at the level of stochastic trajectories, as is the case for nonautonomous demons [33,34].

We introduce the entropy production associated with $x, \sigma_x := -\beta Q_x + s(x_N,T) - s(x_0,0)$, and the mutual information that quantifies the correlation between $x$ and $y$. The stochastic mutual information between the particle and the wall is given by $I_t(x;y) := \ln P(x,y,t)/P(x,t)P(y,t)$ [33,34], whose ensemble average gives the mutual information [35]. The change in the mutual information associated with the dynamics of the particle is given by

$$\Delta I_x := I_{x,\mathrm{jump}} + \int_0^T F_x(x(t),y(t),t)dt. \tag{21}$$

Here, $I_{x,\mathrm{jump}}$ represents the change in the mutual information induced by jumps in $x$:

$$I_{x,\mathrm{jump}} := \sum_{i=1}^{N} I_{t_i}(x_i;y_i) - I_{t_i}(x_{i-1};y_{i-1})\delta_{y_i,y_{i-1}}, \tag{22}$$

where $\delta$ is the Kronecker delta. With the notation $J^y_{x',x}(t) := J((x',y) \to (x,y);t)$, $F_x(x,y,t)$ is defined as

$$F_x(x,y,t) := \frac{1}{P(x,y,t)} \sum_{x'} J^y_{x',x}(t) - \frac{1}{P(x,t)} \sum_{y,x'} J^y_{x',x}(t), \tag{23}$$

which represents the change in the mutual information induced by the time evolution of the probability distribution induced by transitions in $x$. To confirm the meaning of $F_x(x,y,t)$, we transform $F_x(x,y,t)$ into another representation. By abbreviating $P(x,y,t)$ to $p_{x,y}$, the mutual information such as $I_t(0;r)$ can be regarded as a function with three arguments $p_{0,r}, p_{0,l}$, and $p_{1,r}$ such that $I_t(0;r) = \ln[p_{0,r}/(p_{0,r}+p_{0,l})(p_{0,r}+p_{1,r})]$. The time differential of the mutual information is then written as

$$\frac{dI_t(x;y)}{dt} = \sum_{c\in\{0,1\}} \sum_{d\in\{l,r\}} \frac{\partial I_t(x;y)}{\partial p_{c,d}} \frac{dp_{c,d}}{dt}. \tag{24}$$

It is easy to show that $F_x(x,y,t)$ corresponds to the contribution to (24) from the probability flux with $x$:

$$F_x(x,y,t) = \sum_{c\in\{0,1\}} \sum_{d\in\{l,r\}} \frac{\partial I_t(x;y)}{\partial p_{c,d}} \sum_{x'} J^{d}_{x',c}(t). \tag{25}$$

We note that $\Delta I_x$ is also rewritten as

$$\Delta I_x = \int_0^T \iota_x(t)dt, \tag{26}$$

where $\iota_x(t)$ is defined as

$$\iota_x(t) := \lim_{\Delta t\to 0} \frac{1}{\Delta t} I(x(t+\Delta t);y(t)) - I(x(t);y(t)). \tag{27}$$

Defining $\iota_y(t)$ in a similar way, we obtain

$$\iota_x(t) + \iota_y(t) = \frac{dI_t(x;y)}{dt}. \tag{28}$$

Here, the ensemble average of $\iota_x(t)$ is equal to the dynamic information flow given in Refs. [21,25].

We now apply Eq. (15) to this model. We set $\Omega$ to transitions in $x$ [i.e., $\Omega := \{(0,r) \rightleftharpoons (1,r), (0,l) \rightleftharpoons (1,l)\}$]. Then $Q_\Omega$ describes the heat absorbed by the particles (i.e., $Q_x = Q_\Omega$). We also obtain

$$\begin{aligned} -\sum_{w'\to w\in\Omega} s_{w'\to w,\mathrm{jump}} &= I_{x,\mathrm{jump}} - \sum_{i=1}^{N} s(x_i,t_i) - s(x_{i-1},t_i) \\ &= I_{x,\mathrm{jump}} - s(x_N,T) + s(x_0,0) \\ &\quad - \int_0^T \frac{\sum_{y,x'} J^{y}_{x',x(t)}(t)}{P(x(t),t)} dt, \end{aligned} \tag{29}$$

and hence $\sigma_\Omega = \sigma_x - \Delta I_x$. Then Eq. (15) reduces to

$$\langle e^{-\sigma_x + \Delta I_x}\rangle = 1, \tag{30}$$

in which mutual information contributes to the FT on an equal footing with the entropy production associated with the particles.

Notably, for any bipartite system described as $w = (x,y)$ with time-dependent transition rates, Eq. (30) holds with the same derivation. In this sense, Eq. (30) includes a previously obtained FT for nonautonomous demons [33,34] as its particular case (see the Appendix). Thus, Eq. (30) provides a unified view of autonomous and nonautonomous demons, where mutual information is a resource of the entropy decrease of a subsystem.

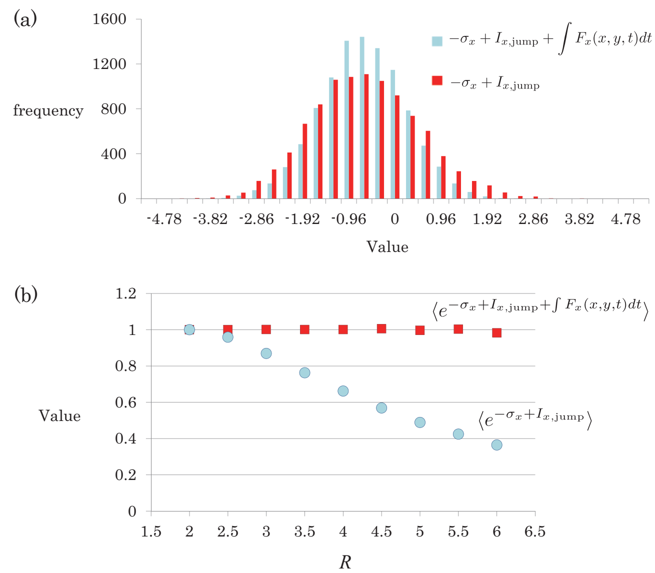


FIG. 4. (Color online) Numerical test of Eq. (30). (a) A histogram of $-\sigma_x + I_{x,\mathrm{jump}}$ [blue (light gray) lines] and $-\sigma_x + I_{x,\mathrm{jump}} + \int F_x(x,y,t)dt$ [red (dark gray) lines] on $R = 3.5$ with 10 000 trials. (b) $\langle e^{-\sigma_x + I_{x,\mathrm{jump}}}\rangle$ (blue circles) and $\langle e^{-\sigma_x + I_{x,\mathrm{jump}} + \int F_x(x,y,t)dt}\rangle$ (red squares) with the change in $R$. The system is in equilibrium at $R = 2$; the larger is $R$, the larger the stationary flux becomes.

Using Jensen's inequality, Eq. (30) leads to a second lawlike inequality

$$\langle \dot{\sigma}_x\rangle - \sum_{x,x',y} J^{y}_{x,x'}(t) I_t(x';y) - I_t(x;y) \geqslant 0, \tag{31}$$

which implies that the entropy production rate of the particles is bounded by the mutual information flow. This inequality has also been obtained in Refs. [21,24,25]. Note that this inequality does not include any contribution from $F_x(x,y,t)$, because the ensemble average of $F_x(x,y,t)$ is equal to zero.

While the ensemble average of $F_x(x,y,t)$ vanishes, this term is needed in Eq. (30). We explicitly show this point with numerical simulation. Set the parameters $P(1\to 0|\mathrm{r}) = P(0\to 1|\mathrm{l}) = 1, P(0\to 1|\mathrm{r}) = P(1\to 0|\mathrm{l}) = 2$, $P(\mathrm{r}\to \mathrm{l}|1) = P(\mathrm{l}\to \mathrm{r}|0) = 1, P(\mathrm{l}\to \mathrm{r}|1) = P(\mathrm{r}\to \mathrm{l}|0) =:$ $R, T = 10$, and set the initial state at its stationary state. We obtain the probability distribution of $-\sigma_x + I_{x,\mathrm{jump}}$ [blue (light gray) lines] and that of $-\sigma_x + I_{x,\mathrm{jump}} + \int F_x(x,y,t)dt$ [red (dark gray) lines] on $R = 3.5$. As shown in Fig. 4(a), the variance of the distribution of $-\sigma_x + I_{x,\mathrm{jump}} + \int F_x(x,y,t)dt$ is larger than that of $-\sigma_x + I_{x,\mathrm{jump}}$. Since the tails of the distributions make a significant contribution in Eq. (30), $\langle e^{-\sigma_x + I_{x,\mathrm{jump}}}\rangle$ deviates from unity as $R$ increases, whereas $\langle e^{-\sigma_x + I_{x,\mathrm{jump}} + \int F_x(x,y,t)dt}\rangle$ stays at unity in agreement with Eq. (30).

## V. FLUCTUATION-DISSIPATION THEOREM

By expanding our general result (15) around equilibrium, we derive a FDT for a pair of transitions. Although this FDT is general, we here discuss it with a specific example, a simple model of kinesin. The kinesin conveys an object by consuming

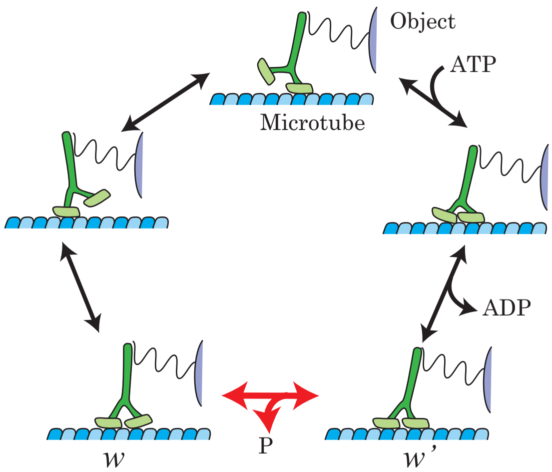


FIG. 5. (Color online) Schematic of the state space of a model of a kinesin. By consuming ATP, the kinesin conveys an object with five steps. We focus on the transition $w' \leftrightarrow w$, which corresponds to the bold arrow colored red, and derive a relation between the nonequilibrium current and two equilibrium fluctuations with $w' \to w$.

the chemical fuel ATP with five cyclic steps (see Fig. 5). In this simple model, the kinesin is in equilibrium with a stall force. If the applied force is slightly varied from the stall force, the kinesin is in a linearly nonequilibrium steady state with stationary current. We will show that the stationary current is characterized by the fluctuations in the equilibrium state.

The perturbation on the transition $w' \to w$ (see Fig. 5) is defined as $a := s(w) - s(w') + \Delta\mu$, where $\Delta\mu$ is the chemical potential difference coupled to reaction $w' \to w$. If the change in the applied force is of order $\epsilon$,$a$ is also of order $\epsilon$. We then introduce two key quantities. First, let $N := \sum_i \delta_{w'\to w}(w_{i-1} \to w_i) - \delta_{w\to w'}(w_{i-1} \to w_i)$ be the empirical current from $w'$ to $w$. The ensemble average of the empirical current equals the probability flux in the steady state with perturbation $\epsilon$ such that $\langle N \rangle = TJ$. Here, we write $J(w' \to w)$ as $J$ for simplicity. Next, we define the degree of the fluctuation of the empirical measures for $w$ and $w'$ as $C := \tau(w')/P(w') - \tau(w)/P(w)$, where $\tau(w) := \int_0^T \delta_{w,w(t)} dt$ is the empirical measure at $w$. The quantity $C$ indicates how the rate of the empirical measure between $w$ and $w'$,$\tau(w)/\tau(w')$, differs from its ensemble average, $P(w)/P(w')$. If the empirical measure is equal to the ensemble average, $C$ is equal to 0. Both the empirical current and the empirical measure are well studied in the context of large deviation theory [36–39].

By setting $\Omega$ to the transitions $w' \leftrightarrow w$, Eq. (15) is written as $\langle e^{-aN+JC} \rangle = 1$. Note that $a$ and $J$ are of order $\epsilon$. Hence, the above equality is expanded as

$$\langle -aN + JC \rangle + \tfrac{1}{2}\langle (-aN + JC)^2 \rangle + O(\epsilon^3) = 0. \quad (32)$$

From $\langle C \rangle = 0$ and $\langle \cdot \rangle = [1 + O(\epsilon)]\langle \cdot \rangle_0$, Eq. (32) is transformed into

$$aTJ = \tfrac{1}{2}\langle (-aN + JC)^2 \rangle_0 + O(\epsilon^3), \quad (33)$$

where $\langle \cdot \rangle_0$ represents the ensemble average in the equilibrium state. Since $NC$ changes its sign for the time-reversal trajectory, and since in equilibrium a trajectory and its time reversal have the same probability, the cross term of the RHS, $\langle NC \rangle_0$, is equal to 0. Substituting $\langle NC \rangle_0 = 0$ into (33) and taking the equality up to $\epsilon^2$ order, we obtain the FDT

$$aTJ = \frac{a^2}{2}\langle N^2 \rangle_0 + \frac{J^2}{2}\langle C^2 \rangle_0. \quad (34)$$

Here, since $\langle N \rangle_0 = \langle C \rangle_0 = 0, \langle N^2 \rangle_0$ and $\langle C^2 \rangle_0$ represent the current fluctuation and the empirical measure fluctuation in the equilibrium state, respectively. The obtained FDT (34) connects the nonequilibrium stationary current $J$ and the two kinds of equilibrium fluctuations. In contrast to the usual FDT, the empirical measure fluctuation appears in this FDT. We note that the fluctuation of $C$ is significant in Eq. (34), while its ensemble average is zero.

In addition, the condition that $J$ is a real number leads to

$$\frac{1}{T^2}\langle N^2 \rangle_0 \langle C^2 \rangle_0 \leqslant 1, \quad (35)$$

which implies that both the current fluctuation and the empirical measure fluctuation cannot be large at the same time.

## VI. CONCLUDING REMARKS

We have derived a FT (15) for partially masked nonequilibrium dynamics. Applying the general result to specific situations, we can obtain both previous results [1,3,5,6,33,34] and additional relations like Eq. (30), Eq. (34), and Eq. (35). Equation (30) clarifies the role of mutual information in both autonomous and nonautonomous Maxwell's demons. Equations (34) and (35) show the features of equilibrium fluctuations with a pair of transition paths. The single-path entropy production (4) is regarded as a building block for constructing various thermodynamic relations for Markov processes. Although we treat only the Markov jump processes in this paper, it is easy to extend our result to the Markov chain and the Langevin dynamics.

We here make a remark on the relationship between our result and a previous work. Although Eq. (15) looks similar to an equality obtained by Hartich *et al.* (Appendix A of Ref. [24]), there is a crucial difference between their result and ours. Their result is a special case of the following equality:

$$\left\langle e^{\beta Q_\Omega - s_{\Omega,\mathrm{jump}} - \int_0^T J_\Omega(w(t),t)/P(w(t),t)dt} \right\rangle = 1, \quad (36)$$

where we assumed that, if $w \to w'$ is in $\Omega$ ($\Omega^c$), $w' \to w$ is also in $\Omega$ ($\Omega^c$). The sign of $J_\Omega(w(t),t)/P(w(t),t)$ in Eq. (36) is opposite to that in Eq. (15). Therefore, the exponent of the left-hand side in Eq. (36) does not satisfy the additivity, whereas the additivity is the crucial characterization of our approach.

The partial entropy production given in Eqs. (4) and (10) satisfies both the additivity and the fluctuation theorem. The additivity implies that the total entropy production can be decomposed into those of the subsets of transitions. The fluctuation theorem leads to a variety of thermodynamic relations. Therefore, our definition of the partial entropy production is a reasonable way to assign the entropy production to individual transitions. This approach would enhance our understanding of stochastic thermodynamics at the level of individual transition paths. For instance, we have derived a FDT for a pair of transition paths. Another possible application of our framework is to biological molecular motors, which are regarded as small heat engines converting fuel into work [15,40,41]. In this approach, for example, we would be able to reveal a bottleneck process in terms of the thermodynamic efficiency of motors, and its connection to the design principle of the molecular structure.

## ACKNOWLEDGMENTS

The authors thank Sosuke Ito and Kyogo Kawaguchi for fruitful discussion and useful comments. They also thank Jukka Pekola for helpful suggestions. This work is supported by JSPS KAKENHI Grants No. 25800217 and No. 22340114.

## APPENDIX A: DERIVATION OF THE FT FOR NONAUTONOMOUS MAXWELL'S DEMONS FROM EQ. (30)

We reproduce the FT for nonautonomous Maxwell's demons [33,34] from Eq. (30). We consider a bipartite system with state $w = (x,y)$. Intuitively, $x$ is the state of the engine and $y$ is the state of the memory of the demon. We assume that the transition rates satisfy

$$P(x \to x'; t|y) = 0 \quad (T_{2i} \leqslant t < T_{2i+1}), \tag{A1}$$

$$P(y \to y'; t|x) = 0 \quad (T_{2i+1} \leqslant t < T_{2i+2}), \tag{A2}$$

with $0 = T_0 < T_1 < T_2 < \cdots < T_{2M} = T$ (see also Fig. 6). In other words, only $y$ can change in the time interval $T_{2i} \leqslant t < T_{2i+1}$, where a measurement is performed by the demon; the measurement outcome is registered in the memory. Whereas only $x$ can change in the time interval $T_{2i+1} \leqslant t < T_{2i+2}$, where feedback control is performed; the engine evolves depending on the outcome registered in the memory.

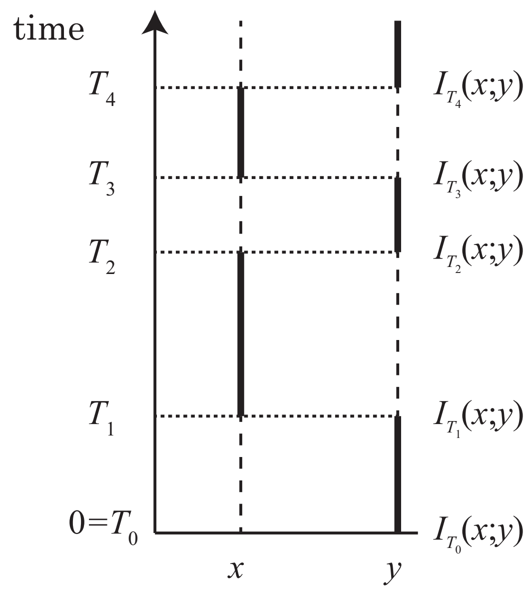


FIG. 6. Schematic of dynamics of the total system. The bold lines indicate the time intervals when the subsystem can evolve, whereas the dashed lines indicate the time intervals when the subsystem is frozen.

We apply Eq. (30) to this situation and calculate $\Delta I_x$. While $\Delta I_x$ is equal to zero for the time interval $T_{2i} \leqslant t < T_{2i+1}$,

$$F_x(x,y,t) = \frac{\partial}{\partial t} I_t(x;y) \tag{A3}$$

holds for the time interval $T_{2i+1} \leqslant t < T_{2i+2}$, because the probability distribution $P(x,y,t)$ changes only by transitions in $x$ during this time interval. Therefore, $\Delta I_x$ for $T_{2i+1} \leqslant t < T_{2i+2}$ becomes

$$\begin{aligned} \Delta I_x &= I_{x,\text{jump}} + \int_{T_{2i+1}}^{T_{2i+2}} \frac{\partial}{\partial t} I_t(x;y) dt \\ &= I_{T_{2i+2}}(x;y) - I_{T_{2i+1}}(x;y). \end{aligned} \tag{A4}$$

We then transform Eq. (30) into

$$\left\langle e^{-\sigma_x + \sum_i I_{T_{2i+2}}(x;y) - I_{T_{2i+1}}(x;y)} \right\rangle = 1, \tag{A5}$$

which is equivalent to the FT obtained in Refs. [33,34].

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[2] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[3] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[4] C. Jarzynski, J. Stat. Phys. **98**, 77 (2000).
[5] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[6] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[7] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **89**, 050601 (2002).
[8] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Science **296**, 1832 (2002).
[9] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Nature (London) **437**, 231 (2005).
[10] K. Hayashi, H. Ueno, R. Iino, and H. Noji, Phys. Rev. Lett. **104**, 218103 (2010).
[11] J. V. Koski, T. Sagawa, O-P. Saira, Y. Yoon, A. Kutvonen, P. Solinas, M. Mottonen, T. Ala-Nissila, and J. P. Pekola, Nat. Phys. **9**, 644 (2013).
[12] V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh, Nature (London) **445**, 523 (2007).
[13] V. Balzani, A. Credi, S. Silvi, and M. Venturi, Chem. Soc. Rev. **35**, 1135 (2006).
[14] K. K. Likharev, Proc. IEEE **87**, 606 (1999).
[15] S. Toyabe, H. Ueno, and E. Muneyuki, Europhys. Lett. **97**, 40004 (2012).
[16] A. Yildiz, M. Tomishige, A. Gennerich, and R. D. Vale, Cell **134**, 1030 (2008).

[17] E. A. Korobkova, T. Emonet, H. Park, and P. Cluzel, Phys. Rev. Lett. **96**, 058105 (2006).
[18] Y. Tu, Proc. Natl. Acad. Sci. USA **105**, 11737 (2008).
[19] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, Nat. Phys. **8**, 422 (2012).
[20] K. Sekimoto, Physica D **205**, 242 (2005).
[21] A. E. Allahverdyan, D. Janzing, and G. Mahler, J. Stat. Mech. (2009) P09011.
[22] M. Esposito, Phys. Rev. E **85**, 041125 (2012).
[23] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Phys. Rev. Lett. **110**, 040601 (2013).
[24] D. Hartich, A. C. Barato, and U. Seifert, J. Stat. Mech. (2014) P02016.
[25] J. M. Horowitz and M. Esposito, Phys. Rev. X **4**, 031015 (2014).
[26] D. Mandal and C. Jarzynski, Proc. Natl. Acad. Sci. USA **109**, 11641 (2012).
[27] J. M. Horowitz, T. Sagawa, and J. M. R. Parrondo, Phys. Rev. Lett. **111**, 010602 (2013).
[28] A. C. Barato, D. Hartich, and U. Seifert, Phys. Rev. E **87**, 042104 (2013).
[29] G. Diana and M. Esposito, J. Stat. Mech. (2014) P04010.
[30] S. Ito and T. Sagawa, Phys. Rev. Lett. **111**, 180603 (2013).
[31] A. C. Barato and U. Seifert, Phys. Rev. Lett. **112**, 090601 (2014).
[32] T. Nishizawa, S. Kita, A. D. Maturana, N. Furuya, K. Hirata, G. Kasuya, S. Ogasawara, N. Dohmae, T. Iwamoto, R. Ishitani, and O. Nureki, Science **341**, 168 (2013).
[33] T. Sagawa and M. Ueda, Phys. Rev. Lett. **109**, 180602 (2012).
[34] T. Sagawa and M. Ueda, New J. Phys. **15**, 125012 (2013).
[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley and Sons, New York, 1991).
[36] T. Bodineau and B. Derrida, Phys. Rev. Lett. **92**, 180601 (2004).
[37] T. Nemoto and S.-I. Sasa, Phys. Rev. E **83**, 030105 (2011).
[38] C. Maes, K. Netocny, and B. Wynants, Physica A **387**, 2675 (2008).
[39] N. Shiraishi, J. Stat. Phys. **152**, 336 (2013).
[40] R. Yasuda, H. Noji, K. Kinosita, and M. Yoshida, Cell **93**, 1117 (1998).
[41] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer Associates, Sunderland, MA, 2001).