Unified formalism for entropy production and fluctuation relations

Ying-Jen Yang [*] and Hong Qian[†]

Department of Applied Mathematics, University of Washington, Seattle, Washington 98195-3925, USA

Stochastic entropy production, which quantifies the difference between the probabilities of trajectories of a stochastic dynamics and its time reversals, has a central role in nonequilibrium thermodynamics. In the theory of probability, the change in the statistical properties of observables due to reversals can be represented by a change in the probability measure. We consider operators on the space of probability measures that induce changes in the statistical properties of a process, and we formulate entropy production in terms of these change-of-probability-measure (CPM) operators. This mathematical underpinning of the origin of entropy production allows us to achieve an organization of various forms of fluctuation relations: All entropy production has a nonnegative mean value, admit the integral fluctuation theorem, and satisfy a rather general fluctuation relation. Other results such as the transient fluctuation theorem and detailed fluctuation theorems then are derived from the general fluctuation relation with more constraints on the operator of entropy production. We use a discrete-time, discrete-state-space Markov process to draw the contradistinction among three reversals of a process: time reversal, protocol reversal, and the dual process. The properties of their corresponding CPM operators are examined, and the domains of validity of various fluctuation relations for entropy production in physics and chemistry are revealed. We also show that our CPM operator formalism can help us rather easily extend other fluctuation relations for excess work and heat, discuss the martingale properties of entropy production, and derive the stochastic integral formulas for entropy production in constant-noise diffusion process with Girsanov theorem. Our formalism provides a general and concise way to study the properties of entropy-related quantities in stochastic thermodynamics and information theory.

I. INTRODUCTION

Stochastic thermodynamics is a milestone extending equilibrium statistical physics to the nonequilibrium realm, and it could provide a general theory for emergent phenomena in mesoscopic systems [1–5]. It studies entropy production (EP), its relation to work and heat done by the system of interest [6], its statistical properties, such as expectations and martingale properties, and also how its probability density functions change after time reversals, called fluctuation theorems or fluctuation relations (FRs) [3,7–9].

Various distinct FRs for different EP have been studied in various settings including discrete-time Markov chains [6,10–12], continuous-time Markov chains (Markov jump processes) [13–18], diffusion processes (as stochastic differential equations or Langevin equations) [10,17–22], and even general stochastic processes [2,8,17,23–25]. To discuss a few, in Ref. [8], Crooks' fluctuation theorem for the total entropy production was introduced for systems with detailed balance and the conditions for it to hold was illustrated; in Ref. [23], the dissipation function from Evans and Searles [26] was rigorously shown to generally admit the transient fluctuation theorem (TFT); in Ref. [15], a detailed fluctuation theorem related to the involutive property of the change in probability (iDFT) was introduced but the correct condition

for total entropy production and nonadiabatic entropy production was not stated; in Ref. [3], the generalized Crooks' fluctuation theorem for nondetail balanced systems (rDFT) and TFT for different EP were discussed thoroughly for diffusion processes. iDFT was also discussed briefly; and in Ref. [27], rDFT was thoroughly discussed in general Markov processes for the three entropy productions except dissipation function.

The extensiveness of FTs calls for an unifying formalism to derive all FTs mentioned above in one general theory, organize their domain of validity comprehensively, and open ways to reveal more properties of EP. It is suggested in Refs. [2,17,23–25,28] that measure-theoretic probability theory pioneered by Komogorov [29] will do the trick. One key is the understanding that EP in physics is the *fluctuating relative entropy* for trajectories between the original process and the reversed process, with different EP given by different reversals or composite of reversals [3,6,8]. In the measure-theoretic formalism, EP is thus mathematically expressed as the negative logarithm of the Radon-Nikodym derivative (RND), which is a reweighting factor in taking expectation to change a probability measure from the original one to the other. It is this mathematical underpinning that enables us to arrive an organization of FRs and further derive/recover other properties with a deeper understanding of EP, e.g., its martingale properties [25,30,31].

Our paper thus serves as a comprehensive overview on how to understand EP and its statistical properties, primarily FRs, from measure-theoretic probability theory and our

---

[*]yangyj@uw.edu
[†]hqian@uw.edu

TABLE I. The five fluctuation relations (FRs) of the EP, S_ν, with the CPM operator ν discussed in this paper. IFT is the integral fluctuation theorem [7]; GFR is the general FR in Eq. (8c); TFT is the transient fluctuation theorem in Refs. [3,20,23,26], iDFT is the detailed fluctuation theorem in Ref. [15], and rDFT is the generalization of Crooks' fluctuation theorem [3,8,22,27]. We have shown in the text that TFT implies iDFT. This summarizes the hierarchical structure of the validity of the five FRs.

FR	Definition	EP involved	Validity and sufficient conditions
GFR	$\mathbb{P}\{S_\nu \in ds\} = e^s \mathbb{P}^\nu\{S_\nu \in ds\}$	$S_\nu(\omega) = \ln \frac{d\mathbb{P}}{d\mathbb{P}^\nu}(\omega)$	Generally valid
IFT	$\mathbb{E}[e^{-S_\nu}] = 1$		
rDFT	$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}^\nu\{\bar{S}_\nu[r(\omega)] \in -ds\}$	$\bar{S}_\nu[r(\omega)] = \ln \frac{d\mathbb{P}^R}{d\mathbb{P}^{R\nu}}(\omega)$	$\bar{S}_\nu[r(\omega)] = -S_\nu(\omega)$
iDFT	$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}^\nu\{\tilde{S}_\nu(\omega) \in -ds\}$	$\tilde{S}_\nu(\omega) = \ln \frac{d\mathbb{P}^\nu}{d\mathbb{P}^{\nu\nu}}(\omega)$	ν involutive on $\mathcal{P}$
TFT	$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}\{S_\nu(\omega) \in -ds\}$	$S_\nu(\omega) = \ln \frac{d\mathbb{P}}{d\mathbb{P}^\nu}(\omega)$	ν realized by an involutive map on Ω

change-of-probability-measure (CPM) operator theory. We revisit discussions on different time reversals, their associated EP in physics and chemistry, and derive various FRs from our general and concise approach. Our goal is to demonstrate that by adopting this CPM operator formalism, one can neatly derive many known results in the literature with more rigor and generality, achieve new understanding on EP and FRs, and reveal more properties of EP.

The outline of this paper is summarized below. In Sec. II, we briefly introduce measure-theoretic probability theory, the notion of CPM operator, and present general statistical properties for a general EP as a fluctuating relative entropy in general stochastic processes *without* Markovian assumption. With more constraints on the properties of the CPM operator, we deduce the general conditions for various known FRs and reveal a hierarchy of the domain of validity for various FRs with new relations recognized between them such as TFT $\Rightarrow$ iDFT. In Sec. III, we further use a discrete-time Markov chain to illustrate the contradistinction of three different time reversals of the dynamics that are prominent in physics and chemistry. The involutive and commutative properties of their corresponding CPM operators are discussed.

In Sec. IV, we apply the results in the previous two sections to discuss the properties of the four EP commonly considered in physics and chemistry. Notably, we discuss the difference between dissipation function and total entropy production and show that the two EP have nonzero difference in expectation for finite time interval in time homogeneous processes but have the same entropy production rate in infinitesimal time interval. The two seemingly contradicting results are resolved by noting the nonadditivity of the dissipation function when connecting time intervals.

We further demonstrate how properties of EP other than its FRs can be derived and extended in a rather straightforward way with this CPM operator formalism. We discuss the martingale properties of the four EP which can lead to more statistics on EP [25,30,31], and we extend the so-called differential FR for work and heat [32,33] to nonequilibrium systems at the end of Sec. IV. We also show how to use our CPM operator formalism and Girsanov theorem [2] to derive the stochastic integral formulas of the four EP for general time inhomogeneous constant-noise diffusion processes in Sec. V. The notations for the five heavily discussed FRs and the EP involved in them are summarized in Table I. Properties of the four EP we have discussed primarily in this paper are

also summarized in Table II. Finally, in Sec. VI, we discuss possible future extensions of our work.

II. GENERAL THEORY

To describe stochastic processes with a measure-theoretic probability theory, we start by specifying a tuple $(\Omega, \mathcal{F})$ called *measurable space*, where the sample space Ω collects all possible trajectories ω and the σ-algebra $\mathcal{F}$ collects all events of interest. Physical quantities, as observables, are then random variables defined on $(\Omega, \mathcal{F})$ [21]. The statistical properties of a stochastic process are further specified by a *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ with a probability measure $\mathbb{P}$ that assigns probabilities to events of interest in $\mathcal{F}$. See in subsection 1 of Appendix for a more thorough introduction.

The collection of all possible probability measures on a given measurable space $(\Omega, \mathcal{F})$ forms an affine space of probability measures $\mathcal{P}$ [35]. Each probability measure $\mathbb{P} \in \mathcal{P}$ corresponds to a stochastic process with specific statistical properties [36]. In this paper, we would assume $\mathcal{P}$ collects probability measures that are absolutely continuous to each other (also called *equivalent* in probability theory); i.e., if an event has zero probability for a stochastic process $\mathbb{P} \in \mathcal{P}$, then the event has zero probability under all $\mathbb{P}$ in $\mathcal{P}$.

A. Change of probability measure

With the statistical properties of stochastic processes specified by probability measures $\mathbb{P} \in \mathcal{P}$, the difference between the statistical properties of two processes is characterized by a change of probability measure (CPM) $\mathbb{P} \to \mathbb{P}^\nu \in \mathcal{P}$, which induces changes in the statistical properties of observables. This change in statistical properties can be mathematically represented by a random variable called the *Radon-Nikodym derivative* (RND), denoted as $\frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega)$ [35,37]. Intuitively, RND serves as a reweighting factor in taking expectation. For an arbitrary random variable $Y(\omega)$ defined on $(\Omega, \mathcal{F})$, it's expectation under $\mathbb{P}^\nu$, denoted as $\mathbb{E}^\nu[Y(\omega)]$, can be expressed by the reweighting factor and the previous expectation $\mathbb{E}[\cdot]$,

$$\mathbb{E}^\nu[Y(\omega)] = \mathbb{E}\left[Y(\omega)\frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega)\right]. \qquad (1)$$

Note that to get the probability density function of a random variable $X(\omega)$, we can let $Y(\omega)$ to be an indicator

TABLE II. Properties of the four entropy production (EP) for Markov processes. S_T is the dissipation function [3,26]; S_{tot} is the total EP [3,14,15,34]; Q_{hk} is the housekeeping heat [3,14] or called the adiabatic EP in Ref. [15]; and S_{na} is the nonadiabatic EP [15]. $\mathbb{E}[\cdot]$ denotes the expectation and the nonnegative expectation of entropy production is the classical second law of thermodynamics. P_0 is the initial distribution of the original process. P_t^R is the terminal distribution of the protocol reversed process, i.e., the distribution one gets by starting with initial distribution P_t and then marching with reversed order of transition proability matrices for t steps. TH stands for time homogeneous; SS stands for steady state; and DB means the steady state has detailed balance. Note that when TH and SS, i.e., when the system is in nonequilibrium steady state, we have $S_T \equiv S_{\text{tot}} \equiv Q_{\text{hk}}$ and $S_{\text{na}} = 0$.

EP	$= 0$ when	$\mathbb{E}[\cdot] \geqslant 0$, IFT, GFR	TFT	iDFT	rDFT	Additive in time	$e^{-[\cdot]}$ a martingale
$S_T = \ln \frac{d\mathbb{P}}{d\mathbb{P}^T}$	TH+SS+DB	Yes	Yes	Yes	TH+SS	TH+SS	TH+SS
$S_{\text{tot}} = \ln \frac{d\mathbb{P}}{d\mathbb{P}^{RT}}$	TH+SS+DB	Yes	TH+SS	TH+SS	if $P_0 = P_t^R$	Yes	TH+SS
$Q_{\text{hk}} = \ln \frac{d\mathbb{P}}{d\mathbb{P}^\dagger}$	TH+DB	Yes	TH+SS	Yes	Yes	Yes	Yes
$S_{\text{na}} = \ln \frac{d\mathbb{P}}{d\mathbb{P}^{R\dagger T}}$	TH+SS	Yes	TH+SS	TH+SS	if $P_0 = P_t^R$	Yes	TH+SS

function of $X(\omega)$ taking values in between x and $x + dx$ (denoted as $X(\omega) \in dx$ for simplicity). That is,

$$\mathbb{P}^\nu\{X(\omega) \in dx\} = \mathbb{E}^\nu[\mathbb{I}_{\{\omega:X(\omega)\in dx\}}] \tag{2a}$$

$$= \mathbb{E}\left[\mathbb{I}_{\{\omega:X(\omega)\in dx\}} \frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega)\right], \tag{2b}$$

with the indicator function $\mathbb{I}_A(\omega)$ returning 1 if ω is in the event A and returning 0 otherwise. Throughout the paper, we use capitalized letters for random variables and their corresponding lower letters for their values for a specific realization.

CPM and RND are key concepts in the theory of fluctuating entropy and EP [18,21,37]. For systems with a discrete sample space Ω (trajectory space when considering processes), RND reduces to the ratio of two probability mass functions, and for those with continuous sample spaces, it reduces to the ratio of two probability density functions. In either case, the RND is obtained from the ratio, which is defined on the codomain of a random variable, with random variables plugged back in [37].

In stochastic thermodynamics, a physical operation such as a time reversal in the dynamics is an operation that changes the probability measure $\mathbb{P} \in \mathcal{P}$ to a new probability measure $\mathbb{P}^\nu \in \mathcal{P}$ based on $\mathbb{P}$ alone. This means we are interested in a transformation ν on the space of all probability measures $\mathcal{P} \to \mathcal{P}$, and for each given measure $\mathbb{P} \in \mathcal{P}$, a physical operation defines a RND that is dependent upon the current process $(\Omega, \mathcal{F}, \mathbb{P})$. Thus, we consider operators ν that operates on $\mathbb{P}$, giving every measure $\mathbb{P} \in \mathcal{P}$ an image $\mathbb{P}^\nu \in \mathcal{P}$ and a corresponding RND. The new probability measure $\mathbb{P}^\nu$ is obtained by

$$\mathbb{P}^\nu\{A\} := \mathbb{E}\left[\frac{d\mathbb{P}^\nu}{d\mathbb{P}} \mathbb{I}_A\right], \tag{3}$$

$\forall A \in \mathcal{F}$ and $\forall \mathbb{P} \in \mathcal{P}$.

As we shall show, only very special operator ν defined on $\mathcal{P}$ can be represented as a result of a map μ from Ω to Ω. In those special cases, the map μ maps an event of interest $A \in \mathcal{F}$ to another, $\mu(A) \in \mathcal{F}$, which is obtained by replacing all the ωs in A by the $\mu(\omega)$s; e.g., if $A = \omega_1 \cup \omega_2$, then $\mu(A) = \mu(\omega_1) \cup \mu(\omega_2)$. The new measure is then

given by

$$\mathbb{P}^\nu\{A\} = \mathbb{P}\{\mu(A)\}. \tag{4}$$

B. Fluctuating entropy production

In stochastic thermodynamics, *fluctuating* EP of a CPM operator ν is defined as the negative natural logarithm of the RND,

$$S_\nu(\omega) := \ln \frac{d\mathbb{P}}{d\mathbb{P}^\nu}(\omega), \tag{5}$$

which is also a random variable [18,21,37–39]. The advantages of working with S_ν instead of the RND can be seen by its additivity, statistical properties and applications in information theory [40–42]. We note that S_ν is always finite given our assumption that $\mathcal{P}$ collects $\mathbb{P}$s that are absolute continuous to each other, i.e., $0 < \frac{d\mathbb{P}}{d\mathbb{P}^\nu}(\omega) < \infty$.

The prominent role of the reference measure $\mathbb{P}^\nu$ in the very definition of EP has a clear physical meaning. As the concept of energy, both entropy and EP are relative to a reference state, for which the choice is often question dependent. It is well understood that various different "free energies," as thermodynamic potentials, are determined by the physical settings of an equilibrium ensemble. In theories of dynamical systems, ergodic stationary measure, with a translational symmetry in time, has been widely used as a "natural" reference in physics and mathematics [43]. In stochastic dynamics, stationarity does not imply local time-reversal symmetry, which is often coupled to certain parity symmetry. In the work below, this is best illustrated as an involutive map on Ω and/or an involutive operation on $\mathcal{P}$.

The EP, $S_\nu(\omega)$, can be understood as the fluctuating relative entropy of trajectories ω with respect to the reference probability measure $\mathbb{P}^\nu$. It reflects the difference between two probability measures. If the stochastic process is symmetric under the operator ν, i.e., $\mathbb{P}^\nu = \mathbb{P}$, then $S_\nu(\omega) = 0, \forall \omega \in \Omega$. With different ν, we can have various different EP that are physically important, e.g. the nonadiabatic EP that is related to work and heat. It is therefore desirable to find the general statistical properties of a given EP given it's definition in Eq. (5).

C. Fluctuation relations

Directly from the definition of EP in Eq. (5), the following three key statistical properties of S_ν can be derived rather straightforwardly.

(a) *Nonnegative expectation:* By Jensen's inequality, the expectation of S_ν with respect to $\mathbb{P}$ is nonnegative, $\mathbb{E}[S_\nu] \geqslant 0$, and equality only holds when $\mathbb{P} = \mathbb{P}^\nu$ due to the strict convexity of negative logarithm. This result for EP in physical processes extends the classical second law of thermodynamics [17].

(b) *Integral Fluctuation Theorem* (IFT), or Jarzynski's equality [7]:

$$\mathbb{E}[e^{-S_\nu(\omega)}] = \mathbb{E}\left[\frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega)\right] = \mathbb{E}^\nu[1] = 1. \qquad (6)$$

(c) *General fluctuation relation* (GFR):

$$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}^\nu\{S_\nu(\omega) \in ds\}, \qquad (7)$$

where ds is a shorthand for the infinitesimal interval $(s, s + ds)$. This GFR states for the EP, $S_\nu = \ln \frac{d\mathbb{P}}{d\mathbb{P}^\nu}$, that quantifies the difference between $\mathbb{P}$ and $\mathbb{P}^\nu$, its probability densities under $\mathbb{P}$ and $\mathbb{P}^\nu$ are up to a exponential factor.

GFR can be derived by considering the probability density of S_ν under the new measure $\mathbb{P}^\nu$, characterizing the statistical properties of S_ν in the ν process:

$$\mathbb{P}^\nu\{S_\nu(\omega) \in ds\} = \mathbb{E}\left[\frac{d\mathbb{P}^\nu}{d\mathbb{P}}\mathbb{I}_{\{\omega:S_\nu(\omega) \in ds\}}\right] \qquad (8a)$$

$$= \mathbb{E}[e^{-S_\nu(\omega)}\mathbb{I}_{\{\omega:S_\nu(\omega) \in ds\}}] \qquad (8b)$$

$$= e^{-s}\mathbb{P}\{S_\nu(\omega) \in ds\}. \qquad (8c)$$

We will see below that most fluctuation relations discussed in the literature come directly from this GFR. We remark that the three properties above holds for *any* fluctuating relative entropy defined as the negative logarithm of a RND.

1. Detailed fluctuation theorems

The detailed fluctuation theorems that were considered in the literature [3,8,15,22] have the form of

$$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}^\nu\{\hat{S}_\nu(\omega) \in -ds\}, \qquad (9)$$

where $\hat{S}_\nu$ can be various different random variables under different considerations. The DFT comes directly from GFR in Eq. (8c) if there is an odd parity between the original random variable S_ν and the new random variable $\hat{S}_\nu$ under consideration,

$$\hat{S}_\nu(\omega) = -S_\nu(\omega). \qquad (10)$$

Two choices of $\hat{S}_\nu$ were discussed in the past, which we will briefly summarize below.

Recall that S_ν serves as a random variable that quantifies the effect of the CPM operator ν acting on $(\Omega, \mathcal{F}, \mathbb{P})$. When considering a different process $(\Omega, \mathcal{F}, \mathbb{P}^\eta)$, the EP that does the same for the η process as S_ν does for the original process should be given by replacing $\mathbb{P}$ with $\mathbb{P}^\eta$,

$$S_{\eta\nu}^\eta(\omega) := \ln \frac{d\mathbb{P}^\eta}{d\mathbb{P}^{\eta\nu}}(\omega), \qquad (11)$$

where $\mathbb{P}^{\eta\nu} = [\mathbb{P}^\eta]^\nu$ is operating η on $\mathbb{P}$ first and then applying ν on $\mathbb{P}^\eta$. The two $\hat{S}_\nu$s considered in the past [3,15] correspond to two different ηs.

A mathematically natural consideration for η is to take η as ν, which will lead to Esposito and Van den Broeck's detailed fluctuation theorem in Ref. [15]. In this setting, the odd parity requirement in Eq. (10) becomes an involutive requirement of the operator ν,

$$S_{\nu\nu}^\nu = \ln \frac{d\mathbb{P}^\nu}{d\mathbb{P}^{\nu\nu}} = -S_\nu = \ln \frac{d\mathbb{P}^\nu}{d\mathbb{P}} \Leftrightarrow \mathbb{P}^{\nu\nu} = \mathbb{P}. \qquad (12)$$

Denoting $S_{\nu\nu}^\nu$ as $\tilde{S}_\nu$, the detailed fluctuation theorem from the involutive property (iDFT) [15] then reads

$$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}^\nu\{\tilde{S}_\nu(\omega) \in -ds\}. \qquad (13)$$

In a physical process, the driving protocol of the system is determined by macroscopic thermodynamics parameters, and thermodynamics quantities such as heat and work are dependent upon the driving protocol. Therefore, in most of the physics literature, the $\hat{S}_\nu$ considered was given by first taking η as the macroscopic, protocol reversal R, as we will defined explicitly in Eq. (23), and then evaluating $\hat{S}_\nu$ at the order reversed trajectory $r(\omega)$ where r is a map from Ω to Ω that reverses the trajectory $\omega \in \Omega$. By denoting $\bar{S}_\nu(\omega) := S_{R\nu}^R(\omega)$ and using $\bar{S}_\nu[r(\omega)]$ for $\hat{S}(\omega)$, we get the generalized Crooks' fluctuation theorem [3,8,11,22,27],

$$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}^\nu\{\bar{S}_\nu[r(\omega)] \in -ds\}. \qquad (14)$$

To fix the terminology, we would refer this detailed fluctuation theorem as the rDFT. The odd parity requirement $\bar{S}_\nu[r(\omega)] = -S_\nu(\omega)$ turns out to be an involutive requirement on the operator R for the CPM operators we are interested in, as shown in Eqs. (39) and (58).

To check the validity of iDFT and rDFT, we should check the necessary and sufficient condition: for S_ν and $-\hat{S}_\nu$ to have the same probability density of under $\mathbb{P}^\nu$. The odd parity condition in Eq. (10) is a stricter condition on S_ν and $\hat{S}_\nu$ [44] and only a sufficient condition. To prove a DFT to be valid, we can show this sufficient condition to be true, but to prove a DFT to be invalid, we would need to show a necessary condition of it to be false. For simplicity, we will only check the the sufficient condition in Eq. (10) for a DFT in this paper. If an EP does not admit this odd-parity condition, then we will leave its DFT inconclusive and leave it for future consideration.

2. Transient fluctuation theorem

If an operator ν is involutive, i.e., $\mathbb{P}^{\nu\nu} = \mathbb{P}$, then we already know that S_ν admits iDFT. Now, if the operator ν is further realized by an involutive map $\mu : \Omega \to \Omega$ on the trajectory space as we have shown in Eq. (4), i.e., $\mu[\mu(\omega)] = \omega$, then we would have $S_\nu[\mu(\omega)] = \ln \frac{d\mathbb{P}^\nu}{d\mathbb{P}^{\nu\nu}}(\omega) = \tilde{S}_\nu(\omega) = \ln \frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega) = -S_\nu(\omega)$. Then, by Eq. (4), we have $\mathbb{P}^\nu\{S_\nu(\omega) \in ds\} = \mathbb{P}\{S_\nu[\mu(\omega)] \in ds\}$. With our GFR, we obtain the so-called *transient fluctuation theorem* (TFT) [1–3,23,26,45],

$$\mathbb{P}\{S_\nu(\omega) \in ds\} = e^s \mathbb{P}\{S_\nu(\omega) \in -ds\}. \qquad (15)$$

TFT is particularly important since it provides explicitly the asymmetry between having a positive and negative EP in the same process. The probability (density) of finding a positive EP is exponentially higher than the probability of finding a negative one.

The validity of TFT is easy to check since involutive property $\mathbb{P}^{\nu\nu} = \mathbb{P}$ is an necessary condition for TFT. If $\mathbb{P}^{\nu\nu} \neq \mathbb{P}$, then $\mathbb{E}[\ln\frac{d\mathbb{P}}{d\mathbb{P}^{\nu\nu}}] > 0 \Rightarrow \mathbb{E}[S_\nu] > \mathbb{E}[-S_{\nu\nu}^\nu]$ by Jensen's inequality. We then know S_ν and $-S_{\nu\nu}^\nu$ have different probability densities with respect to $\mathbb{P} \Rightarrow$ TFT is false. Here, we see that TFT is a sufficient condition for S_ν to have iDFT but not necessary. This is because not all involutive operator $\nu : \mathcal{P} \to \mathcal{P}$ can be realized by an involutive map $\mu : \Omega \to \Omega$ [46]

D. Summary

Treating a physical operation on stochastic processes as a CPM operator ν on probability space $\mathcal{P}$, we have characterized the change in the statistical properties of a physical operation via the negative natural logarithm of the RND, which one defines it as the EP S_ν. In fact, a hierarchy of the validity for FRs in general stochastic processes is revealed from our work as summarized in Table I:

(a) Nonnegative expectation, IFT [3,7], and GFR are generally true from the definition of S_ν.

(b) With $\hat{S}_\nu(\omega) = -S_\nu(\omega)$, we have DFTs. In the literature, $\hat{S}_\nu(\omega)$ was chosen to be $\tilde{S}_\nu(\omega) := S_{\nu\nu}^\nu(\omega)$ to get iDFT [15] if the CPM operator ν is involutive, or $\bar{S}_\nu[r(\omega)] := S_{R\nu}^R[r(\omega)]$ to get rDFT [3,8,11] if the protocol reversal operator R is involutive (for the ν we considered).

(c) Further with the CPM operator ν as an involutive map on the trajectory space, from $\Omega \to \Omega$, we have TFT [3,23,26].

The results above are true no matter the stochastic process has discrete or continuous state space $\mathcal{X}$, is with discrete or continuous time, is time homogeneous or not, or has any specific initial distribution such as the invariant distribution. The Markovian assumption is not even imposed except the definition of the protocol reversal R. Our derivation only relies on assuming all $\mathbb{P} \in \mathcal{P}$ to be absolute continuous to each other, the notation of CPM operator, the definition of EP, and conditions for more restricted FRs such as DFTs and TFT.

With these general results in hand, we shall consider EP in physics and chemistry, the reversal operators it corresponds to, and its fluctuation relations as examples in the following sections. We will start by introducing different reversals in Sec. III and then various EP with their fluctuation relations in Sec. IV. As we will see, our CPM operator notion clarifies the difference between different time reversals and between EP, especially between the dissipation function and the total entropy production, which are easily confused quantities.

III. DIFFERENT TYPES OF REVERSAL

EP in nonequilibrium physics and chemistry is introduced by comparing the original process to its "time reversal" [3]. The definition of a time reversal, however, is inevitably based on our understanding of the physics of time. See Ref. [47] for a discussion of "overdamped" versus "underdamped" thermodynamics.

Here, for general Markov processes, we would consider three different reversals. We use a discrete-time Markov chain with t time steps and discrete state space $\mathcal{X}$ as a paradigm. Markov processes in continuous time and continuous space will be discussed in Sec. V.

We use the colon notation $X_{0:t}$ to represent a sequence of random variables $(X_0, X_1, \cdots, X_t)$ and a specific trajectory $x_{0:t} = (x_0, x_1, x_2, \cdots, x_t)$. In our consideration, ω is a specific trajectory $x_{0:t}$ and our trajectory space Ω is given by the outer product of $(t+1)$ state spaces, $\mathcal{X} \otimes \mathcal{X} \otimes \cdots \otimes \mathcal{X}$ or simply $\mathcal{X}^{t+1}$. The full probabilistic description of the state variables $X_{0:t}$ is given by their joint probability denoted as

$$P_{0:t}(x_{0:t}) := \mathbb{P}\{X_0 = x_0, X_1 = x_1, \cdots, X_t = x_t\}. \quad (16)$$

The marginal probabilities $P_n(x_n) := \mathbb{P}\{X_n = x_n\}$ and the conditional probabilities $P_{m|n}(x_m|x_n) := \mathbb{P}\{X_m = x_m | X_n = x_n\}$ can be computed from the joint probability. We would denote the transition matrix at the nth time step as

$$M_n(x_n|x_{n-1}) := P_{n|n-1}(x_n|x_{n-1}). \quad (17)$$

With this notion, the joint probability for a time-inhomogeneous Markov process,

$$P_{0:t}(x_{0:t}) = P_0(x_0) \prod_{n=1}^{t} M_n(x_n|x_{n-1}), \quad (18)$$

is determined by the *driving protocol*, which constitutes the initial distribution P_0 and all the transition matrices M_n for $n = 1, 2, \cdots, t$. For each transition matrix, we also assumed the existence of a unique invariant distribution π_n satisfying $\sum_{i \in \mathcal{X}} \pi_n(i) M_n(j|i) = \pi_n(j)$.

A. Time reversal

The time reversal of a Markov chain for $n = 0, 1, \cdots, t$ is conventionally defined by a change of random variable

$$X_n^T(\omega) = X_{t-n}(\omega), \quad (19)$$

where we use superscript T to represent time reversal. This definition of X_n^T can be treated as the random variable induced by a map on the trajectory space, $r : \Omega \to \Omega$,

$$X_n^T(\omega) = X_n[r(\omega)], \quad (20)$$

where the map r reverses the order of a trajectory ω, $r(x_{0:t}) = x_{t:0}$. Given a specific trajectory $\omega = x_{0:t}$, the state variable $X_n(\omega)$ is understood as the observed state of the system at time n. We can then clearly see the equivalence between these two definitions:

$$X_n^T(x_{0:t}) = X_{t-n}(x_{0:t}) = x_{t-n} = X_n(x_{t:0}). \quad (21)$$

By the equivalence between change of random variable and change of probability measure [37], instead of regarding time reversal as a change of random variable, we can also characterize the time reversal as a change of probability measure with a CPM operator T. The CPM operator T is realized by the map $r : \Omega \to \Omega$ on the trajectory space. The

joint probability after time reversal is thus given by

$$P_{0:t}^T(x_{0:t}) = \mathbb{P}^T\{X_0(\omega) = x_0, \cdots, X_t(\omega) = x_t\} \quad (22a)$$

$$= \mathbb{P}\{X_0[r(\omega)] = x_0, \cdots X_t[r(\omega)] = x_t\} \quad (22b)$$

$$= \mathbb{P}\{X_t(\omega) = x_0, \cdots, X_0(\omega) = x_t\} \quad (22c)$$

$$= P_{t:0}(x_{0:t}) \quad (22d)$$

$$= \mathbb{P}\{X_0(\omega) = x_t, \cdots X_t(\omega) = x_0\} \quad (22e)$$

$$= P_{0:t}(x_{t:0}). \quad (22f)$$

We see that the joint probability of finding $x_{0:t}$ in the time reversed process is the same as the joint probability of finding the order-reversed trajectory $x_{t:0}$ in the original process. The assumption that the order-reversed trajectory has a nonzero probability in the original process is the *microscopic reversible* assumption required in Ref. [8].

When using a change of probability measure perspective, the meaning of the random variables X_n is preserved as the nth state of the process. The changes in its statistical properties are due to the change in probability measure. An interesting analog to these two equivalence ways of characterizing the change is the Schrödinger's and Heisenberg's pictures of quantum mechanics [37]. We also note that, from the results above, it can be mathematically shown that the time reversed Markov chain is still Markovian but will be time inhomogeneous even if the original process is time homogeneous.

B. Protocol reversal

The joint probability of a Markov chain is determined by the *driving protocol*, P_0 and M_n, $n = 1, 2, \cdots, t$. Thus, we can consider the process where we used the terminal distribution P_t as our new initial distribution and reverse the temporal order of the transition matrices. We shall call this reversal the *protocol reversal* of the process and denote the corresponding CPM operator as R. The new joint distribution is then given by

$$P_{0:t}^R(x_{0:t}) = P_t(x_0) \prod_{n=1}^{t} M_{t+1-n}(x_n | x_{n-1}). \quad (23)$$

Compare to the time reversal T which is a "time reversal" at the microscopic/trajectory level, protocol reversal R is rather a "time reversal" at the macroscopic/thermodynamics level.

The "time reversal" that was considered by most of the previous studies on fluctuation relations [15,18,22] is in fact the composition of the two reversals we have just introduced: R and T, denoted as $\mathbb{P}^{RT}$. The joint distribution is given by

$$\left[P_{0:t}^R\right]^T(x_{0:t}) = P_{0:t}^R(x_{t:0}) \quad (24a)$$

$$= P_t(x_t) \prod_{n=1}^{t} M_{t+1-n}(x_{t-n} | x_{t-n+1}). \quad (24b)$$

This computation in Eq. (24b) actually gives us a convenient result when working on composite CPMs with time reversal as the last operation, $\mathbb{P} \to \mathbb{P}^{\nu T}$. The joint probability for such composite operators is given by evaluating at the order-reversed trajectory:

$$P_{0:t}^{\nu T}(x_{0:t}) = P_{0:t}^{\nu}(x_{t:0}). \quad (25)$$

Note that the two operators R and T do not generally commute, $P_{0:t}^{RT} \neq P_{0:t}^{TR}$ [48].

C. The dual process

The last reversal we consider in this paper is by introducing the driving protocol that reverses the probability flux in the invariant steady state *at each time step*. This new process is called the *dual* process [3,11] of the original process. For a time homogeneous process, the dual process is equivalent to the time reversal of the process if the process starts and stays in the invariant steady state. However, for a general time inhomogeneous process, the correspondence between the dual and the time reversal can only be drawn within each given time step.

For the nth time step where the transition matrix is M_n and the invariant distribution is π_n, the probability flux from state i to state j is given by the joint probability difference between $i \to j$ and $j \to i$:

$$J_n(i, j) := \pi_n(i)M_n(j|i) - \pi_n(j)M_n(i|j). \quad (26)$$

The probability flux can then be reversed, $J_n^\dagger(i, j) = -J_n(i, j)$, by replacing M_n with its dual matrix

$$M_n^\dagger(j|i) = \frac{\pi_n(j)}{\pi_n(i)} M_n(i|j). \quad (27)$$

The definition of a dual process is thus given by replacing all the M_n with $M_n^\dagger$:

$$P_{0:t}^\dagger(x_{0:t}) = P_0(x_0) \prod_{n=1}^{t} M_n^\dagger(x_n | x_{n-1}). \quad (28)$$

It can be shown from Eq. (27) that M_n and $M_n^\dagger$ have the same invariant distribution π_n.

Recall the detailed balance condition is given by

$$\pi_n(i)M_n(j|i) = \pi_n(j)M_n(i|j), \forall i, j \in \mathcal{X}, \quad (29)$$

which is equivalent to $J_n^\dagger = J_n = 0$ and $M_n^\dagger = M_n$. Therefore, comparing the dual process to the original one directly reveals whether the system possesses detailed balance or not. Detailed balance systems are invariant under the CPM operator †.

D. Involutive properties of the reversals

Considering reversals of a process, it is natural to ask whether we can recover the original process by applying the reversal twice or not, i.e., in mathematical terms, whether the operator is involutive or not. As we have shown above, the involutive properties of the CPM operator for a EP are the keys for the EP to have FRs.

It is rather straightforward to show that both T and † are involutive. The time reversal T is involutive since the map $r : \Omega \to \Omega$ is involutive. We can verify this by computing $P_{0:t}^{TT}(x_{0:t}) = P_{0:t}^T(x_{t:0}) = P_{0:t}(x_{0:t})$. The dual reversal † is involutive by computing $P_{0:t}^{\dagger\dagger}(x_{0:t}) = P_0^\dagger(x_0) \prod_{n=1}^{t} M_n^{\dagger\dagger}(x_n | x_{n-1})$. From the joint probability in Eq. (28), we get $P_0^{\dagger\dagger} = P_0$ and one can show $M_n^{\dagger\dagger} = M_n$ by $\pi_n^\dagger = \pi_n$.

Finally, the protocol reversal R is *not* involutive in general. To see this, we start by

$$P_{0:t}^{RR}(x_{0:t}) = P_t^R(x_0) \prod_{n=1}^{t} M_{t+1-n}^R(x_n|x_{n-1}). \tag{30}$$

We thus need to compute P_t^R and M_{t+1-n}^R from the joint probability given in Eq. (23). It is straightforward to check that the latter is given simply by

$$M_{t+1-n}^R(j|i) = M_n(j|i). \tag{31}$$

However, the terminal distribution of the protocol reversed process is generally not the initial distribution of the original process, $P_t^R \neq P_0$ [16]. This can be seen by a time homogeneous Markov chain where P_t^R is given by P_t further evolved by t more steps, which gives us P_{2t} not P_0. Thus, we have

$$P_{0:t}^{RR}(x_{0:t}) = \frac{P_t^R(x_0)}{P_0(x_0)} P_{0:t}(x_{0:t}). \tag{32}$$

From this, it is also clear that if $P_t^R = P_0$, then the protocol reversal R becomes involutive [49].

IV. ENTROPY PRODUCTION IN PHYSICS AND CHEMISTRY

With different reversals and their corresponding CPM operators introduced, we are now ready to consider different EP in physics and chemistry and their fluctuation relations. We already knew that every EP as a fluctuating relative entropy has a nonnegative expectation and admit both IFT and GFR. Thus, we would mainly discuss the rDFT, iDFT, and TFT for various different EP in physics and chemistry.

A. Dissipation function S_T

The EP that corresponds to the time reversal T is historically called the *dissipation function* by Evans and Searles [3,26], a term goes back to Onsager,

$$S_T(\omega) := \ln \frac{d\mathbb{P}}{d\mathbb{P}^T}(\omega) = \ln \frac{P_{0:t}[X_{0:t}(\omega)]}{P_{0:t}[X_{t:0}(\omega)]}. \tag{33}$$

We note that the dissipation function does *not* satisfy additive properties when connecting two time intervals, i.e., for $0 < s < t$, we have

$$S_T(x_{0:t}) \neq S_T(x_{0:s}) + S_T(x_{s:t}). \tag{34}$$

Since the CPM operator T is realized by an involutive map $r : \Omega \to \Omega$. We thus know S_T admits both TFT and iDFT. The TFT of S_T has been discussed in Refs. [3,23,26]. However, S_T does not satisfy the odd-parity, sufficient condition for rDFT. We note that the dissipation function S_T on the protocol reversed process R is given by $\bar{S}_T(\omega) = \ln \frac{d\mathbb{P}^R}{d\mathbb{P}^{RT}}(\omega)$, which gives

$$\bar{S}_T[r(\omega)] = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^{RTT}}(\omega) = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^R}(\omega) \tag{35a}$$

$$= -\bar{S}_T(\omega) \neq -S_T(\omega). \tag{35b}$$

Unless pathologically $\bar{S}_T[r(\omega)] \neq -S_T(\omega)$ but $\mathbb{P}^T\{\bar{S}_T[r(\omega)] \in -ds\} = \mathbb{P}^T\{S_T(\omega) \in -ds\}$, the dissipation function S_T would not admit rDFT.

B. Total entropy production S_{tot}

The *total entropy production* discussed in Refs. [3,14,15,34] is given by composing protocol reversal R and then the time reversal T,

$$S_{\text{tot}}(x_{0:t}) = \ln \frac{d\mathbb{P}}{d\mathbb{P}^{RT}}(x_{0:t}) = \ln \frac{d\mathbb{P}}{d\mathbb{P}^C}(x_{0:t}) \tag{36a}$$

$$= \ln \frac{P_0(x_0)M_1(x_1|x_0) \cdots M_t(x_t|x_{t-1})}{P_t(x_t)M_t(x_{t-1}|x_t) \cdots M_1(x_0|x_1)}, \tag{36b}$$

where we have denoted the composition of R and T as a composite operator C, i.e., $\mathbb{P}^C = [\mathbb{P}^R]^T$. We note that the total entropy production satisfies additive property when connecting two time intervals, i.e., for $0 < s < t$,

$$S_{\text{tot}}(x_{0:t}) = S_{\text{tot}}(x_{0:s}) + S_{\text{tot}}(x_{s:t}). \tag{37}$$

The composite operator C is generally not involutive [50]. Thus, S_{tot} does not admit TFT and the odd-parity, sufficient condition for iDFT. For rDFT, we note

$$\bar{S}_{\text{tot}}(\omega) := \ln \frac{d\mathbb{P}^R}{d\mathbb{P}^{RC}}(\omega) = \ln \frac{d\mathbb{P}^R}{d\mathbb{P}^{RRT}}(\omega), \tag{38}$$

and thus

$$\bar{S}_{\text{tot}}[r(\omega)] = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^{RRTT}}(\omega) = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^{RR}}(\omega), \tag{39}$$

which becomes $-S_{\text{tot}}(\omega)$ if R is involutive, i.e., $\mathbb{P}^{RR} = \mathbb{P}$. Hence, if R is involutive, S_{tot} admits rDFT. Recall that the requirement for R to be involutive is the terminal distribution of the protocol reversed process to recover the initial distribution of the original process, i.e., $P_t^R = P_0$. This condition was discussed in Refs. [3,8,18].

C. Difference between S_T and S_{tot}

The physical meanings of the two EP discussed above are clearly different. For a given trajectory $\omega = x_{0:t}$, the dissipation function S_T quantifies the probability difference between observing the trajectory $\omega = x_{0:t}$ and the order reversal of it $r(\omega) = x_{t:0}$ in the original process. However, the total EP, S_{tot}, quantifies the probability difference between observing a trajectory $x_{0:t}$ in the original process and observing the order-reversed trajectory $x_{t:0}$ in the protocol-reversed process R.

In time homogeneous processes where all $M_n = M$, $\forall n = 1, 2, \cdots, t$, their difference gives another EP,

$$D := S_T - S_{\text{tot}} = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^T} \tag{40a}$$

$$= \ln \frac{P_t(X_t)M_t(X_{t-1}|X_t) \cdots M_1(X_0|X_1)}{P_0(X_t)M_1(X_{t-1}|X_t) \cdots M_t(X_0|X_1)} \tag{40b}$$

$$= \ln \frac{P_t(X_t)}{P_0(X_t)} = \ln \frac{d\mathbb{P}}{d\mathbb{P}^d}, \tag{40c}$$

where the time homogeneous assumption kicks in to eliminate all the Ms. The corresponding joint probability in the measure $\mathbb{P}^d$ would be

$$P_{0:t}^d(x_{0:t}) = \frac{P_0(x_t)}{P_t(x_t)} P_{0:t}(x_{0:t}). \tag{41}$$

This implies that the expectation of the dissipation function is bigger than the total EP,

$$\mathbb{E}[D] \geqslant 0 \Leftrightarrow \mathbb{E}[S_T] \geqslant \mathbb{E}[S_{\text{tot}}], \tag{42}$$

for time homogeneous processes. By Jensen's inequality, we also know the equality holds if and only if $P_t = P_0$.

In the time homogeneous cases, the expectation of D is actually the Kullback-Leibler divergence between the terminal distribution P_t and initial distribution P_0. Without time homogeneity, D is generally not an EP, since $\frac{P_0(x_t)M_1(x_{t-1}|x_t)\cdots M_t(x_0|x_1)}{P_t(x_t)M_t(x_{t-1}|x_t)\cdots M_1(x_0|x_1)}P_{0:t}(x_{0:t})$ is not generally normalizable and thus not a joint probability. This also indicates that the expectation of D in time inhomogeneous systems does not generally have a definite sign.

The infinitesimal time interval limit $t \to 0$ can be taken to consider the *entropy production rate* (EPR) [14,17]. In such a limit, we see that $\mathbb{E}[D]$ is in $O(t^2)$, whereas both $\mathbb{E}[S_T]$ and $\mathbb{E}[S_{\text{tot}}]$ are $e_{\text{p}}t + O(t^2)$ with the same entropy production rate, e_{p}. This is one of the reason why the contradistinction between S_T and S_{tot} was not clear in the past literature. See in subsection 2 of Appendix for derivation.

Recall that when connecting two time intervals, S_{tot} satisfies the additive property whereas S_T does not! Thus, when one integrate the EPR of S_T and S_{tot} over time, the value one gets is $\mathbb{E}[S_{\text{tot}}]$ not $\mathbb{E}[S_T]$. This resolves the seemly contradicting results that the expectations of the two EP are different for any finite time interval but with the same rate in infinitesimal time interval.

D. Total heat, excess heat, and housekeeping heat dissipation

One of the most important breakthrough in nonequilibrium thermodynamics is the discovery of the heat dissipation in nonequilibrium steady state (NESS) and its statistical properties [2,10,13,14,19,20,23,45,51,52]. By the energy conservation, this amount is also the amount of energy required to sustain the NESS, historically called the *housekeeping heat* Q_{hk} and is conventionally chosen to be positive as a heat *dissipated by* the system.

To understand the housekeeping heat with our Markov chain paradigm, we consider a time inhomogeneous t-step Markov chain with a transition matrix M_n. The total heat dissipation for a trajectory $\omega = x_{0:t}$ is given by the transition matrices,

$$Q(\omega) = \ln \prod_{n=1}^{t} \frac{M_n(x_n|x_{n-1})}{M_n(x_{n-1}|x_n)}, \tag{43}$$

for general systems even nondetailed balanced [8,10]. Without detailed balance, the probability flux between two states i, j at NESS at time n is nonzero:

$$0 \neq J_n(j|i) = \pi_n(i)M_n(j|i) - \pi_n(j)M_n(i|j). \tag{44}$$

This is physically originated from the nonconservative force that sustains NESS [3]. With the invariant distribution π_n in nondetailed balanced systems, we can use the so called (fluctuating) *nonequilibrium potential* $\Phi_n(x)$ based on the invariant distribution at each time step [3],

$$\Phi_n(x) = -\ln \pi_n(x), \tag{45}$$

which would be the potential of mean force if one chooses the free energy of the entire system (the system of interest and the environment) as the zero potential energy reference point [5]. It is worth noting that for time homogeneous diffusion processes in the thermodynamic limit, this *definition* of the nonequilibrium potential, with proper scaling, gives us a Lyapunov function of the emerged, dissipative deterministic dynamics [53].

The changes of the nonequilibrium potential due to a transition in microstate would then be *the excess heat dissipation*:

$$Q_{\text{ex}}(x_{0:t}) = \sum_{n=1}^{t} \Phi_n(x_{n-1}) - \Phi_n(x_n) \tag{46a}$$

$$= \ln \prod_{n=1}^{t} \frac{\pi_n(x_n)}{\pi_n(x_{n-1})}. \tag{46b}$$

The housekeeping heat (dissipation) is given by the difference between the total heat dissipation Q and the excess heat dissipation Q_{ex}:

$$Q_{\text{hk}} = Q - Q_{\text{ex}}. \tag{47}$$

This will become

$$Q_{\text{hk}}(\omega) = \sum_{n=1}^{t} \ln \frac{M_n(X_n|X_{n-1})}{M_n^{\dagger}(X_n|X_{n-1})} \tag{48a}$$

$$= \ln \frac{P_0(X_0)}{P_0(X_0)} \prod_{n=1}^{t} \frac{M_n(X_n|X_{n-1})}{M_n^{\dagger}(X_n|X_{n-1})} \tag{48b}$$

$$= \ln \frac{d\mathbb{P}}{d\mathbb{P}^{\dagger}}(\omega), \tag{48c}$$

which shows that it is an EP with corresponding the CPM operator $\dagger$, thus has a nonnegative expectation, and admits both IFT and GFR. If the system possesses detailed balance, then $Q_{\text{hk}} = 0$. Straight from the definition in Eq. (48a) that, similar to S_{tot}, the housekeeping heat is also additive when connecting time intervals:

$$Q_{\text{hk}}(x_{0:t}) = Q_{\text{hk}}(x_{0:s}) + Q_{\text{hk}}(x_{s:t}). \tag{49}$$

With detailed balance, $Q_{\text{hk}} = 0$ and the excess heat dissipation Q_{ex} reduces to the total heat dissipation Q [19].

Since $\dagger$ is involutive, the housekeeping heat Q_{hk} generally admits iDFT. To show whether it admits rDFT or not, we compute the housekeeping heat in the protocol-reversed process evaluated at the order-reversed trajectory:

$$\bar{Q}_{\text{hk}}[r(\omega)] = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^{R\dagger T}}(\omega) \tag{50a}$$

$$= \ln \prod_{n=1}^{t} \frac{M_{t+1-n}(X_{t-n}|X_{t+1-n})}{M_{t+1-n}^{\dagger}(X_{t-n}|X_{t+1-n})} \tag{50b}$$

$$= \ln \prod_{n=1}^{t} \frac{M_{t+1-n}^{\dagger}(X_{t+1-n}|X_{t-n})}{M_{t+1-n}(X_{t+1-n}|X_{t-n})} \tag{50c}$$

$$= -Q_{\text{hk}}(\omega), \tag{50d}$$

where we have used $\frac{M_k(j|i)}{M_k^{\dagger}(j|i)} = \frac{M_k^{\dagger}(i|j)}{M_k(i|j)}$. Hence, the housekeeping heat $Q_{\text{hk}}(\omega)$ generally admits rDFT.
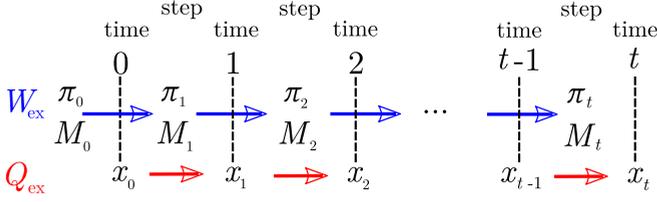
FIG. 1. Excess heat and excess work definitions in a discrete time Markov chain. x_n is the state of the system at time n. M_n is the transition matrices at the nth time step, and π_n is the corresponding unique invariant distribution.

Last, for time homogeneous processes, the housekeeping heat is the dissipation function S_T starting at the invariant distribution π since, as we have discussed, the dual process is equivalent to the time reversal if the system reaches the invariant state. Since S_T admits TFT for arbitrary initial distribution, this also means that Q_{hk} also admits TFT for time homogeneous processes starting at the invariant steady state [2,13,20,23,51].

E. Exergy, excess work, and nonadiabatic entropy production

With the notion of the nonequilibrium potential Φ_n defined in Eq. (45) at each step for general nondetailed balance systems, the *excess work* done *by* the system for a trajectory $x_{0:t}$ is then the difference between the nonequilibrium potential dissipation $-\Delta\Phi$ and the excess heat dissipated Q_{ex} as illustrated in Fig. 1,

$$W_{ex}(x_{0:t}) = -\Delta\Phi(x_{0:t}) - Q_{ex}(x_{0:t}) \tag{51a}$$

$$= \ln \prod_{n=0}^{t-1} \frac{\pi_{n+1}(x_n)}{\pi_n(x_n)}, \tag{51b}$$

which is the change in the nonequilibrium potential due to change in the transition matrices (and thus the corresponding invariant distribution).

The fluctuating relative entropy between P_n and π_n is given by

$$F_n(x) = \Phi_n(x) - S_n(x) = \ln \frac{P_n(x)}{\pi_n(x)}. \tag{52}$$

For systems with detailed balance, the sum of this fluctuating relative entropy and the free energy defined in classical equilibrium thermodynamics was called *nonequilibrium free energy* in Ref. [54]. For this reason, the relative entropy also got a name *nonsteady-state addition* (to free energy) in Ref. [12]. Furthermore, it was shown in Ref. [55] that this relative entropy itself could be understood physically as a "free energy" as well. To avoid possible confusion on the terminology in this paper, we would follow Ref. [56] and call this (fluctuating) *exergy* from now on. For a trajectory $\omega = x_{0:t}$, the exergy that got absorbed by the system is then

$$\Delta F(x_{0:t}) = \ln \frac{P_t(x_t)}{\pi_t(x_t)} - \ln \frac{P_0(x_0)}{\pi_0(x_0)}. \tag{53}$$

The difference between the exergy dissipation $-\Delta F$ and the excess work done *by* the system W_{ex} then gives us the *nonadiabatic* EP, S_{na}, defined in Ref. [15]:

$$S_{na}(\omega) = -\Delta F(\omega) - W_{ex}(\omega) \tag{54a}$$

$$= Q_{ex}(\omega) + \Delta S(\omega) \tag{54b}$$

$$= \ln \frac{d\mathbb{P}}{d\mathbb{P}^{R\dagger T}}(\omega). \tag{54c}$$

The equivalence between the last two lines can be seen by a direct computation of $P_{0:t}^{R\dagger T}(x_{0:t})$. We note that S_{na} becomes the exergy dissipation $-\Delta F$ in time homogeneous processes since $W_{ex} = 0$ in time homogeneous processes. We also note that the nonadiabatic EP is also additive when connecting time interval,

$$S_{na}(x_{0:t}) = S_{na}(x_{0:s}) + S_{na}(x_{s:t}), \tag{55}$$

which is obvious from its relation to W_{ex} and ΔF. Note that S_{na} reduces to the dissipative work W_d defined in Refs. [6,7] for systems with detailed balance.

Similar to the total EP S_{tot}, the nonadiabatic EP S_{na} admit neither the TFT nor the odd-parity, sufficient condition for iDFT since the composite operator $R\dagger T$ is not involutive in general. For rDFT, we compute

$$\bar{S}_{na}[r(\omega)] = \ln \frac{d\mathbb{P}^{RT}}{d\mathbb{P}^{RR\dagger}}(\omega) \tag{56a}$$

$$= \ln \frac{P_t(X_t)}{P_t^R(X_0)} \prod_{n=1}^{t} \frac{M_n(X_{n-1}|X_n)}{M_n^\dagger(X_n|X_{n-1})}, \tag{56b}$$

where the denominator can be obtained by using Eq. (32) for $\mathbb{P}^{RR}$ and apply $\dagger$ on it. Now, since

$$\frac{M_n(X_{n-1}|X_n)}{M_n^\dagger(X_n|X_{n-1})} = \frac{\pi_n(X_{n-1})}{\pi_n(X_n)}, \tag{57}$$

we get

$$\bar{S}_{na}[r(\omega)] = \ln \frac{P_0(X_0)}{P_t^R(X_0)} - S_{na}(\omega). \tag{58}$$

We see that the condition for the odd parity to hold, $\bar{S}_{na}[r(\omega)] = -S_{na}(\omega)$, is $P_0 = P_t^R$, i.e., R to be involutive! Hence, similar to the total EP S_{tot}, the nonadiabatic EP S_{na} admits rDFT if the CPM operator R is involutive. The rDFT of S_{na} is an extension to Crooks' fluctuation theorem [6,8,11].

F. Martingale properties of entropy production

With our measure-theoretic understanding of EP, more statistical properties of EP could be found by revealing more on the mathematical properties of its corresponding RND and CPM operators. For example, by recognizing the RND of Q_{hk}, $\exp(-Q_{hk})$, is a martingale, a statistics of the infimum of Q_{hk} is introduced in Refs. [25,30,31]. Here, we shall discuss the martingale properties of the four EP and the conditions for the exponential of the negative of them to be a martingale.

In our discrete time Markov chain paradigm, a functional $M(X_{0:t})$ is a martingale if it satisfies

$$\mathbb{E}[M(X_{0:t})|X_{0:s}] = M(X_{0:s}), \tag{59}$$

$\forall s \in \{0, 1, \cdots, t\}$. It can be shown rather straightforwardly that since S_T is not additive in time, $\exp(-S_T)$ would not

generally be a martingale. Thus, in the following discussion, we will focus on S_{tot}, Q_{hk}, and S_{na}.

We note that all of S_{tot}, Q_{hk}, and S_{na} are additive in time. Therefore, for $\exp(-S_\nu)$ to be a martingale where $\nu = RT$, $\dagger$, or $R\dagger T$, we need

$$\mathbb{E}[e^{-S_\nu(X_{0:t})}|X_{0:s}] = e^{-S_\nu(X_{0:s})}\mathbb{E}[e^{-S_\nu(X_{s:t})}|X_{0:s}] \quad (60a)$$

$$= e^{-S_\nu(X_{0:s})}\mathbb{E}[e^{-S_\nu(X_{s:t})}|X_s] \quad (60b)$$

$$\overset{!}{=} e^{-S_\nu(X_{0:s})}. \quad (60c)$$

We thus want to have

$$\mathbb{E}[e^{-S_\nu(X_{s:t})}|X_s] = 1, \quad (61)$$

$\forall s \in \{0, 1, \cdots, t\}$.

By using the definition for $\nu = RT$, $\dagger$, and $R\dagger T$, one will find

$$\mathbb{E}[e^{-S_{\text{tot}}(X_{s:t})}|X_s] = \frac{P_{\text{end}}^R(X_s)}{P_s(X_s)}, \quad (62a)$$

$$\mathbb{E}[e^{-Q_{\text{hk}}(X_{s:t})}|X_s] = 1, \text{ and} \quad (62b)$$

$$\mathbb{E}[e^{-S_{\text{na}}(X_{s:t})}|X_s] = \frac{P_{\text{end}}^{R\dagger}(X_s)}{P_s(X_s)}, \quad (62c)$$

where

$$P_{\text{end}}^R(x_s) = \sum_{x_{s+1:t}} P_t(x_t) \prod_{n=0}^{t-s-1} M_{t-n}(x_{t-n-1}|x_{t-n}) \text{ and}$$

$$P_{\text{end}}^{R\dagger}(x_s) = \sum_{x_{s+1:t}} P_t(x_t) \prod_{n=0}^{t-s-1} M_{t-n}^{\dagger}(x_{t-n-1}|x_{t-n}) \quad (63)$$

are the terminal distributions of the processes R and $R\dagger$ defined on the time interval $s : t$.

This shows that $\exp(-Q_{\text{hk}})$ is always a martingale which implies that Q_{hk} is a submartingale satisfying

$$\mathbb{E}[Q_{\text{hk}}(X_{0:t})|X_{0:s}] \geqslant Q_{\text{hk}}(X_{0:s}), \quad (64)$$

by the convexity of negative logarithm. Since s is arbitrary in $0 : t$, the right-hand side of Eqs. (62a) and (62c) needs to be 1 for all x_s and s. This means that $\exp(-S_{\text{tot}})$ or $\exp(-S_{\text{na}})$ are only martingale when the reversal R or $R\dagger$ recovers all the marginal distributions in a reversed order: $P_t \to P_{t-1} \to \cdots \to P_0$ in the reversed process, which is not generally true. One exception is when the dynamics is time homogeneous and also starts with the invariant distribution such that $P_n = \pi$, $\forall n \in \{0, 1, \cdots, t\}$.

G. Summary

The properties of the four EP in physics and chemistry including various FRs we have discussed above have been summarized in Table II. Here, we note that the three reversals T, R, and $\dagger$ are actually related. By direct computation, one can get

$$\frac{d\mathbb{P}}{d\mathbb{P}^{RT}} = \frac{d\mathbb{P}}{d\mathbb{P}^{\dagger}}\frac{d\mathbb{P}}{d\mathbb{P}^{R\dagger T}}, \quad (65)$$

which leads to the famous decomposition of the total EP introduced in Ref. [17]:

$$S_{\text{tot}} = S_{\text{na}} + Q_{\text{hk}}. \quad (66)$$

Since $Q_{\text{hk}} = 0$ when the system possesses detailed balance, we also see that $S_{\text{tot}} \equiv S_{\text{na}}$ in detailed balance systems.

H. Two fluctuation relations for heat and work

As an another demonstration on how the formalism can help us obtain statistical properties of EP-related quantities, we note that there is another generally valid FR called *differential fluctuation theorem* for the work done *by* the system W, derived in Refs. [32,33] and experimental verified in Ref. [57] for detailed balanced systems. Here we provide a more general derivation to extend it to nondetailed balanced systems. The key observation is that the excess work defined in Eq. (51b) always has odd parity under the composite CPM operator $C \equiv RT$, i.e., $\bar{W}_{\text{ex}}[r(\omega)] = -W_{\text{ex}}(\omega)$. Knowing this, we can consider the joint probability of the work W_{ex}, initial state X_0, and the terminal state X_t,

$$P_{W_{\text{ex}},X_0,X_t}(w, x_0, x_t) = \mathbb{E}[\mathbb{I}_{\{W_{\text{ex}} \in dw, X_0 = x_0, X_t = x_t\}}]. \quad (67)$$

The joint probability under the measure $\mathbb{P}^{R\dagger T}$ is then given by

$$P_{W_{\text{ex}},X_0,X_t}^{R\dagger T}(w, x_0, x_t) = \mathbb{E}[e^{-S_{\text{na}}}\mathbb{I}_{\{W_{\text{ex}} \in dw, X_0 = x_0, X_t = x_t\}}] \quad (68a)$$

$$= e^{w+\Delta F(x_0,x_t)}\mathbb{E}[\mathbb{I}_{\{W_{\text{ex}} \in dw, X_0 = x_0, X_t = x_t\}}] \quad (68b)$$

$$= e^{w+\Delta F(x_0,x_t)}P_{W_{\text{ex}},X_0,X_t}(w, x_0, x_t), \quad (68c)$$

where we have used the fact that the exergy increment $\Delta F(\omega) = \ln \frac{P_t(X_t)}{\pi_t(X_t)} - \ln \frac{P_0(X_0)}{\pi_0(X_0)}$ is a function of X_0 and X_t.

By using $\bar{W}_{\text{ex}}[r(\omega)] = -W_{\text{ex}}(\omega)$, we thus have

$$\frac{P_{W_{\text{ex}},X_0,X_t}(w, x_0, x_t)}{P_{\bar{W}_{\text{ex}},X_0,X_t}^{R\dagger}(-w, x_t, x_0)} = e^{-\Delta F(x_0,x_t)-w}. \quad (69)$$

Also, a similar differential fluctuation theorem for the excess heat dissipated $Q_{\text{ex}}(\omega)$ can also be derived,

$$\frac{P_{Q_{\text{ex}},X_0,X_t}(q, x_0, x_t)}{P_{\bar{Q}_{\text{ex}},X_0,X_t}^{R\dagger}(-q, x_t, x_0)} = e^{q+\Delta S(x_0,x_t)}, \quad (70)$$

since $\bar{Q}_{\text{ex}}[r(\omega)] = -Q_{\text{ex}}(\omega)$ and $S_{\text{na}} = -\Delta F - W_{\text{ex}} = Q_{\text{ex}} + \Delta S$.

V. ENTROPY PRODUCTION IN CONSTANT-NOISE DIFFUSION PROCESSES

We will now briefly go through how to use measure-theoretic probability theory and the CPM operator formalism to derive the stochastic integral formulas for the four EP in time inhomogeneous constant-noise diffusion processes. With the formulas, one can further derive the expression for the moments of EP with Ito calculus. In a constant-noise diffusion process in $\mathbb{R}^n$, the probability density of the microstate variable $\mathbf{X}_t$, $p(\mathbf{x}, t)$, is governed by the Fokker-Planck equation,

$$\frac{\partial}{\partial t}p(\mathbf{x}, t) = -\nabla \cdot \mathbf{J}[p(\mathbf{x}, t)], \quad (71)$$

with probability flux given by

$$\mathbf{J}[p(\mathbf{x}, t)] = \mathbf{b}(\mathbf{x}, t)p(\mathbf{x}, t) - \mathbf{D}\nabla p(\mathbf{x}, t). \qquad (72)$$

and the stochastic trajectory governed by the stochastic differential equation,

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t, t)dt + \mathbf{\Gamma}d\mathbf{W}_t, \qquad (73)$$

where $\mathbf{D} = \mathbf{\Gamma}\mathbf{\Gamma}^{\mathsf{T}}/2$ is a constant diffusion matrix (T denoting transpose) and $\mathbf{W}_t$ is the n dimensional Wiener processes (Brownian motions) with each component independent and having unit strength of noise.

To derive the stochastic integral formula for the four EP from their RND definitions, we rely on Girsanov theorem [2] to give us the RND to "kill" the drift $\mathbf{b}(\mathbf{X}_t, t)$ in the dynamics of $\mathbf{X}_t$. Given a time interval $t \in [0, \tau]$ and the RND,

$$\frac{d\mathbb{P}^B}{d\mathbb{P}} = e^{-\frac{1}{2}\int_0^\tau \mathbf{D}^{-1}\mathbf{b}\cdot d\mathbf{X}_t + \frac{1}{4}\int_0^\tau \mathbf{b}\cdot\mathbf{D}^{-1}\mathbf{b}dt} \qquad (74a)$$

$$= e^{-\frac{1}{2}\int_0^\tau \mathbf{D}^{-1}\mathbf{b}\circ d\mathbf{X}_t + \frac{1}{2}\int_0^\tau \nabla\cdot\mathbf{b}dt + \frac{1}{4}\int_0^\tau \mathbf{b}\cdot\mathbf{D}^{-1}\mathbf{b}dt}, \qquad (74b)$$

the probability density of the process $\mathbf{X}_t$ under the measure $\mathbb{P}^B$ satisfies $\frac{\partial}{\partial t}p^B(\mathbf{x}, t) = \nabla \cdot \mathbf{D}\nabla p^B(\mathbf{x}, t)$, i.e., $\mathbf{X}_t$ is a Brownian motion with strength $\mathbf{\Gamma}$ under $\mathbb{P}^B$. The CPM operator B "kills" the drift by changing of the probability measure from $\mathbb{P}$ to $\mathbb{P}^B$. We note that the first stochastic integral in Eq. (74a) is an Ito integral and the first stochastic integral in Eq. (74b) is a Stratonovich integral. We rewrote the stochastic integral into a Stratonovich form for later convenience when considering time reversals.

A. Dissipation function

To use Eq. (74b) to get S_T, we perform the decomposition of its corresponding RND:

$$\frac{d\mathbb{P}}{d\mathbb{P}^T} = \frac{d\mathbb{P}}{d\mathbb{P}^B}\frac{d\mathbb{P}^B}{d\mathbb{P}^{BT}}\frac{d\mathbb{P}^{BT}}{d\mathbb{P}^T}. \qquad (75)$$

We already have the first term $\frac{d\mathbb{P}}{d\mathbb{P}^B} = -\frac{d\mathbb{P}^B}{d\mathbb{P}}$ by Eq. (74b). For the third term, it can be rewritten as

$$\frac{d\mathbb{P}^B}{d\mathbb{P}}[r(\omega)] = \exp\left\{\frac{1}{2}\int_0^\tau \mathbf{D}^{-1}\mathbf{b}(\mathbf{X}_t, \tau - t) \circ d\mathbf{X}_t \right.$$
$$+ \frac{1}{2}\int_0^\tau \nabla \cdot \mathbf{b}(\mathbf{X}_t, \tau - t)dt$$
$$\left. + \frac{1}{4}\int_0^\tau \mathbf{b} \cdot \mathbf{D}^{-1}\mathbf{b}(\mathbf{X}_t, \tau - t)dt\right\}, \qquad (76)$$

where we have used the change of integration variable for the three integrals in the exponent using

$$\int_0^\tau \mathbf{b}(\mathbf{X}_{\tau-t}, t) \circ d\mathbf{X}_{\tau-t} = -\int_0^\tau \mathbf{b}(\mathbf{X}_t, \tau - t) \circ d\mathbf{X}_t \text{ and}$$

$$\int_0^\tau f(\mathbf{X}_{\tau-t}, t)dt = \int_0^\tau f(\mathbf{X}_t, \tau - t)dt. \qquad (77)$$

We note that the first equality in Eq. (77) is true since it is a Stratonovich integral. If the stochastic integral is in any other

integration scheme, then the change of integration variable will lead to a change of integral scheme.

The second term in Eq. (75) is the RND of the dissipation function in Brownian motion. Since $\mathbf{X}_t$ is drift-free under $\mathbb{P}^B$ and the noise strength is a constant, the conditional probability for having a path $\omega = x_{0:\tau}$ conditioning that it starts at x_0 is the same as the one for the reversed path $r(\omega) = x_{\tau:0}$ conditioning that it starts at x_t due to the symmetry of Brownian motion. We thus see that the T operator only changes $\mathbb{P}^B$ by the initial probability density p_0 from evaluating at $\mathbf{X}_0$, $p_0(\mathbf{X}_0)$, to evaluating at $\mathbf{X}_\tau$, $p_0(\mathbf{X}_\tau)$, and the change of measure is completed by

$$\frac{d\mathbb{P}^B}{d\mathbb{P}^{BT}} = \frac{p_0(\mathbf{X}_0)}{p_0(\mathbf{X}_\tau)}, \qquad (78)$$

where we have also used that the operator B leaves the initial probability density invariant, $p_0^B = p_0$.

Putting Eqs. (75)–(78) together, we arrive the dissipation function for a time inhomogeneous constant-noise diffusion:

$$S_T = \ln\frac{p_0(\mathbf{X}_0)}{p_0(\mathbf{X}_\tau)} + \int_0^\tau \mathbf{D}^{-1}\bar{\mathbf{b}} \circ d\mathbf{X}_t$$
$$+ \frac{1}{2}\int_0^\tau \delta(\nabla \cdot \mathbf{b})dt + \frac{1}{4}\int_0^\tau \delta(\mathbf{b} \cdot \mathbf{D}^{-1}\mathbf{b})dt, \qquad (79)$$

where we have used the notation

$$\bar{\mathbf{b}}(\mathbf{X}_t, t) = [\mathbf{b}(\mathbf{X}_t, t) + \mathbf{b}(\mathbf{X}_t, \tau - t)]/2, \text{ and} \qquad (80a)$$

$$\delta f(\mathbf{X}_t, t) = f(\mathbf{X}_t, \tau - t) - f(\mathbf{X}_t, t). \qquad (80b)$$

The stochastic integral expression of the dissipation function S_T in Eq. (79) in general time inhomogeneous constant-noise diffusion process, as far as we know, is a new result.

We note that if the drift is *symmetric in the interval* $t \in [0, \tau]$, i.e., $\mathbf{b}(\cdot, t) = \mathbf{b}(\cdot, \tau - t)$, then the two terms with δ become zero. The expression for dissipation function then reduces to

$$S_T = \ln\frac{p_0(\mathbf{X}_0)}{p_0(\mathbf{X}_\tau)} + Q \qquad (81)$$

where $Q = \int_0^\tau \mathbf{D}^{-1}\mathbf{b} \circ d\mathbf{X}_t$ is the heat dissipation. As time homogeneous process being symmetric in the interval, this is consistent with Ref. [3]. It can also be seen from this expression that, consistent with our understanding in our Markov chain paradigm, S_T is not generally additive when connecting time intervals.

B. Total entropy production

With $S_{\text{tot}} = \ln\frac{d\mathbb{P}}{d\mathbb{P}^{RT}}$, we derive the formula of it in a similar decomposition of the corresponding RND,

$$\frac{d\mathbb{P}}{d\mathbb{P}^{RT}} = \frac{d\mathbb{P}}{d\mathbb{P}^B}\frac{d\mathbb{P}^B}{d\mathbb{P}^{RBT}}\frac{d\mathbb{P}^{RBT}}{d\mathbb{P}^{RT}}. \qquad (82)$$

We can find the second and the third RNDs in a similar way. One has

$$\frac{d\mathbb{P}^B}{d\mathbb{P}^{RBT}} = \frac{p_0(\mathbf{X}_0)}{p_\tau(\mathbf{X}_\tau)}, \qquad (83)$$

since RT only changes the initial distribution for $\mathbb{P}^B$ and

$$
\frac{d\mathbb{P}^{RBT}}{d\mathbb{P}^{RT}}(\omega) = \exp\left\{\frac{1}{2}\int_0^\tau \mathbf{D}^{-1}\mathbf{b}(\mathbf{X}_t, t) \circ d\mathbf{X}_t \right.
$$
$$
+ \frac{1}{2}\int_0^\tau \nabla \cdot \mathbf{b}(\mathbf{X}_t, t)dt
$$
$$
\left. + \frac{1}{4}\int_0^\tau \mathbf{b}\cdot\mathbf{D}^{-1}\mathbf{b}(\mathbf{X}_t, t)dt\right\}, \quad (84)
$$

by replacing $\mathbf{b}(\cdot, t)$ with $\mathbf{b}^R(\cdot, t) = \mathbf{b}(\cdot, \tau - t)$ in Eq. (76).

Putting these together, we can get the famous entropy change decomposition in time inhomogeneous constant-noise diffusion processes,

$$
S_{\text{tot}} = \ln\frac{p_0(\mathbf{X}_0)}{p_\tau(\mathbf{X}_\tau)} + Q = \Delta S + Q. \quad (85)
$$

With this, it is obvious that S_{tot} is additive in time.

We also note that the difference between S_T and S_{tot} in processes with time symmetric $\mathbf{b}$, i.e., when $\mathbf{b}^R = \mathbf{b}$, is

$$
S_T - S_{\text{tot}} = \ln\frac{p_\tau(\mathbf{X}_\tau)}{p_0(\mathbf{X}_\tau)}. \quad (86)
$$

This is consistent with what we had in Eq. (40c), leading one to conclude $\mathbb{E}[S_T] \geqslant \mathbb{E}[S_{\text{tot}}]$ when $\mathbf{b}^R = \mathbf{b}$.

C. Housekeeping heat and nonadiabatic entropy production

Using the CPM perspective to get Q_{hk} has already been rigorously studied in Ref. [2], here we shall briefly revisit the derivation and rely on the relation $S_{\text{tot}} = Q_{\text{hk}} + S_{\text{na}}$ to get the formula for S_{na}.

In time inhomogeneous diffusion process, we could consider the instantaneous stationary probability density π_t such that

$$
\nabla \cdot \mathbf{J}[\pi_t] = 0, \quad (87)
$$

where $\mathbf{J}[\pi_t] = \mathbf{b}(\mathbf{X}_t, t)\pi_t - \mathbf{D}\nabla\pi_t$. From the fact that the adjoint process is given by reversing $\mathbf{J}[\pi_t]$, we see that the effect of the operator † on $\mathbf{b}$ is given by

$$
\mathbf{b}^\dagger = -\mathbf{b} - 2\mathbf{D}\nabla(-\ln\pi_t). \quad (88)
$$

Then, we use the decomposition

$$
\frac{d\mathbb{P}}{d\mathbb{P}^\dagger} = \frac{d\mathbb{P}}{d\mathbb{P}^B}\frac{d\mathbb{P}^B}{d\mathbb{P}^{\dagger B}}\frac{d\mathbb{P}^{\dagger B}}{d\mathbb{P}^\dagger} \quad (89)
$$

and note that the second term is 1 since neither † nor B changes the initial distribution and B kills the drift no matter it is $\mathbf{b}$ or $\mathbf{b}^\dagger$. Also, the third term can be evaluated by Eq. (74b) by substituting $\mathbf{b}$ with $\mathbf{b}^\dagger$. Further, using Eq. (87) to simplify the expression one got from above, one would arrive at

$$
Q_{\text{hk}} = \int_0^\tau [\mathbf{D}^{-1}\mathbf{b} + \nabla(-\ln\pi_t)] \circ d\mathbf{X}_t
$$
$$
= Q - Q_{\text{ex}}, \quad (90)
$$

where $Q_{\text{ex}} = -\int_0^\tau \nabla(-\ln\pi_t) \circ d\mathbf{X}_t$ is the excess heat dissipation in diffusion. The nonadiabatic entropy production S_{na}

would then be given by

$$
S_{\text{na}} = S_{\text{tot}} - Q_{\text{hk}} = W_{\text{ex}} - \Delta F
$$
$$
= \int_0^\tau \frac{\partial(-\ln\pi_t)}{\partial t}dt - \ln\frac{p_\tau(X_\tau)\pi_0(X_0)}{\pi_\tau(X_\tau)p_0(X_0)}. \quad (91)
$$

VI. DISCUSSION

In this paper, we characterize the difference between the statistical properties of the original stochastic process and the one after reversal by a change of probability measure, an analog to Schrödinger's picture on quantum mechanics [37]. A change in statistical properties from a physical operation is represented by an operator ν operating on a probability measure $\mathbb{P}$ in the probability measure space $\mathcal{P}$. With our mathematically more general and concise CPM formalism, we have presented a comprehensive study of the properties of EP, including FRs. Sufficient conditions for the FRs of the four EP are summarized in Table II. Importantly, a hierarchy of the generality for FRs in general stochastic processes can be revealed from our work: both IFT and GFR are generally true; rDFT and iDFT require odd parity symmetry with different $\hat{S}_\nu$ as stated in Ref. [3]; and TFT further requires the CPM operator to be realized by an involutive map on the trajectory space Ω. This hierarchical structure of the domain of validity for FRs reveals relation between FRs such as TFT implies iDFT.

We further demonstrate how to obtain other properties of EP from its logarithm RND definitions, such as its martingale properties, and distinguish the difference between dissipation function introduced by Evans and Searles [26] and the total entropy production. The "paradox" that the two EP have the same entropy production rate but with nonnegative difference in expectation for finite time interval in time homogeneous processes is resolved by noting the failure of time additivity for the dissipation function. Stochastic integration expressions for the two EP are also derived in general time inhomogeneous constant-noise diffusion to better see the contradistinction of their physical meaning and properties.

It is important to note that throughout this paper, we have assumed the state variables X_n to have even parity under the time reversal, i.e., they are position-like physical quantities. One extension to our work is to consider variables that have odd parity under time reversal such as velocity [58–60].

The unit of EP S_ν is chosen to be in the natural unit of information (nat) throughout the paper [8,42]. And temperature of the heat reservoir is assumed to be constant. When considering diffusion processes in Sec. V, we have also restricted ourselves to constant-noise diffusion processes. If the strength of noise $\mathbf{\Gamma}$ varies in spacetime, then the use of Girsanov theorem to derive the formula for S_T becomes more involved since $\frac{d\mathbb{P}^B}{d\mathbb{P}^{BT}}$ becomes less straightforward. We note that Jiang, Qian, and Qian have used Girsanov theorem to derive the housekeeping heat Q_{hk} for autonomous, nonconstant noise diffusion in Ref. [2]. In there, we have a similar form of Eq. (74) for non-constant noise diffusion and can thus use the same derivation to argue that Eq. (85) still holds for autonomous, nonconstant noise diffusion processes. To obtain the integral equation for S_T, however, will need other methods. One may need the path integral formalism to obtain

the probability "density" of a diffusion path [61], or make use of the Riemann geometry introduced by $\mathbf{\Gamma}$ and $\mathbf{D}$ to consider a constant-noise diffusion on a curved space [62].

The general theory we presented in Sec. II is in fact a general result for *fluctuating relative entropy* and its statistical properties. It can also be applied to entropies defined in information theory [42]. For example, suppose we have a finite state space $\mathcal{X} \equiv \Omega$, i.e., there is a fundamental state random variable X that labels every $\omega \in \Omega$ with a unique real number, we can then choose P_X^ν to be the uniform distribution, i.e., $\mathbb{P}^\nu$ as the Lebesgue measure, to get an entropy corresponding to the maximum entropy $\ln \|\Omega\|$ minuses the fluctuating Shannon entropy of the system,

$$H(\omega) = \ln \|\Omega\| - \{-\ln P_X[X(\omega)]\}, \qquad (92)$$

where $\|\Omega\|$ represents the size of the sample space. Another example would be to have $\Omega = \mathcal{X} \otimes \mathcal{Y}$ and consider the fluctuating mutual information between two random variables X and Y,

$$I(\omega) = \ln \frac{P_{X,Y}[X(\omega), Y(\omega)]}{P_X[X(\omega)] P_Y[Y(\omega)]}. \qquad (93)$$

Our theory immediately implies that both $H(\omega)$ and $I(\omega)$ have nonnegative expectation, and admit IFT and GFR.

The combination of information theory and stochastic thermodynamics is a natural application and extension to the theory [54,63–65]. One example would be to consider random transition matrices for stochastic driving protocol. With randomness in transition matrices, the second law of thermodynamics is refined by incorporating the mutual information between the past and present, and the mutual information between the present and the future, giving us a thermodynamics of prediction [65]. In our change-of-measure formalism, we could extend $(\Omega, \mathcal{F})$ to include all possible transition matrices. Such future work could be conducted by considering the theories of random dynamical system for Markov chains [66,67].

APPENDIX

1. Modern probability theory and Radon-Nikodym derivative

Here we briefly introduce several key concepts of modern measure-theoretic probability theory pioneered by Kolmogorov [29]. For recent textbook introductions, see Refs. [68,69]. One of the most important concept in the theory is that one specifies the sample space Ω, the σ-algebra $\mathcal{F}$, and a probability (measure) $\mathbb{P}$ to start any probabilistic discussion on stochastic processes.

An outcome ω is a particular result of a random trial. A sample space Ω is the collection of all possible outcomes. We can think of Ω as the state space of our system of interest. In this paper, a "state" of the system of interest would be a trajectory of a given finite time interval t in a stochastic process. Ω then becomes the space of trajectories.

An event of interest A is a subset of the sample space that we seek the probability of. For example, in this paper, it can be the set of periodic trajectories $\omega \in \Omega$, i.e., we might ask "What is the probability of having a periodic trajectory?" The collection of all events of interest of Ω is called σ-algebra [70], usually denoted as $\mathcal{F}$.

Note that when we are interested in an event, the complement of it, A^c, i.e., when A does not happen, is also under our interest. Moreover, if we are already interested in a bunch of events $A_1, A_2^c, \ldots$, then the countable union of them, i.e., the event when at least one of it happen, is also interesting to us. Note that with complement and countable union, the countable intersections are automatically included in $\mathcal{F}$.

As a result, an σ-algebra of Ω as a collection of all events of interest must have the following three properties in its very definition: (1) containing the empty set $\emptyset$ where nothing happen, (2) being closed under countable union, and (3) being closed under complement. One of the many reasons to introduce σ-algebra $\mathcal{F}$ along with the sample space Ω is that since we are interested in knowing the probability of every events of interest, which is given by a probability measure $\mathbb{P}$, we need the collection of all events of interest $\mathcal{F}$ to actually define a $\mathbb{P}$.

A probability measure $\mathbb{P}$ measures the probability of an event of interest $A \in \mathcal{F}$ by assigning the event a real value, $\mathbb{P}\{A\}$, between 0 and 1. Since an event $A \in \mathcal{F}$ is a subset of Ω, the probability of an event A can also be represented as an expectation of the indicator function of an event A, $\mathbb{I}_A(\omega)$, i.e.,

$$\mathbb{P}\{A\} = \mathbb{E}[\mathbb{I}_A(\omega)], \qquad (A1)$$

where $\mathbb{I}_A(\omega)$ gives 1 if $\omega \in A$ and 0 otherwise, and $\mathbb{E}[\cdot]$ denotes the expectation. This expectation expression of $\mathbb{P}\{A\}$ is convenient in many calculations, especially when we consider a change of probability measure as shown in Eqs. (2b) and (3).

Physical observable as a random variable $X(\omega)$ is not just a function from Ω to $\mathbb{R}$. Since we are interested in the statistical properties of $X(\omega)$, we would like to use the probability measure $\mathbb{P}$ defined on the $\mathcal{F}$ of Ω to get the probability of an event such as $\{X(\omega) \in (a, b); a, b \in \mathbb{R}\}$ or for $X(\omega)$ in any countable union/complements of open intervals. The collection of any set that can be formed from open intervals in $\mathbb{R}$, complement of them, and/or countable union of them is called the *Borel sets* of $\mathbb{R}$, denoted as $\mathcal{B}(\mathbb{R})$. The requirement of being able to use a probability measure $\mathbb{P}$ to get that statistical properties of $X(\omega)$ is then the requirement of *measurability*. That is, for $\mathbb{P}\{X(\omega) \in B\}$ to make sense, we need that $\forall B \in \mathcal{B}(\mathbb{R})$, $\{X(\omega) \in B\}$ as a set of ω to be in $\mathcal{F}$. Hence, a random variable is not a function of Ω but a *measurable function* $X(\omega): \Omega \mapsto \mathbb{R}$ defined on $(\Omega, \mathcal{F})$ so that $\{X(\omega) \in B\} \in \mathcal{F}$, $\forall B \in \mathcal{B}(\mathbb{R})$.

The very fact that X is defined on a pair $(\Omega, \mathcal{F})$, without a specific $\mathbb{P}$, is a manifestation of the Schrödinger's picture of changes in statistical properties. For two different stochastic processes, the change in the statistical properties of a random

variable $X(\omega)$, the difference in its distribution P_X and P_X^ν, is due to the change of probability measure $\mathbb{P} \to \mathbb{P}^\nu$ [37]. Without $\mathbb{P}$, the pair $(\Omega, \mathcal{F})$ is called *measurable space*. With a $\mathbb{P}$ defined, the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called the *probability space*. In this paper, a stochastic process is specified by a probability space.

With a given measurable space $(\Omega, \mathcal{F})$, there are many $\mathbb{P}$ that can be considered. Radon-Nikodym derivative (RND) is a special random variable that can be used to change one probability measure $\mathbb{P}$ to another $\mathbb{P}^\nu$, denoted as $\frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega)$. As explained in the main context, the probability of an event $A \in \mathcal{F}$ in the new probability measure is given by the expectation representation shown in Eq. (A1),

$$\mathbb{P}^\nu\{A\} = \mathbb{E}^\nu[\mathbb{I}_A(\omega)] = \mathbb{E}\left[\frac{d\mathbb{P}^\nu}{d\mathbb{P}}(\omega)\mathbb{I}_A(\omega)\right]. \quad \text{(A2)}$$

For Ω whose ω in Ω can be 1-1 labeled by $x \in \mathbb{Z}$, we have the RND reduced to the ratio of probability mass functions $P^\nu(x)/P(x)$, since

$$\mathbb{P}^\nu\{A\} = \mathbb{E}^\nu[\mathbb{I}_A] = \sum_{x\in\mathbb{Z}} \mathbb{I}_A(x)P^\nu(x) \quad \text{(A3a)}$$

$$= \sum_{x\in\mathbb{Z}} \mathbb{I}_A(x)\frac{P^\nu(x)}{P(x)}P(x) \quad \text{(A3b)}$$

$$= \mathbb{E}\left[\mathbb{I}_A(x)\frac{P^\nu(x)}{P(x)}\right]. \quad \text{(A3c)}$$

For those Ω whose ω in Ω can be 1-1 labeled by $x \in \mathbb{R}$ such as diffusion processes, we have the RND reduced to the ratio of probability density functions $\rho^\nu(x)/\rho(x)$, since

$$\mathbb{P}^\nu\{A\} = \mathbb{E}^\nu[\mathbb{I}_A] = \int_{\mathbb{R}} \mathbb{I}_A(x)\rho^\nu(x)dx \quad \text{(A4a)}$$

$$= \int_{\mathbb{R}} \mathbb{I}_A(x)\frac{\rho^\nu(x)}{\rho(x)}\rho(x)dx \quad \text{(A4b)}$$

$$= \mathbb{E}\left[\mathbb{I}_A(x)\frac{\rho^\nu(x)}{\rho(x)}\right]. \quad \text{(A4c)}$$

2. Entropy production rates for S_T and S_{tot}

With our trajectory-based definitions for dissipation function S_T and total entropy production S_{tot} given as

$$S_T(\omega) = \ln\frac{d\mathbb{P}}{d\mathbb{P}^T}(\omega) \text{ and } S_{\text{tot}}(\omega) = \ln\frac{d\mathbb{P}}{d\mathbb{P}^{RT}}(\omega), \quad \text{(B1)}$$

where $\omega = x_0 x_1 \cdots x_t$, here we show that although the two EP differ in finite time interval as shown in our main context, they have the same rate in expectation in infinitesimal time interval $t \to 0$.

With an infinitesimal time interval $t \to 0$, we only need to consider one infinitesimal time step where the system state goes from i to j,

$$S_T(i, j) = \ln\frac{P_0(i)M(j|i)}{P_0(j)M(i|j)} \quad \text{(B2)}$$

and

$$S_{\text{tot}}(i, j) = \ln\frac{P_0(i)M(j|i)}{P_t(j)M(i|j)}. \quad \text{(B3)}$$

The difference between the two EP is

$$D(i, j) := S_T(i, j) - S_{\text{tot}}(i, j) = \ln\frac{P_t(j)}{P_0(j)}. \quad \text{(B4)}$$

With $t \to 0$, M is approaching to an identity matrix:

$$M(j|i) = \delta_{i,j} + q(j|i)t + O(t^2). \quad \text{(B5)}$$

Up to the linear order, we can write

$$P_t(j) = \sum_k P_0(k)M(j|k) \quad \text{(B6a)}$$

$$= P_0(j) + t\sum_k P_0(k)q(j|k) + O(t^2). \quad \text{(B6b)}$$

In the literature, entropy production rates of S_T and S_{tot} are defined as their limiting re-scaled expectation: $\lim_{t\to 0}\frac{\mathbb{E}[S_T]}{t}$ and $\lim_{t\to 0}\frac{\mathbb{E}[S_{\text{tot}}]}{t}$. Therefore, let us compute the expectation of S_T, S_{tot}, and D as an asymptotic series of small t. Since $D = S_T - S_{\text{tot}}$, we have

$$\mathbb{E}[D] = \mathbb{E}[S_T] - \mathbb{E}[S_{\text{tot}}], \quad \text{(B7)}$$

and we will compute $\mathbb{E}[D]$ and $\mathbb{E}[S_T]$.

For $\mathbb{E}[S_T]$, we have

$$\mathbb{E}[S_T] = \sum_{i,j} P_0(i)M(j|i)\ln\frac{P_0(i)M(j|i)}{P_0(j)M(j|i)} \quad \text{(B8a)}$$

$$= 0 + t\sum_{i,j;i\neq j} P_0(i)q(j|i)\ln\frac{P_0(i)q(j|i)}{P_0(j)q(j|i)}. \quad \text{(B8b)}$$

And for $\mathbb{E}[D]$, we have

$$\mathbb{E}[D] = \sum_{i,j} P_0(i)M(j|i)\ln\frac{P_t(j)}{P_0(j)} \quad \text{(B9a)}$$

$$= \sum_j P_t(j)\ln\frac{P_t(j)}{P_0(j)}. \quad \text{(B9b)}$$

With $P_t(j) = P_0(j) + t\sum_k P_0(k)q(j|k) + O(t^2)$, we compute

$$\ln\frac{P_t(j)}{P_0(j)} = \ln\frac{P_0(j) + t\sum_k P_0(k)q(j|k) + O(t^2)}{P_0(j)} \quad \text{(B10a)}$$

$$= t\frac{\sum_k P_0(k)q(j|k)}{P_0(j)} + O(t^2). \quad \text{(B10b)}$$

We thus see that

$$\mathbb{E}[D] = t\sum_j P_t(j)\frac{\sum_k P_0(k)q(j|k)}{P_0(j)} + O(t^2) \quad \text{(B11a)}$$

$$= t\sum_k P_0(k)\sum_j q(j|k) + O(t^2) \quad \text{(B11b)}$$

$$= 0 + O(t^2). \quad \text{(B11c)}$$

Hence, the two EP have the same entropy production rate

$$e_{\text{p}} = \lim_{t\to 0}\frac{\mathbb{E}[S_T]}{t} \quad \text{(B12a)}$$

$$= \lim_{t\to 0}\frac{\mathbb{E}[S_{\text{tot}}] + \mathbb{E}[D]}{t} = \lim_{t\to 0}\frac{\mathbb{E}[S_{\text{tot}}]}{t}. \quad \text{(B12b)}$$

in the infinitesimal time interval limit.

[1] D. J. Searles and D. J. Evans, Fluctuation theorem for stochastic systems, Phys. Rev. E **60**, 159 (1999).

[2] D.-Q. Jiang, M. Qian, and M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems* (Springer, Berlin/New York, 2004).

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

[4] H. Qian, Nonlinear stochastic dynamics of complex systems, I: A chemical reaction kinetic perspective with mesoscopic nonequilibrium thermodynamics, arXiv:1605.08070.

[5] L. F. Thompson and H. Qian, Nonlinear stochastic dynamics of complex systems, II: Potential of entropic force in Markov systems with nonequilibrium steady state, generalized Gibbs function, and criticality, Entropy **18**, 309 (2016).

[6] G. E. Crooks, Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, J. Stat. Phys. **90**, 1481 (1998).

[7] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Phys. Rev. Lett. **78**, 2690 (1997).

[8] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[9] C. Maes, On the origin and the use of fluctuation relations for the entropy, in *Proceedings of the Poincaré Seminar 2003*, Progress in Mathematical Physics Vol. 38 (Birkhäuser Verlag, 2004), p. 145.

[10] H. Qian, Nonequilibrium steady-state circulation and heat dissipation functional, Phys. Rev. E **64**, 022101 (2001).

[11] G. E. Crooks, Path-ensemble averages in systems driven far from equilibrium, Phys. Rev. E **61**, 2361 (2000).

[12] P. M. Riechers and J. P. Crutchfield, Fluctuations when driving between nonequilibrium steady states, J. Stat. Phys. **168**, 873 (2017).

[13] J. L. Lebowitz and H. Spohn, A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. **95**, 333 (1999).

[14] H. Ge and H. Qian, Physical origins of entropy production, free energy dissipation, and their mathematical representations, Phys. Rev. E **81**, 051133 (2010).

[15] M. Esposito and C. Van den Broeck, Three Detailed Fluctuation Theorems, Phys. Rev. Lett. **104**, 090601 (2010).

[16] R. Rao and M. Esposito, Detailed fluctuation theorems: A unifying perspective, Entropy **20**, 635 (2018).

[17] H. Ge, Extended forms of the second law for general time-dependent stochastic processes, Phys. Rev. E **80**, 021137 (2009).

[18] U. Seifert, Entropy Production Along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).

[19] T. Hatano and S. I. Sasa, Steady-State Thermodynamics of Langevin Systems, Phys. Rev. Lett. **86**, 3463 (2001).

[20] J. Kurchan, Fluctuation theorem for stochastic dynamics, J. Phys. A **31**, 3719 (1998).

[21] H. Qian, Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation, Phys. Rev. E **65**, 016102 (2001).

[22] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, Path-integral analysis of fluctuation theorems for general Langevin processes, J. Stat. Mech. (2006) P08001.

[23] H. Ge and D.-Q. Jiang, The transient fluctuation theorem of sample entropy production for general stochastic processes, J. Phys. A **40**, F713 (2007).

[24] B. H. Shargel, The measure-theoretic identity underlying transient fluctuation theorems, J. Phys. A **43**, 135002 (2010).

[25] R. Chétrite and S. Gupta, Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale, J. Stat. Phys. **143**, 543 (2011).

[26] D. J. Evans and D. J. Searles, The fluctuation theorem, Adv. Phys. **51**, 1529 (2002).

[27] R. García-García, V. Lecomte, A. B. Kolton, and D. Domínguez, Joint probability distributions and fluctuation theorems, J. Stat. Mech. (2012) P02009.

[28] M. P. Wojtkowski, Abstract fluctuation theorem, Ergod. Th. Dynam. Sys. **29**, 273 (2009).

[29] A. N. Kolmogorov, *Foundations of the Theory of Probability*, 2nd ed. (Dover Publications, Mineola, NY, 2018).

[30] I. Neri, É. Roldán, and F. Jülicher, Statistics of Infima and Stopping Times of Entropy Production and Applications to Active Molecular Processes, Phys. Rev. X **7**, 011019 (2017).

[31] R. Chétrite, S. Gupta, I. Neri, and É. Roldán, Martingale theory for housekeeping heat, Europhys. Lett. **124**, 60006 (2019).

[32] C. Jarzynski, Hamiltonian derivation of a detailed fluctuation theorem, J. Stat. Phys. **98**, 77 (2000).

[33] P. Maragakis, M. Spichty, and M. Karplus, A differential fluctuation theorem, J. Phys. Chem. B **112**, 6168 (2008).

[34] R. J. Harris and G. M. Schutz, Fluctuation theorems for stochastic dynamics, J. Stat. Mech. (2007) P07020.

[35] L. Hong, H. Qian, and L. F. Thompson, Representations and metrics in the space of probability measures and stochastic thermodynamics, arXiv:1902.09766.

[36] Strictly speaking, a stochastic process can be defined without a probability measure [71]. However, in this paper, we are interested in the statistical properties of observables in the process. We thus say processes with different statistical properties are different processes.

[37] H. Qian, Y.-C. Cheng, and L. F. Thompson, Ternary representation of stochastic change and the origin of entropy and its fluctuations, arXiv:1902.09536.

[38] H. Ge, D.-Q. Jiang, and M. Qian, Reversibility and entropy production of inhomogeneous Markov chains, J. Appl. Probab. **43**, 1028 (2006).

[39] H. Ge and M. Qian, Generalized Jarzynski equality in inhomogeneous Markov chains, J. Math. Phys. **48**, 053302 (2007).

[40] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[41] A. Ya Khinchin, *Mathematical Foundations of Information Theory*, 1st ed. (Dover Publications, Mineola, NY, 1957).

[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, Hoboken, NJ, 2006).

[43] L.-S. Young, What are SRB measures and which dynamical systems have them? J. Stat. Phys. **108**, 733 (2002).

[44] Note that saying two random variables to be the same $X(\omega) = Y(\omega)$ usually means that they return the same value for every $\omega \in \Omega$. This is a stronger statement than saying two random variables $X(\omega)$ and $Y(\omega)$ to have the same probability distribution.

[45] D. J. Evans and D. J. Searles, Equilibrium microstates which generate second law violating steady states, Phys. Rev. E **50**, 1645 (1994).

[46] Here is an explicit example for an involutive operator $\nu : \mathcal{P} \to \mathcal{P}$ not realized by an involutive map $\mu : \Omega \to \Omega$. Consider two binary random variables X_1 and X_2 that can take values 0 or 1. Suppose the joint probabilities for four possible realizations are $P_{1,2}(0,0) = p$, $P_{1,2}(0,1) = q$, $P_{1,2}(1,0) = r$, $P_{1,2}(1,1) = s$, where $p + q + r + s = 1$. The joint probabilities in the new measure $\mathbb{P}^\nu$ is given by $P_{1,2}^\nu(0,0) = \sqrt{(1 - r - s) - p^2}$, $P_{1,2}^\nu(0,1) = (1 - r - s) - \sqrt{(1 - r - s) - p^2}$, $P_{1,2}^\nu(1,0) = r$, $P_{1,2}^\nu(1,1) = s$. One can show that $\mathbb{P}^{\nu\nu} = \mathbb{P}$. And clearly the involutive operator ν can not be realized by an involutive map $\mu : \Omega \to \Omega$.

[47] H. Qian, The zeroth law of thermodynamics and volume-preserving conservative system in equilibrium with stochastic damping, Phys. Lett. A **378**, 609 (2014).

[48] From definition, $P_{0:t}^{TR}(x_{0:t}) = [P_{0:t}^T]^R(x_{0:t})$ is $P_t^T(x_0) \prod_{n=1}^t P_{t+1-n|t-n}^T(x_n|x_{n-1})$. From the fact that T changes the random variable, we get the right-hand side equals $P_0(x_0) \prod_{n=1}^t P_{n-1|n}(x_n|x_{n-1})$ and becomes $\frac{P_0(x_0)}{P_t(x_t)} \prod_{n=1}^t \frac{P_{n-1}(x_n)}{P_n(x_{n-1})} P_{0:t}^{RT}(x_{0:t})$ by Bayes' rule.

[49] An example for an involutive R is to start at $P_0 = \pi_1$ and to fix the driving protocol at M_1 for enough time steps so that the protocol reversed process have enough time to relax back to π_1 by the end of the reversed process.

[50] This can be easily seen by considering the case for only one time step. We have $P_{0,1}^C(x_0, x_1) = P_1(x_1) P_{1|0}(x_0|x_1)$ from Eq. (24b). Thus, we have $P_{0,1}^{CC}(x_0, x_1)$ equal to $P_1^C(x_1) P_{1|0}^C(x_0|x_1)$. With $P_{0,1}^C$ given above, we get $P_1^C = P_1$ and $P_{1|0}^C = P_1(x_0) P_{1|0}(x_1|x_0) / P_0^C(x_1)$. Therefore, $P_{0,1}^{CC}(x_0, x_1) = P_1(x_1) P_1(x_0) P_{1|0}(x_1|x_0) / P_0^C(x_1)$, which is not $P_{0,1}(x_0, x_1)$ unless all P_0, P_1, and P_0^C are the invariant distribution of the transition matrix $P_{1|0}$.

[51] G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in stationary states, J. Stat. Phys. **80**, 931 (1995).

[52] Y. Oono and M. Paniconi, Steady state thermodynamics, Prog. Theor. Phys. **130**, 29 (1998).

[53] H. Qian, Kinematic basis of emergent energetic descriptions of general stochastic dynamics, arXiv:1704.01828.

[54] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

[55] H. Qian, Relative entropy: Free energy associated with equilibrium fluctuations and nonequilibrium deviations, Phys. Rev. E **63**, 042103 (2001).

[56] B. Altaner, Nonequilibrium thermodynamics and information theory: Basic concepts and relaxing dynamics, J. Phys. A **50**, 454001 (2017).

[57] T. M. Hoang, R. Pan, J. Ahn, J. Bang, H. T. Quan, and T. Li, Experimental Test of the Differential Fluctuation Theorem and a Generalized Jarzynski Equality for Arbitrary Initial States, Phys. Rev. Lett. **120**, 080602 (2018).

[58] R. E. Spinney and I. J. Ford, Nonequilibrium Thermodynamics of Stochastic Systems with Odd and Even Variables, Phys. Rev. Lett. **108**, 170603 (2012).

[59] H. Ge, Time reversibility and nonequilibrium thermodynamics of second-order stochastic processes, Phys. Rev. E **89**, 022127 (2014).

[60] G. Li and Z. C. Tu, Stochastic thermodynamics with odd controlling parameters, Phys. Rev. E **100**, 012127 (2019).

[61] L. Onsager and S. Machlup, Fluctuations and irreversible processes, Phys. Rev. **91**, 1505 (1953).

[62] T. Fujita and S. Kotani, The Onsager-Machlup function for diffusion processes, J. Math. Kyoto. Univ. **22**, 115 (1982).

[63] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).

[64] C. H. Bennett, The thermodynamics of computation—A review, Int. J. Theor. Phys. **21**, 905 (1982).

[65] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, Thermodynamics of Prediction, Phys. Rev. Lett. **109**, 120604 (2012).

[66] F. X.-F. Ye, Y. Wang, and H. Qian, Stochastic dynamics: Markov chains and random transformations, Discrete Contin. Dyn. Syst. Ser. B **21**, 2337 (2016).

[67] F. X.-F. Ye and H. Qian, Stochastic dynamics II: Finite random dynamical systems, linear representation, and entropy production, Discrete Contin. Dyn. Syst. Ser. B **24**, 4341 (2019).

[68] R. Durrett, *Probability: Theory and Examples*, 4th ed. (Cambridge University Press, Cambridge, 2010).

[69] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd ed. (Oxford University Press, Oxford/New York, 2001).

[70] It is called σ-algebra because it considers countable union instead of union of a finite number of sets. If the latter case, then it is called an algebra instead of an σ-algebra.

[71] M. Nutz, Pathwise construction of stochastic integrals, Electron. Commun. Probab. **17** (2012).