Jarzynski equality, Crooks fluctuation theorem, and the fluctuation theorems of heat for arbitrary initial states

Zongping Gong[1] and H. T. Quan[1,2,*]
[1]*School of Physics, Peking University, Beijing 100871, China*
[2]*Collaborative Innovation Center of Quantum Matter, Beijing 100871, China*
(Received 17 March 2015; published 23 July 2015)

By taking full advantage of the dynamic property imposed by the detailed balance condition, we derive a new refined unified fluctuation theorem (FT) for general stochastic thermodynamic systems. This FT involves the joint probability distribution functions of the final phase-space point and a thermodynamic variable. Jarzynski equality, Crooks fluctuation theorem, and the FTs of heat as well as the trajectory entropy production can be regarded as special cases of this refined unified FT, and all of them are generalized to arbitrary initial distributions. We also find that the refined unified FT can easily reproduce the FTs for processes with the feedback control, due to its unconventional structure that separates the thermodynamic variable from the choices of initial distributions. Our result is heuristic for further understanding of the relations and distinctions between all kinds of FTs and might be valuable for studying thermodynamic processes with information exchange.

DOI: 10.1103/PhysRevE.92.012131 PACS number(s): 05.70.Ln, 05.20.−y, 05.40.−a, 05.10.Gg

I. INTRODUCTION

There has been great progress in the study of nonequilibrium statistical physics of small systems since the mid-1990s. Compared with classical statistical physics where relative thermal fluctuations are generally Gaussian and vanishingly small, fluctuations become much more prominent in small systems that undergo processes arbitrarily far from equilibrium due to some nonequilibrium external drivings. Despite the complexity that originates from the arbitrariness of the driving protocols, these fluctuations turn out to satisfy some strong, useful, and elegant universal properties, exactly depicted by the fluctuation theorems (FTs) [1–6]. For example, the Jarzynski equality (JE) [1], which connects the work done in nonequilibrium processes and the free-energy difference of the system at the initial and the final stages, is an integral fluctuation theorem (IFT) of work, while its stronger version, the Crooks fluctuation theorem (CFT) [3], is a detailed fluctuation theorem (DFT) of work. It should be emphasized that both the JE and the CFT require the system initially prepared in a canonical distribution and the validity of the detailed balance (DB) condition. As a more universal IFT, the entropy production identity (EPI) [5] holds for arbitrary initial distributions even without the DB condition. It is worth mentioning that a DFT of heat has been discovered quite recently, where the distribution of the initial state in the phase space is required to be uniform rather than canonical [6].

We posit several fundamental questions: Why are there strict requirements for the distributions of the initial state in these FTs and can they be released? Can these existing FTs be traced back to the same root? Besides the FTs of the work, the heat, and the trajectory entropy production, are there any other FTs associated with other thermodynamic variables? These questions were partially answered by the unified FTs explored by Seifert [7,8]. For a stochastic system undergoing a nonequilibrium process within the time interval $[0,\tau]$, driven by a temporally varying work parameter λ_t $(0 \leqslant t \leqslant \tau)$, the unified IFT reads [7,8]

$$\langle e^{-R}\rangle = \left\langle \frac{p^a_\tau(\Gamma_\tau)}{p_0(\Gamma_0)} e^{-\Delta s_m} \right\rangle = 1, \tag{1}$$

where the trajectory-dependent functional $R[\Gamma_t] \equiv \ln[\mathcal{P}(\Gamma_t)/\bar{\mathcal{P}}(\bar{\Gamma}_t)] = \Delta s_m - \ln \bar{p}_0(\Gamma^\dagger_\tau) + \ln p_0(\Gamma_0)$. Here Δs_m is the entropy production in the medium [5]; $p^a_\tau(\Gamma)$ is the auxiliary final distribution in the phase space; $p_0(\Gamma)$ $[\bar{p}_0(\Gamma)]$ is the initial distribution of the system in the phase space for the forward (backward) process; $\mathcal{P}(\Gamma_t)$ $[\bar{\mathcal{P}}(\bar{\Gamma}_t)]$ is the probability density of a trajectory Γ_t $(\bar{\Gamma}_t)$ in the trajectory space due to the protocol λ_t (time-reverse protocol $\bar{\lambda}_t \equiv \lambda_{\tau-t}$); $\bar{\Gamma}_t \equiv \Gamma^\dagger_{\tau-t}$ is the time-reverse trajectory of Γ_t with $\Gamma^\dagger$ to be the time-reversal of the phase-space point Γ [e.g., for the underdamped Langevin dynamics, $\Gamma_t = (\boldsymbol{r}_t, \boldsymbol{p}_t)$, then $\Gamma^\dagger_t = (\boldsymbol{r}_t, -\boldsymbol{p}_t)$ and $\bar{\Gamma}_t = (\boldsymbol{r}_{\tau-t}, -\boldsymbol{p}_{\tau-t})$]. In fact, Eq. (1) is a unification of the IFT of heat, the JE, and the EPI, since they can be respectively generated by setting the initial distribution $p_0(\Gamma_0)$ and the auxiliary final distributions $p^a_\tau(\Gamma_\tau)$ to be both uniform and both canonical and the distributions connected by the real dynamic evolution. For those processes with the thermodynamic variable S_α with a definite time-reversal parity, say, $S_\alpha[\bar{\Gamma}_t] = \epsilon_\alpha S_\alpha[\Gamma_t]$, $\epsilon_\alpha = \pm 1$, we further have the following unified DFT [8]:

$$\frac{\bar{P}(\{S_\alpha = \epsilon_\alpha s_\alpha\})}{P(\{S_\alpha = s_\alpha\})} = \langle e^{-R} | \{S_\alpha = s_\alpha\}\rangle, \tag{2}$$

where $P(\{S_\alpha\})[\bar{P}(\{S_\alpha\})]$ denotes the joint distribution of multiple variables S_α for the forward (backward) process [9,10] and the right-hand side denotes the conditional expectation of e^{-R} when $S_\alpha = s_\alpha$. To generate the DFT of heat or the CFT, we simply choose the single odd parity quantity S to be the heat (Q) or the work (W) and then set $p_0(\Gamma)$ and $\bar{p}_0(\Gamma)$ to be both uniform distributions or both canonical distributions. This choice leads to $R = \beta Q$ or $\beta(W - \Delta F)$ and thus makes the right-hand side of Eq. (2) simply $e^{-\beta Q}$ or $e^{-\beta(W-\Delta F)}$, since R is uniquely determined by S and the conditional distribution is a δ one. From this point of view, we can say that the FTs mentioned in the last paragraph do share the same root, which might be

*htquan@pku.edu.cn

summarized as a combination of the microscopic reversibility (MR) and the dynamic property. The MR ensures the validity of Eq. (1) [and Eq. (2), though not obviously], since it is no more than the probability-normalization relation of all the time-reverse trajectories whose forward ones are of nonzero probability. In fact, the conventional FTs are inapplicable to absolutely irreversible processes because of the breakdown of the MR, as has been highlighted in recent investigations [11]. The dynamic property, such as the DB condition, is necessary to relate certain thermodynamic variables to R by properly choosing the two initial distributions. This is important for endowing specific physical meaning to the abstract identities (1) and (2) as merely the corollaries of the MR.

On the other hand, despite of their universal validity, the unified IFT (1) will be physically meaningless if R cannot be related to certain thermodynamic variables, while the right-hand side of the unified DFT (2) is usually difficult to either calculate or be given a transparent physical interpretation, unless R uniquely depends on S. Such entanglement between the thermodynamic variable and the initial distributions is the reason why the EPI holds for arbitrary distributions while the FTs of the heat and the work require specific initial distributions: The functional of entropy production naturally contains the distribution functions of the initial and final phase-space points, but the heat and the work have nothing to do with them.

If the initial distribution is not canonical (uniform) distribution, then can we still construct a fluctuation theorem for the work (heat)? According to the above analysis, the answer seems to be no. However, in this article, we propose a new refined unified FT which achieves the separation of the thermodynamic variable from the choices of the initial distributions. As a result of the DB condition, this refined unified FT is even more "detailed" than DFT (2), because it involves the joint distributions with the phase-space point. In the studies of FTs, the first encounter of the joint distributions with the phase-space point is in the Hummer-Szabo relation [2,4,12]. However, the Hummer-Szabo relation still requires the distribution of the initial state to be a canonical distribution and is valid for the work only. In our current investigation, we extend the FTs of work and heat to an arbitrary initial distribution and even to other variables, such as the entropy production. The cost is that we need to know the joint distribution function with the phase-space point, such as $P_\tau(W,\Gamma)$, which is more detailed than the usual distribution function, such as the work distribution $P_\tau(W)$. We show that the new refined unified FT can generate many existing FTs as well as many new FTs that are not known previously by choosing proper initial states. It may also be of potential values in investigating information thermodynamics, where the initial distributions might be quite irregular due to the extra distribution rectification by the information.

We notice that similar problems, i.e., the FTs for arbitrary initial states, are investigated in a recent work [13]. However, the systems considered therein are discrete and automatous (not externally driven) and may not satisfy the DB condition, and the FTs are associated with the local currents in a general dynamic network. Thus both aspects of their study are complementary to ours. This issue is also discussed in Ref. [14] but is not its main focus.

This paper is organized as follows. In Sec. II, we derive the refined unified FT by analyzing its close relation to the DB condition for general stochastic thermodynamic systems. In Sec. III, we reproduce the existing FTs and generate some new ones of the work, the heat and the entropy production as some examples of the refined unified FT. Some applications of the refined unified FT are explored in Sec. IV. In Sec. V, we summarize the paper. Six appendices are added at the end of the paper, which we think are helpful for understanding the details of the paper and the relevant issues.

II. MAIN RESULT AND ITS DERIVATION

A. Main result: The refined unified FT

We focus on general stochastic thermodynamic systems with the DB condition. The system is coupled to a heat bath with the inverse temperature β, which thus undergoes isothermal processes within the time interval $[0,\tau]$ due to an external driving protocol denoted by $\lambda_t, t \in [0,\tau]$. We consider a trajectory-dependent thermodynamic quantity $A[\Gamma_t]$ in the following form:

$$A[\Gamma_t] = \beta Q[\Gamma_t] + a_\tau(\Gamma_\tau) - a_0(\Gamma_0). \quad (3)$$

Here the heat functional $Q[\Gamma_t] \equiv -\int_0^\tau dt\dot{\Gamma}_t\partial_\Gamma U_t(\Gamma_t)$ ($Q > 0$ corresponds to the release of the heat from the system to the heat bath), with $U_t(\Gamma)$ to be the energy of phase-space point Γ at time t; $a_t(\Gamma)$ can be an arbitrary time-dependent function of the phase-space point Γ, for which we can generally define its time-reversal $\bar{a}_t(\Gamma) \equiv a_{\tau-t}(\Gamma^\dagger)$ [15] and, correspondingly,

$$\bar{A}[\Gamma_t] = \beta Q[\Gamma_t] + \bar{a}_\tau(\Gamma_\tau) - \bar{a}_0(\Gamma_0). \quad (4)$$

Throughout the paper, if not particularly indicated, we always stipulate that the energy possesses the property of the time-reversal invariance, i.e., $U_t(\Gamma^\dagger) = U_t(\Gamma)$. Our main result reads

$$\frac{\int_{\mathfrak{S}_0} d\Gamma\,\bar{P}_\tau(-A,\Gamma)p_0(\Gamma^\dagger)e^{\bar{a}_\tau(\Gamma)}}{\int_{\mathfrak{S}_\tau} d\Gamma\,P_\tau(A,\Gamma)\bar{p}_0(\Gamma^\dagger)e^{a_\tau(\Gamma)}} = e^{-A}, \quad (5)$$

where the accessible phase space at time t $(\tau - t)$ for the forward (backward) process is denoted by $\mathfrak{S}_t$, thus the integral of Γ in the denominator (numerator) goes over the whole accessible phase space of Γ_τ (Γ_0 or $\bar{\Gamma}_\tau \equiv \Gamma_0^\dagger$) [16]; $p_0(\Gamma)$ [$\bar{p}_0(\Gamma)$] denotes the distribution of the initial state in the phase space for the forward (backward) process; $P_\tau(A,\Gamma)$ [$\bar{P}_\tau(\bar{A},\Gamma)$] is the joint distribution function of the thermodynamic variable A ($\bar{A}$) accumulated until time τ and the final (at time τ) phase-space point, starting from the initial distribution $p_0(\Gamma)$ [$\bar{p}_0(\Gamma)$] and driven under the protocol λ_t ($\bar{\lambda}_t \equiv \lambda_{\tau-t}$). Equation (5) is valid for arbitrary initial distributions $p_0(\Gamma)$ and $\bar{p}_0(\Gamma)$ and arbitrary state functions $a_t(\Gamma)$. Also, similar to the JE and the CFT, it is valid for an arbitrary protocol λ_t and an arbitrary driving time τ.

B. Dynamic property equivalent to the DB condition

In order to demonstrate the close relation between our main result (5) and the DB condition transparently, we use a somewhat complicated method to carry out the derivation, though a relatively simple but much more mathematical proof is available by using the path-integral approach (see Appendix A).

We first write down the Fokker-Planck equation (FPE) of the stochastic system,

$$\partial_t p_t(\Gamma) = \mathcal{L}_t p_t(\Gamma), \tag{6}$$

with the generator $\mathcal{L}_t$ to be a time-dependent linear operator corresponding to the protocol λ_t and only acting on Γ. For later use, we define the transpose operator of $\mathcal{L}_t$ (denoted by $\mathcal{L}_t^T$), which satisfies $\int_{\mathfrak{S}_t} d\Gamma g(\Gamma)\mathcal{L}_t f(\Gamma) = \int_{\mathfrak{S}_t} d\Gamma f(\Gamma)\mathcal{L}_t^T g(\Gamma)$ for arbitrary normalizable (i.e., $|\int d\Gamma f(\Gamma)| < +\infty$) functions $f(\Gamma)$ and $g(\Gamma)$ defined in the phase space. One can see that the normalization of $p_t(\Gamma)$ will impose the property $\mathcal{L}_t^T 1 = 0$ to $\mathcal{L}_t$ if we integrate the FPE (6) over Γ, though this is not rigorous for the systems with infinite phase space where 1 is not normalizable. We can also define the time-reversed operator of $\mathcal{L}_t$ (denoted by $\mathcal{L}_t^\dagger$). It is obtained by adding minus signs to all the components in $\mathcal{L}_t$ with the odd time-reversal parity (e.g., $\boldsymbol{p} \to -\boldsymbol{p}$, $\partial_{\boldsymbol{p}} \to -\partial_{\boldsymbol{p}}$, $\boldsymbol{p}$ is momentum).

In terms of the generator, the DB condition manifests itself in the following algebraic symmetry:

$$e^{\beta U_t(\Gamma)}\mathcal{L}_t^\dagger e^{-\beta U_t(\Gamma)} = \mathcal{L}_t^T. \tag{7}$$

Such a dynamic property comes directly from the combination of the original definition of the DB condition $e^{-\beta U_t(\Gamma_1)} w_t(\Gamma_1 \to \Gamma_2) = e^{-\beta U_t(\Gamma_2)} w_t(\Gamma_2^\dagger \to \Gamma_1^\dagger)$ (for any $\Gamma_1, \Gamma_2 \in \mathfrak{S}_t$) and the transition rate formula $w_t(\Gamma_1 \to \Gamma_2) = \int_{\mathfrak{S}_t} d\Gamma \delta(\Gamma - \Gamma_2)\mathcal{L}_t \delta(\Gamma - \Gamma_1)$ [17]. Conversely, the original DB condition follows if Eq. (7) is previously assumed, thus the dynamic property is actually equivalent to the DB condition. We emphasize that Eq. (7) should be valid at any time $t \in [0,\tau]$ despite the fact that the work parameter is temporally varied. This is the straightforward dynamic generalization of the common static version (where the work parameter is fixed), which can be found in the standard literature [17]. We also want to mention that Eq. (7) ensures when we suddenly stop the external driving at time t, the system will always relax to the canonical distribution determined by λ_t due to the property $\mathcal{L}_t^T 1 = 0$. In addition, Eq. (7) is much more stronger than merely imposing $\mathcal{L}_t e^{-\beta U_t(\Gamma)} = 0$, and that is why the balance is called "detailed."

As an example, it can be noted that Eq. (7) is true for the Langevin dynamics (in both the underdamped and the overdamped regimes, see Appendix B). It is also notable that a discrete but more general (may be without the DB condition) version of Eq. (7) has appeared in Ref. [18] [see Eq. (3.24) therein].

C. Derivation of Eq. (5) based on the characteristic function

According to the FPE (6) as well as the definition of the functional A (3), it can be proved (see Appendix C for a more general formula) that the joint distribution function $P_t(A,\Gamma)$ satisfies the following equation of motion:

$$\partial_t P_t(A,\Gamma) = \{e^{-[a_t(\Gamma)-\beta U_t(\Gamma)]\partial_A}\mathcal{L}_t e^{[a_t(\Gamma)-\beta U_t(\Gamma)]\partial_A} - \partial_t a_t(\Gamma)\partial_A\} P_t(A,\Gamma). \tag{8}$$

Notice that the operator acting on $P_t(A,\Gamma)$ on the right-hand side only contains ∂_A but is independent of A. It is natural to (partly) perform an integral transformation, as is always done when we construct the Feynman-Kac formula [e.g., $a_t(\Gamma) = \beta U_t(\Gamma)$ in Eq. (8)]. We take the inverse Fourier transformation to change $P_t(A,\Gamma)$ into its characteristic function $G_t(\mu,\Gamma) \equiv \int_{-\infty}^{+\infty} dA e^{i\mu A} P_t(A,\Gamma)$. For further simplicity, we define the modified characteristic function $G_t^m(\mu,\Gamma) \equiv e^{-i\mu a_t(\Gamma)} G_t(\mu,\Gamma)$. Based on Eq. (8), it is found that $G_t^m(\mu,\Gamma)$ satisfies the following equation of motion:

$$\partial_t G_t^m(\mu,\Gamma) = \mathcal{L}_t(\mu) G_t^m(\mu,\Gamma), \tag{9}$$

where $\mathcal{L}_t(\mu) \equiv e^{-i\mu\beta U_t(\Gamma)}\mathcal{L}_t e^{i\mu\beta U_t(\Gamma)}$. For the backward process, we define $\bar{\mathcal{L}}_t(\mu) \equiv \mathcal{L}_{\tau-t}(\mu)$, so the equation of motion of the modified characteristic function of the backward process, $\bar{G}_t^m(\mu,\Gamma) \equiv e^{-i\mu\bar{a}_t(\Gamma)}\bar{G}_t(\mu,\Gamma)$, can be expressed as

$$\partial_t \bar{G}_t^m(\mu,\Gamma) = \bar{\mathcal{L}}_t(\mu)\bar{G}_t^m(\mu,\Gamma). \tag{10}$$

In terms of $\mathcal{L}_t(\mu)$, the relation (7) can be rewritten as

$$\mathcal{L}_t^\dagger(\mu + i) = \bar{\mathcal{L}}_{\tau-t}^T(-\mu) \tag{11}$$

by using the identity $\mathcal{L}_t^T(\mu) = e^{i\mu\beta U_t(\Gamma)}\mathcal{L}_t^T e^{-i\mu\beta U_t(\Gamma)}$. In order to make full use of Eq. (11), we write in the following special forms the equations of motion for both the forward (9) and the backward (10) processes:

$$\partial_t G_t^m(\mu + i,\Gamma^\dagger) = \mathcal{L}_t^\dagger(\mu + i) G_t^m(\mu + i,\Gamma^\dagger), \tag{12}$$

$$-\partial_t \bar{G}_{\tau-t}^m(-\mu,\Gamma) = \bar{\mathcal{L}}_{\tau-t}(-\mu)\bar{G}_{\tau-t}^m(-\mu,\Gamma). \tag{13}$$

Combining these two equations of motion with Eq. (11), after a straightforward calculation, we deduce that

$$\partial_t \int_{\mathfrak{S}_t} d\Gamma G_t^m(\mu + i,\Gamma^\dagger)\bar{G}_{\tau-t}^m(-\mu,\Gamma) = 0. \tag{14}$$

This means the integral $\int_{\mathfrak{S}_t} d\Gamma G_t^m(\mu + i,\Gamma^\dagger)\bar{G}_{\tau-t}^m(\mu,\Gamma)$ is a conserved quantity during the dynamic evolution. Particularly at the terminal time points ($t = 0,\tau$), we have

$$\int_{\mathfrak{S}_\tau} d\Gamma G_\tau^m(\mu + i,\Gamma)\bar{G}_0^m(-\mu,\Gamma^\dagger) = \int_{\mathfrak{S}_0} d\Gamma G_0^m(\mu + i,\Gamma^\dagger)\bar{G}_\tau^m(-\mu,\Gamma). \tag{15}$$

After substituting the expressions of the modified characteristic functions into Eq. (15), we obtain

$$\int_{\mathfrak{S}_\tau} d\Gamma G_\tau(\mu + i,\Gamma)\bar{p}_0(\Gamma^\dagger)e^{a_\tau(\Gamma)} = \int_{\mathfrak{S}_0} d\Gamma \bar{G}_\tau(-\mu,\Gamma)p_0(\Gamma^\dagger)e^{\bar{a}_\tau(\Gamma)}. \tag{16}$$

By taking the Fourier transformation of Eq. (16), we finally obtain our main result (5).

It is worth mentioning that a similar method based on the symmetry of the generator has appeared in a recent work [19] [see Eq. (11) here and Eq. (10) therein]. But their focus was on the Bochkov-Kuzovlev equality [20] for the open quantum systems described by Lindblad master equations. Therefore, we believe there is a quantum generalization of the main result (5) (and its extensive corollaries) for at least Lindblad-type open quantum systems, which we leave for our future work.

III. REFINEMENT OF THE EXISTING FTS AND GENERATING OF NEW FTS

A. Refined FT of work

The simplest specialization of the refined unified FTs (5) is when we choose $a_t(\Gamma) = \beta U_t(\Gamma)$. In this manner, both the quantity A (3) and $\bar{A}$ (4) turn out to be the dimensionless work βW, due to the first law of thermodynamics at the level of individual trajectories [21]. The refined FT of work reads

$$\frac{\int_{\mathfrak{S}_0} d\Gamma\, \bar{P}_\tau(-W,\Gamma) p_0(\Gamma^\dagger) e^{\beta U_0(\Gamma^\dagger)}}{\int_{\mathfrak{S}_\tau} d\Gamma\, P_\tau(W,\Gamma) \bar{p}_0(\Gamma^\dagger) e^{\beta \bar{U}_0(\Gamma^\dagger)}} = e^{-\beta W}. \quad (17)$$

Similarly to the CFT, this relation is valid for an arbitrary driving protocol λ_t and an arbitrary driving time τ. Moreover, this relation is more general than the CFT because it is valid for arbitrary initial distributions $p_0(\Gamma)$ and $\bar{p}_0(\Gamma)$. Obviously, if we want to construct the existing work FTs from Eq. (17), the choices of the distributions of the initial state in the phase space should be the canonical ones for both the forward and the backward processes. That is, $p_0(\Gamma) = p_0^{\rm eq}(\Gamma) \equiv e^{-\beta U_0(\Gamma)}/Z_0(\beta)$ and $\bar{p}_0(\Gamma) = \bar{p}_0^{\rm eq}(\Gamma) \equiv e^{-\beta \bar{U}_0(\Gamma)}/\bar{Z}_0(\beta)$, where the partition functions $Z_0(\beta) \equiv \int_{\mathfrak{S}_0} d\Gamma e^{-\beta U_0(\Gamma)}$ and $\bar{Z}_0(\beta) \equiv \int_{\mathfrak{S}_\tau} d\Gamma e^{-\beta \bar{U}_0(\Gamma)}$. For such choices, Eq. (17) reduces to the well-known CFT [3],

$$\frac{\bar{P}_\tau(-W)}{P_\tau(W)} = e^{-\beta(W-\Delta F)}, \quad (18)$$

where $P_\tau(W) \equiv \int_{\mathfrak{S}_\tau} d\Gamma\, P_\tau(W,\Gamma)$ and $\bar{P}_\tau(W) \equiv \int_{\mathfrak{S}_0} d\Gamma\, \bar{P}_\tau(W,\Gamma)$, $\Delta F \equiv -\beta^{-1} \ln[Z_\tau(\beta)/Z_0(\beta)]$ is the free-energy difference [22]. The integral version of the CFT (18) is the celebrated JE [1]

$$\langle e^{-\beta(W-\Delta F)} \rangle = 1. \quad (19)$$

We can also easily reproduce the Hummer-Szabo relation [12]

$$\langle \delta(\Gamma_\tau - \Gamma') e^{-\beta W} \rangle = \frac{e^{-\beta U_\tau(\Gamma')}}{Z_0(\beta)}, \quad (20)$$

as long as we set the initial distribution for the forward process to be a canonical distribution $p_0^{\rm eq}(\Gamma)$ and the initial distribution for the backward process to be a δ distribution $\bar{p}_0(\Gamma) = \delta(\Gamma - \Gamma'^\dagger)$ in Eq. (17).

The discrete version of Eq. (17) is

$$\frac{\sum_{\bar{m}} \bar{P}_\tau(-W,\bar{m}) p_0(\bar{m}^\dagger) e^{\beta \bar{E}_\tau^{\bar{m}}}}{\sum_n P_\tau(W,n) \bar{p}_0(n^\dagger) e^{\beta E_\tau^n}} = e^{-\beta W}, \quad (21)$$

where E_t^n ($\bar{E}_t^{\bar{m}}$) is the energy of the state n ($\bar{m}$) at time t for the forward (backward) process; $n^\dagger$ denotes the time-reversed state of n, e.g., the spin-down state, if n denotes spin-up. Certainly, this relation holds for discrete-level stochastic systems. We emphasize that Eq. (21) is also valid for an isolated quantum system (not necessarily with time-reversal symmetry), where the initial density matrix is generally $\varrho_0 = \sum_m p_0(m)|m\rangle\langle m|$ (and $\bar{\varrho}_0 = \sum_{\bar{n}} \bar{p}_0(\bar{n})|\bar{n}\rangle\langle\bar{n}|$ for the time-reversed process), as long as we admit the two-point projection measurement definition of quantum work (see Appendix D). Furthermore, β in Eq. (21) can be arbitrarily chosen and may even be a complex number, since the quantum system is isolated from any heat bath.

B. Generating new FTs of work

In addition, from the refined FT of work (17) one can derive several FTs of work that were previously not known to researchers in this field. For that purpose, let us first choose both $p_0(\Gamma) = \delta(\Gamma - \Gamma_0)$ and $\bar{p}_0(\Gamma) = \delta(\Gamma - \Gamma'^\dagger)$ to be δ distributions. Therefore Eq. (17) becomes

$$P_\tau(W,\Gamma'|\Gamma_0) e^{-\beta(W-\Delta F)} = \frac{\bar{p}_0^{\rm eq}(\Gamma'^\dagger)}{p_0^{\rm eq}(\Gamma_0)} \bar{P}_\tau(-W,\Gamma_0^\dagger|\Gamma'^\dagger), \quad (22)$$

where $P_\tau(W,\Gamma'|\Gamma_0)$ is the conditional joint probability distribution. It means the sum of the probabilities of all the forward trajectories that end at Γ' and the work accumulated along each of these trajectories are equal to W conditioned on the given initial phase-space point Γ_0. $\bar{P}_\tau(-W,\Gamma_0^\dagger|\Gamma'^\dagger)$ can be understood in a similar way but for the backward process. This relation can be regarded as a generalization of CFT to initial δ distributions and has previously been obtained in Refs. [23,24]. We perform the integral with respect to W on both sides of Eq. (22) and obtain

$$\langle \delta(\Gamma_\tau - \Gamma') e^{-\beta(W-\Delta F)} | \Gamma_0 \rangle = \frac{\bar{p}_0^{\rm eq}(\Gamma'^\dagger)}{p_0^{\rm eq}(\Gamma_0)} \bar{p}_\tau(\Gamma_0^\dagger|\Gamma'^\dagger), \quad (23)$$

where $\langle F[\Gamma_t] | \Gamma_0 \rangle$ indicates the ensemble average of the functional $F[\Gamma_t]$ over all trajectories that start from the phase-space point Γ_0 and $\bar{p}_\tau(\Gamma_0^\dagger|\Gamma'^\dagger)$ is the conditional probability distribution in the phase space for the backward process. It describes the final probability distribution of the backward process at $\Gamma_0^\dagger$ given the initial state at $\Gamma'^\dagger$. This relation (23) can be regarded as a generalization of the Hummer-Szabo relation (20) to the δ initial distribution. If we further perform the integral with respect to Γ' on both sides of Eq. (23), we obtain a JE-like FT:

$$\langle e^{-\beta(W-\Delta F)} | \Gamma_0 \rangle = \frac{\bar{p}_\tau(\Gamma_0^\dagger)}{p_0^{\rm eq}(\Gamma_0)}, \quad (24)$$

where $\bar{p}_\tau(\Gamma)$ is the final probability distribution in the phase space evolved from the initial canonical distribution $\bar{p}_0^{\rm eq}(\Gamma)$ in the time-reversed process. This relation can be regarded as the generalization of JE to the δ initial distribution. Similar mathematical relations have been obtained in Ref. [25] for isolated classical systems and in Ref. [26] for open quantum systems. But in both these two references, their focus is on different problems, and the mathematical relations are not interpreted in this way.

Besides the initial δ distribution, the JE (19) and the Hummer-Szabo relation (20) can actually be extended to an arbitrary initial distribution in the forward process. We multiply an arbitrary initial distribution $p_0(\Gamma_0)$ to both sides of Eq. (22) before we do the integral with respect to Γ_0, and we obtain the extended Hummer-Szabo relation for an arbitrary initial distribution $p_0(\Gamma)$,

$$\langle \delta(\Gamma_\tau - \Gamma') e^{-\beta(W-\Delta F)} \rangle_{p_0(\Gamma)} = \int_{\mathfrak{S}_0} d\Gamma\, \bar{p}_\tau(\Gamma^\dagger|\Gamma'^\dagger) \bar{p}_0^{\rm eq}(\Gamma'^\dagger) \frac{p_0(\Gamma)}{p_0^{\rm eq}(\Gamma)}, \quad (25)$$

where $\langle \ldots \rangle_{p_0(\Gamma)}$ represents the average over all trajectories when the initial distribution is $p_0(\Gamma)$ for the forward process. If we further perform the integral with respect to Γ' on both sides of Eq. (25) we obtain the generalized JE for an arbitrary initial distribution:

$$\langle e^{-\beta(W-\Delta F)}\rangle_{p_0(\Gamma)} = \int_{\mathfrak{S}_0} d\Gamma\, \bar{p}_\tau(\Gamma^\dagger)\frac{p_0(\Gamma)}{p_0^{\rm eq}(\Gamma)}. \quad (26)$$

In this equality, the initial distribution of the forward process can be an arbitrary distribution $p_0(\Gamma)$, while the initial distribution of the backward process must be a canonical distribution. It is worth mentioning that in the FT of the total entropy production along individual trajectories [5], the initial distribution can be an arbitrary distribution (38). It was assumed previously that only when the initial distribution is a canonical distribution (globally thermal equilibrium) or a partially thermal equilibrium distribution [27,28] can one construct the FT of work. Here we generalize the FT of work [mainly JE (19), CFT (18), and the Hummer-Szabo relation (20)] to an arbitrary initial state. The generalized JE for an arbitrary initial state (26) may have potential applications in recovering free energy via the nonequilibrium work measurement and numerical free-energy computation.

C. Refined FT of heat

Another specialization of the refined unified FT (5) is the FT of heat when $a_t(\Gamma)$ equals to a constant (e.g., zero), thus $A = \bar{A} = \beta Q$. The refined FT of heat reads

$$\frac{\int_{\mathfrak{S}_0} d\Gamma\, \bar{P}_\tau(-Q,\Gamma) p_0(\Gamma^\dagger)}{\int_{\mathfrak{S}_\tau} d\Gamma\, P_\tau(Q,\Gamma)\bar{p}_0(\Gamma^\dagger)} = e^{-\beta Q}. \quad (27)$$

Similarly to the DFT of heat [6], this relation is valid for an arbitrary driving protocol λ_t and driving time τ. Moreover, this relation is more general than the DFT of heat, because it is valid for arbitrary initial distributions $p_0(\Gamma)$ and $\bar{p}_0(\Gamma)$. If we want to reduce the joint distribution function to the heat distribution function, the only choice is to make both $p_0(\Gamma)$ and $\bar{p}_0(\Gamma)$ independent of the phase-space point, i.e., the uniform distribution, as was found in Ref. [6]. However, a uniform distribution can never be truly realized in continuous systems, where the entropy has no upper bound due to the infinite volume of the phase space. So we have to consider finite-level systems, such as spin systems. Similarly to Eq. (21), the discrete version of Eq. (27) is

$$\frac{\sum_{\bar{m}} \bar{P}_\tau(-Q,\bar{m}) p_0(\bar{m}^\dagger)}{\sum_n P_\tau(Q,n)\bar{p}_0(n^\dagger)} = e^{-\beta Q}. \quad (28)$$

Suppose that the system has totally N states. By setting $\bar{p}_0(n^\dagger) = p_0(\bar{m}^\dagger) \equiv 1/N$ (maximum entropy state) in Eq. (28), we obtain

$$\frac{\bar{P}_\tau(-Q)}{P_\tau(Q)} = e^{-\beta Q}, \quad (29)$$

where $P_\tau(Q) \equiv \sum_n P_\tau(Q,n)$ and $\bar{P}_\tau(Q) \equiv \sum_{\bar{m}} \bar{P}_\tau(Q,\bar{m})$ are respectively the heat distribution functions for the forward and the backward processes. Equation (29) is nothing but the DFT of heat, whose integral version reads [6]

$$\langle e^{-\beta Q}\rangle = 1. \quad (30)$$

It is worth mentioning that Eqs. (29) and (30) are likely to be experimentally tested in analogy to the verification of the EPI for a finite-level system [29]. More generally, we can test Eq. (27) in whatever systems with finite phase space.

One may imagine that the uniform initial distribution of the DFT or the IFT of heat might be approached via the limit $\beta' \to 0$ ($T' \to \infty$), with β' to be the inverse temperature of the initial canonical distribution. Hence, one may ask whether

$$\lim_{\beta'\to 0} \frac{1}{\beta Q} \ln \frac{P_\tau(Q)}{\bar{P}_\tau(-Q)} = 1 \quad (31)$$

holds true even for a continuous system with its initial distribution to be $p_0(\Gamma) = e^{-\beta' U_0(\Gamma)}/Z_0(\beta')$. In fact, it has been demonstrated [6] that for a driven Brownian harmonic oscillator, if we rewrite Q in $w\tau p$, with w as a positive quantity that characterizes the driving speed, and take the limit $\tau \to \infty$ (so $Q \to \infty$ for finite p) before $\beta' \to 0$, then Eq. (31) will be indeed true. It might be an intriguing but involved subject to investigate the validity range of Eq. (31) for more general kinds of stochastic systems with infinite phase space.

D. Generating new FTs of heat

Similarly to JE for an arbitrary initial distribution, we can extend the integral FT of heat (30) to an arbitrary initial state. For the sake of a well-defined uniform distribution, we write down the discrete version for an N-state system:

$$\langle e^{-\beta Q}\rangle_{p_0(m)} = \sum_m \bar{p}_\tau(m^\dagger)\frac{p_0(m)}{p_0^u(m)}, \quad (32)$$

where $p_0^u(m) \equiv 1/N$ is the initial uniform distribution of the forward process, and the final distribution of the microscopic state for the backward process $\bar{p}_\tau(m^\dagger) = \int dQ\, \bar{P}_\tau(Q,m^\dagger)$ in Eq. (32) must correspond to the uniform initial distribution. One can see that the right-hand side of Eq. (32) usually deviates from 1 if the initial distribution of the forward process is not prepared in the uniform distribution. Equation (32) is valid for an arbitrary protocol and an arbitrary driving time. Moreover, it does not require the initial distribution of the forward process to be the uniform distribution. Hence, if Eq. (27) can be regarded as the generalization of the detailed version of the heat FT [6] to an arbitrary initial distribution, Eq. (32) can be regarded as the generalization of the integral version of heat FT [6] to an arbitrary initial distribution.

For the heat FT, one can also derive a Hummer-Szabo–like relation. The initial distribution for the forward process and the reverse process are chosen to be the uniform distribution and the δ distribution respectively. Let m_τ be the final (at time τ) state of the forward process; this relation then reads

$$\langle \delta_{m_\tau n} e^{-\beta Q}\rangle = \frac{1}{N}. \quad (33)$$

Furthermore, we can also extend this Hummer-Szabo–like relation for heat (33) to an arbitrary initial distribution $p_0(m)$,

$$\langle \delta_{m_\tau n} e^{-\beta Q}\rangle_{p_0(m)} = \sum_m p_0(m)\bar{p}_\tau(m^\dagger|n^\dagger). \quad (34)$$

Until now, we have successfully generalized the FTs of work and heat to arbitrary initial distributions. For a comparison between previous results and our current results, please see Table I.

TABLE I. Comparison between previous results and our results on the requirements on the initial state for different FTs. Please note that in previous results only distribution functions of certain thermodynamic variables, such as $P_\tau(W)$ and $P_\tau(Q)$, are required. But in our current results, more detailed joint distribution functions, such as $P_\tau(W,\Gamma)$ and $P_\tau(Q,\Gamma)$, are required.

	Fluctuation theorems	Previous requirements on the distributions of initial state	Our requirements on the distributions of initial state
Work FTs	CFT [3]	Canonical distribution (18)	Arbitrary distribution (17)
	JE [1]	Canonical distribution (19)	Arbitrary distribution (26)
	Hummer-Szabo relation [12]	Canonical distribution (20)	Arbitrary distribution (25)
Heat FTs	Detailed heat FT [6]	Uniform distribution (29)	Arbitrary distribution (27)
	Integral heat FT [6]	Uniform distribution (30)	Arbitrary distribution (32)
	Hummer-Szabo like relation	N/A	Arbitrary distribution (34)
FTs with feedback control	Sagawa-Ueda equalities [35,36]	Canonical distribution (51),(53); arbitrary distribution (54)	Arbitrary distribution (56)

E. Unified IFT and EPI

Besides reproducing the JE, the CFT, and the heat FTs, the refined unified FT (5) can reproduce the EPI. Let us again have a look at the main result (5). We find that if $a_\tau(\Gamma) = -\ln \bar{p}_0(\Gamma^\dagger)$ and $\bar{a}_\tau(\Gamma) = -\ln p_0(\Gamma^\dagger)$, which are sufficient to determine $a_0(\Gamma) = a_\tau(\Gamma^\dagger) = -\ln p_0(\Gamma)$ and $\bar{a}_0(\Gamma) = a_\tau(\Gamma^\dagger) = -\ln \bar{p}_0(\Gamma)$ (thus $A[\Gamma_t]$ and $\bar{A}[\bar{\Gamma}_t]$ are completely determined) due to the definition of the time-reversal $\bar{a}_t(\Gamma) = a_{\tau-t}(\Gamma^\dagger)$, Eq. (5) will simply reduce to

$$\frac{\bar{P}_\tau(-A)}{P_\tau(A)} = e^{-A}. \tag{35}$$

Here $P_\tau(A) \equiv \int_{\mathfrak{S}_\tau} d\Gamma\, P_\tau(A,\Gamma)$ and $\bar{P}_\tau(A) \equiv \int_{\mathfrak{S}_0} d\Gamma\, \bar{P}_\tau(A,\Gamma)$. The quantity $A[\Gamma_t] = \beta Q[\Gamma_t] - \ln \bar{p}_0(\Gamma_\tau^\dagger) + \ln p_0(\Gamma_0)$ depends on the initial distributions of both the forward and the backward processes, which differs markedly from the cases of the heat and the work. The integral version of Eq. (35) reads

$$\langle e^{-A}\rangle = \left\langle \frac{p_\tau^a(\Gamma_\tau)}{p_0(\Gamma_0)} e^{-\beta Q}\right\rangle = 1, \tag{36}$$

where the auxiliary final distribution in the phase space $p_\tau^a(\Gamma) \equiv \bar{p}_0(\Gamma^\dagger)$ can be an arbitrary one, since the choice of $\bar{p}_0(\Gamma)$ has no restriction. The generalized version of Eq. (36) to the cases without the DB condition, which is merely to replace βQ with Δs_m, is the unified IFT (1) we mentioned in the beginning.

Despite the fact that the quantity A defined in such way always satisfies the DFT (35), usually it cannot be related to any thermodynamic variable we are familiar with. To see this, we consider the average of A over all the trajectories

$$\langle A\rangle = \beta\langle Q\rangle + D[p_\tau(\Gamma)||\bar{p}_0(\Gamma^\dagger)] + \langle s\rangle_\tau - \langle s\rangle_0, \tag{37}$$

where $D[p_1(\Gamma)||p_2(\Gamma)] \equiv \int d\Gamma p_1(\Gamma)\ln[p_1(\Gamma)/p_2(\Gamma)]$ is the Kullback-Leibler divergence between two probability distributions $p_1(\Gamma)$ and $p_2(\Gamma)$; $\langle s\rangle_t$ is the ensemble-average entropy of the system at time t; and $p_\tau(\Gamma)$ is the final distribution in the phase space determined by the real dynamic evolution. A special case in which A has a transparent physical meaning is when $\bar{p}_0(\Gamma^\dagger) = p_\tau(\Gamma)$. In this case the second term in Eq. (37) vanishes. For such a choice of the initial distribution for the backward process, A is the trajectory-dependent total entropy production $\Delta s_{\rm tot}$ for the forward process, and $\langle A\rangle$ consists with the ensemble-average value $\langle \Delta s_{\rm tot}\rangle$. However, since $\bar{a}_\tau(\Gamma)$ has been determined by $p_0(\Gamma)$ as $-\ln p_0(\Gamma^\dagger)$, $p_0(\Gamma^\dagger)$ usually differs from the final phase-space point distribution $\bar{p}_\tau(\Gamma)$ starting from the initial distribution $\bar{p}_0(\Gamma) = p_\tau(\Gamma^\dagger)$ and driven by the time-reversed protocol [30]. As a result, the functional $\bar{A}[\bar{\Gamma}_t] = \beta Q[\bar{\Gamma}_t] - \ln p_0(\bar{\Gamma}_\tau^\dagger) + \ln \bar{p}_0(\bar{\Gamma}_0)$ cannot be regarded as the total entropy production of the time-reversed trajectory $\bar{\Gamma}_t$. Hence, even when we choose $\bar{p}_0(\Gamma^\dagger) = p_\tau(\Gamma)$, Eq. (35) cannot be regarded as a detailed EPI. On the other hand, $\bar{P}_\tau(A)$ is still a normalized distribution function. So we can obtain the EPI [5] by setting A to be $\Delta s_{\rm tot}$ without any problem,

$$\langle e^{-\Delta s_{\rm tot}}\rangle = 1. \tag{38}$$

Nevertheless, what if $p_0(\Gamma^\dagger)$ happens to be the final distribution of the time-reversed process? In fact, this kind of specific distribution, named the *echo state*, has been recently studied by Van den Broeck and collaborators [31], which clarifies the initial conditions of both the detailed EPI and the other DFTs first proposed in Ref. [32]. Similarly to the discrete case that Van den Broeck *et al.* discussed, for any given FPE (6), the corresponding echo state $p_0^{\rm echo}(\Gamma)$ for the detailed EPI can be uniquely determined by solving

$$p_0^{\rm echo}(\Gamma) = \mathcal{T}\{e^{\int_0^\tau dt\bar{\mathcal{L}}_t^\dagger}\}\mathcal{T}\{e^{\int_0^\tau dt\mathcal{L}_t}\}p_0^{\rm echo}(\Gamma), \tag{39}$$

where $\mathcal{T}\{\dots\}$ denotes the time-ordered product. For such a choice of the initial distribution, the detailed EPI

$$\frac{\bar{P}_\tau(-\Delta s_{\rm tot})}{P_\tau(\Delta s_{\rm tot})} = e^{-\Delta s_{\rm tot}} \tag{40}$$

holds unambiguously. This can be regarded as a generalization of the detailed EPI for the steady state [5], where $\mathcal{L}_t$ should be time independent, i.e., the work parameter should be fixed.

IV. APPLICATIONS OF THE REFINED UNIFIED FT

A. Calculating distribution functions of work and heat for arbitrary initial states

Let us come back to the main result (5). If we only set the initial distribution of the backward process to be

$\bar{p}_0(\Gamma) = e^{-\bar{a}_0(\Gamma)}/\int_{\mathfrak{S}_\tau} d\Gamma e^{-\bar{a}_0(\Gamma)}$, we will obtain

$$P_\tau(A) = e^A \int_{\mathfrak{S}_\tau} d\Gamma e^{-\bar{a}_0(\Gamma)} \int_{\mathfrak{S}_0} d\Gamma \bar{P}_\tau(-A,\Gamma) p_0(\Gamma^\dagger) e^{\bar{a}_\tau(\Gamma)}. \tag{41}$$

This equation implies that $P_\tau(A)$ can be easily computed for arbitrary initial distribution $p_0(\Gamma)$ as long as we know the joint distribution function $\bar{P}_\tau(A,\Gamma)$ corresponding to the particular initial distribution $\bar{p}_0(\Gamma)$ for the time-reversed process. Specially, if $p_0(\Gamma)$ is a δ function, Eq. (41) will become

$$P_\tau(A|\Gamma_0) = e^{A+a_0(\Gamma_0)} \bar{P}_\tau(-A,\Gamma_0^\dagger) \int_{\mathfrak{S}_\tau} d\Gamma e^{-\bar{a}_0(\Gamma)}, \tag{42}$$

where $P_\tau(A|\Gamma_0)$ denotes the distribution of A in the condition where the initial state is known to be Γ_0. In fact, Eq. (41) and Eq. (42) are equivalent to each other, similarly to the equivalence between the dynamic property (7) and the DB condition, due to the additivity of the probability for exclusive events, $P_\tau(A) = \int_{\mathfrak{S}_0} d\Gamma P_\tau(A|\Gamma) p_0(\Gamma)$. As two examples, we can calculate the work and the heat distributions for any initial distributions $p_0(\Gamma)$ $[p_0(m)]$ via the following two formulas:

$$P_\tau(W) = e^{\beta(W-\Delta F)} \int_{\mathfrak{S}_0} d\Gamma \bar{P}_\tau(-W,\Gamma^\dagger) \frac{p_0(\Gamma)}{p_0^{\rm eq}(\Gamma)}, \tag{43}$$

$$P_\tau(Q) = e^{\beta Q} \sum_m \bar{P}_\tau(-Q,m^\dagger) \frac{p_0(m)}{p_0^u(m)}. \tag{44}$$

These relations provide an alternative approach to obtain the work (heat) statistics besides, e.g., directly solving the Feynman-Kac formula for the forward process with an arbitrary initial condition. Such approach may be advantageous in certain cases due to the specific initial condition of the backward process.

As an example, let us consider an overdamped breathing Brownian harmonic oscillator in one dimension, of which the generator in the FPE reads

$$\mathcal{L}_t = \frac{1}{\gamma} \partial_x (k_t x + \beta^{-1} \partial_x), \tag{45}$$

with the protocol chosen to be $k_t = k_0 + \kappa t$. According to Ref. [33], using the Gaussian ansatz, the Feynman-Kac formula (partial differential equation) of this model can be reduced to a set of first-order ordinary differential equations, which can be easily solved numerically. Furthermore, if the initial state is the equilibrium state and the driving speed is slow, we can even perturbatively work out some analytical results, such as the mean work correction in the linear response regime. However, the direct perturbative analysis breaks down if the initial state deviates significantly from the equilibrium one. To bypass the difficulty, we make use of the following equivalent form of Eq. (43):

$$G_\tau(\mu) = e^{-\beta\Delta F} \int_{\mathfrak{S}_0} d\Gamma \bar{G}_\tau(-\mu + i\beta,\Gamma^\dagger) \frac{p_0(\Gamma)}{p_0^{\rm eq}(\Gamma)}, \tag{46}$$

where $\bar{G}_\tau(\mu,\Gamma)$ has an exact perturbative solution due to its equilibrium initial state. In particular, if we choose $p_0(x) = (\beta' k_0/2\pi)^{\frac{1}{2}} e^{-\beta' k_0 x^2/2}$, the canonical distribution of another temperature β'^{-1}, using the conclusions in Ref. [33], then we will obtain a Gaussian distribution in the slow driving limit, of which the mean reads (see Appendix E for details)

$$\langle W \rangle = \Delta F + \frac{\gamma}{8\beta} \left[\left(\frac{2\beta}{\beta'} - 1 \right) \frac{1}{k_0^2} - \frac{1}{k_\tau^2} \right] \kappa + O(\kappa^2), \tag{47}$$

while the variance σ_W^2 is the same as that for the equilibrium initial state. Here the free-energy difference $\Delta F = \ln \sqrt{k_\tau/k_0}$ and $k_\tau = k_0 + \kappa\tau$. A notable inference follows that for a squeezing process ($k_\tau > k_0$), $\langle W \rangle$ will be less than ΔF when $\beta'^{-1} < (k_0^2/k_\tau^2 + 1)\beta^{-1}/2$, which implies the necessity of an equilibrium initial state (with the same temperature of the heat bath) for the validity of the minimum work principle. Actually, the above analysis can be generalized to arbitrary slowly driven overdamped Langevin systems, whose Feynman-Kac formulas have a unified analytical treatment [34].

B. Deriving FTs for feedback control processes

Although the canonical distribution is much more common than other ones in real cases, the initial distribution corresponding to a specific protocol can be rather irregular in feedback-control processes. In the extreme cases, we can exactly measure the initial phase-space point and then choose its unique protocol. More generally, we assign a protocol to a region in the phase space, while the region can be arbitrary in principle. According to these observations, we infer that our main result may have potential applications in thermodynamics with the feedback control.

Actually, we can easily derive one of the Sagawa-Ueda equalities [35] based on our main result or its side products. We apply Eq. (43) to a protocol λ_t^y which corresponds to the measurement result y,

$$P_\tau(W|y) = e^{\beta(W-\Delta F^y)} \int_{\mathfrak{S}_0} d\Gamma \bar{P}_\tau(-W,\Gamma|y) \frac{p_0(\Gamma^\dagger|y)}{p_0^{\rm eq}(\Gamma^\dagger)}. \tag{48}$$

Here the conditional initial distribution in the phase space $p_0(\Gamma|y)$ satisfies

$$p_0(\Gamma|y)p(y) = p(y|\Gamma)p_0^{\rm eq}(\Gamma) \tag{49}$$

as long as the system is initially prepared in a thermal equilibrium state, with $p(y) \equiv \int_{\mathfrak{S}_0} d\Gamma p(y|\Gamma) p_0^{\rm eq}(\Gamma)$ as the probability density whereby the measurement result turns out to be y. $p(y|\Gamma)$ is the conditional probability density that the measurement result is y conditioned on the phase-space point Γ. For simplicity, we assume that the measurement has the property $p(y^\dagger|\Gamma^\dagger) = p(y|\Gamma)$ [35]. This relation is obviously satisfied for the error-free measurement $p(y|\Gamma) = \delta(y-\Gamma)$, and the Gaussian-error measurement $p(y|\Gamma) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-|y-\Gamma|^2/2\sigma^2}$, with $\Gamma = (\boldsymbol{r},\boldsymbol{p})$. Combining Eq. (48) and Eq. (49), we obtain

$$p(y)P_\tau(W|y)e^{-\beta(W-\Delta F^y)} = \int_{\mathfrak{S}_0} d\Gamma \bar{P}_\tau(-W,\Gamma|y) p(y|\Gamma^\dagger). \tag{50}$$

After integrating Eq. (50) over both y and W, we get its integral version,

$$\langle e^{-\beta(W-\Delta F)} \rangle = \eta, \tag{51}$$

where

$$\eta = \int_{\mathfrak{S}_0} dy d\Gamma p(y^\dagger|\Gamma)\bar{p}_\tau(\Gamma|y) = \int_{\mathfrak{S}_0} dy \bar{p}(y^\dagger|y). \quad (52)$$

Here $\bar{p}_\tau(\Gamma|y) \equiv \int dW \bar{P}_\tau(W,\Gamma|y)$ is the final distribution of the phase-space point for the backward process, and $\bar{p}(y'|y)$ is the probability density that the final phase-space point is measured to be y', after being driven by the protocol $\bar{\lambda}_t^y$ from the canonical time-reversed initial distribution.

The other main result in Ref. [35] reads

$$\langle e^{-\beta(W-\Delta F)-I}\rangle = 1, \quad (53)$$

where I is the initial mutual information between the system and the measurement device. However, in both Eq. (51) and Eq. (53) it is required that the initial distribution must be a canonical distribution. To construct a more general information-involved FT, Sagawa and Ueda proposed

$$\langle e^{-\sigma+\Delta I}\rangle = 1 \quad (54)$$

in Ref. [36] as a generalization of the EPI. Here ΔI is the difference of the initial and the final mutual information, $\sigma \equiv \beta Q + \Delta s$ and $\Delta s \equiv -\ln p_\tau(\Gamma_\tau) + \ln p_0(\Gamma_0)$, with $p_\tau(\Gamma) = \int dy p_\tau(\Gamma|y)p(y)$. The above two Sagawa-Ueda equalities (53) and (54) are actually contained in the unified IFT (1) if we use the point of view in Ref. [36] to regard the combination of the original system and the device as a composite system (see Appendix F). However, the counterpart of Eq. (51) as another generalization of the EPI seems to be unexplored so far. Now we can derive it quite straightforwardly in analogy to the derivation of Eq. (51). Following the same procedure as that from Eq. (48) to Eq. (50), we obtain

$$p(y)P_\tau(\sigma|y)e^{-\sigma} = \int_{\mathfrak{S}_0} d\Gamma \bar{P}_\tau(-\sigma,\Gamma|y)p(y|\Gamma^\dagger), \quad (55)$$

which indicates

$$\langle e^{-\sigma}\rangle = \eta. \quad (56)$$

The expression of η is exactly the same as Eq. (52), but $\bar{p}(y'|y)$ must be associated with the initial distribution of the time-reversed process, which is the real time reversal of the forward final distribution, i.e., $\bar{p}_0(\Gamma) = p_\tau(\Gamma^\dagger)$, instead of a canonical one.

As an illustrative example of Eq. (56), we consider the Szilard engine [37], as was used in Ref. [35] to illustrate Eqs. (51) and (53). If the process is reversible and the measurement is error free, we will always have $\Delta s = 0$ and $\sigma = \beta Q = -\ln 2$ (δ distribution), so $\langle e^{-\sigma}\rangle = 2$. On the other hand, if $y = l$ (r), i.e., the particle is found in the left (right) half chamber, λ_t^y will be the rightward (leftward) expansion. Thus $\bar{\lambda}_t^y$ will be the leftward (rightward) compression. It is easy to see that $\bar{p}(l|l)$ or $\bar{p}(r|r)$ ($l^\dagger = l$, $r^\dagger = r$, since l or r is related to the position that is invariant under time reversal) is simply 1, due to the fact that the particle must be always in the left (right) half chamber after the leftward (rightward) compression. Hence, $\eta = \bar{p}(l|l) + \bar{p}(r|r) = 2$, and Eq. (56) is indeed valid for such feedback control processes. Moreover, even if the process is far from the quasistatic one, we will still have $\eta = 2$ due to the same analysis. This result generally implies $-\langle\sigma\rangle \leqslant \ln 2$, which is one aspect of the Landauer's principle [38,39].

V. CONCLUSION

In summary, we propose a refined unified FT (5) under the DB condition. The refined unified FT is applicable to several thermodynamic variables, such as the heat, the work, and the trajectory entropy production, and is even more refined than the DFT. Compared with the previous unified IFT and DFT [7,8], our refined unified FT (5) eliminates the entanglement between the thermodynamic variable and the choice of the initial distributions in the phase space and thus is physically more natural and comprehensible. In particular, our refined unified FT generalizes the FTs of the work and the heat, such as the JE and the CFT, to arbitrary initial distributions (see Table I), as well as generates several new FTs that, as far as we know, were previously not known to researchers in this field, for example, the Jarzynski equality to an arbitrary initial state (26). We also revisit the validity of the DFT of entropy production [30–32]. The price we pay for the generalization to arbitrary initial distributions is that we need to know the joint distribution functions with the phase-space point. Due to such kinds of generalizations, our results might be valuable for the free-energy recovering experiment and free-energy computing as well as studying thermodynamics with information exchange, where the initial distributions are, to some extent, irregular. Based on the refined unified FT, we reproduce one (51) of the Sagawa-Ueda equalities and derive a new generalized EPI (56) for the feedback control processes.

ACKNOWLEDGMENTS

We thank Fei Liu for helpful discussions. H.T.Q. gratefully acknowledges support from the National Science Foundation of China under Grant No. 11375012 and The Recruitment Program of Global Youth Experts of China.

APPENDIX A: PATH-INTEGRAL DERIVATION OF THE MAIN RESULT (5)

In the path-integral representation, the DB condition is reflected by the following relation:

$$\beta Q[\Gamma_t] = \ln \frac{\mathcal{P}(\Gamma_t|\Gamma_0)}{\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)}, \quad (A1)$$

which is a crucial building block in the later derivation of the main result. Here $\mathcal{P}(\Gamma_t|\Gamma_0)$ [$\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)$] is the conditional probability density of the trajectory Γ_t ($\bar{\Gamma}_t$) when the initial phase-space point is known to be Γ_0 ($\bar{\Gamma}_0$) for the forward (backward) process. While Eq. (A1) has appeared in various references, such as its original discrete version in Ref. [40], to make the paper self-contained, we briefly derive it from the basic trajectory definition of heat and the dynamic property (7). To do this, we first consider a sufficiently short time interval $[t,t+dt]$ in the forward process, during which

$$p(\Gamma_{t+dt}|\Gamma_t) = \int_{\mathfrak{S}_{t+\frac{dt}{2}}} d\Gamma \delta(\Gamma-\Gamma_{t+dt})(1+\mathcal{L}_{t+\frac{dt}{2}}dt)\delta(\Gamma-\Gamma_t). \quad (A2)$$

Here $p(\Gamma_{t_2}|\Gamma_{t_1}) \equiv \int_{\mathfrak{S}_{t_2}} d\Gamma \delta(\Gamma-\Gamma_{t_2})\mathcal{T}\{e^{\int_{t_1}^{t_2} dt \mathcal{L}_t}\}\delta(\Gamma-\Gamma_{t_1})$ is the conditional probability density that the phase-space point at t_2 is Γ_{t_2} starting from Γ_{t_1} (a δ distribution) at $t_1(< t_2)$, and the terms of the order of magnitude of $(dt)^2$ are neglected. Similarly, for a short time interval $[\tau - t - dt, \tau - t]$ in the backward process, we have

$$\begin{aligned}\bar{p}(\bar{\Gamma}_{\tau-t}|\bar{\Gamma}_{\tau-t-dt}) &= \int_{\mathfrak{S}_{t+\frac{dt}{2}}} d\Gamma\delta(\Gamma - \bar{\Gamma}_{\tau-t})(1 + \bar{\mathcal{L}}_{\tau-t-\frac{dt}{2}}dt)\delta(\Gamma - \bar{\Gamma}_{\tau-t-dt}) \\ &= \int_{\mathfrak{S}_{t+\frac{dt}{2}}} d\Gamma\delta(\Gamma - \Gamma_t)(1 + \mathcal{L}^\dagger_{t+\frac{dt}{2}}dt)\delta(\Gamma - \Gamma_{t+dt}).\end{aligned} \tag{A3}$$

Combining Eqs. (A2) and (A3) with the dynamic property (7), we deduce that

$$\ln\frac{p(\Gamma_{t+dt}|\Gamma_t)}{\bar{p}(\bar{\Gamma}_{\tau-t}|\bar{\Gamma}_{\tau-t-dt})} = -\beta\partial_\Gamma U_{t+\frac{dt}{2}}(\Gamma_{t+\frac{dt}{2}})\dot{\Gamma}_{t+\frac{dt}{2}}dt. \tag{A4}$$

Due to the Markovianity of the dynamics, the probability density of a trajectory can be expressed as $\mathcal{P}[\Gamma_t|\Gamma_0] = \lim_{K\to\infty}\prod_{k=1}^{K} p(\Gamma_{\frac{k}{K}\tau}|\Gamma_{\frac{k-1}{K}\tau})$. Therefore, we obtain

$$\begin{aligned}\ln\frac{\mathcal{P}[\Gamma_t|\Gamma_0]}{\bar{\mathcal{P}}[\bar{\Gamma}_t|\bar{\Gamma}_0]} &= \lim_{K\to\infty}\sum_{k=1}^{K}\ln\frac{p(\Gamma_{\frac{k}{K}\tau}|\Gamma_{\frac{k-1}{K}\tau})}{\bar{p}(\bar{\Gamma}_{\frac{K-k+1}{K}\tau}|\bar{\Gamma}_{\frac{K-k}{K}\tau})} \\ &= -\beta\int_0^\tau dt\,\partial_\Gamma U(\Gamma_t)\dot{\Gamma}_t = \beta Q[\Gamma_t].\end{aligned} \tag{A5}$$

Now we start the actual derivation of Eq. (5). Recalling the definition of $P_\tau(A,\Gamma)$, in the path-integral representation we have

$$P_\tau(A,\Gamma) = \int\mathcal{D}[\Gamma_t]\mathcal{P}(\Gamma_t|\Gamma_0)p_0(\Gamma_0)\delta(A[\Gamma_t] - A)\delta(\Gamma_\tau - \Gamma), \tag{A6}$$

with $\mathcal{D}[\Gamma_t] \equiv \lim_{K\to\infty}\prod_{k=0}^{K} d\Gamma_{k\tau/K}$. In this manner, the denominator on the left-hand side of Eq. (5), denoted by L for convenience, can be rewritten as

$$L = \int\mathcal{D}[\Gamma_t]\int_{\mathfrak{S}_\tau} d\Gamma e^{a_\tau(\Gamma)}\mathcal{P}(\Gamma_t|\Gamma_0)p_0(\Gamma_0)\bar{p}_0(\Gamma^\dagger) \times\delta(A[\Gamma_t] - A)\delta(\Gamma_\tau - \Gamma). \tag{A7}$$

By making use of Eq. (A1) and the property of the δ function, we obtain

$$L = \int\mathcal{D}[\Gamma_t]e^{\beta Q[\Gamma_t]+a_\tau(\Gamma_\tau)}\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)\bar{p}_0(\Gamma^\dagger_\tau) \times p_0(\Gamma_0)\delta(A[\Gamma_t] - A). \tag{A8}$$

Then we use the definition of A (3) to replace $\beta Q[\Gamma_t] + a_\tau(\Gamma_\tau)$ with $A[\Gamma_t] + a_0(\Gamma_0)$, which leads to

$$L = \int\mathcal{D}[\Gamma_t]e^{A[\Gamma_t]+a_0(\Gamma_0)}\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)\bar{p}_0(\Gamma^\dagger_\tau) \times p_0(\Gamma_0)\delta(A[\Gamma_t] - A). \tag{A9}$$

Again we use the property of the δ function, getting

$$L = e^A\int\mathcal{D}[\Gamma_t]e^{a_0(\Gamma_0)}\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)\bar{p}_0(\Gamma^\dagger_\tau) \times p_0(\Gamma_0)\delta(A[\Gamma_t] - A). \tag{A10}$$

According to the definition of the time reversal, it can be checked that $\bar{A}[\bar{\Gamma}_t] = -A[\Gamma_t]$, e.g., $Q[\bar{\Gamma}_t] = -Q[\Gamma_t]$, so

$$L = e^A\int\mathcal{D}[\bar{\Gamma}_t]\int_{\mathfrak{S}_0} d\Gamma e^{\bar{a}_\tau(\bar{\Gamma}_\tau)}\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)\bar{p}_0(\bar{\Gamma}_0) \times p_0(\bar{\Gamma}^\dagger_\tau)\delta(\bar{A}[\bar{\Gamma}_t] + A)\delta(\bar{\Gamma}_\tau - \Gamma). \tag{A11}$$

The insertion of the δ function aims at constructing the joint distribution function of $\bar{A}$ for the time-reversed process and the final phase-space point. Particularly, we rewrite Eq. (A11) as

$$L = e^A\int\mathcal{D}[\bar{\Gamma}_t]\int_{\mathfrak{S}_0} d\Gamma e^{\bar{a}_\tau(\Gamma)}\bar{\mathcal{P}}(\bar{\Gamma}_t|\bar{\Gamma}_0)\bar{p}_0(\bar{\Gamma}_0) \times p_0(\Gamma^\dagger)\delta(\bar{A}[\bar{\Gamma}_t] - (-A))\delta(\bar{\Gamma}_\tau - \Gamma). \tag{A12}$$

Recall the path-integral representation of the joint distribution function (A6), Eq. (A12) actually leads to

$$L = e^A\int_{\mathfrak{S}_0} d\Gamma\bar{P}_\tau(-A,\Gamma)p_0(\Gamma^\dagger)e^{\bar{a}_\tau(\Gamma)}, \tag{A13}$$

which is nothing but the numerator on the left-hand side of Eq. (5) multiplied by e^A.

APPENDIX B: VALIDITY OF THE DYNAMIC PROPERTY (7) FOR THE LANGEVIN DYNAMICS

We focus on the one-dimensional case here. The generalization to higher dimensions is straightforward, though an anisotropic effect may emerge. For the overdamped Langevin dynamics, the generator $\mathcal{L}_t$ in the FPE (6) reads [41]

$$\mathcal{L}_t = \frac{1}{\gamma}\partial_x(\partial_x U_t + \beta^{-1}\partial_x). \tag{B1}$$

Here γ is the viscous friction coefficient; $U_t \equiv V_t(x)$ only depends on the position, indicating

$$\mathcal{L}^\dagger_t = \mathcal{L}_t. \tag{B2}$$

Notice that $\int_{-\infty}^{+\infty} dx f(x)\frac{d}{dx}g(x) = -\int_{-\infty}^{+\infty} dx g(x)\frac{d}{dx}f(x)$ for any normalizable functions $f(x)$ and $g(x)$ [so $\lim_{x\to\pm\infty} f(x)g(x) = 0$]; the transpose operator of $\mathcal{L}_t$ should be

$$\mathcal{L}^T_t = \frac{1}{\gamma}(-\partial_x U_t + \beta^{-1}\partial_x)\partial_x. \tag{B3}$$

Before checking the dynamic property (7), we introduce the following two useful relations:

$$\begin{aligned}e^{\beta U_t}\partial_x e^{-\beta U_t} &= \partial_x - \beta\partial_x U_t, \\ e^{\beta U_t}\partial_x^2 e^{-\beta U_t} &= (\partial_x - \beta\partial_x U_t)^2 \\ &= \partial_x^2 - 2\beta\partial_x U_t\partial_x - \beta\partial_x^2 U_t + \beta^2(\partial_x U_t)^2.\end{aligned} \tag{B4}$$

With these relations in hand, we start to calculate the left-hand side of Eq. (7):

$$\begin{aligned}&e^{\beta U_t}\mathcal{L}^\dagger_t e^{-\beta U_t} \\ &= \frac{1}{\gamma}\big[\partial_x^2 U_t + \partial_x U_t\partial_x - \beta(\partial_x U_t)^2 \\ &\quad + \beta^{-1}\partial_x^2 - 2\partial_x U_t\partial_x - \partial_x^2 U_t + \beta(\partial_x U_t)^2\big] \\ &= \frac{1}{\gamma}\big(-\partial_x U_t\partial_x + \beta^{-1}\partial_x^2\big) = \mathcal{L}^T_t.\end{aligned} \tag{B5}$$

Thus, Eq. (7) has been confirmed to be valid for the overdamped Langevin dynamics.

Let us move on to the underdamped Langevin dynamics. The generator $\mathcal{L}_t$ reads [41]

$$\mathcal{L}_t = -\frac{p}{m}\partial_x + \partial_p\left(\partial_x U_t + \gamma\frac{p}{m}\right) + \gamma\beta^{-1}\partial_p^2, \tag{B6}$$

based on which we can obtain its time-reversed operator and its transpose operator

$$\begin{aligned}\mathcal{L}_t^\dagger &= \frac{p}{m}\partial_x + \partial_p\left(-\partial_x U_t + \gamma\frac{p}{m}\right) + \gamma\beta^{-1}\partial_p^2,\\ \mathcal{L}_t^T &= \frac{p}{m}\partial_x - \left(\partial_x U_t + \gamma\frac{p}{m}\right)\partial_p + \gamma\beta^{-1}\partial_p^2.\end{aligned} \tag{B7}$$

Here $U_t \equiv p^2/2m + V_t(x)$ depends on both the position and the momentum. Again we introduce two useful relations first:

$$\begin{aligned}e^{\beta U_t}\partial_p e^{-\beta U_t} &= \partial_p - \beta\frac{p}{m}\\ e^{\beta U_t}\partial_p^2 e^{-\beta U_t} &= \left(\partial_p - \beta\frac{p}{m}\right)^2\\ &= \partial_p^2 - 2\beta\frac{p}{m}\partial_p - \frac{\beta}{m} + \left(\frac{\beta p}{m}\right)^2.\end{aligned} \tag{B8}$$

Then we calculate the left-hand side of Eq. (7):

$$\begin{aligned}&e^{\beta U_t}\mathcal{L}_t^\dagger e^{-\beta U_t}\\ &= \frac{p}{m}(\partial_x - \beta\partial_x U_t)\\ &\quad + \frac{\gamma}{m} + \left(-\partial_x U_t + \gamma\frac{p}{m}\right)\left(\partial_p - \beta\frac{p}{m}\right)\\ &\quad + \gamma\beta^{-1}\partial_p^2 - 2\gamma\frac{p}{m}\partial_p - \frac{\gamma}{m} + \gamma\beta\left(\frac{p}{m}\right)^2\\ &= \frac{p}{m}\partial_x - \partial_x U_t\partial_p - \gamma\frac{p}{m}\partial_p + \gamma\beta^{-1}\partial_p^2 = \mathcal{L}_t^T.\end{aligned} \tag{B9}$$

Thus, Eq. (7) has also been confirmed to be valid for the underdamped Langevin dynamics.

APPENDIX C: GENERAL EQUATION OF MOTION FOR THE JOINT DISTRIBUTION FUNCTION

Instead of the definition in the main text (3), let u¡¯s consider a thermodynamic variable associated with a process generally expressed as

$$A[\Gamma_t] = \int_0^\tau dt\partial_t w_t(\Gamma_t) + \int_0^\tau dt\dot{\Gamma}_t\partial_\Gamma q_t(\Gamma_t). \tag{C1}$$

Here $w_t(\Gamma)$ and $q_t(\Gamma)$ can be arbitrary time-dependent functions with respect to the phase-space point. We will show that if the generator is $\mathcal{L}_t$, then the equation of motion for $P_t(A,\Gamma)$ will be

$$\partial_t P_t(A,\Gamma) = [e^{-q_t(\Gamma)\partial_A}\mathcal{L}_t e^{q_t(\Gamma)\partial_A} - \partial_t w_t(\Gamma)\partial_A]P_t(A,\Gamma). \tag{C2}$$

It is instructive to first consider a simple case where $q_t(\Gamma)=0$. In this case, $P_{t+dt}(A,\Gamma)$ is related to $P_t(A,\Gamma)$ via the following relation [41] (terms with the magnitude of $(dt)^2$ are ignored):

$$\begin{aligned}P_{t+dt}(A,\Gamma) &= (1+\mathcal{L}_t dt)P_t[A - \partial_t w_t(\Gamma)dt,\Gamma]\\ &= \mathcal{L}_t P_t(A,\Gamma)dt + P_t[A - \partial_t w_t(\Gamma)dt,\Gamma],\end{aligned} \tag{C3}$$

which implies

$$\partial_t P_t(A,\Gamma) = [\mathcal{L}_t - \partial_t w_t(\Gamma)\partial_A]P_t(A,\Gamma). \tag{C4}$$

This result is familiar to us since it is the precursor of the Feynman-Kac formula before performing the integral transformation. In order to generalize to the case with $q_t(\Gamma)\neq 0$, we rewrite the first term in the rightmost of Eq. (C3) as

$$\begin{aligned}\mathcal{L}_t P_t(A,\Gamma)dt = \int_{\mathfrak{S}_t} d\Gamma' d\Gamma'' P_t(A,\Gamma'')\\ \times\delta(\Gamma'-\Gamma)\mathcal{L}_t'\delta(\Gamma'-\Gamma'')dt,\end{aligned} \tag{C5}$$

where $\mathcal{L}_t'$ acts only on Γ'. Now the physical meaning is transparent: This is the contribution to $P_{t+dt}(A,\Gamma)$ due to motions in the phase space, consisting of both parts that come from other phase-space points ($\Gamma''\neq\Gamma$) and leave Γ [$\Gamma''=\Gamma$, this is necessary to correct the second term in the rightmost of Eq. (C3), the contribution due to the temporal variation of $w_t(\Gamma)$]. Based on such an interpretation, when a nonzero $q_t(\Gamma)$ appears, we simply modify Eq. (C3) as

$$\begin{aligned}&P_{t+dt}(A,\Gamma)\\ &= \int_{\mathfrak{S}_t} d\Gamma' d\Gamma'' P_t[A - q_t(\Gamma) + q_t(\Gamma''),\Gamma'']\\ &\quad\times\delta(\Gamma'-\Gamma)\mathcal{L}_t'\delta(\Gamma'-\Gamma'')dt + P_t[A-\partial_t w_t(\Gamma)dt,\Gamma],\end{aligned} \tag{C6}$$

due to the second term in Eq. (C1). Since the first term in Eq. (C6) is already of the order of the magnitude of dt, we do not have to further add modifications like $\partial_t w_t(\Gamma)dt$ or $\partial_t q_t(\Gamma)dt$ in addition to $-q_t(\Gamma)+q_t(\Gamma'')$, which will merely result in differences of the order of the magnitude of $(dt)^2$. By using the identity $e^{a\frac{d}{dx}}f(x) = f(x+a)$ as well as the property of the δ function, we can simplify Eq. (C6) as

$$\begin{aligned}P_{t+dt}(A,\Gamma) &= e^{-q_t(\Gamma)\partial_A}\mathcal{L}_t e^{q_t(\Gamma)\partial_A}P_t(A,\Gamma)dt\\ &\quad + P_t(A-\partial_t w_t(\Gamma)dt,\Gamma),\end{aligned} \tag{C7}$$

which finally leads to Eq. (C2).

Now let us return to the functional in the main text (3). By using the identity

$$a_\tau(\Gamma_\tau) - a_0(\Gamma_0) = \int_0^\tau dt\partial_t a_t(\Gamma_t) + \int_0^\tau dt\dot{\Gamma}_t\partial_\Gamma a_t(\Gamma_t), \tag{C8}$$

we find that Eq. (3) can be rewritten as

$$A[\Gamma_t] = \int_0^\tau dt\partial_t a_t(\Gamma_t) + \int_0^\tau dt\dot{\Gamma}_t\partial_\Gamma[a_t(\Gamma_t) - \beta U_t(\Gamma_t)]. \tag{C9}$$

Comparing Eq. (C9) with Eq. (C1), we have $w_t(\Gamma) = a_t(\Gamma)$ and $q_t(\Gamma) = a_t(\Gamma) - \beta U_t(\Gamma)$. Substituting them into Eq. (C2), we get the equation of motion in the main text (8).

APPENDIX D: FT OF WORK (21) FOR DRIVEN ISOLATED QUANTUM SYSTEMS

We generally denote the Hamiltonian of an isolated quantum system by $\mathcal{H}_t$, whose time reversal is determined by

$\bar{\mathcal{H}}_t \equiv \Theta\mathcal{H}_{\tau-t}\Theta^{-1}$, with Θ to be the antiunitary time-reversal operator. To be specific, the quantum version of Eq. (21) can be written as

$$\sum_n P_\tau(W,|n\rangle)\bar{p}_0(\Theta|n\rangle)e^{\beta E_\tau^n} = e^{\beta W}\sum_{\bar{m}} \bar{P}_\tau(-W,|\bar{m}\rangle)p_0(\Theta^{-1}|\bar{m}\rangle)e^{\beta \bar{E}_\tau^{\bar{m}}}. \tag{D1}$$

Here $|n\rangle$ ($|\bar{m}\rangle$) is an eigenstate of $\mathcal{H}_\tau$ ($\bar{\mathcal{H}}_\tau$) with the eigenenergy E_τ^n ($\bar{E}_\tau^{\bar{m}}$); $p_0(\Theta^{-1}|\bar{m}\rangle)$ [$\bar{p}_0(\Theta|n\rangle)$] is the probability that the initial state for the forward (backward) process is measured to be $\Theta^{-1}|\bar{m}\rangle$ ($\Theta|n\rangle$), which is obviously an eigenstate of $\mathcal{H}_0$ ($\bar{\mathcal{H}}_0$) due to the former definitions. Such probability can be evaluated in terms of the initial density operator ϱ_0 ($\bar{\varrho}_0$) via

$$p_0(\Theta^{-1}|\bar{m}\rangle) = \langle\bar{m}|\Theta\varrho_0\Theta^{-1}|\bar{m}\rangle, \qquad \bar{p}_0(\Theta|n\rangle) = \langle n|\Theta^{-1}\bar{\varrho}_0\Theta|n\rangle, \tag{D2}$$

where $[\varrho_0,\mathcal{H}_0] = [\bar{\varrho}_0,\bar{\mathcal{H}}_0] = 0$ due to its structure assumed in the main text. Accordingly, $[\Theta\varrho_0\Theta^{-1},\bar{\mathcal{H}}_\tau] = [\Theta^{-1}\bar{\varrho}_0\Theta,\mathcal{H}_\tau] = 0$ holds subsequently. Based on the two-point projection measurement definition of quantum work, the joint distribution functions should be

$$P_\tau(W,|n\rangle) = \sum_m |\langle n|\mathcal{U}_{\tau,0}|m\rangle|^2 p_0(|m\rangle)\delta\big(W - E_\tau^n + E_0^m\big),$$
$$\bar{P}_\tau(W,|\bar{m}\rangle) = \sum_{\bar{n}} |\langle \bar{m}|\bar{\mathcal{U}}_{\tau,0}|\bar{n}\rangle|^2 \bar{p}_0(|\bar{n}\rangle)\delta\big(W - E_\tau^{\bar{m}} + E_0^{\bar{n}}\big), \tag{D3}$$

where $p_0(|m\rangle)$ [$\bar{p}_0(|\bar{n}\rangle)$] can also be related to ϱ_0 ($\bar{\varrho}_0$) by $\langle m|\varrho_0|m\rangle$ ($\langle\bar{n}|\bar{\varrho}_0|\bar{n}\rangle$).

To prove Eq. (D1), we again take the characteristic function-based approach, which has been widely used in the studies of quantum thermodynamics [42,43]. We take the inverse Fourier transformation on both sides of Eq. (D1) and obtain

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dW e^{i\mu W}\sum_n P_\tau(W,|n\rangle)\bar{p}_0(\Theta|n\rangle)e^{\beta E_\tau^n}\\
&= \sum_n\sum_m |\langle n|\mathcal{U}_{\tau,0}|m\rangle|^2 p_0(|m\rangle)\bar{p}_0(\Theta|n\rangle)e^{(i\mu+\beta)E_\tau^n - i\mu E_0^m}\\
&= \sum_n\sum_m \langle n|e^{i\nu\mathcal{H}_\tau}\mathcal{U}_{\tau,0}\varrho_0|m\rangle\langle m|e^{i\mu\mathcal{H}_0}\mathcal{U}_{0,\tau}\Theta^{-1}\bar{\varrho}_0\Theta|n\rangle\\
&= \mathrm{Tr}[e^{i\nu\mathcal{H}_\tau}\mathcal{U}_{\tau,0}\varrho_0 e^{-i\mu\mathcal{H}_0}\mathcal{U}_{0,\tau}\Theta^{-1}\bar{\varrho}_0\Theta]\ ,\\
&\int_{-\infty}^{+\infty} dW e^{\beta W}e^{i\mu W}\sum_{\bar{m}} \bar{P}_\tau(-W,|\bar{m}\rangle)\bar{p}_0(\Theta^{-1}|\bar{m}\rangle)e^{\beta \bar{E}_\tau^{\bar{m}}}\\
&= \sum_{\bar{m}}\sum_{\bar{n}} |\langle \bar{m}|\bar{\mathcal{U}}_{\tau,0}|\bar{n}\rangle|^2 \bar{p}_0(|\bar{n}\rangle)p_0(\Theta^{-1}|\bar{m}\rangle)e^{(i\mu+\beta)\bar{E}_0^{\bar{n}} - i\mu \bar{E}_\tau^{\bar{m}}}\\
&= \sum_{\bar{m}}\sum_{\bar{n}} \langle\bar{m}|e^{-i\mu\bar{\mathcal{H}}_\tau}\bar{\mathcal{U}}_{\tau,0}\bar{\varrho}_0|\bar{n}\rangle\langle\bar{n}|e^{i\nu\bar{\mathcal{H}}_0}\bar{\mathcal{U}}_{0,\tau}\Theta\varrho_0\Theta^{-1}|\bar{m}\rangle\\
&= \mathrm{Tr}[e^{-i\mu\bar{\mathcal{H}}_\tau}\bar{\mathcal{U}}_{\tau,0}\bar{\varrho}_0 e^{i\nu\bar{\mathcal{H}}_0}\bar{\mathcal{U}}_{0,\tau}\Theta\varrho_0\Theta^{-1}].
\end{aligned} \tag{D4}$$

Here $\mathcal{U}_{t,t'}$ ($\bar{\mathcal{U}}_{t,t'}$) is the time-evolution operator for the forward (backward) process, governed by the Schrödinger equation $i\hbar\partial_t\mathcal{U}_{t,t'} = \mathcal{H}_t\mathcal{U}_{t,t'}$, $\mathcal{U}_{t',t'} \equiv \mathcal{I}$ ($i\hbar\partial_t\bar{\mathcal{U}}_{t,t'} = \bar{\mathcal{H}}_t\bar{\mathcal{U}}_{t,t'}$, $\bar{\mathcal{U}}_{t',t'} \equiv \mathcal{I}$), $\mathcal{I}$ is the identity operator; both μ and $\nu \equiv \mu - i\beta$ are generally complex numbers, i.e., μ and ν are unnecessarily their complex conjugates μ^*, ν^* and are also independent of each other, due to the arbitrariness of β. By making use of the algebraic properties of Θ and the trace (see Ref. [43] for details), especially $\bar{\mathcal{U}}_{t,t'} = \Theta\mathcal{U}_{\tau-t,\tau-t'}\Theta^{-1}$, $\Theta e^{-i\kappa\mathcal{H}_t}\Theta^{-1} = e^{i\kappa^*\bar{\mathcal{H}}_{\tau-t}}$, $\mathrm{Tr}[\Theta^{-1}\mathcal{A}\Theta] = \mathrm{Tr}[\mathcal{A}^\dagger]$ and $\mathrm{Tr}[\mathcal{A}\mathcal{B}] = \mathrm{Tr}[\mathcal{B}\mathcal{A}]$ (be careful that this may be invalid if $\mathcal{A} = \Theta$), as well as the commutation relations $[\bar{\varrho}_0,\bar{\mathcal{H}}_0] = 0$ and $[\Theta\varrho_0\Theta^{-1},\bar{\mathcal{H}}_\tau] = 0$, we have

$$\begin{aligned}
&\mathrm{Tr}[e^{i\nu\mathcal{H}_\tau}\mathcal{U}_{\tau,0}\varrho_0 e^{-i\mu\mathcal{H}_0}\mathcal{U}_{0,\tau}\Theta^{-1}\bar{\varrho}_0\Theta]\\
&= \mathrm{Tr}[e^{i\nu\mathcal{H}_\tau}\Theta^{-1}\bar{\mathcal{U}}_{0,\tau}\Theta\varrho_0 e^{-i\mu\mathcal{H}_0}\Theta^{-1}\bar{\mathcal{U}}_{\tau,0}\bar{\varrho}_0\Theta]\\
&= \mathrm{Tr}[\Theta^{-1}e^{-i\nu^*\bar{\mathcal{H}}_0}\bar{\mathcal{U}}_{0,\tau}\Theta\varrho_0\Theta^{-1}e^{i\mu^*\bar{\mathcal{H}}_\tau}\bar{\mathcal{U}}_{\tau,0}\bar{\varrho}_0\Theta]\\
&= \mathrm{Tr}[\bar{\varrho}_0\bar{\mathcal{U}}_{0,\tau}e^{-i\mu\bar{\mathcal{H}}_\tau}\Theta\varrho_0\Theta^{-1}\bar{\mathcal{U}}_{\tau,0}e^{i\nu\bar{\mathcal{H}}_0}]\\
&= \mathrm{Tr}[e^{i\nu\bar{\mathcal{H}}_0}\bar{\mathcal{U}}_{0,\tau}\Theta\varrho_0\Theta^{-1}e^{-i\mu\bar{\mathcal{H}}_\tau}\bar{\mathcal{U}}_{\tau,0}\bar{\varrho}_0]\\
&= \mathrm{Tr}[e^{-i\mu\bar{\mathcal{H}}_\tau}\bar{\mathcal{U}}_{\tau,0}\bar{\varrho}_0 e^{i\nu\bar{\mathcal{H}}_0}\bar{\mathcal{U}}_{0,\tau}\Theta\varrho_0\Theta^{-1}].
\end{aligned} \tag{D5}$$

Thus the inverse Fourier transformation of the two sides of Eq. (D1) turns out to be the same. So far, the discrete version of the refined FT of work (21) has been confirmed to be generally valid for driven isolated quantum systems.

As an example, if $\varrho_0 = e^{-\beta\mathcal{H}_0}/Z_0(\beta)$ and $\bar{\varrho}_0 = e^{-\beta\bar{\mathcal{H}}_0}/\bar{Z}_0(\beta)$, $Z_0(\beta) \equiv \mathrm{Tr}[e^{-\beta\mathcal{H}_0}]$, $\bar{Z}_0(\beta) \equiv \mathrm{Tr}[e^{-\beta\bar{\mathcal{H}}_0}] = \mathrm{Tr}[e^{-\beta\mathcal{H}_\tau}] = Z_\tau(\beta)$ (due to $\mathrm{Tr}[\Theta^{-1}\mathcal{A}\Theta] = \mathrm{Tr}[\mathcal{A}^\dagger]$ and $\mathcal{H}_\tau^\dagger = \mathcal{H}_\tau$. This is the quantum analogy of footnote [22]), Eq. (D5) will become $G(\mu)/Z_\tau(\beta) = \bar{G}(-\mu + i\beta)/Z_0(\beta)$. This means that the quantum CFT holds even for systems without the time-reversal symmetry, e.g., a charged particle subjected to a time-dependent magnetic field. This is a generalization of Ref. [43], where $[\Theta,\mathcal{H}_t] = 0$ is assumed. However, we should be careful that the time-reversed Hamiltonian must be $\Theta\mathcal{H}_{\tau-t}\Theta^{-1}$ but usually not $\mathcal{H}_{\tau-t}$.

APPENDIX E: DETAILED CALCULATIONS ON THE BREATHING BROWNIAN OSCILLATOR

To be consistent with Ref. [33], we use the generating function $\rho_t(\lambda,x) \equiv G_t(i\lambda,x)$ instead of the characteristic function $G_t(\mu,x)$. We first present the main result in Ref. [33], which focused on an arbitrary process driven by the protocol k_t starting from the equilibrium state $p_0^{\mathrm{eq}}(x) \equiv (k_0\beta/2\pi)^{\frac{1}{2}}e^{-\beta k_0 x^2/2}$. By making the Gaussian ansatz

$$\rho_t(\lambda,x) = \sqrt{\frac{[\psi_\lambda(t)]^3}{2\pi\phi_\lambda(t)}}e^{-\frac{x^2\psi_\lambda(t)}{2\phi_\lambda(t)}}, \tag{E1}$$

the Feynman-Kac formula can be self-consistently reduced to the following two first-order ordinary differential equations:

$$\dot{\psi}_\lambda(t) = -\frac{\lambda\dot{k}_t}{2}\phi_\lambda(t), \tag{E2}$$

$$\dot{\phi}_\lambda(t) = -\frac{2k_t}{\gamma}\phi_\lambda(t) + \frac{2}{\beta\gamma}\psi_\lambda(t) - \frac{3\lambda\dot{k}_t}{2}\frac{\phi_\lambda^2(t)}{\psi_\lambda(t)}, \tag{E3}$$

with the initial condition to be $\psi_\lambda(0) = 1$ and $\phi_\lambda(0) = (\beta k_0)^{-1}$. To perform perturbative analysis, it is convenient to define $g_\lambda(t) \equiv \beta k_t\phi_\lambda(t)/\psi_\lambda(t)$. Based on Eqs. (E2) and (E3), it can be checked that $g_\lambda(t)$ satisfies the following Ricatti equation:

$$\dot{g}_\lambda(t) = -\frac{2k_t}{\gamma}[g_\lambda(t)-1] + \frac{\dot{k}_t}{k_t}g_\lambda(t)\left[1-\frac{\lambda}{\beta}g_\lambda(t)\right], \quad \text{(E4)}$$

with the initial condition to be $g_\lambda(0)=1$. Once $g_\lambda(t)$ is determined, $\psi_\lambda(\tau)=\langle e^{-\lambda W}\rangle$ can be obtained via $\ln\psi_\lambda(\tau) = -(\lambda/\beta)\int_0^\tau dt g_\lambda(t)\dot{k}_t/2k_t$. If we regard $\dot{k}_t$ in Eq. (E4) as a small quantity, the zeroth-order solution will simply read $g_\lambda^{(0)}(t)=1$, while the first-order correction should be $g_\lambda^{(1)}(t) = \gamma(1-\lambda/\beta)\dot{k}_t/2k_t^2$. Accordingly, $\ln\psi_\lambda(\tau)$ can be expressed as follows up to the first-order accuracy:

$$\ln\psi_\lambda(\tau) = -\frac{\lambda}{2\beta}\ln\frac{k_\tau}{k_0} - \frac{\gamma\lambda}{4\beta}\left(1-\frac{\lambda}{\beta}\right)\int_0^\tau dt\frac{\dot{k}_t^2}{k_t^3} + O(\dot{k}_t^2). \quad \text{(E5)}$$

In particular, for the protocol $k_t = k_0+\kappa t$, we have

$$\ln\psi_\lambda^{(1)}(\tau) = -\frac{\lambda}{2\beta}\ln\frac{k_\tau}{k_0} - \frac{\lambda}{8\beta}\left(1-\frac{\lambda}{\beta}\right)\left(\frac{1}{k_0^2}-\frac{1}{k_\tau^2}\right)\gamma\kappa. \quad \text{(E6)}$$

Now let us start to calculate the generating function of the work distribution for the nonequilibrium initial state $p_0(x) = (k_0\beta'/2\pi)^{\frac{1}{2}}e^{-\frac{1}{2}\beta' k_0 x^2}$. Using Eq. (46), we obtain

$$\begin{aligned}
&\langle e^{-\lambda W}\rangle_{p_0(x)} \\
&= e^{-\beta\Delta F}\int_{-\infty}^{+\infty}dx\,\bar{\rho}_\tau(\beta-\lambda,x)\frac{p_0(x)}{p_0^{\rm eq}(x)} \\
&= \sqrt{\frac{\beta' k_0[\bar{\psi}_{\beta-\lambda}(\tau)]^3}{2\pi\beta k_\tau\bar{\phi}_{\beta-\lambda}(\tau)}}\int_{-\infty}^{+\infty}dx\,e^{-\frac{x^2}{2}[\frac{\bar{\psi}_{\beta-\lambda}(\tau)}{\bar{\phi}_{\beta-\lambda}(\tau)}+k_0(\beta'-\beta)]} \\
&= \sqrt{\frac{k_0}{k_\tau}}\left[\frac{\beta}{\beta'}+\left(1-\frac{\beta}{\beta'}\right)\bar{g}_{\beta-\lambda}(\tau)\right]^{-\frac{1}{2}}\bar{\psi}_{\beta-\lambda}(\tau), \quad \text{(E7)}
\end{aligned}$$

where all the quantities with an overline must be associated with the time-reversed protocol $\bar{k}_t \equiv k_{\tau-t}$. After the first-order approximation, Eq. (E7) becomes

$$\begin{aligned}
&\ln\langle e^{-\lambda W}\rangle_{p_0(x)} \\
&= \frac{1}{2}\ln\frac{k_0}{k_\tau} - \frac{1}{2}\left(1-\frac{\beta}{\beta'}\right)\bar{g}_{\beta-\lambda}^{(1)}(\tau) + \ln\bar{\psi}_{\beta-\lambda}^{(1)}(\tau) + O(\kappa^2) \\
&= \ln\psi_\lambda^{(1)}(\tau) - \frac{1}{2}\left(1-\frac{\beta}{\beta'}\right)\bar{g}_{\beta-\lambda}^{(1)}(\tau) + O(\kappa^2), \quad \text{(E8)}
\end{aligned}$$

where $\bar{g}_{\beta-\lambda}^{(1)}(\tau) = -\gamma\kappa\lambda/(2k_0^2\beta)$, and the fluctuation-dissipation relation $\ln\psi_\lambda^{(1)}(\tau) = -\beta\Delta F + \ln\bar{\psi}_{\beta-\lambda}^{(1)}(\tau)$ is used. Since the generating function in the form $\ln\langle e^{-\lambda W}\rangle = -\lambda\langle W\rangle + \lambda^2\sigma_W^2/2$ must correspond to a Gaussian distribution centered at $\langle W\rangle$ and with variance σ_W^2, we finally obtain the results (47) in the main text.

APPENDIX F: DERIVATION OF THE SAGAWA-UEDA EQUALITIES (53) AND (54) FROM THE UNIFIED IFT

We consider the extended phase space as the direct (Cartesian) product of the phase spaces of the system and the measurement device. The phase-space point in such extended space can be denoted by $\Sigma_t \equiv (\Gamma_t, y)$, with Γ_t and y to be the components of the system and the device respectively [44]. The unified IFT of the composite system reads

$$\left\langle\frac{p_\tau^a(\Sigma_\tau)}{p_0(\Sigma_0)}e^{-\Delta s_m}\right\rangle = 1. \quad \text{(F1)}$$

This is the key relation that we will use.

To derive Eq. (53), we choose

$$\begin{aligned}
p_0(\Sigma_0) &= p(y|\Gamma_0)\frac{e^{-\beta U_0(\Gamma_0)}}{Z_0(\beta)}, \\
p_\tau^a(\Sigma_\tau) &= p(y)\frac{e^{-\beta U_\tau^y(\Gamma_\tau)}}{Z_\tau^y(\beta)}.
\end{aligned} \quad \text{(F2)}$$

By substituting Eq. (F2) into Eq. (F1), we obtain

$$\left\langle\frac{Z_0(\beta)}{Z_\tau^y(\beta)}e^{-\beta[Q+U_\tau^y(\Gamma_\tau)-U_0(\Gamma_0)]-\ln[p(y|\Gamma_0)/p(y)]}\right\rangle = 1. \quad \text{(F3)}$$

Here the DB condition, and thus $\Delta s_m = \beta Q$, has been assumed. To simplify Eq. (F3), we make use of the first law $W[\Gamma_t] = Q[\Gamma_t] + U_\tau(\Gamma_\tau) - U_0(\Gamma_0)$ as well as the expression of the free-energy difference $\Delta F^y \equiv -\beta^{-1}\ln[Z_\tau^y(\beta)/Z_0(\beta)]$, and further define the state function of the mutual information $I(\Sigma_0) \equiv \ln[p(y|\Gamma_0)/p(y)]$. Combining these relations we obtain

$$\langle e^{-\beta(W-\Delta F)-I}\rangle = 1. \quad \text{(F4)}$$

To derive Eq. (54), we choose $p_\tau^a(\Sigma_\tau)$ to be the real final distribution $p_\tau(\Sigma_\tau)$. So we get

$$\langle e^{-\Delta s_{\rm tot}}\rangle = 1, \quad \text{(F5)}$$

which is nothing but the EPI for the composite system. Concretely, we have

$$\Delta s_{\rm tot} = \beta Q - \ln p_\tau(\Sigma_\tau) + \ln p_0(\Sigma_0), \quad \text{(F6)}$$

as long as the DB condition holds. To distinguish two contributions to $\Delta s_{\rm tot}$: (i) the correlation between the system and the device and (ii) the entropy production of the system and in the medium, we define the mutual information at the initial and the final stages,

$$I_i \equiv \ln\frac{p_0(\Sigma_0)}{p_0(\Gamma_0)p(y)}, I_f \equiv \ln\frac{p_\tau(\Sigma_\tau)}{p_\tau(\Gamma_\tau)p(y)}. \quad \text{(F7)}$$

Then we can obtain another decomposition identity of $\Delta s_{\rm tot}$ as follows:

$$\Delta s_{\rm tot} = \beta Q - \ln p_\tau(\Gamma_\tau) + \ln p_0(\Gamma_0) - I_f + I_i = \sigma - \Delta I, \quad \text{(F8)}$$

which leads to

$$\langle e^{-\sigma+\Delta I}\rangle = 1. \quad \text{(F9)}$$

It should be mentioned that $p_0(\Sigma)$ and $p_\tau^a(\Sigma)$ are tacitly assumed to be nonzero for any Σ to guarantee the validity of the two Sagawa-Ueda equalities in this Appendix. However, this assumption is usually not satisfied for error-free measurements or a rigorously localized initial state, thus the two Sagawa-Ueda equalities may break down in these cases [11]. Recent researches have shown that by adding an extra modification term on the exponential, the Sagawa-Ueda equalities can be generalized to be applicable to the feedback control processes with error-free measurements [45].

[1] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[2] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).
[3] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[4] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[5] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[6] K. Kim, C. Kwon, and H. Park, Phys. Rev. E **90**, 032117 (2014).
[7] U. Seifert, Eur. Phys. J. B **64**, 423 (2008).
[8] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[9] R. García-García, D. Domínguez, V. Lecomte, and A. B. Kolton, Phys. Rev. E **82**, 030104(R) (2010).
[10] R. García-García, V. Lecomte, A. B. Kolton, and D. Domínguez, J. Stat. Mech. (2012) P02009.
[11] Y. Murashita, K. Funo, and M. Ueda, Phys. Rev. E **90**, 042110 (2014).
[12] G. Hummer and A. Szabo, Proc. R. Acad. Sci. **98**, 3658 (2001).
[13] M. Polettini and M. Esposito, J. Stat. Mech. (2014) P10033.
[14] S. Lahiri and A. M. Jayannavar, Indian J. Phys. **89**, 515 (2014).
[15] To see why we choose such a definition instead of $\bar{a}_t(\Gamma) \equiv a_{\tau-t}(\Gamma)$, we simply consider an example that $a_t(\Gamma) = p_t(\Gamma)$ is the distribution function of the phase-space point. If we want to obtain $\bar{p}_0(\Gamma) = p_\tau(\Gamma^\dagger)$, the standard preparation of the initial distribution for the time-reversed process, we must follow the definition in the main text. Such a definition may also be heuristic for the potential generalization to the cases with nonconservative force, such as Lorentz force, which will break the symmetry $U_t(\Gamma) = U_t(\Gamma^\dagger)$.
[16] It is necessary to emphasize that the accessible phase space $\mathfrak{S}_t$ may change with time t, e.g., a molecule confined in an expanding piston. However, at any given time t, the accessible phase space of Γ_t is always assumed to be the same as that of $\Gamma_t^\dagger$, namely $\Gamma_t \in \mathfrak{S}_t$ if and only if $\Gamma_t^\dagger \in \mathfrak{S}_t$.
[17] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer, Berlin, 1989).
[18] R. J. Harris and G. M. Schütz, J. Stat. Mech. (2007) P07020.
[19] M. Silaev, T. T. Heikkilä, and P. Virtanen, Phys. Rev. E **90**, 022103 (2014).
[20] F. Liu, Phys. Rev. E **89**, 042122 (2014).
[21] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[22] Remember that we have assumed that $\Gamma \in \mathfrak{S}$ if and only if $\Gamma^\dagger \in \mathfrak{S}$. Since the time-reversal operation of the phase-space point does not change the size of volume element, we have $\int_{\mathfrak{S}} d\Gamma f(\Gamma) = \int_{\mathfrak{S}} d\Gamma f(\Gamma^\dagger)$ for any normalizable function $f(\Gamma)$ defined in the phase space $\mathfrak{S}$. In particular, $Z_\tau(\beta) = \bar{Z}_0(\beta)$, where $\mathfrak{S} = \mathfrak{S}_\tau$, $f(\Gamma) = e^{-\beta U_\tau(\Gamma)}$ and $f(\Gamma^\dagger) = e^{-\beta U_\tau(\Gamma^\dagger)} = e^{-\beta \bar{U}_0(\Gamma)}$. This is true even if $U_t(\Gamma) \neq U_t(\Gamma^\dagger)$, though we assume $U_t(\Gamma) = U_t(\Gamma^\dagger)$ througout the paper.
[23] C. Jarzynski, J. Stat. Phys. **98**, 77 (2000).
[24] P. Maragakis, M. Spichty, and M. Karplus, J. Phys. Chem. B **112**, 6168 (2008).
[25] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).
[26] F. Liu, Phys. Rev. E **90**, 032121 (2014).
[27] I. Junier, A. Mossa, M. Manosas, and F. Ritort, Phys. Rev. Lett. **102**, 070602 (2009).
[28] A. Alemany, M. Ribezzi, and F. Ritort, in *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*, edited by R. Klages, W. Just, and C. Jarzynski (Wiley-VCH, Weinheim, 2012).
[29] S. Schuler, T. Speck, C. Tietz, J. Wrachtrup, and U. Seifert, Phys. Rev. Lett. **94**, 180602 (2005).
[30] C. Van den Broeck and M. Esposito, Physica A **418**, 6 (2015). For example, consider a free thermalization process from a nonequilibrium initial distribution $p_0(\Gamma)$ without any external driving. The time-reversed process is simply to continue the thermal relaxation after the instantaneous time reversal at time τ. For the overdamped Langevin dynamics, the momentum degree of freedom is ignored, thus $\Gamma^\dagger = \Gamma$, implying that nothing should be done for the time-reversal at time τ, i.e., $\bar{p}_0(\Gamma) = p_\tau(\Gamma)$. If the final phase-space point distribution for the backward process $\bar{p}_\tau(\Gamma)$ is $p_0(\Gamma)$, we will get a ridiculous result where the effect of the thermalization after 2τ time turns out to be returning to the nonequilibrium initial distribution.
[31] T. Becker, T. Willaert, B. Cleuren, and C. Van den Broeck, Phys. Rev. E **91**, 012101 (2015).
[32] M. Esposito, C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
[33] T. Speck, J. Phys. A: Math. Theo. **44**, 305001 (2011).
[34] T. Speck and U. Seifert, Phys. Rev. E **70**, 066112 (2004).
[35] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010). Rigorously speaking, what we derive here is a special case that the measurement is carried out initially. The original Sagawa-Ueda equalities are applicable to an arbitrary measurement time $t_m \in [0,\tau]$.
[36] T. Sagawa and M. Ueda, Phys. Rev. Lett. **109**, 180602 (2012).
[37] L. Szilard, Z. Phys. **53**, 840 (1929).
[38] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[39] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).
[40] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
[41] K. Sekimoto, *Stochastic Energetics* (Springer, Berlin, 2010).
[42] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102 (2007).
[43] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).
[44] According to Ref. [36], if we exchange the physical interpretation of Γ_t and y, namely Γ_t referring to the measurement device while y referring to the system, we can talk about the measurement processes.
[45] Y. Ashida, K. Funo, Y. Murashita, and M. Ueda, Phys. Rev. E **90**, 052125 (2014).