Quantum Jarzynski equality of measurement-based work extraction

Yohei Morikuni[*]

Department of Physics, The University of Tokyo, Komaba, Meguro, Tokyo 153-8505, Japan

Hiroyasu Tajima[†]

Department of Physics, The University of Tokyo, Komaba, Meguro, Tokyo 153-8505, Japan
and Center for Emergent Matter Science (CEMS), RIKEN, Wako, Saitama 351-0198, Japan

Naomichi Hatano[‡]

Institute of Industrial Science, The University of Tokyo, Komaba, Meguro, Tokyo 153-8505, Japan
(Received 24 October 2016; published 31 March 2017)

Many studies of quantum-size heat engines assume that the dynamics of an internal system is unitary and that the extracted work is equal to the energy loss of the internal system. Both assumptions, however, should be under scrutiny. In the present paper, we analyze quantum-scale heat engines, employing the measurement-based formulation of the work extraction recently introduced by Hayashi and Tajima [M. Hayashi and H. Tajima, arXiv:1504.06150]. We first demonstrate the inappropriateness of the unitary time evolution of the internal system (namely, the first assumption above) using a simple two-level system; we show that the variance of the energy transferred to an external system diverges when the dynamics of the internal system is approximated to a unitary time evolution. Second, we derive the quantum Jarzynski equality based on the formulation of Hayashi and Tajima as a relation for the work measured by an external macroscopic apparatus. The right-hand side of the equality reduces to unity for "natural" cyclic processes but fluctuates wildly for noncyclic ones, exceeding unity often. This fluctuation should be detectable in experiments and provide evidence for the present formulation.

I. INTRODUCTION

Thermodynamics was first introduced as a practical study to clarify the optimal performance of heat engines [1] and has become one of the most important fields in physics [2]. Recently, however, the development of experimental technology is realizing heat engines which are out of the scope of the standard thermodynamics, i.e., small-size heat engines. Quantum-scale and mesoscopic thermomotors, which used to be imaginary devices, are being realized in the laboratory [3–7]. The functions of biomolecules, which are micromachines in nature, are being clarified too [8].

We cannot apply the standard thermodynamics to these nanometer-size heat engines as it is, because it is a phenomenology for macroscopic systems. Statistical mechanics, another fundamental field of physics, has been applied to the small-size heat engines whose small-size working body is connected to the infinitely large heat baths, and is achieving a splendid success [9–33].

Many studies of such engines [17–33] adopt a model of a microscopic internal quantum system connected to a macroscopic external agent. They use the following setup and assumption:

(i) Thermodynamic operation of the internal quantum system by an external system is represented by a unitary operator of the time-dependent Hamiltonian of the internal quantum system which is controlled by external parameters.

(ii) The work performed on the external system is equal to the energy loss of the internal quantum system.

This approach has a practical advantage in that we can formulate thermodynamic relations by analyzing only the internal quantum system.

However, there are always two concerns about the validity of this approach. One is about the assumption of the unitary dynamics and the other is about the definition of the work. In actual situations of heat engines, the internal system, whether quantum or not, is always attached to an external macroscopic system [34], and hence the dynamics of the internal quantum system cannot be exactly unitary. It is not guaranteed either that all energy loss of the internal system becomes the work done to the external system. Is it accurate enough to approximate the true dynamics with a unitary dynamics? Is it legitimate to regard the energy loss as the extracted work? In spite of these concerns, the approach has been accepted by many researchers, because there have been results out of this approach which seem to be consistent with thermodynamics. For example, we can derive some thermodynamic relations [19,26–28] and a kind of quantum extension of the Jarzynski equality [17,18,24,25].

In the present paper, we face these two concerns. We claim here that the work out of a heat engine should be measured by means of a macroscopic system, for example, as movement of a macroscopic piston or wheel. From this point of view, the work extraction from the internal quantum system by the macroscopic system can be described as a standard quantum measurement process, in which the measurement result is evaluated on the side of the macroscopic system. In other words, the time evolution of the internal system is described as a quantum measurement process and the extracted work is regarded as a measurement outcome of this process.

―――――――
[*]morikuni@iis.u-tokyo.ac.jp
[†]hiroyasu.tajima@riken.jp
[‡]hatano@iis.u-tokyo.ac.jp

This formulation, recently introduced by Hayashi and Tajima [35–37], raises a serious problem of the conventional approach in the form of a general trade-off relation. The relation shows that when the time evolution of the internal system is approximated to a unitary, it becomes difficult to fix the amount of extracted work.

In the present paper, we first demonstrate the inappropriateness of the unitary time evolution of the internal system [namely, assumption (i) above]. It was pointed out by Hayashi and Tajima that the amount of the extracted work is not able to be evaluated on the side of the external system when the time evolution of the internal system is approximated to a unitary one [35]. It was also illustrated with a plain example by Tasaki, who is one of the advocates of the conventional approach [38]. In order to demonstrate this fact, we show for a simple model that the conventional approach has a problem regarding the fluctuation of the work. More specifically, we show that the variance of the energy transferred to the external system diverges when the time evolution of the internal system is approximated to a unitary, and it is completely different from that of the energy loss of the internal system. This result demands us to change the derivation of the quantum Jarzynski equality from the previous ones [17,18,24,25] based on the conventional approach, which employed the *internal* unitary time evolution and regarded the measured value of the energy loss of the internal system as the "work," whereas the energy gain of the *external* system is the actual work that we can use. The Jarzynski equality derived from the previous approach does not contain relevant information about the fluctuation of the actual work.

In order to resolve this problem, we next derive the quantum Jarzynski equality based on the measurement-based work extraction of Hayashi and Tajima [35,36]. This quantum Jarzynski equality is the relation of the work measured by the external macroscopic apparatus, which is equal to the measured value of the energy gain of the external system. Therefore, our derivation correctly contains the information about the fluctuation of the actual work, namely, the energy gain of the external system.

Our derivation of the quantum Jarzynski equality is essentially different from other derivations [17,18,24,25,39–42] which used the conventional two-measurement formulation. Note that the formulation of the measurement-based work extraction describes the operation of the work extraction as a measurement process itself. It is based on the fact that the standard thermodynamic operations give the amount of the extracted work and the change of the state of the internal system at the same time. It pointed out that we do not need to separate the dynamics of extracting the work from the measurement for evaluating the work, as has been done in the two-measurement formulation.

There has been another approach to the two concerns, in which they [43,44] used an internal quantum system connected to an external quantum system. This approach assumes that the time evolution of the total system is unitary and the work extracted from the internal system is defined as the energy gain of the external system. We claim that the setup in our approach is more realistic than in this approach [43,44] in the sense that our external system is a macroscopic measurement apparatus.

II. INAPPROPRIATENESS OF THE UNITARY TIME EVOLUTION

In general, the time evolution of an internal system interacting with an external system is not unitary. However, many studies of statistical physics assume that the time evolution of the internal system can be approximated to a unitary one generated by a time-dependent Hamiltonian of the internal system.

In this section, using a toy model, we consider the problem as to whether we can represent the time evolution of an internal system in terms of a unitary one. More specifically, we show that the variance of the energy transferred to an external system E becomes large when we approximate the time evolution of the two-level system I by means of a unitary one.

Let H_I and H_E denote the Hamiltonians of the internal system I and the external system E, respectively:

$$H_\mathrm{I} := \frac{\Delta}{2}(|1\rangle_\mathrm{I}\langle 1| - |0\rangle_\mathrm{I}\langle 0|), \tag{1}$$

$$H_\mathrm{E} := \sum_{n=-\infty}^{\infty} n\Delta|n\rangle_\mathrm{E}\langle n|, \tag{2}$$

where Δ is the level spacing, while $|n\rangle_\mathrm{E}$ with any integer n are the energy eigenstates of E with the eigenvalue $n\Delta$. The fact that n runs from negative infinity to positive infinity represents that the external system is macroscopic.

We consider the case in which the time evolution operator U_IE of the total system is unitary and conserves the energy as in $[H_\mathrm{I} + H_\mathrm{E}, U_\mathrm{IE}] = 0$. We can therefore decompose U_IE into the form

$$U_\mathrm{IE} := \sum_{n,n'} K_{n,n'} \otimes |n\rangle_\mathrm{E}\langle n'| \tag{3}$$

with

$$\begin{aligned} K_{n,n'} := {} & \delta_{n,n'}(a_n|0\rangle_\mathrm{I}\langle 0| + b_n|1\rangle_\mathrm{I}\langle 1|) \\ & + \delta_{n+1,n'}c_n|1\rangle_\mathrm{I}\langle 0| + \delta_{n-1,n'}d_n|0\rangle_\mathrm{I}\langle 1| \end{aligned} \tag{4}$$

for all integers n and n', where the coefficients a_n, b_n, c_n, and d_n are complex numbers and satisfy

$$|a_n|^2 + |c_{n-1}|^2 = 1, \tag{5}$$

$$|b_n|^2 + |d_{n+1}|^2 = 1, \tag{6}$$

$$a_n d_n^* + b_{n-1}^* c_{n-1} = 0, \tag{7}$$

for all integers n. The time evolution of the internal system I is hence given by

$$\mathcal{K}(\rho_\mathrm{I}) := \mathrm{Tr}_\mathrm{E}[U_\mathrm{IE}(\rho_\mathrm{I} \otimes \rho_\mathrm{E})U_\mathrm{IE}^\dagger] \tag{8}$$

$$= \sum_{n,n',n''} {}_\mathrm{E}\langle n' | \rho_\mathrm{E} | n''\rangle_\mathrm{E} K_{n,n'} \rho_\mathrm{I} K_{n,n''}^\dagger, \tag{9}$$

where ρ_I and ρ_E are the initial states of I and E, respectively, and Tr_E denotes the trace operation with respect to the external system E. The state ρ_I of I must be written in the form

$$\rho_\mathrm{I} := p|0\rangle_\mathrm{I}\langle 0| + (1-p)|1\rangle_\mathrm{I}\langle 1| + q^*|1\rangle_\mathrm{I}\langle 0| + q|0\rangle_\mathrm{I}\langle 1| \tag{10}$$

with $0 \leqslant p \leqslant 1$, $q \in \mathbb{C}$. The positivity of ρ_I dictates the coefficient q to satisfy $|q|^2 \leqslant p(1-p) \leqslant 1/4$.

We now define the energy transferred to the external system E during the time evolution as

$$w_j := h_j - \mathrm{Tr}\,[\rho_{\mathrm{E}} H_{\mathrm{E}}], \tag{11}$$

where Tr denotes the trace operation with respect to the total system. We here use the expectation value of the energy of the external system because we have to use a mixed state for the initial state ρ_{E}.

The probability of w_j is given by

$$p_j := \mathrm{Tr}[|j\rangle_{\mathrm{E}}\langle j|U_{\mathrm{IE}}(\rho_{\mathrm{I}} \otimes \rho_{\mathrm{E}})U_{\mathrm{IE}}^{\dagger}] \tag{12}$$

$$= \sum_{n',n''}{}_{\mathrm{E}}\langle n' \,|\, \rho_{\mathrm{E}} \,|\, n''\rangle_{\mathrm{E}}\,\mathrm{Tr}[K_{j,n'}\rho_{\mathrm{I}}K_{j,n''}^{\dagger}], \tag{13}$$

where $h_j := j\Delta$ is an eigenvalue of the Hamiltonian H_{E} of the external system E. The energy conservation $[H_{\mathrm{I}} + H_{\mathrm{E}}, U_{\mathrm{IE}}] = 0$ leads to

$$\langle w \rangle_{\mathrm{E}} := \sum_{j=-\infty}^{\infty} w_j p_j \tag{14}$$

$$= \mathrm{Tr}[H_{\mathrm{E}}U_{\mathrm{IE}}(\rho_{\mathrm{I}} \otimes \rho_{\mathrm{E}})U_{\mathrm{IE}}^{\dagger}] - \mathrm{Tr}[\rho_{\mathrm{E}} H_{\mathrm{E}}] \tag{15}$$

$$= \mathrm{Tr}[(H_{\mathrm{I}} + H_{\mathrm{E}})(\rho_{\mathrm{I}} \otimes \rho_{\mathrm{E}})] - \mathrm{Tr}[H_{\mathrm{I}}U_{\mathrm{IE}}(\rho_{\mathrm{I}} \otimes \rho_{\mathrm{E}})U_{\mathrm{IE}}^{\dagger}]$$
$$- \mathrm{Tr}[\rho_{\mathrm{E}} H_{\mathrm{E}}] \tag{16}$$

$$= \mathrm{Tr}[\rho_{\mathrm{I}} H_{\mathrm{I}}] - \mathrm{Tr}[\mathcal{K}(\rho_{\mathrm{I}})H_{\mathrm{I}}]. \tag{17}$$

The variance $V(w)$ of the energy transfer is given by

$$V_{\mathrm{E}}(w) := \langle w^2 \rangle_{\mathrm{E}} - \langle w \rangle_{\mathrm{E}}^2 = \langle h^2 \rangle_{\mathrm{E}} - \langle h \rangle_{\mathrm{E}}^2, \tag{18}$$

where $\langle h \rangle_{\mathrm{E}}$ and $\langle h^2 \rangle_{\mathrm{E}}$ are the average and the mean square of h_j, respectively, with respect to p_j.

So far, everything has been exact. Now, we try to approximate the time evolution of I to a specific unitary matrix:

$$U_{\mathrm{I}} := \sum_{k,l=0,1} u_{k,l}|k\rangle_{\mathrm{I}}\langle l|. \tag{19}$$

Here, the approximation to the unitary matrix U_{I} means the following condition: for any $\epsilon > 0$, the inequality $d(\mathcal{K}(\rho_{\mathrm{I}}), U_{\mathrm{I}}\rho_{\mathrm{I}}U_{\mathrm{I}}^{\dagger}) < \epsilon$ holds for all states ρ_{I}, where d is a distance function.

Using the example given in Refs. [36,45], we now choose the initial state of the external system E and the coefficients of Eq. (4) in the form

$$\rho_{\mathrm{E}} := |\psi\rangle_{\mathrm{E}}\langle\psi|, \quad |\psi\rangle_{\mathrm{E}} := \frac{1}{\sqrt{M}}\sum_{m=1}^{M}|m\rangle_{\mathrm{E}}, \tag{20}$$

$$K_{n,n'} = \delta_{n,n'}(u_{0,0}|0\rangle_{\mathrm{I}}\langle 0| + u_{1,1}|1\rangle_{\mathrm{I}}\langle 1|)$$
$$+ \delta_{n+1,n'}u_{1,0}|1\rangle_{\mathrm{I}}\langle 0| + \delta_{n-1,n'}u_{0,1}|0\rangle_{\mathrm{I}}\langle 1|, \tag{21}$$

with $M := 8\lceil\epsilon^{-2}\rceil$ and all integrals n and n'. Then, we obtain [36]

$$b(\mathcal{K}(\rho_{\mathrm{I}}), U_{\mathrm{I}}\rho_{\mathrm{I}}U_{\mathrm{I}}^{\dagger}) < \epsilon, \tag{22}$$

where b is the Bures distance. Because the initial state is far from an energy eigenstate, we can indeed approximate the time evolution of I to U_{I}.

A problem arises as follows, however. Combining Eqs. (3) and (20), we obtain Eq. (13) in the form

$$p_j = \frac{1}{M}\mathrm{Tr}[\bar{K}_j^{\dagger}\bar{K}_j\rho_{\mathrm{I}}] \tag{23}$$

with

$$\bar{K}_j := \sum_{m=1}^{M}K_{j,m}. \tag{24}$$

The expectation and the mean square of h_j are respectively given by

$$\langle h \rangle_{\mathrm{E}} = \sum_j h_j p_j = \frac{\Delta}{M}\sum_j j\,\mathrm{Tr}[\bar{K}_j^{\dagger}\bar{K}_j\rho_{\mathrm{I}}], \tag{25}$$

$$\langle h^2 \rangle_{\mathrm{E}} = \sum_j (h_j)^2 p_j = \frac{\Delta^2}{M}\sum_j j^2\,\mathrm{Tr}[\bar{K}_j^{\dagger}\bar{K}_j\rho_{\mathrm{I}}]. \tag{26}$$

After the algebra in Appendix A, we obtain

$$\langle h \rangle_{\mathrm{E}} = \frac{\Delta}{2}(M + 1) + \langle w \rangle_{\mathrm{E}}, \tag{27}$$

$$\langle h^2 \rangle_{\mathrm{E}} = \frac{\Delta^2}{6}(M + 1)(2M + 1) + \Delta(M + 1)\langle w \rangle_{\mathrm{E}}$$
$$+ \Delta^2(|u_{0,1}|^2(1 - p) + |u_{1,0}|^2 p), \tag{28}$$

where $\langle w \rangle_{\mathrm{E}}$ is the expectation of w_j in the form

$$\langle w \rangle_{\mathrm{E}} = \Delta[|u_{0,1}|^2(1 - p) - |u_{1,0}|^2 p$$
$$+ 2(1 - M^{-1})\mathrm{Re}(u_{0,0}u_{0,1}^* q)]. \tag{29}$$

Therefore, we obtain

$$V_{\mathrm{E}}(w) = \frac{\Delta^2}{12}(M^2 - 1) - \langle w \rangle_{\mathrm{E}}^2$$
$$+ \Delta^2(|u_{0,1}|^2(1 - p) + |u_{1,0}|^2 p). \tag{30}$$

Because $\langle w \rangle_{\mathrm{E}} = O(1)$, the variance of the energy transfer diverges as $M = 8\lceil\epsilon^{-2}\rceil \to \infty$ in the limit $\epsilon \to 0$. In other words, we cannot fix the amount of the energy transfer when we approximate the dynamics of I by a unitary one. This demonstrates that it is not appropriate to use the unitary dynamics for the internal system I and define the work as the energy loss of I.

Let us compare this with the variance of the energy loss of the internal system I. We measure the energy of the internal system I before and after the time evolution (9) and regard the difference of the two measurement outcomes as the energy loss of the internal system, which the previous derivation of the quantum Jarzynski equalities [17,18,24,25] defined as the work. Because the time evolution of the internal system I is given by Eq. (9), the probability of the energy loss of the internal system

$$\upsilon_{\alpha,\beta} := e_{\alpha} - e_{\beta} \tag{31}$$

is given by

$$q_{\alpha,\beta} := {}_{\mathrm{I}}\langle\alpha \,|\, \rho_{\mathrm{I}} \,|\, \alpha\rangle_{\mathrm{I}}\,\mathrm{Tr}[|\beta\rangle_{\mathrm{I}}\langle\beta|\mathcal{K}(|\alpha\rangle_{\mathrm{I}}\langle\alpha|)], \tag{32}$$

where $e_0 := -\Delta/2$ and $e_1 := \Delta/2$ are the eigenvalues of the Hamiltonian H_{I}. The variance $V_{\mathrm{I}}(\upsilon)$ of the energy loss of the

internal system is hence given by

$$V_I(\upsilon) := \langle \upsilon^2 \rangle_I - \langle \upsilon \rangle_I^2, \tag{33}$$

where $\langle \upsilon \rangle_I$ and $\langle \upsilon^2 \rangle_I$ are the average and the mean square of $\upsilon_{\alpha,\beta}$, respectively, with respect to $q_{\alpha,\beta}$. Since $e_0 := -\Delta/2$ and $e_1 := \Delta/2$, the energy loss $\upsilon_{\alpha,\beta}$ is $-\Delta$, zero, or Δ. Therefore, we obtain

$$\langle \upsilon^2 \rangle_I := \sum_{\alpha,\beta=0,1} q_{\alpha,\beta} \upsilon_{\alpha,\beta}^2 \leqslant \Delta^2, \tag{34}$$

and hence

$$V_I(\upsilon) := \langle \upsilon^2 \rangle_I - \langle \upsilon \rangle_I^2 \leqslant \Delta^2. \tag{35}$$

We thus see that $V_I(\upsilon)$ appears to be completely different from $V_E(w)$ in Eq. (30) when we approximate the time evolution of the internal system to a unitary one.

The above demonstration raises a problem of the quantum Jarzynski equalities derived in the previous approach [17,18,24,25], which employed a unitary for the time evolution of the internal system and regarded the energy loss of the internal system as the work. Because the energy gain of the external system is the actual work that we can use, the work defined in the previous approach as well as the Jarzynski equalities derived thereby do not contain relevant information about the fluctuation of the actual work. In order to resolve this problem, we derive in Sec. III the quantum Jarzynski equality using the measurement-based work extraction of Hayashi and Tajima [35,36].

Incidentally, in the case of $M = 1$, we cannot approximate the time evolution to a unitary one. To show it, we consider the quantity

$$\min_{U_I:\text{unitary}} d_{\text{Tr}}(\mathcal{K}(\rho_I), U_I \rho_I U_I^\dagger), \tag{36}$$

where d_{Tr} is the trace distance. For $M = 1$, the initial state (20) of E is a pure energy eigenstate with a fixed energy level n_0, namely, $\rho_E = |n_0\rangle_E\langle n_0|$. Then, Eq. (9) reduces to

$$\mathcal{K}(\rho_I) = \sum_n K_{n,n_0} \rho_I K_{n,n_0}^\dagger$$
$$= [|a_{n_0}|^2 p + |d_{n_0+1}|^2(1-p)]|0\rangle_I\langle 0| + [|b_{n_0}|^2(1-p)$$
$$+ |c_{n_0-1}|^2 p]|1\rangle_I\langle 1| + a_{n_0}^* b_{n_0} q^*|1\rangle_I\langle 0|$$
$$+ a_{n_0} b_{n_0}^* q|0\rangle_I\langle 1|. \tag{37}$$

From Eq. (17), we have the average of the transferred energy in the form

$$\langle w \rangle_E = \Delta[|d_{n_0+1}|^2(1-p) - |c_{n_0-1}|^2 p]. \tag{38}$$

Since the states $\mathcal{K}(\rho_I)$ and ρ_I are Hermitian operators, we can diagonalize them using unitary operators. In other words, there exist unitary operators V and V' such that

$$\rho_I = V \Lambda V^\dagger, \tag{39}$$

$$\mathcal{K}(\rho_I) = V' \Lambda' V'^\dagger, \tag{40}$$

where Λ and Λ' are diagonal matrices. Because of the unitary invariance of the trace distance, Eq. (36) becomes

$$\min_{U_I:\text{unitary}} d_{\text{Tr}}(\mathcal{K}(\rho_I), U_I \rho_I U_I^\dagger) = \min_{\tilde{U}_I:\text{unitary}} d_{\text{Tr}}(\tilde{U}_I \Lambda' \tilde{U}_I^\dagger, \Lambda). \tag{41}$$
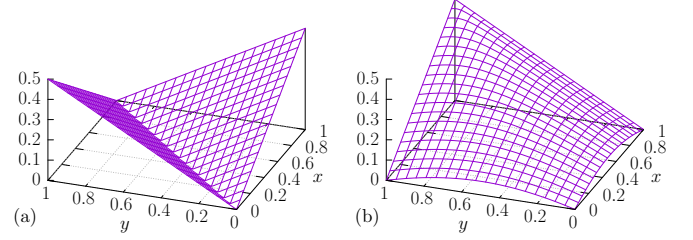

FIG. 1. Plots of Eq. (43). We take the parameters (a) $(p, |q|) = (1/2, 0)$ and (b) $(1/2, 1/2)$.

The calculation in Appendix B then gives

$$\min_{\tilde{U}_I:\text{unitary}} d_{\text{Tr}}(\tilde{U}_I \Lambda' \tilde{U}_I^\dagger, \Lambda) = \frac{1}{2} \min_{\tilde{U}_I:\text{unitary}} \text{Tr}\,|\tilde{U}_I \Lambda' \tilde{U}_I^\dagger - \Lambda| \tag{42}$$

$$= |R' - R| \tag{43}$$

with

$$R := \sqrt{\left(p - \frac{1}{2}\right)^2 + |q|^2}, \tag{44}$$

$$R' := \sqrt{\left((1-x-y)p + x - \frac{1}{2}\right)^2 + (1-x)(1-y)|q|^2}, \tag{45}$$

$$x := |d_{n_0+1}|^2, \quad y := |c_{n_0-1}|^2. \tag{46}$$

In the case of $x = y = 0$, the distance of Eq. (43) is equal to zero. However, the energy transfer (38) is trivially equal to zero in this case. For that reason, we consider cases other than $x = y = 0$. When we choose the initial state of I as $p = 1/2$ and $|q| = 0$, the distance of Eq. (43) is not equal to zero [Fig. 1(a)]. When we choose the initial state of I as $p = 1/2$ and $|q| = 1/2$, the distance of Eq. (43) is also not equal to zero for $x = y \neq 0$ [Fig. 1(b)]. Consequently, we cannot approximate the time evolution to a unitary when $M = 1$.

From the results above, we suspect that the variance of the energy transfer generally diverges when the time evolution of the internal system is approximated to a unitary one. This is because the divergence is caused by the superposition of the energy eigenstates of the initial state of the external system and because the time evolution of the internal system cannot be approximated to a unitary one when the initial state of the external system is an energy eigenstate.

III. JARZYNSKI EQUALITY

As the main result of the present paper, we here derive the Jarzynski equality for the measurement-based work extraction. In the previous section, we showed that the approach using the unitary time evolution of the internal system is inappropriate for describing the quantum heat engine. Previous quantum versions of the Jarzynski equality [17,18,24,25] also have this problem. To resolve this problem, we introduce a new derivation of the quantum Jarzynski equality based on the formulation of Hayashi and Tajima [35,36].

Note that this formulation, called measurement-based work extraction, points out that operations of the work extraction are measurement processes. It is essentially different from the

conventional two-measurement formulation [17,18,24,25,39–41] which separates the dynamics of the extracting work from the measurement of one.

Also note that the fluctuations in the previous and present sections are generated from different causes. While the fluctuation in the previous section is generated from the uncertainty of the initial energy of the external system, one in the present section is generated from a mixture of the state of the internal system.

A. Cyclic process

We first consider a cyclic process, in which the final Hamiltonian is equal to the initial one. The external system E receives energy from the internal system I from time $t = 0$ to T. The time evolution of the total system between time $t = 0$ and T is unitary.

Let H_I and H_E denote the time-independent Hamiltonians of I and E, respectively, and $H_{int}(t)$ denote the time-dependent Hamiltonian interacting between I and E; the total system is given by

$$H_{tot}(t) := H_I + H_E + H_{int}(t). \qquad (47)$$

The case in which the Hamiltonian of the internal system depends on time is treated by the discussion in Sec. III B.

The eigenvalue decompositions of H_I and H_E are denoted by

$$H_I := \sum_x h_x |h_x\rangle_I \langle h_x|, \qquad (48)$$

$$H_E := \sum_i e_i |e_i\rangle_E \langle e_i|, \qquad (49)$$

where h_x and e_i are the eigenvalues of H_I and H_E, respectively. We assume that the interaction Hamiltonian satisfies the condition

$$[H_I + H_E, U_{IE}] = 0, \qquad (50)$$

where U_{IE} is the time evolution of total system given by [35]

$$U_{IE} := \mathcal{T} \exp\left(-\frac{i}{\hbar} \int_0^T H_{tot}(\tau)\, d\tau\right), \qquad (51)$$

where $\mathcal{T}$ is a time-ordered product. The condition of Eq. (50) means that the total energy from $H_I + H_E$ is the same before and after U_{IE}, and hence the net energy from the interaction Hamiltonian is equal to zero. We further assume that the initial states of I and E are the canonical distribution at an inverse temperature β and a pure eigenstate of energy e_0, respectively:

$$\rho_{I,can} := \sum_x \frac{e^{-\beta h_x}}{Z_I} |h_x\rangle_I \langle h_x|, \qquad (52)$$

$$\rho_E := |e_0\rangle_E \langle e_0|, \qquad (53)$$

with $Z_I := \mathrm{Tr}[e^{-\beta H_I}]$.

We then consider the following process:

(i) We set the initial states given by Eqs. (52) and (53).

(ii) We then let the total system evolve under the unitary operator U_{IE}. The key here is to consider the unitary time evolution of the total system, not of the internal system I.

(iii) We finally measure the energy of the external system E using the projection operator $|e_j\rangle_E \langle e_j|$ and define

$$\Delta e_j := e_j - e_0 \qquad (54)$$

as the energy gain. It is essential that at this point the "work" Δe_j is not a fixed value but given probabilistically.

We stress that this process is an indirect measurement with respect to the internal system I and modeling the measurement-based work extraction.

In the above process, the time evolution of I and the measurement process of the specific energy gain Δe_j are defined as

$$\mathcal{E}(\rho_I) := \mathrm{Tr}_E[U_{IE}(\rho_I \otimes |e_0\rangle_E \langle e_0|)U_{IE}^\dagger] \qquad (55)$$

with $\mathcal{E} = \sum_j \mathcal{E}_j$ for all density operators ρ_I of I, where

$$\mathcal{E}_j(\rho_I) := \mathrm{Tr}_E[|e_j\rangle_E \langle e_j| U_{IE}(\rho_I \otimes |e_0\rangle_E \langle e_0|)U_{IE}^\dagger] \qquad (56)$$

and Tr_E denotes the trace operation with respect to E.

We now introduce the probability distribution of the extracted work W as

$$P(W) := \sum_j \delta(W - \Delta e_j)\, \mathrm{Tr}[\mathcal{E}_j(\rho_{I,can})], \qquad (57)$$

where $\delta(x)$ is the delta function and Tr denotes the trace operation with respect to the total system. Because $\mathcal{E}_j$ is a linear operator, we have

$$P(W) = \sum_{j,x,y} \delta(W - \Delta e_j)$$
$$\times \frac{e^{-\beta h_x}}{Z_I} \mathrm{Tr}[|h_y\rangle_I \langle h_y| \mathcal{E}_j(|h_x\rangle_I \langle h_x|)], \qquad (58)$$

where we inserted a resolution of unity $\sum_y |h_y\rangle_I \langle h_y|$. Since the total energy $H_I + H_E$ does not change after U_{IE} and the external system gains the energy Δe_j after the process $\mathcal{E}_j$, the energy of the internal system must be $h_y = h_x - \Delta e_j$ outside the operator $\mathcal{E}_j$. We therefore find

$$P(W) = e^{-\beta W} \sum_{j,x,y} \delta(W - \Delta e_j) \frac{e^{-\beta h_y}}{Z_I}$$
$$\times \mathrm{Tr}[|h_y\rangle_I \langle h_y| \mathcal{E}_j(|h_x\rangle_I \langle h_x|)] \qquad (59)$$
$$= e^{-\beta W} \sum_j \delta(W - \Delta e_j) \mathrm{Tr}[\rho_{I,can} \mathcal{E}_j(\mathbb{1}_I)]. \qquad (60)$$

Let us denote the average with respect to $P(W)$ by

$$\langle f(W) \rangle := \int dW f(W) P(W), \qquad (61)$$

where $f(W)$ is an arbitrary function of the extracted work W. From Eq. (60), we can therefore obtain the Jarzynski equality under the cyclic process in the form

$$\langle e^{\beta W} \rangle = \gamma \qquad (62)$$

with

$$\gamma := \mathrm{Tr}[\mathcal{E}^\dagger(\rho_{I,can})], \qquad (63)$$

where $\mathcal{E}^\dagger$ is the adjoint map of $\mathcal{E}$, given by $\mathrm{Tr}[\mathcal{E}^\dagger(A^\dagger)B] = \mathrm{Tr}[A^\dagger \mathcal{E}(B)]$. Note that we did not use the details of the external

system E but the energy conservation of the time evolution, Eq. (50).

Applying Jensen's inequality $\langle e^f \rangle \geqslant e^{\langle f \rangle}$ to Eq. (62), we obtain

$$\langle W \rangle \leqslant \beta^{-1} \ln \gamma. \tag{64}$$

This inequality is the second law of thermodynamics under the measurement-based work extraction.

In the high-temperature limit $\beta = 0$, the initial state (52) reduces to $\rho_{\mathrm{I,can}} = \mathbb{1}_\mathrm{I}/\mathcal{N}$, where $\mathcal{N}$ is the dimensionality of the Hilbert space of the internal system I. Thus, owing to the linearity and the trace preserving of $\mathcal{E}$, the quantity γ reduces to unity.

For general values of β, let us assume that the time evolution of I is a "natural" thermodynamic process; that is, the measurement process is not a feedback process and satisfies the first and second laws of thermodynamics for an arbitrary initial state. Hayashi and Tajima introduced and called it the standard completely positive work extraction [35,36]. The first law is satisfied by the measurement process $\mathcal{E}_j$ which changes the energy eigenstate $|h_x\rangle_\mathrm{I}$ of I to the state of the energy $h_x - \Delta e_j$ when the external system gains the energy Δe_j. The second law corresponds to the time evolution $\mathcal{E} = \sum_j \mathcal{E}_j$ which satisfies

$$S(\rho_\mathrm{I}) \leqslant S(\mathcal{E}(\rho_\mathrm{I})) \tag{65}$$

for all initial states ρ_I of I, where $S(\rho_\mathrm{I}) := -\mathrm{Tr}[\rho_\mathrm{I} \ln \rho_\mathrm{I}]$ is the von Neumann entropy. As a necessary and sufficient condition of Eq. (65) for all initial states, the time evolution $\mathcal{E}$ must be a unital map [46]:

$$\mathcal{E}(\mathbb{1}_\mathrm{I}) = \mathbb{1}_\mathrm{I}. \tag{66}$$

Then, the quantity γ is unity, and Eqs. (62) and (64) reduce to

$$\langle e^{\beta W} \rangle = 1, \tag{67}$$

$$\langle W \rangle \leqslant 0, \tag{68}$$

respectively. Hence, we obtain the same form as the previous derivations [17,18,24,25] of the Jarzynski equality under the cyclic process.

The difference of the Jarzynski equality from unity is known for feedback processes [29,47,48] and/or absolutely irreversible processes [33,49]. Because the time evolution $\mathcal{E}$ includes the feedback process, Eq. (62) also applies to the feedback process, for which the quantity γ denotes the efficiency [47]. On the other hand, because we assume that the initial state (52) of the internal system is the canonical distribution, Eq. (62) does not apply to the absolutely irreversible process.

We stress, however, that the present result is essentially different from the previous ones [17,18,24,25]. The previous derivation of the Jarzynski equality did not contain information about the fluctuation of the energy gain of the *external* system appropriately, defining the measured value of the energy loss of the *unitarily evolving internal* system as the random variable W. As we have shown in Sec. II, however, under the approximation of the unitary dynamics of the internal system, the variance of W in the previous derivation is completely different from that of the energy gain of the external system, which is the actual work that we can use. The Jarzynski equality derived from the previous formulation therefore does not give relevant information about the fluctuation of the actual work.

Our derivation of the Jarzynski equality is different in this point. We also define the measured value of the energy loss of the internal system as a random variable W, but it is equal to the measured value of the energy gain of the external system, because now we employ the unitary time evolution of the total system satisfying Eq. (50). Therefore, our derivation correctly contains the information about the fluctuation of the actual work, namely, the energy gain of the external system. For a possible extension of the present formulation to the measurement process with error, see Sec. V.

B. Noncyclic process

Next, we consider the Jarzynski equality under a noncyclic process, extending the case of the cyclic process in Sec. III A. For a noncyclic process, the energy spectrum of the internal system is different between the initial and final Hamiltonians. To apply the formalism for the cyclic process to the noncyclic one, we divide the internal system I into two subsystems, namely, a (further) internal system S and a control system C [36,43]. The internal system S is a working substance, such as a gas, while the control system C controls the Hamiltonian of the internal system S as a piston. We consider the work extracted from the internal system S.

We assume the initial Hamiltonian (48) of the internal system I in the form

$$
\begin{aligned}
H_\mathrm{I} &:= \sum_\lambda H_\mathrm{S}(\lambda) \otimes |\lambda\rangle_\mathrm{C}\langle\lambda| \\
&= \sum_\lambda \sum_{x(\lambda)} h_{x(\lambda)} |h_{x(\lambda)}, \lambda\rangle_\mathrm{SC}\langle h_{x(\lambda)}, \lambda|,
\end{aligned} \tag{69}
$$

where $H_\mathrm{S}(\lambda)$ is the Hamiltonian of S, whose eigenstate and the corresponding eigenvalue are denoted as $|h_{x(\lambda)}\rangle_\mathrm{S}$ and $h_{x(\lambda)}$, respectively; $\{|\lambda\rangle_\mathrm{C}\}$ is an orthonormal basis of the control system C; and $|h_{x(\lambda)}, \lambda\rangle_\mathrm{SC}$ denotes $|h_{x(\lambda)}\rangle_\mathrm{S} \otimes |\lambda\rangle_\mathrm{C}$. We vary the control parameter λ, making the process noncyclic. Note that we made the λ dependence of the index $x(\lambda)$ explicit, because the set of the eigenvalues of $H_\mathrm{S}(\lambda)$ depends on λ. The energy of S changes from $h_{x(\lambda)}$ to $h_{x(\lambda)} - \Delta e_j$ after the measurement process $\mathcal{E}_j$ of the specific energy gain Δe_j.

We set the initial state of I to be the canonical distribution of S with a pure state $|\lambda_\mathrm{i}\rangle_\mathrm{C}$ of C:

$$
\begin{aligned}
\rho_\mathrm{I}(\lambda_\mathrm{i}) &:= \sum_{x(\lambda_\mathrm{i})} \frac{e^{-\beta h_{x(\lambda_\mathrm{i})}}}{Z_\mathrm{S}(\lambda_\mathrm{i})} |h_{x(\lambda_\mathrm{i})}, \lambda_\mathrm{i}\rangle_\mathrm{SC}\langle h_{x(\lambda_\mathrm{i})}, \lambda_\mathrm{i}| \\
&= \rho_{\mathrm{S,can}}(\lambda_\mathrm{i}) \otimes |\lambda_\mathrm{i}\rangle_\mathrm{C}\langle\lambda_\mathrm{i}|
\end{aligned} \tag{70}
$$

with $Z_\mathrm{S}(\lambda) := \mathrm{Tr}[e^{-\beta H_\mathrm{S}(\lambda)}]$ and $\rho_{\mathrm{S,can}}(\lambda) := e^{-\beta H_\mathrm{S}(\lambda)}/Z_\mathrm{S}(\lambda)$. This means that the internal system S starts from the equilibrium with the fixed parameter λ_i. The free energy of S for a specific value of λ is given by

$$F_\mathrm{S}(\lambda) := -\beta^{-1} \ln Z_\mathrm{S}(\lambda). \tag{71}$$

We define the probability distribution of the extracted work W during the process in which the state of C changes from λ_i

to λ_f as

$$P_{\lambda_i \to \lambda_f}(W) := \sum_j \frac{\delta(W - \Delta e_j)}{p_{\lambda_i \to \lambda_f}} \text{Tr}[|\lambda_f\rangle_C\langle\lambda_f|\mathcal{E}_j(\rho_{S,\text{can}}(\lambda_i)$$

$$\otimes |\lambda_i\rangle_C\langle\lambda_i|)], \quad (72)$$

where

$$p_{\lambda_i \to \lambda_f} := \text{Tr}[|\lambda_f\rangle_C\langle\lambda_f|\mathcal{E}(\rho_{S,\text{can}}(\lambda_i) \otimes |\lambda_i\rangle_C\langle\lambda_i|)] \quad (73)$$

is the transition probability that the state of C changes from λ_i to λ_f. In the same way as in Eq. (60) of Sec. III A, we obtain

$$P_{\lambda_i \to \lambda_f}(W) = e^{-\beta W} e^{-\beta \Delta F_S(\lambda_i,\lambda_f)}$$

$$\times \sum_j \frac{\delta(W - \Delta e_j)}{p_{\lambda_i \to \lambda_f}} \text{Tr}[\rho_{S,\text{can}}(\lambda_f)$$

$$\otimes |\lambda_f\rangle_C\langle\lambda_f|\mathcal{E}_j(\mathbb{1}_S(\lambda_i) \otimes |\lambda_i\rangle_C\langle\lambda_i|)] \quad (74)$$

with $\Delta F_S(\lambda_i,\lambda_f) := F_S(\lambda_f) - F_S(\lambda_i)$, while $\mathbb{1}_S(\lambda_i) := \sum_{x(\lambda_i)} |h_{x(\lambda_i)}\rangle_S\langle h_{x(\lambda_i)}|$ is the identity operator of S with fixed parameter λ_i.

We modify the average (61) to

$$\langle f(W)\rangle_{\lambda_i \to \lambda_f} := \int dW f(W) P_{\lambda_i \to \lambda_f}(W). \quad (75)$$

We therefore arrive at the Jarzynski equality under a noncyclic process in the form

$$\langle e^{\beta W}\rangle_{\lambda_i \to \lambda_f} = \gamma_{\lambda_i \to \lambda_f} e^{-\beta \Delta F_S(\lambda_i,\lambda_f)} \quad (76)$$

with

$$\gamma_{\lambda_i \to \lambda_f} := \frac{q_{\lambda_f \to \lambda_i}}{p_{\lambda_i \to \lambda_f}}, \quad (77)$$

$$q_{\lambda_f \to \lambda_i} := \text{Tr}[\rho_{S,\text{can}}(\lambda_f) \otimes |\lambda_f\rangle_C\langle\lambda_f|\mathcal{E}(\mathbb{1}_S(\lambda_i) \otimes |\lambda_i\rangle_C\langle\lambda_i|)]. \quad (78)$$

We note that the state of C is measured only in the initial and final states. During the dynamics between these states, we cannot tell the path of the change of physical quantities of C, such as the position of a piston, nor can we tell the motion of S. It is in contrast with the fact that, in the previous derivation of the Jarzynski equality [17,18,24,25], the motion of the system is fully determined by a given path of a parameter.

In the high-temperature limit $\beta = 0$, Eq. (70) reduces to $\rho_I(\lambda) = \mathbb{1}_S(\lambda)/\mathcal{N}(\lambda) \otimes |\lambda\rangle_C\langle\lambda|$, where $\mathcal{N}(\lambda)$ is the dimensionality of the Hilbert space of S with fixed parameter λ. Thus, Eq. (77) reduces to $\gamma_{\lambda_i \to \lambda_f} = \mathcal{N}(\lambda_f)/\mathcal{N}(\lambda_i)$. In particular, when the dimensionality of the Hilbert space of S with fixed parameter λ_f is equal to one with fixed parameter λ_i, Eq. (77) reduces to unity whether $\mathcal{E}$ is unital or not.

We now argue for general values of β that the quantity $\gamma_{\lambda_i \to \lambda_f}$ is not necessarily unity for a unital map as γ was for the cyclic process. When the time evolution $\mathcal{E}$ is unital, namely, the "natural" thermodynamic process defined in the previous section, the quantity $\gamma_{\lambda_i \to \lambda_f}$ gives the ratio of the forward and the backward transition probabilities. When the time evolution $\mathcal{E}$ is unital, completely positive, and trace preserving, so is its adjoint $\mathcal{E}^\dagger$. Therefore, we can regard the adjoint map $\mathcal{E}^\dagger$ as another time evolution. Equation (78) indeed gives the backward transition probability that the state of C changes

from λ_f to λ_i:

$$q_{\lambda_f \to \lambda_i} = \text{Tr}[|\lambda_i\rangle_C\langle\lambda_i|\mathcal{E}^\dagger(\rho_{S,\text{can}}(\lambda_f) \otimes |\lambda_f\rangle_C\langle\lambda_f|)]. \quad (79)$$

As can be seen from the calculation of a simple model in Sec. IV, the backward transition probability (79) is not necessarily equal to the forward one (73). Therefore, the quantity $\gamma_{\lambda_i \to \lambda_f}$ is not necessarily unity for a unital map.

When the time evolution $\mathcal{E}$ is not unital, incidentally, we cannot regard Eq. (78) as a transition probability; because the adjoint $\mathcal{E}^\dagger$ of a nonunital map $\mathcal{E}$ is not trace preserving, the sum of Eq. (78) over λ_i is not unity:

$$\sum_{\lambda_i} q_{\lambda_f \to \lambda_i} = \text{Tr}[\rho_{S,\text{can}}(\lambda_f) \otimes |\lambda_f\rangle_C\langle\lambda_f|\mathcal{E}(\mathbb{1}_I)] \quad (80)$$

$$= \text{Tr}[\mathcal{E}^\dagger(\rho_{S,\text{can}}(\lambda_f) \otimes |\lambda_f\rangle_C\langle\lambda_f|)] \neq 1, \quad (81)$$

where $\mathbb{1}_I := \sum_\lambda \mathbb{1}_S(\lambda) \otimes |\lambda\rangle_C\langle\lambda|$ is the identity operator of I.

Finally, we show that Eq. (76) reduces to the case of the cyclic process (62) when the control system C has only one eigenstate. In this case, the state of C cannot change from the initial state, and we thereby obtain $p_{\lambda_i \to \lambda_i} = 1$ and $\Delta F_I(\lambda_i,\lambda_i) = 1$. Therefore, Eq. (74) reduces to

$$P_{\lambda_i \to \lambda_i}(W) = e^{-\beta W} \sum_j \delta(W - \Delta e_j) \text{Tr}[\rho_{S,\text{can}}(\lambda_i)$$

$$\otimes |\lambda_i\rangle_C\langle\lambda_i|\mathcal{E}_j(\mathbb{1}_S(\lambda_i) \otimes |\lambda_i\rangle_C\langle\lambda_i|)]. \quad (82)$$

Since $\mathbb{1}_I = \mathbb{1}_S(\lambda_i) \otimes |\lambda_i\rangle_C\langle\lambda_i|$ and $\rho_{I,\text{can}} = \rho_{S,\text{can}}(\lambda_i) \otimes |\lambda_i\rangle_C\langle\lambda_i|$, this equation is equivalent to Eq. (60) in Sec. III A.

IV. COEFFICIENT $\gamma_{\lambda_i \to \lambda_f}$ FOR A SIMPLE MODEL

In this section, we evaluate the quantity $\gamma_{\lambda_i \to \lambda_f}$ of Sec. III B using a simple system. We suppose that the Hamiltonian of the simple system is given by

$$H_I(\omega) = \sum_{\lambda=0,1} H_S(\lambda; \omega) \otimes |\lambda\rangle_C\langle\lambda|, \quad (83)$$

$$H_S(\lambda; \omega) := \frac{(\lambda + 1)\omega}{2} \sigma_z^S, \quad (84)$$

where ω is level spacing and $\sigma_z^S := |1\rangle_S\langle 1| - |0\rangle_S\langle 0|$. The canonical distribution $\rho_S(\lambda)$ of S at the inverse temperature β is given by

$$\rho_S(\tilde{\beta},\lambda) = \frac{1}{1 + e^{(\lambda+1)\tilde{\beta}}} |1\rangle_S\langle 1| + \frac{1}{1 + e^{-(\lambda+1)\tilde{\beta}}} |0\rangle_S\langle 0|, \quad (85)$$

where $\tilde{\beta} := \beta\omega$ is the dimensionless inverse temperature.

We consider the following measurement process $\mathcal{E}_j$:

$$\mathcal{E}_j(\rho_I) := M_j \rho_I M_j^\dagger, \quad (86)$$

$$M_j := \sum_{\lambda,\lambda'} \sum_{x(\lambda),y(\lambda')} \delta_{\Delta e_j, h_{x(\lambda)} - h_{y(\lambda')}} \Pi_{h_{y(\lambda')}} U_{\text{eff}} \Pi_{h_{x(\lambda)}}, \quad (87)$$

where $\Pi_{h_{x(\lambda)}}$ is a projection on the eigenvalue of H_I. The effective time-evolution operator U_{eff} is given by tracing out the external system from the time evolution of the total system. It defines the effective Hamiltonian H_{eff} as in $U_{\text{eff}} = e^{-iH_{\text{eff}}T}$, where T is the time duration of the measurement process $\mathcal{E}_j$. The corresponding time evolution $\mathcal{E} := \mathcal{E}_j$ is unital and,
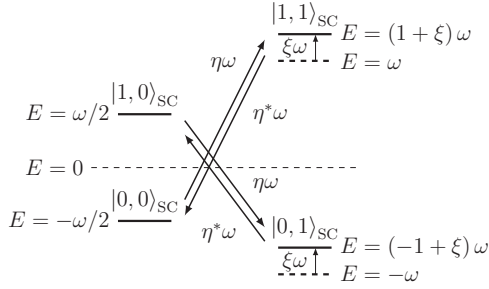
FIG. 2. Illustration of the effective Hamiltonian of Eq. (88). The parameter ξ is a real number representing the shift of the potential energy and the parameter η is a complex number representing the hopping amplitude.

therefore, is not a feedback process. For simplicity, let us suppose that the effective Hamiltonian H_{eff} is

$$H_{\text{eff}}(\omega,\xi,\eta) := H_I(\omega) + V_{\text{eff}}(\xi,\omega) + H_{\text{hop}}(\eta,\omega), \quad (88)$$

$$V_{\text{eff}}(\xi,\omega) := \xi\omega\mathbb{1}_S \otimes |1\rangle_C\langle 1|, \quad (89)$$

$$H_{\text{hop}}(\eta,\omega) := \sigma_x^S \otimes \omega(\eta|1\rangle_C\langle 0| + \eta^*|0\rangle_C\langle 1|), \quad (90)$$

where $V_{\text{eff}}(\xi,\omega)$ and $H_{\text{hop}}(\eta,\omega)$ are potential and hopping terms, the parameters ξ and η are real and complex numbers, respectively (Fig. 2), and $\sigma_x^S := |1\rangle_S\langle 0| + |0\rangle_S\langle 1|$. In other words, the unitary operator U_{eff} is given by

$$U_{\text{eff}}(\xi,\eta,\tilde{T}) := \exp(-i\tilde{T}H_{\text{eff}}(1,\xi,\eta)), \quad (91)$$

where $\tilde{T} := \omega T/\hbar$ is the dimensionless time duration.

Then, Eqs. (73) and (78) are given by

$$p_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{\beta},\tilde{T}) = \frac{p^{(1)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T})}{1 + e^{(\lambda_i+1)\tilde{\beta}}} + \frac{p^{(0)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T})}{1 + e^{-(\lambda_i+1)\tilde{\beta}}}, \quad (92)$$

$$q_{\lambda_f \to \lambda_i}(\xi,\eta,\tilde{\beta},\tilde{T}) = \frac{q^{(1)}_{\lambda_f \to \lambda_i}(\xi,\eta,\tilde{T})}{1 + e^{(\lambda_f+1)\tilde{\beta}}} + \frac{q^{(0)}_{\lambda_f \to \lambda_i}(\xi,\eta,\tilde{T})}{1 + e^{-(\lambda_f+1)\tilde{\beta}}}, \quad (93)$$

with

$$p^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T}) := \text{Tr}[|\lambda_f\rangle_C\langle\lambda_f|U_{\text{eff}}(\xi,\eta,\tilde{T})(|n\rangle_S\langle n|$$
$$\otimes |\lambda_i\rangle_C\langle\lambda_i|)U^\dagger_{\text{eff}}(\xi,\eta,\tilde{T})], \quad (94)$$

$$q^{(n)}_{\lambda_f \to \lambda_i}(\xi,\eta,\tilde{T}) := \text{Tr}[|\lambda_i\rangle_C\langle\lambda_i|U^\dagger_{\text{eff}}(\xi,\eta,\tilde{T})(|n\rangle_S\langle n|$$
$$\otimes |\lambda_f\rangle_C\langle\lambda_f|)U_{\text{eff}}(\xi,\eta,\tilde{T})], \quad (95)$$

for $n = 0,1$.

We first show $\gamma_{0\to0} = \gamma_{1\to1} = 1$ and $\gamma_{1\to0} = 1/\gamma_{0\to1}$. The projective operators $|n\rangle_S\langle n|$ ($n = 0,1$) and $|\lambda\rangle_C\langle\lambda|$ ($\lambda = 0,1$) are invariant with respect to a unitary operator $S_1 := \mathbb{1}_S \otimes (|1\rangle_C\langle 1| + e^{-i\chi}|0\rangle_C\langle 0|)$ for any real number χ and the time reversal operator Θ, which is an antiunitary operator, and the unitary operator (91) satisfies

$$S_1^\dagger U_{\text{eff}}(\xi,\eta,\tilde{T})S_1 = U_{\text{eff}}(\xi,e^{i\chi}\eta,\tilde{T}), \quad (96)$$

$$\Theta^\dagger U_{\text{eff}}(\xi,\eta,\tilde{T})\Theta = U^\dagger_{\text{eff}}(\xi,\eta^*,\tilde{T}). \quad (97)$$

Applying S_1 and Θ in Eqs. (94) and (95), we obtain

$$p^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T}) = p^{(n)}_{\lambda_i \to \lambda_f}(\xi,e^{i\chi}\eta,\tilde{T}), \quad (98)$$

$$p^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T}) = q^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta^*,\tilde{T}). \quad (99)$$

For a unitary operator $S_2 := \sigma_x^S \otimes \mathbb{1}_C$, we also have

$$S_2^\dagger U_{\text{eff}}(\xi,\eta,\tilde{T})S_2 = U^\dagger_{\text{eff}}(-\xi, -\eta,\tilde{T}), \quad (100)$$

$$S_2^\dagger(|n\rangle_S\langle n| \otimes |\lambda\rangle_C\langle\lambda|)S_2 = |1-n\rangle_S\langle 1-n| \otimes |\lambda\rangle_C\langle\lambda|, \quad (101)$$

for $n,\lambda = 0,1$, and hence

$$p^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T}) = q^{(1-n)}_{\lambda_i \to \lambda_f}(-\xi, -\eta,\tilde{T}). \quad (102)$$

Combining Eqs. (98), (99), and (102), we thus obtain

$$p^{(n)}_{\lambda_i \to \lambda_f}(\xi,|\eta|,\tilde{T}) = p^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T}) \quad (103)$$

$$= q^{(n)}_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{T}) \quad (104)$$

$$= p^{(1-n)}_{\lambda_i \to \lambda_f}(-\xi,\eta,\tilde{T}). \quad (105)$$

Therefore, we obtain

$$p_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{\beta},\tilde{T}) = q_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{\beta},\tilde{T})$$
$$= \frac{p^{(0)}_{\lambda_i \to \lambda_f}(-\xi,|\eta|,\tilde{T})}{1 + e^{(\lambda_i+1)\tilde{\beta}}}$$
$$+ \frac{p^{(0)}_{\lambda_i \to \lambda_f}(\xi,|\eta|,\tilde{T})}{1 + e^{-(\lambda_i+1)\tilde{\beta}}}, \quad (106)$$

and hence

$$\gamma_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{\beta},\tilde{T}) = \frac{q_{\lambda_f \to \lambda_i}(\xi,\eta,\tilde{\beta},\tilde{T})}{p_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{\beta},\tilde{T})} \quad (107)$$

$$= \frac{p_{\lambda_f \to \lambda_i}(\xi,\eta,\tilde{\beta},\tilde{T})}{p_{\lambda_i \to \lambda_f}(\xi,\eta,\tilde{\beta},\tilde{T})}. \quad (108)$$

As can be seen from Eq. (108), the quantity $\gamma_{\lambda_i \to \lambda_f}$ is always equal to unity if $\lambda_i = \lambda_f$; that is, $\gamma_{0\to0} = \gamma_{1\to1} = 1$, and $\gamma_{1\to0} = 1/\gamma_{0\to1}$.

Let us now find $\gamma_{1\to0}$. Calculating $p^{(0)}_{1\to0}(\xi,|\eta|,\tilde{T})$ and $p^{(0)}_{0\to1}(\xi,|\eta|,\tilde{T})$, we obtain

$$p^{(0)}_{1\to0}(\xi,|\eta|,\tilde{T}) = p^{(0)}_{0\to1}(-\xi,|\eta|,\tilde{T}) \quad (109)$$

$$= |\eta|^2\tilde{T}^2 \text{sinc}^2(f(\xi,|\eta|)\tilde{T}), \quad (110)$$

with

$$\text{sinc}(x) := \frac{\sin(x)}{x}, \quad (111)$$

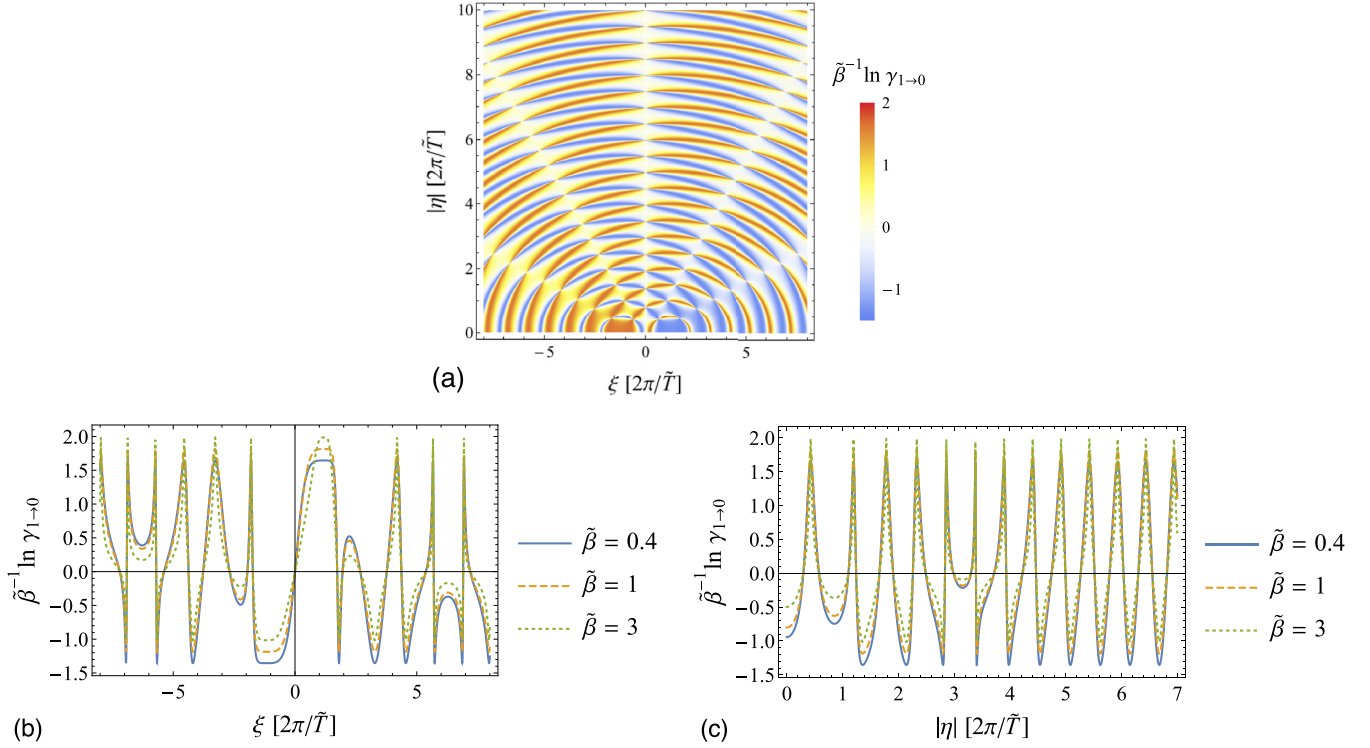$$f(\xi,|\eta|) := \sqrt{|\eta|^2 + \left(\frac{3-2\xi}{4}\right)^2}, \quad (112)$$

FIG. 3. (a) Dependence of $\tilde{\beta}^{-1} \ln \gamma_{1 \to 0}$ on the parameters ξ and $|\eta|$ for $\tilde{T} = 5$ and $\tilde{\beta} = 0.4$. Its cross sections for (b) $|\eta| = 4\pi/5$ and (c) $\xi = 6\pi/5$. The unity of $\gamma_{1 \to 0}$ (namely, the zero of $\ln \gamma_{1 \to 0}$) does not depend on the inverse temperature $\tilde{\beta}$.

and hence

$$p_{1 \to 0}(\xi, |\eta|, \tilde{\beta}, \tilde{T}) = \frac{|\eta|^2 \tilde{T}^2 \, \mathrm{sinc}^2(f(-\xi, |\eta|)\tilde{T})}{1 + e^{2\tilde{\beta}}}$$

$$+ \frac{|\eta|^2 \tilde{T}^2 \, \mathrm{sinc}^2(f(\xi, |\eta|)\tilde{T})}{1 + e^{-2\tilde{\beta}}}, \quad (113)$$

$$p_{0 \to 1}(\xi, |\eta|, \tilde{\beta}, \tilde{T}) = \frac{|\eta|^2 \tilde{T}^2 \, \mathrm{sinc}^2(f(\xi, |\eta|)\tilde{T})}{1 + e^{\tilde{\beta}}}$$

$$+ \frac{|\eta|^2 \tilde{T}^2 \, \mathrm{sinc}^2(f(-\xi, |\eta|)\tilde{T})}{1 + e^{-\tilde{\beta}}}. \quad (114)$$

We thereby plotted $\gamma_{1 \to 0}$ in Fig. 3. The quantity $\gamma_{1 \to 0}$ fluctuates around unity wildly, even exceeding unity often. This fluctuation of $\gamma_{\lambda_i \to \lambda_f}$ should be detectable in experiments and provide evidence for the present approach.

For $\xi = 0$, we find $p_{1 \to 0}(0, \eta, \tilde{\beta}, \tilde{T}) = p_{0 \to 1}(0, \eta, \tilde{\beta}, \tilde{T})$ and $\gamma_{1 \to 0}(0, \eta, \tilde{\beta}, \tilde{T}) = 1$. The other conditions for $\gamma_{1 \to 0} = 1$ are $\tilde{\beta} = 0$ or

$$\mathrm{sinc}^2(f(\xi, |\eta|)\tilde{T}) = \mathrm{sinc}^2(f(-\xi, |\eta|)\tilde{T}), \quad (115)$$

which does not depend on the inverse temperature $\tilde{\beta}$ [Figs. 3(b) and 3(c)].

V. CONCLUSION

In the present paper, we have shown that the variance of the energy transferred to the external system diverges when the dynamics of the internal quantum system is approximated to a unitary. Because of this, the work extraction under the assumption of a unitary dynamics of the internal system is unsuitable for the thermodynamics of a microscopic quantum system. We claim that the work extraction from the internal quantum system should be described as a quantum measurement process introduced by Hayashi and Tajima [35,36] and have applied this formulation to the quantum Jarzynski equality.

In the present paper, we have assumed that the measurement process can measure the energy loss of the internal system without errors; in other words, it can convert all of the energy loss into the work. However, since the actual external system is a thermodynamic system, the energy transfer may be separated into the work and the heat. The measurement process corresponding to this should be an incomplete process in which the extracted work is not equal to the energy loss; the heat generated into the external system should be regarded as the difference of the energy loss and the extracted work. It will be an important future work to consider the incomplete measurement process in order to understand the work and the heat in quantum thermodynamics.

APPENDIX A: CALCULATION OF EQS. (27) AND (28)

We here show details of the calculation of Eqs. (25) and (26). To calculate them, we introduce the matrix representation of Eqs. (10) and (21):

$$\rho_{\mathrm{I}} := \begin{pmatrix} p & q \\ q^* & 1-p \end{pmatrix}, \tag{A1}$$

$$K_{n,n'} = \begin{pmatrix} \delta_{n,n'} u_{0,0} & \delta_{n-1,n'} u_{0,1} \\ \delta_{n+1,n'} u_{1,0} & \delta_{n,n'} u_{1,1} \end{pmatrix}. \tag{A2}$$

We find

$$\bar{K}_j := \sum_{m=1}^{M} K_{j,m} = \begin{cases} \begin{pmatrix} 0 & 0 \\ u_{1,0} & 0 \end{pmatrix} & \text{for } j = 0 \\[2ex] \begin{pmatrix} u_{0,0} & 0 \\ u_{1,0} & u_{1,1} \end{pmatrix} & \text{for } j = 1 \\[2ex] \begin{pmatrix} u_{0,0} & u_{0,1} \\ u_{1,0} & u_{1,1} \end{pmatrix} & \text{for } 2 \leqslant j \leqslant M-1 \\[2ex] \begin{pmatrix} u_{0,0} & u_{0,1} \\ 0 & u_{1,1} \end{pmatrix} & \text{for } j = M \\[2ex] \begin{pmatrix} 0 & u_{0,1} \\ 0 & 0 \end{pmatrix} & \text{for } j = M+1 \\[2ex] 0 & \text{otherwise,} \end{cases} \tag{A3}$$

and thereby obtain

$$\bar{K}_j^\dagger \bar{K}_j = \begin{cases} \begin{pmatrix} |u_{1,0}|^2 & 0 \\ 0 & 0 \end{pmatrix} & \text{for } j = 0 \\[2ex] \begin{pmatrix} |u_{0,0}|^2 + |u_{1,0}|^2 & u_{1,0}^* u_{1,1} \\ u_{1,0} u_{1,1}^* & |u_{1,1}|^2 \end{pmatrix} & \text{for } j = 1 \\[2ex] \begin{pmatrix} |u_{0,0}|^2 + |u_{1,0}|^2 & u_{0,0}^* u_{0,1} + u_{1,0}^* u_{1,1} \\ u_{0,0} u_{0,1}^* + u_{1,0} u_{1,1}^* & |u_{0,1}|^2 + |u_{1,1}|^2 \end{pmatrix} & \text{for } 2 \leqslant j \leqslant M-1 \\[2ex] \begin{pmatrix} |u_{0,0}|^2 & u_{0,0}^* u_{0,1} \\ u_{0,0} u_{0,1}^* & |u_{0,1}|^2 + |u_{1,1}|^2 \end{pmatrix} & \text{for } j = M \\[2ex] \begin{pmatrix} 0 & 0 \\ 0 & |u_{0,1}|^2 \end{pmatrix} & \text{for } j = M+1 \\[2ex] 0 & \text{otherwise.} \end{cases} \tag{A4}$$

Using Eqs. (5), (6), and (7), we obtain

$$\sum_j j \bar{K}_j^\dagger \bar{K}_j = \frac{1}{2} M(M+1) + \begin{pmatrix} -|u_{1,0}|^2 M & u_{0,0}^* u_{0,1}(M-1) \\ u_{0,0} u_{0,1}^*(M-1) & |u_{0,1}|^2 M \end{pmatrix}, \tag{A5}$$

$$\sum_j j^2 \bar{K}_j^\dagger \bar{K}_j = \frac{1}{6} M(M+1)(2M+1) + \begin{pmatrix} -|u_{1,0}|^2 M^2 & u_{0,0}^* u_{0,1}(M^2-1) \\ u_{0,0} u_{0,1}^*(M^2-1) & |u_{0,1}|^2(M^2+2M) \end{pmatrix}. \tag{A6}$$

Inserting Eqs. (A1), (A5), and (A6) into Eqs. (25) and (26), we obtain

$$\langle h \rangle = \Delta \left[\frac{1}{2}(M+1) + |u_{0,1}|^2(1-p) - |u_{1,0}|^2 p + 2(1-M^{-1})\mathrm{Re}(u_{0,0} u_{0,1}^* q) \right], \tag{A7}$$

$$\langle h^2 \rangle = \Delta^2 \left[\frac{1}{6}(M+1)(2M+1) + |u_{0,1}|^2(M+2)(1-p) - |u_{1,0}|^2 Mp + 2(M-M^{-1})\mathrm{Re}(u_{0,0} u_{0,1}^* q) \right]. \tag{A8}$$

Now, we arrange Eqs. (A7) and (A8) using the average of the energy transfer $\langle w \rangle$. The expectation of energy with respect to the initial state (20) is given by

$$\mathrm{Tr}\,[H_{\mathrm{E}} \rho_{\mathrm{E}}] = \frac{\Delta}{2}(M+1). \tag{A9}$$

Combining Eqs. (11), (A7), and (A9), we obtain the expectation of w in the form

$$\langle w \rangle = \langle h \rangle - \mathrm{Tr}\,[H_{\mathrm{E}}\rho_{\mathrm{E}}] \tag{A10}$$

$$= \Delta[|u_{0,1}|^2(1-p) - |u_{1,0}|^2 p + 2(1 - M^{-1})$$

$$\times \mathrm{Re}(u_{0,0}u_{0,1}^* q)]. \tag{A11}$$

Therefore, Eqs. (A7) and (A8) reduce to Eqs. (27) and (28).

APPENDIX B: CALCULATION OF EQ. (43)

Let us find Eq. (43). The eigenvalues of $\mathcal{K}(\rho_{\mathrm{I}})$ and ρ_{I} are given by

$$\lambda_{\pm} := \tfrac{1}{2} \pm R, \quad \lambda'_{\pm} := \tfrac{1}{2} \pm R', \tag{B1}$$

where R and R' are defined in Eqs. (44) and (45), respectively. The diagonal matrices Λ and Λ' are given by

$$\Lambda := \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \quad \Lambda' := \begin{pmatrix} \lambda'_+ & 0 \\ 0 & \lambda'_- \end{pmatrix}. \tag{B2}$$

We parametrize the arbitrary unitary matrix $\tilde{U}_{\mathrm{I}}$ as

$$\tilde{U}_{\mathrm{I}} := e^{i\phi} \begin{pmatrix} e^{i\psi_1}\cos\theta & e^{i\psi_2}\sin\theta \\ -e^{-i\psi_2}\sin\theta & e^{-i\psi_1}\cos\theta \end{pmatrix}, \tag{B3}$$

where ϕ, θ, ψ_1, and ψ_2 are real parameters. Combining Eqs. (B2) and (B3), we obtain

$$\tilde{U}_{\mathrm{I}}\Lambda'\tilde{U}_{\mathrm{I}}^\dagger - \Lambda = \begin{pmatrix} R'\cos 2\theta - R & -e^{i(\psi_1+\psi_2)}R'\sin 2\theta \\ -e^{-i(\psi_1+\psi_2)}R'\sin 2\theta & -(R'\cos 2\theta - R) \end{pmatrix}, \tag{B4}$$

where eigenvalues are given by $\pm\sqrt{R'^2 + R^2 - 2RR'\cos 2\theta}$. We thereby obtain

$$\tfrac{1}{2}\,\mathrm{Tr}\,|\tilde{U}_{\mathrm{I}}\Lambda'\tilde{U}_{\mathrm{I}}^\dagger - \Lambda| = \sqrt{R'^2 + R^2 - 2RR'\cos 2\theta}. \tag{B5}$$

Therefore, we obtain Eq. (43).

[1] S. Carnot, *Reflections on the Motive Power of Fire and on Machines Fitted to Develop that Power* (Bachelier, Paris, 1824).
[2] E. Fermi, *Thermodynamics* (Dover, New York, 1956).
[3] J. Rousselet, L. Salome, A. Ajdari, and J. Prostt, Nature (London) **370**, 446 (1994).
[4] L. P. Faucheux, L. S. Bourdieu, P. D. Kaplan, and A. J. Libchaber, Phys. Rev. Lett. **74**, 1504 (1995).
[5] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Nat. Phys. **6**, 988 (2010).
[6] S. An, J.-N. Zhang, M. Um, D. Lv, Y. Lu, J. Zhang, Z.-Q. Yin, H. T. Quan, and K. Kim, Nat. Phys. **11**, 193 (2014).
[7] T. B. Batalhão, A. M. Souza, L. Mazzola, R. Auccaise, R. S. Sarthour, I. S. Oliveira, J. Goold, G. De Chiara, M. Paternostro, and R. M. Serra, Phys. Rev. Lett. **113**, 140601 (2014).
[8] Y. Ishii and T. Yanagida, Single Mol. **1**, 5 (2000).
[9] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
[10] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[11] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).
[12] J. Kurchan, J. Phys. A **31**, 3719 (1998).
[13] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[14] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[15] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[16] S. Ito and T. Sagawa, Phys. Rev. Lett. **111**, 180603 (2013).
[17] J. Kurchan, arXiv:cond-mat/0007360.
[18] H. Tasaki, arXiv:cond-mat/0009244.
[19] C. M. Bender, D. C. Brody, and B. K. Meister, Proc. R. Soc. London Ser. A **458**, 1519 (2002).
[20] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102 (2007).
[21] T. Sagawa and M. Ueda, Phys. Rev. Lett. **100**, 080403 (2008).
[22] T. Sagawa and M. Ueda, Phys. Rev. Lett. **102**, 250602 (2009).
[23] K. Jacobs, Phys. Rev. A **80**, 012322 (2009).
[24] M. Campisi, P. Talkner, and P. Hänggi, Phys. Rev. Lett. **102**, 210401 (2009).
[25] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).
[26] M. Esposito and C. Van den Broeck, Europhys. Lett. **95**, 40004 (2011).
[27] S. De Liberato and M. Ueda, Phys. Rev. E **84**, 051122 (2011).
[28] T. Sagawa, in *Lectures on Quantum Computing, Thermodynamics, and Statistical Physics*, edited by M. Nakahara and S. Tanaka (World Scientific, Singapore, 2012), Chap. 3, pp. 125–190.
[29] K. Funo, Y. Watanabe, and M. Ueda, Phys. Rev. E **88**, 052121 (2013).
[30] H. Tajima, Phys. Rev. E **88**, 042143 (2013).
[31] H. Tajima, arXiv:1311.1285.
[32] H. Tajima, in Proceedings of the 12th Asia Pacific Physics Conference [J. Phys. Soc. Jpn. 012129, 1 (2014)].
[33] K. Funo, Y. Murashita, and M. Ueda, New J. Phys. **17**, 075005 (2015).
[34] K. Yamamoto and N. Hatano, Phys. Rev. E **92**, 042165 (2015).
[35] M. Hayashi and H. Tajima, arXiv:1504.06150.
[36] H. Tajima, Ph.D. thesis, The University of Tokyo, 2013 (unpublished).
[37] H. Tajima and M. Hayashi, arXiv:1405.6457.
[38] H. Tasaki, Phys. Rev. Lett. **116**, 170402 (2016).
[39] D. Kafri and S. Deffner, Phys. Rev. A **86**, 044302 (2012).
[40] T. Albash, D. A. Lidar, M. Marvian, and P. Zanardi, Phys. Rev. E **88**, 032146 (2013).
[41] J. Goold and K. Modi, arXiv:1407.4618.
[42] J. Goold, M. Paternostro, and K. Modi, Phys. Rev. Lett. **114**, 060602 (2015).
[43] M. Horodecki and J. Oppenheim, Nat. Commun. **4**, 1 (2013).
[44] P. Skrzypczyk, A. J. Short, and S. Popescu, Nat. Commun. **5**, 4185 (2014).
[45] J. Åberg, Phys. Rev. Lett. **113**, 150402 (2014).
[46] G. Kimura, in *Decoherence Suppression in Quantum Systems 2008*, edited by M. Nakahara, R. Rahimi, and A. SaiToh (World Scientific, Singapore, 2009), pp. 1–51.
[47] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010).
[48] Y. Morikuni and H. Tasaki, J. Stat. Phys. **143**, 1 (2011).
[49] Y. Murashita, K. Funo, and M. Ueda, Phys. Rev. E **90**, 042110 (2014).