学位論文

Theoretical Study of Controlling Nonequilibrium Systems using Fluctuation Theorems and Trade-off Relations

(ゆらぎの定理とトレードオフ関係を用いた非平衡系の制御の理論的研究)

平成２７年１２月博士（理学）申請

東京大学大学院 理学系研究科 物理学専攻
布能　謙

Abstract

Recent developments of the analysis of the second law of thermodynamics by using the information theoretic quantities has led to the discovery of the so-called nonequilibrium free energy and has allowed us to formulate thermodynamic relations for arbitrary nonequilibrium initial and final states. By using the nonequilibrium free energies, we can quantify thermodynamic costs of information encoding and erasure processes as well as to quantify the extractable work from information heat engines in a unified way. In those general situations, reducing dissipation allows us to increase the extractable work from the system (or to reduce the work cost), and reducing work fluctuation allows us to narrow the work probability distribution and to extract a fluctuation-free work from the system. Suppressing both work fluctuation and dissipation as much as possible is vital to control mesoscopic and nano systems that work at the level of thermal fluctuations. Previous studies have searched for a vanishing work fluctuation protocol by using techniques of quantum information theory, known as the single-shot statistical mechanics (and sometimes called the deterministic work extraction protocol) and the vanishing dissipation protocol, known as the thermodynamically reversible protocol by using the second law of thermodynamics.

In this thesis, we address issues concerning a method to control nonequilibrium systems by using fluctuation theorems and trade-off relation between work fluctuation and fluctuation in dissipation. We derive quantum fluctuation theorems under measurement and feedback control, which allow us to study how to suppress dissipation during a general quantum measurement and a feedback control. We then consider a method to suppress work fluctuation and fluctuation in dissipation simultaneously as much as possible by showing the trade-off relation between work and dissipation in terms of their standard deviations and derive an explicit protocol that achieves the lower bound of the trade-off relation. We show that the explicit protocol reproduces the deterministic work extraction protocol in the limit of vanishing work fluctuation and the thermodynamically reversible protocol in the limit of vanishing fluctuation in dissipation.

Acknowledgment

I would like to express my gratitude to my supervisor, Prof. Masahito Ueda for his sincere guidance, supports, and encouragement. I would like to thank all the group members for an delightful time in Ueda group.

I was financially supported from JSPS (Grant No. 254105) and through Advanced Leading Graduate Course for Photon Science (ALPS).

Publication List

1. K. Funo, Y. Watanabe, and M. Ueda, *Integral quantum fluctuation theorems under measurement and feedback control*, Physical Review E **88**, 052121 (2013).
2. K. Funo, Y. Watanabe, and M. Ueda, *Thermodynamic work gain from entanglement*, Physical Review A **88**, 052319 (2013).
3. Y. Murashita K. Funo, and M. Ueda, *Nonequilibrium equalities in absolutely irreversible processes*, Physical Review E **90**, 042110 (2014).
4. Y. Ashida, K. Funo Y. Murashita, and M. Ueda *General achievable bound of extractable work under feedback control*, Physical Review E **90**, 052125 (2014).
5. M. Hotta, J. Matsumoto, and K. Funo, *Black hole firewalls require huge energy of measurement* Physical Review D **89**, 124023 (2014).
6. M. Frey, K. Funo, and M. Hotta, *Strong local passivity in finite quantum systems*, Physical Review E **90**, 012127 (2014).
7. K. Funo, Y. Murashita, and M. Ueda, *Quantum nonequilibrium equalities with absolute irreversibility*, New Journal of Physics **17**, 075005 (2015).
8. K. Funo and M. Ueda, *Work fluctuation-dissipation trade-off in heat engines*, Physical Review Letters **115**, 260601 (2015).

Contents

Chapter 1

Introduction

The information-theoretic analysis of thermodynamic processes in nonequilibrium statistical mechanics and stochastic thermodynamics enable us to formulate the second law of thermodynamics in arbitrary nonequilibrium situations [1–5]. Owing to this advancement, we can precisely quantify the work cost of information encoding and erasure as well as the extractable work from the information heat engines [6–22]. These studies on information processing in small systems are of extreme importance to experimentally control nanomachines and biological molecules that work at the level of thermodynamic fluctuations. They are also related to the fundamental role of information in thermodynamics, which may elucidate the law of thermodynamic entropy increase and the arrow of time in macroscopic systems [1].

Recently, information heat engines (the Szilard engine) and the Landauer's information erasure process have been demonstrated in mesoscopic and nano systems [23–28]. In those experiments, work fluctuates in general due to thermal fluctuations, and the work probability distribution is obtained by repeating the experiments in the same setup many times. For general nonequilibrium processes, including the experiments above, we want to narrow the width of the work probability distribution and reduce the work fluctuation, and also shift the center of the work probability distribution to extract the maximum average work from the system (which is equivalent to reducing the amount of dissipation during the processes). If we can obtain a method to reduce work fluctuation and dissipation, we can maximize the efficiency of given thermodynamic tasks and also we can extract a fluctuation-free work from the system. Thus, suppressing both work fluctuation and dissipation as much as possible is vital to control mesoscopic and nano systems in a small (quantum) world.

Previous studies have searched for protocols which suppress either work fluctuation or dissipation. Vanishing work fluctuation protocols are explored by using the single-shot statistical mechanics [29–39]. Quantum information-theoretic tools such as the noisy operations and majorization [40–42] were used to study the transfer of information subjected to noise. Recently, they are generalized to thermodynamic settings such as the thermal operation and thermo-majorization [30]. By using these information-theoretic tools, we can search for a general state transformation of the (quantum) system interacting with a heat bath and study the deterministic work extraction protocol. Also, the vanishing dissipation protocol is known as the thermodynamically reversible protocol in nonequilibrium statistical mechanics, which exactly brings the final state of the system back to the initial one via the time-reversal of the original (forward) protocol. The thermodynamically reversible protocol is obtained by searching for a protocol which satisfies the equality condition of the second law of thermodynamics [1]. Moreover the fluctuation theorems and the Jarzynski equality can be used for a wide range of classical [43–57] and quantum systems [58–69], including feedback control [17–19, 70–73] and they provide with a method to analyze nonequilibrium dynamics of the system.

In this thesis, we derive quantum fluctuation theorems under measurement and feedback control, which can be used to analyze the extractable work from the system by utilizing the information about the system via the feedback control. It can also be used to analyze the extractable work from the measurement device when we perform measurement and obtain the information. The obtained equalities enable us to study an efficient way of performing a general quantum measurement and a feedback control and reduce dissipation during those processes.

We then focus on the problem of reducing both the amounts of work fluctuation and dissipation. We prove a trade-off relation between work fluctuation and fluctuation in dissipation, which gives a constraint on fluctuating thermodynamic quantities due to the nonequilibriumness of the state of the system. The explicit protocol which achieves the lower bound of the trade-off relation offers us a method to reduce both work fluctuation and fluctuation in dissipation as much as possible. By analyzing the obtained protocol, we recover the deterministic work extraction protocol in the limit of vanishing work fluctuation and the thermodynamically reversible protocol in the limit of vanishing fluctuation in dissipation. In particular, we obtain a method to derive the single-shot results by using the detailed fluctuation theorem and thus give a connection between the single-shot statistical mechanics and nonequilibrium statistical mechanics.

This thesis is organized as follows. The main results in this thesis are presented in Chap. 4 and Chap. 6.

In Chap. 2, we review a few selected topics concerning information processing protocols which are examples of thermodynamic processes starting and/or ending at arbitrary nonequilibrium states. In Sec. 2.1, we start by formulating the second law of thermodynamics for arbitrary initial and final states. This is carried out by introducing the nonequilibrium free energy [1–3], which plays the role of the reversible work in nonequilibrium situation. We use the obtained second law to quantify the extractable work or the work cost of information processing protocols in the following sections. In Sec. 2.2, we discuss the Szilard engine [9], which is the simplest model of the information heat engine. We explain how the acquired knowledge of the system via the measurement is utilized to extract work from the system. The thermodynamic cost of performing measurement, i.e., the storage of the measurement outcome in the measurement device is discussed in Sec. 2.3. The thermodynamic cost of resetting the stored information, known as the information erasure [11] is discussed in Sec. 2.4. In Sec. 2.5, we discuss some information-theoretic properties of the nonequilibrium free energy. We show that the nonequilibrium free energy can be related to an information distance between the state of the system and the canonical distribution. We consider a quantum system in Sec. 2.6 and review the second law of thermodynamics under measurement and feedback control.

In Chap. 3, we review the nonequilibrium equalities such as the Jarzynski equality [50, 51] and the detailed fluctuation theorems [52, 53]. These exact equality relations on thermodynamic quantities offer strong theoretical tools to analyze the nonequilibrium dynamics of the system. We use these relations to derive one of the main results in this thesis in Chap. 6. In Sec. 3.1, we review the detailed fluctuation theorem which relates the ratio of the path probabilities of the forward and backward paths to the total entropy production. Next, we relate the total entropy production to the extractable work from the system by using the nonequilibrium free energy discussed in Sec. 3.2. Using the detailed fluctuation theorems, we derive the integral fluctuation theorem and a generalized Jarzynski equality for arbitrary initial and final states in Sec. 3.3. From the obtained equalities, we derive the second law of thermodynamics discussed in Chap. 2. In Sec. 3.4, we derive nonequilibrium equalities which are applicable to information heat engines.

In Chap. 4, we derive quantum fluctuation theorems under measurement and feedback control, which constitute one of the main results of this thesis. In Sec. 4.2, we derive detailed fluctuation theorems for both the measurement process and the feedback control process, by introducing path probability distributions corresponding to the measurement and the feedback control. In Sec. 4.3, we use the obtained detailed fluctuation theorems and derive the integrated versions of the quantum fluctuation theorems and also the quantum Jarzynski equalities. From the obtained equalities, the second law of thermodynamics under measurement and feedback control is reproduced. We also discuss the difference between the classical system and the quantum system in Sec. 4.4.

In Chap. 5, we review the single-shot statistical mechanics, which provides a method to construct a vanishing work fluctuation protocol (namely, the deterministic work extraction protocol). In Sec. 5.1, we introduce the thermal operation which describes a general state transformation of the system which interacts with a heat bath. The allowed state transformation via the thermal operation can be decided by using the thermo-majorization criteria, which is discussed in Sec. 5.2. In Sec. 5.3, we introduce the so-called work qubit which stores the extractable work from the system as the excited energy of the work qubit. We show that the deterministic work extraction protocol can be obtained by considering the thermal operation of the combined system composed of the system and the work qubit.

In Chap. 6, we derive the trade-off relations between work fluctuation and dissipation in fluctuation, which is the second main result of this thesis as discussed in Sec. 6.2. We show that work and dissipation cannot take definite values simultaneously, and the degrees of their fluctuations are bounded from below by the fluctuation in the nonequilibrium free-energy difference, which quantifies the nonequilibriumness of the initial and final states. The explicit protocol which achieves the lower bound of the trade-off relations is discussed in Sec. 6.5. We also discuss the amount of dissipation and the extractable work during the explicit protocol and show that they are given by the information distances which quantify the nonequilibriumness of the state of the system. An extension to information heat engines is made in Sec. 6.7. In Chap. 7, we summarize this thesis and discuss future studies.

The results in Chap. 4 are based on Refs. [72,73] which was done in collaboration with M. Ueda, Y. Watanabe and Y. Murashita. The results in Chap. 6 are based on Ref. [74] which was done in collaboration with M. Ueda.

Chapter 2

Thermodynamics and information processing

In this chapter, we review some examples of thermodynamic processes starting from and ending in arbitrary nonequilibrium distributions. As we discuss in the main results, we derive the fundamental bounds on work fluctuation and dissipation in those nonequilibrium situations, and give a method to control those systems. We start from a general setup and derive second law-like inequalities, and move to some specific examples such as information erasure and feedback control protocols. We introduce the nonequilibrium free energy, which is a key information-theoretic quantity to analyze the nonequilibrium situations described above.

2.1 Second law of thermodynamics

In this section, we consider a general setup and derive second law-like inequalities.

Let us consider a classical system which interacts with a single heat bath whose inverse temperature is β. We denote $p_{\rm ini}(x)$ and $p_{\rm fin}(y)$ as the initial and final distributions of the system, respectively. We also denote $\langle Q \rangle$ as the heat absorbed by the system due to the interaction with the heat bath. At this moment, we do not specify the dynamics of the total system composed with the system and the heat bath. Some examples of the dynamics of the composite system and also the definition of heat are presented in Sec. 3.1. Let us introduce the total entropy production which is defined as

$$\langle \sigma \rangle := S(p_{\rm fin}) - S(p_{\rm ini}) - \beta \langle Q \rangle \geq 0, \tag{2.1}$$

where

$$S(p) := -\sum_x p(x) \ln p(x) \tag{2.2}$$

is the Shannon entropy. The right-hand side of Eq. (2.1) is composed of a change in the Shannon entropy of the system and a change in the entropy of the heat bath, if we identify $-\beta \langle Q \rangle$ as the change in the thermodynamic entropy of the heat bath $\Delta S_{\rm bath}$. Thus, the total entropy production (2.1) quantifies the total amount of entropy produced during the thermodynamic process, which is always nonnegative as we can derive for general settings if we specify the dynamics of the composite system. This is carried out in a general way by using the so-called nonequilibrium equalities in Sec. 3.3.

Let us denote the initial and final Hamiltonians of the system as H_{λ_0} and H_{λ_1}, respectively, where λ_t is a control parameter of the Hamiltonian. For example, λ_t specifies a time dependence of an external potential of the system. If the initial and final states are given by the canonical distributions, i.e., $p_{\rm ini} = {\rm e}^{-\beta H_{\lambda_0}}/Z_{\lambda_0}$ and $p_{\rm fin} = {\rm e}^{-\beta H_{\lambda_1}}/Z_{\lambda_1}$, the Shannon entropies satisfy the thermodynamic relations:

$$S\left(\frac{{\rm e}^{-\beta H_{\lambda_0}}}{Z_{\lambda_0}}\right) = \beta \langle H_{\lambda_0}\rangle - \beta F^{\rm eq}_{\lambda_0}, \tag{2.3}$$

$$S\left(\frac{{\rm e}^{-\beta H_{\lambda_1}}}{Z_{\lambda_1}}\right) = \beta \langle H_{\lambda_1}\rangle - \beta F^{\rm eq}_{\lambda_1}, \tag{2.4}$$

where

$$\langle H_{\lambda_0}\rangle := \sum_x p_{\rm ini}(x) H_{\lambda_0}(x) \tag{2.5}$$

$$= \sum_x \frac{{\rm e}^{-\beta H_{\lambda_0}(x)}}{Z_{\lambda_0}} H_{\lambda_0}(x) \tag{2.6}$$

is the initial energy of the system and

$$F^{\rm eq}_{\lambda_0} := -\beta^{-1} \ln Z_{\lambda_0} \tag{2.7}$$

is the initial equilibrium free energy, and similar relations hold for the final state. By substituting Eqs. (2.3) and (2.4) to (2.1), we obtain the second-law like inequality

$$\langle W_{\rm ext}\rangle \leq -\Delta F^{\rm eq}, \tag{2.8}$$

where $\Delta F^{\rm eq} := F^{\rm eq}_{\lambda_1} - F^{\rm eq}_{\lambda_0}$ is the equilibrium free-energy difference and the extractable work from the system is defined as a change in the internal energy of the system plus the heat absorbed by the system, given by

$$\langle W_{\rm ext}\rangle := \langle H_{\lambda_0}\rangle - \langle H_{\lambda_1}\rangle + \langle Q\rangle\,. \tag{2.9}$$

For nonequilibrium initial and final states, the second-law like inequality (2.8) can be generalized to nonequilibrium situations by introducing the nonequilibrium free energy defined as [1–3]

$$\mathcal{F}(p, H) := \langle H\rangle_p - \beta^{-1} S(p), \tag{2.10}$$

where $\langle H\rangle_p := \sum_x p(x)H(x)$. The nonequilibrium free energy generalizes the equilibrium free energy to nonequilibrium distributions. By using Eq. (2.10), the total entropy production (2.1) can be expressed as

$$\begin{aligned}\langle\sigma\rangle &= \beta(\mathcal{F}(p_{\rm ini}, H_{\lambda_0}) - \mathcal{F}(p_{\rm fin}, H_{\lambda_1})) - \beta(\langle H_{\lambda_0}\rangle - \langle H_{\lambda_1}\rangle + \langle Q\rangle) \\ &= \beta(\mathcal{F}(p_{\rm ini}, H_{\lambda_0}) - \mathcal{F}(p_{\rm fin}, H_{\lambda_1})) - \beta\langle W_{\rm ext}\rangle\,,\end{aligned} \tag{2.11}$$

Then, from (2.1), we can derive the second law-like inequality for arbitrary initial and final states as follows:

$$\langle W_{\rm ext}\rangle \leq \mathcal{F}(p_{\rm ini}, H_{\lambda_0}) - \mathcal{F}(p_{\rm fin}, H_{\lambda_1}). \tag{2.12}$$

We find from (2.12) that the extractable work from the system is bounded from above by the nonequilibrium free-energy difference. We can interpret the nonequilibrium free energy as the reversible work in nonequilibrium situations as a counterpart of the equilibrium free energy which is the reversible work for transitions between thermal equilibrium states.

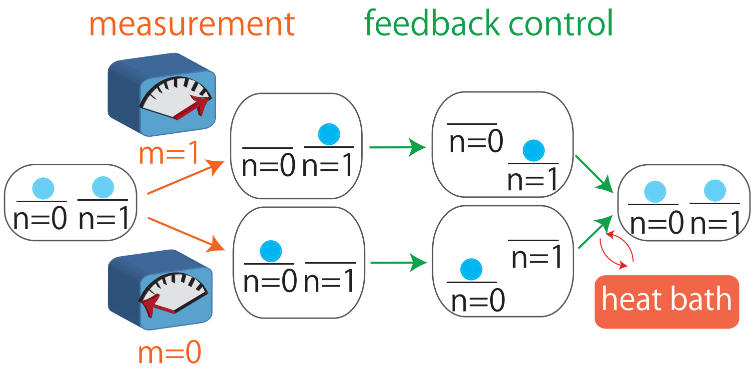

Figure 2.1: Szilard engine. We initially prepare a two-level degenerate Hamiltonian. We assume that the initial state is given by the canonical distribution (i.e., a maximally mixed state). A measurement is performed to acquire information about the state n of the system, and the measurement outcome m is recorded in the measurement apparatus. A feedback is performed by the quench of the Hamiltonian, lowering the energy level of the state $n = m$ relative to that of the other state. Finally, we let the system interact with the heat bath and quasi-statically bring the Hamiltonian back to the initial one.

2.2 Szilard engine

The information heat engines utilize the acquired knowledge obtained by the measurement on the microscopic degrees of freedom and extract work from the system via feedback control. This work extraction can be done because an increase of the knowledge of the system decreases the Shannon entropy, and hence the nonequilibrium free energy (2.10) increases. This increase in the nonequilibrium free energy can be extracted as work by feedback control of the system. In this section, we consider the simplest model of the information heat engine, called the Szilard engine. We first explain the explicit protocol (see also Fig. 2.1).

Let us consider a two-level system with an initially degenerate Hamiltonian. We label the states of the system by $n = 0$ and $n = 1$. If we start from a canonical distribution, the initial state is given by a maximally mixed state

$$p^S(n=0) = \frac{1}{2},\ p^S(n=1) = \frac{1}{2}. \tag{2.13}$$

Next, a measurement is performed to distinguish the state n of the system. We denote m as the measurement outcome, and the probability of obtaining the measurement outcome m is given by

$$p^M(m=0) = \frac{1}{2},\ p^M(m=1) = \frac{1}{2}. \tag{2.14}$$

We use the acquired information via measurement and perform the following feedback control. Depending on the measurement outcome, we quench the Hamiltonian of the system and lower the energy level of the state $n = m$ relative to that of the other state. After the quench process, we let the system interact with the heat bath, and quasi-statically change the energy levels of the Hamiltonian back to the initial degenerate ones. Because of the energy change of the system during the quench process and also the isothermal work extraction during the quasi-static process, a positive amount of work can be extracted from the system. In the ideal limit, we can extract

$$\langle W^S_{\rm ext} \rangle = \beta^{-1} \ln 2 \tag{2.15}$$

from the Szilard engine, where $\ln 2$ corresponds to the acquired $1-$bit of information via the measurement.

We explain the maximum amount of extractable work (2.15) from the Szilard engine by using the nonequilibrium free energy (2.10). Here, we introduce a classical two-level system which records the measurement outcome m and denote this system as the measurement apparatus or the memory M. If the state of the system is $n = 0$ ($n = 1$), the measurement outcome $m = 0$ ($m = 1$) is recorded in M by transforming the state of M to the $m = 0$ ($m = 1$) state. After the measurement outcome is recorded in M, the joint probability distribution of the system S and the measurement apparatus M is given by

$$p^{SM}(n=0, m=0) = p^{SM}(n=1, m=1) = \frac{1}{2}, \tag{2.16}$$

$$p^{SM}(n=0, m=1) = p^{SM}(n=1, m=0) = 0, \tag{2.17}$$

and also the postmeasurement state of the system conditioned on the measurement outcome is given by

$$p^S(n|m) := \frac{p^{SM}(n,m)}{p^M(m)}, \tag{2.18}$$

where

$$p^S(n|m) = 1 \quad \text{if } n = m, \tag{2.19}$$

$$p^S(n|m) = 0 \quad \text{if } n \neq m. \tag{2.20}$$

Then, the nonequilibrium free energy of the post-measurement state conditioned on the measurement outcome m is given by

$$\sum_m p^M(m)\mathcal{F}(p^{S|m}, H^S) = \sum_m p^M(m) \left\langle H^S \right\rangle_{p^{S|m}} - \beta^{-1} \sum_m p^M(m) S(p^{S|m}), \tag{2.21}$$

where we take average over m by multiplying $p^M(m)$ given by (2.14). In Eq. (2.21), we denote H^S as the initial degenerate Hamiltonian of the system and the conditional Shannon entropy as

$$\begin{aligned} \sum_m p^M(m) S(p^{S|m}) &= -\sum_m p^M(m) \sum_n p^S(n|m) \ln p^S(n|m) \\ &= S(p^{SM}) - S(p^M). \end{aligned} \tag{2.22}$$

The initial equilibrium free energy is given by

$$F^S_{\rm eq} = \mathcal{F}(p^S, H^S) = \left\langle H^S \right\rangle_{p^S} - \beta^{-1} S(p^S). \tag{2.23}$$

The measurement process allows us to acquire information about the system and increases the free energy of the system by

$$\sum_m p^M(m)\mathcal{F}(p^{S|m}, H^S) - F^S_{\rm eq} = \beta^{-1} \ln 2, \tag{2.24}$$

where we use Eqs. (2.13), (2.19) and (2.20) and the relation

$$\sum_m p^M(m) \left\langle H^S \right\rangle_{p^{S|m}} = \sum_{n,m} p^{SM}(n,m) H^S(n) = \sum_n p^S(n) H^S(n) = \left\langle H^S \right\rangle_{p^S} \tag{2.25}$$

to obtain the right-hand side of Eq. (2.24). This additional amount of the free energy (2.23) can be extracted as work which is given by Eq. (2.15). For a more detailed derivation of Eq. (2.15),

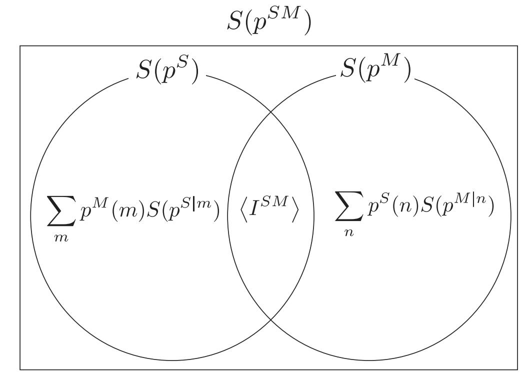

Figure 2.2: Venn diagram illustration of the relation between the Shannon entropies $S(p^S)$, $S(p^M)$, the mutual information $\langle I^{SM}\rangle$, and the conditional entropies $\sum_m p^M(m)S(p^{S|m})$, $\sum_n p^S(n)S(p^{M|n})$.

we use the second-law like inequality (2.12), which states that the upperbound of the extractable work is quantified by the nonequilibrium free-energy difference, given by

$$\begin{aligned}\langle W^S_{\rm ext}\rangle &\leq \sum_m p^M(m)\mathcal{F}(p^{S|m}, H^S) - \mathcal{F}(p^S, H^S) \\ &= -\beta^{-1}\sum_m p^M(m)\sum_n p^S(n|m)\ln p^S(n|m) + \beta^{-1}\sum_n p^S(n)\ln p^S(n) \\ &= \beta^{-1}\sum_{n,m} p^{SM}(n,m)\ln\frac{p^{SM}(n,m)}{p^S(n)p^M(m)} \\ &= \beta^{-1}\langle I^{SM}\rangle, \end{aligned} \tag{2.26}$$

where we use the fact that the final state of the system is also given by the maximally mixed state (2.13). Here, we introduce the mutual information content in Eq. (2.26), defined as

$$\begin{aligned}\langle I^{SM}\rangle &:= S(p^S) + S(p^M) - S(p^{SM}) \\ &= S(p^S) - \sum_m p^M(m)S(p^{S|m}) \\ &= \sum_{n,m} p^{SM}(n,m)\ln\frac{p^{SM}(n,m)}{p^S(n)p^M(m)}. \end{aligned} \tag{2.27}$$

The mutual information quantifies the correlation between two subsystems S and M. The conditional entropy (2.22) quantifies the entropy of the system when we have an additional knowledge of the system labeled by m. Then, the difference between the entropy of the system $S(p^S)$ and the conditional entropy $\sum_m p^M(m)S(p^{S|m})$ quantifies the obtained information via measurement (see Fig. 2.2). From (2.12) and (2.26), the second-law like inequality takes the following form

$$\langle W^S_{\rm ext}\rangle \leq \beta^{-1}\langle I^{SM}\rangle = \beta^{-1}\ln 2, \tag{2.28}$$

where the equality condition in (2.28) gives the maximally extractable work (2.15).

2.3 Measurement

Next, let us analyze the measurement process in the previous section. We assume that the measurement apparatus M is initially in the "standard state" $m = 0$, which means

$$p^M_{\rm std}(m=0) = 1,\ p^M_{\rm std}(m=1) = 0. \tag{2.29}$$

The Hamiltonian of M, denoted by H^M is assumed to be degenerate. After the measurement, the measurement outcome is stored in M, and the probability distribution is given by Eq. (2.14). The measurement process in this setup can be viewed as a thermodynamic process starting from a nonequilibrium distribution and ending in a canonical distribution. The measurement cost can be quantified by applying the second-law-like inequality (2.12) to the joint system SM, given as

$$\begin{aligned}\left\langle W^M_{\rm ext}\right\rangle_{\rm meas} &\leq \mathcal{F}(p^S \otimes p^M_{\rm std}, H^S + H^M) - \mathcal{F}(p^{SM}, H^S + H^M)\\ &= -\beta^{-1}(S(p^S) + S(p^M) - S(p^{SM})) + \mathcal{F}(p^M_{\rm std}, H^M) - \mathcal{F}(p^M, H^M)\\ &= -\beta^{-1}\left\langle I^{SM}\right\rangle + \mathcal{F}(p^M_{\rm std}, H^M) - \mathcal{F}(p^M, H^M). \end{aligned}\tag{2.30}$$

In (2.30), we use the relation $\left\langle W^S_{\rm ext}\right\rangle_{\rm meas} = 0$ because the state of the system does not change during the measurement process. If we substitute Eqs. (2.14) and (2.29) to (2.30), we obtain

$$\left\langle W^M_{\rm ext}\right\rangle_{\rm meas} \leq 0, \tag{2.31}$$

because $\mathcal{F}(p^M_{\rm std}, H^M) - \mathcal{F}(p^M, H^M) = \beta^{-1}\ln 2$ and $\left\langle I^{SM}\right\rangle = \beta^{-1}\ln 2$ cancel with each other in this setup. Note that for a general setup, the minimum amount of the measurement cost $-\left\langle W^M_{\rm ext}\right\rangle_{\rm meas}$ does not necessarily vanish.

2.4 Information erasure

Here, we consider the cost of resetting the measurement apparatus (information erasure). The initial state of this process is given by Eq. (2.14) and the final state is given by Eq. (2.29). The erasure process in this setup can be viewed as a thermodynamic process starting from a canonical distribution and ending in a nonequilibrium distribution. By directly applying the second-law-like inequality (2.12), the extractable work by the erasure protocol is given by

$$\left\langle W^M_{\rm ext}\right\rangle_{\rm ers} \leq \mathcal{F}(p^M, H^M) - \mathcal{F}(p^M_{\rm std}, H^M). \tag{2.32}$$

If we substitute Eqs. (2.14) and (2.29) to (2.32), we obtain

$$\left\langle W^M_{\rm ext}\right\rangle_{\rm ers} \leq -\beta^{-1}\ln 2, \tag{2.33}$$

From Eqs. (2.9) and (2.33), the minimum amount of heat $-\langle Q\rangle_{\rm ers}$ which needs to be dissipated into the heat bath is given by

$$-\langle Q\rangle_{\rm ers} = \beta^{-1}\ln 2, \tag{2.34}$$

which is known as the Landauer's bound. Equivalently, the minimum amount of entropy of the system which is needed to dump into the heat bath is given by $\ln 2$ (see Fig. 2.3).

2.5 Some properties of the nonequilibrium free energy

In this section, we discuss some properties of the nonequilibrium free energy. In particular, we show that the nonequilibrium free energy gives an information distance between the state of the system and the canonical distribution.

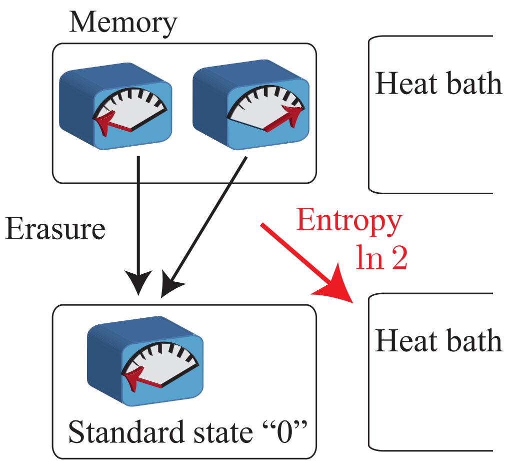

Figure 2.3: Schematic illustration of the erasure process. We start from a maximally mixed state (2.14) which stores the information obtained by the measurement. We erase this information by resetting the memory to the "standard state" (2.29). During this erasure protocol, we need to dump the entropy of $\ln 2$ to the heat bath; equivalently, we dissipate heat of $\beta^{-1}\ln 2$ to the heat bath.

Let us introduce the Kullback-Liebler divergence (relative entropy) which gives the information divergence between two states $p(x)$ and $q(x)$ as

$$D(p||q) := \sum_x p(x) \ln \frac{p(x)}{q(x)}. \tag{2.35}$$

Then, the nonequilibrium free energy (2.10) can be expressed as the equilibrium free energy F^{eq} plus the relative entropy between the probability distribution of the system p and the canonical distribution:

$$\mathcal{F}(p, H) = F^{\text{eq}} + \beta^{-1} D(p||\frac{\mathrm{e}^{-\beta H}}{Z}), \tag{2.36}$$

where $F^{\text{eq}} := -\beta^{-1}\ln Z = -\beta^{-1}\ln\sum_x \exp(-\beta H(x))$ is the equilibrium free energy. Note that we use the relation

$$\begin{aligned}\beta^{-1} D(p||\frac{\mathrm{e}^{-\beta H}}{Z}) &= \beta^{-1}\sum_x p(x)\ln p(x) - \beta^{-1}\sum_x p(x)\ln\frac{\mathrm{e}^{-\beta H(x)}}{Z} \\ &= -\beta^{-1}S(p) - \beta^{-1}\left(-\ln Z - \sum_x p(x)\beta H(x)\right) \\ &= -F^{\text{eq}} + \langle H\rangle_p - \beta^{-1}S(p),\end{aligned} \tag{2.37}$$

in deriving Eq. (2.36). Then, a difference between the nonequilibrium free energy and the equilibrium free energy can be expressed in terms of the relative entropy as

$$\beta^{-1} D(p||\frac{\mathrm{e}^{-\beta H}}{Z}) = \mathcal{F}(p, H) - F^{\text{eq}}. \tag{2.38}$$

In the following, we review some properties of the relative entropy, and use them to show an information-theoretic meaning of the nonequilibrium free energy.

The relative entropy is a measure of the difference between two distributions p and q. It is always nonnegative:

$$D(p||q) \geq 0, \tag{2.39}$$

where the equality holds if and only if $p(x) = q(x)$. We note that the relative entropy is not symmetric in general, i.e., $D(p||q) \neq D(q||p)$. Because of this, the relative entropy is not a true distance. By using (2.39), the nonequilibrium free energy satisfies the following property

$$\mathcal{F}(p, H) \geq F^{\mathrm{eq}}, \tag{2.40}$$

and the equality holds if and only if the state of the system is equal to the canonical distribution, i.e., $p = \mathrm{e}^{-\beta H}/Z$.

As we introduce in detail in Sec. 5.1, the thermal operation gives a general description of the transformation of the state of the system when it interacts with a heat bath[1]. We also show in Sec. 5.2 that the thermal operation transforms the state of the system closer to the canonical distribution[2]. This observation can be expressed in terms of the nonequilibrium free energy as

$$\mathcal{F}(p, H) \geq \mathcal{F}(\mathcal{E}_{\mathrm{thermal}}(p), H), \tag{2.41}$$

where the derivation of (2.41) is given in Sec. 5.7. From (2.40) and (2.41), we find that the nonequilibrium free energy gives an information distance between the state of the system and the canonical distribution. If the state of the system is far from equilibrium, the nonequilibrium free energy takes a large value.

2.6 Quantum case

In this section, we consider a quantum system and review the second law of thermodynamics under information processing [15, 18]. Let us denote the initial density operator of the system S and that of the measurement device M as ρ^S_{ini} and ρ^M_{ini}, respectively. A general quantum measurement is implemented by a joint unitary transformation U^{SM} on SM followed by a projective measurement P^M_k on M, where the projection operator satisfies $P^M_k P^M_l = \delta_{k,l} P^M_k$. Here, we denote the total Hamiltonian of the system as $H^S_{\lambda_0} + H^M + V^{SM}(t)$, where $H^S_{\lambda_0}$ is the initial Hamiltonian of the system which is fixed during the measurement process, H^M is the Hamiltonian of the measurement device, and $V^{SM}(t)$ is the interaction between the S and M. The interaction $V^{SM}(t)$ is turned on during the measurement process. Then, the joint unitary transformation is obtained by the time evolution $U^{SM} = \mathrm{T}\exp(-\mathrm{i}\int \mathrm{d}t' H^S_{\lambda_0} + H^M + V^{SM}(t'))$, where T denotes the time ordering. The probability for the measurement outcome k being obtained is given by

$$p_k := \mathrm{Tr}_{SM}\left[\left(P^M_k \otimes \mathbb{1}^S\right) U^{SM}\left(\rho^S_{\mathrm{ini}} \otimes \rho^M_{\mathrm{ini}}\right) U^{\dagger SM}\left(P^M_k \otimes \mathbb{1}^S\right)\right], \tag{2.42}$$

and the postmeasurement state for the measurement outcome being k is given by

$$\rho^{SM}(k) := \frac{1}{p_k}\left(P^M_k \otimes \mathbb{1}^S\right) U^{SM}\left(\rho^S_{\mathrm{ini}} \otimes \rho^M_{\mathrm{ini}}\right) U^{\dagger SM}\left(P^M_k \otimes \mathbb{1}^S\right). \tag{2.43}$$

Let us introduce an orthonormal basis set of M as $\{|\phi^M_k(b)\rangle\}_{k,b}$. Then, the projection operator can be expressed as

$$P^M_k = \sum_b |\phi^M_k(b)\rangle\langle\phi^M_k(b)|. \tag{2.44}$$

Let us also denote an orthonormal basis set of M as $\{|\psi^M(a)\rangle\}_a$ which diagonalizes ρ^M_{ini} as

$$\rho^M_{\mathrm{ini}} = \sum_a p^M_{\mathrm{ini}}(a)\,|\psi^M(a)\rangle\langle\psi^M(a)|. \tag{2.45}$$

[1] The thermal operation is defined to be a state transformation of a quantum system. However, we only consider quantum states which are diagonal in the energy eigenbasis. Since the off-diagonal elements do not enter into the argument, we can also use the properties of the thermal operation for classical states.

[2] To be precise, this is true for a quantum state which is diagonal in the energy eigenbasis.

By knowing the measurement outcome k, we acquire information about the reduced density matrix of S after the measurement:

$$\begin{aligned}\rho^S(k) &= \frac{1}{p_k}\mathrm{Tr}_M\left[\left(P_k^M\otimes \mathbb{1}^S\right)U^{SM}\left(\rho_{\mathrm{ini}}^S\otimes\rho_{\mathrm{ini}}^M\right)U^{\dagger SM}\left(P_k^M\otimes\mathbb{1}^S\right)\right]\\ &= \frac{1}{p_k}\sum_{a,b}M_{k,a,b}^S\rho_{\mathrm{ini}}^S M_{k,a,b}^{\dagger S},\end{aligned} \tag{2.46}$$

where

$$M_{k,a,b}^S := \left\langle\phi_k^M(b)\right| U^{SM}\left|\psi^M(a)\right\rangle \tag{2.47}$$

is a measurement operator satisfying the normalization condition $\sum_{k,a,b}M_{k,a,b}^{\dagger S}M_{k,a,b}^S = \mathbb{1}^S$.

Next, we change the protocol of the Hamiltonian depending on the measurement outcome k and perform a feedback control on the system. Let us denote the Hamiltonian of the system as $H_{\lambda_t(k)}$ and we also consider an interaction with the heat bath: H^B+V^{SB}. Here, we assume that the interaction between S and B is turned on at the beginning of the feedback control process. The composite system SB evolves in time according to the unitary time evolution $U_k^{SB} = \mathrm{T}\exp(-\mathrm{i}\int \mathrm{d}t'(H_{\lambda_{t'}(k)} + H^B + V^{SB}))$, where T denotes the time ordering. We assume that the initial state of B is given by the canonical distribution

$$\rho_{\mathrm{can}}^B := \frac{\mathrm{e}^{-\beta H^B}}{Z^B}, \tag{2.48}$$

where $Z^B = \mathrm{Tr}[\exp(-\beta H^B)]$ is the partition function. Then, the final state is given by

$$\rho_{\mathrm{fin}}^{SB}(k) := U_k^{SB}(\rho^S(k)\otimes\rho_{\mathrm{can}}^B)U_k^{\dagger SB}. \tag{2.49}$$

Now let us define the extractable work from the system in a manner similar to that of the classical system:

$$\left\langle W_{\mathrm{ext}}^S\right\rangle := \mathrm{Tr}[\rho_{\mathrm{ini}}^S H_{\lambda_0}^S] - \sum_k p_k\mathrm{Tr}[\rho_{\mathrm{fin}}^S(k)H_{\lambda_1}^S(k)] + \langle Q\rangle\,, \tag{2.50}$$

where $H_{\lambda_1}^S(k)$ is the final Hamiltonian of the system and $\langle Q\rangle$ is the heat absorbed by the system, defined as

$$\langle Q\rangle := \mathrm{Tr}[\rho_{\mathrm{can}}^B H^B] - \sum_k p_k\mathrm{Tr}[\rho_{\mathrm{fin}}^B(k)H^B]. \tag{2.51}$$

In Eq. (2.50), we assume that the interaction V^{SB} is either turned off at the end of the feedback control or the strength of the interaction is weak such that the interaction energy V^{SB} is negligible for the final state. The case of the strong coupling is discussed in Ref. [63]. By rewriting the extractable work, we derive the second law of thermodynamics under feedback control. Let us denote the initial Hamiltonian of the system as $H_{\lambda_0}^S$. Then, the initial nonequilibrium free energy of the system is given by

$$\mathcal{F}(\rho_{\mathrm{ini}}^S, H_{\lambda_0}^S) := \mathrm{Tr}[\rho_{\mathrm{ini}}^S H_{\lambda_0}^S] - kTS(\rho_{\mathrm{ini}}^S), \tag{2.52}$$

where $S(\rho) = -\mathrm{Tr}[\rho\ln\rho]$ is the von Neumann entropy. The final nonequilibrium free energy of the system is given by

$$\sum_k p_k\mathcal{F}(\rho^S(k), H_{\lambda_0}^S) := \sum_k p_k\left(\mathrm{Tr}[\rho_{\mathrm{fin}}^S(k)H_{\lambda_1}^S(k)] - kTS(\rho_{\mathrm{fin}}^S(k))\right). \tag{2.53}$$

Let us introduce the information gain defined as

$$\langle I\rangle := S(\rho_{\mathrm{ini}}^S) - \sum_k p_k S(\rho^S(k)), \tag{2.54}$$

which gives a difference in the von Neumann entropy of the system during the measurement process. Then, from Eq. (2.50), we obtain

$$\begin{aligned}
&\left\langle W_{\rm ext}^S\right\rangle + \left\langle \Delta\mathcal{F}^S\right\rangle - kT\left\langle I\right\rangle \\
=\ & kT\sum_k p_k\left(S(\rho(k)) - S(\rho^S_{\rm fin}(k))\right) + {\rm Tr}[\rho^B_{\rm can}H^B] - \sum_k p_k{\rm Tr}[\rho^B_{\rm fin}(k)H^B] \\
=\ & kT\sum_k p_k\left[S\left(\rho^S(k)\otimes\rho^B_{\rm can}\right) + {\rm Tr}[\rho^{SB}_{\rm fin}(k)\ln(\rho^S_{\rm fin}(k)\otimes\rho^B_{\rm can})]\right] \\
=\ & kT\sum_k p_k\left[S(\rho^S(k)\otimes\rho^B_{\rm can}) + {\rm Tr}[\rho^S(k)\otimes\rho^B_{\rm can}\ln\tilde\rho^{SB}_{\rm fin}(k)]\right] \\
=\ & -kT\sum_k p_k S(\rho^S(k)\otimes\rho^B_{\rm can}||\tilde\rho^{SB}_{\rm fin}(k)) \le 0,
\end{aligned} \tag{2.55}$$

where

$$\left\langle \Delta\mathcal{F}^S\right\rangle := \sum_k p_k\mathcal{F}(\rho^S(k), H^S_{\lambda_0}) - \mathcal{F}(\rho^S_{\rm ini}, H^S_{\lambda_0}) \tag{2.56}$$

is the nonequilibrium free-energy difference of the system. We use $S(\rho) = S(U\rho U^\dagger)$ and the nonnegativity of the quantum relative entropy $S(\rho||\sigma) = -S(\rho) - {\rm Tr}[\rho\ln\sigma] \ge 0$ in (2.55). We also introduce the final density matrix of SB for the backward process as

$$\tilde\rho^{SB}_{\rm fin}(k) := U_k^{\dagger SB}(\rho^S_{\rm fin}(k)\otimes\rho^B_{\rm can})U_k^{SB}. \tag{2.57}$$

From (2.55), we obtain the second law of thermodynamics under feedback control [18]:

$$\left\langle W^S_{\rm ext}\right\rangle \le -\left\langle \Delta\mathcal{F}^S\right\rangle + kT\left\langle I\right\rangle. \tag{2.58}$$

If the initial state of M is given by a pure state ($\rho^M_{\rm ini} = \left|\psi^M_0\right\rangle\left\langle\psi^M_0\right|$) and the projection P^M_k is rank-one, i.e., $P^M_k = |\phi_k\rangle\langle\phi_k|$ and $\{|\phi_k\rangle\}_k$ is an orthonormal basis set, the measurement M^S_k is said to be a pure measurement. In this case, the information gain is nonnegative ($\langle I\rangle \ge 0$) and by obtaining the knowledge of the measurement outcome, we can increase the free energy of the system by $kT\langle I\rangle$. Note that this amount of the increased free energy can be extracted as work during the feedback control.

Next, let us define the extractable work from the measurement device during the measurement process. Here, we assume that the interaction between the heat bath and M during the measurement process is turned off. In such a case, the extractable work is given by the internal energy difference:

$$\left\langle W^M_{\rm ext}\right\rangle := {\rm Tr}[\rho^M_{\rm ini}H^M] - \sum_k p_k{\rm Tr}[\rho^M(k)H^M]. \tag{2.59}$$

Let us introduce the initial and the final nonequilibrium free-energies of M and derive a second law of thermodynamics for the measurement process. The initial nonequilibrium free energy of M is given by

$$\mathcal{F}(\rho^M_{\rm ini}, H^M) := {\rm Tr}[\rho^M_{\rm ini}H^M] - kTS(\rho^M_{\rm ini}), \tag{2.60}$$

and the final nonequilibrium free energy of M is given by

$$\mathcal{F}(\rho^M_{\rm meas}, H^M) := {\rm Tr}[\rho^M_{\rm meas}H^M] - kTS(\rho^M_{\rm meas}), \tag{2.61}$$

where $\rho^M_{\rm meas} := \sum_k p_k{\rm Tr}_S[\rho^{SM}(k)]$ is the reduced density matrix of M after the measurement.

Combining Eqs (2.59), (2.60) and (2.61), we obtain

$$
\begin{aligned}
& \left\langle W_{\rm ext}^M \right\rangle + \left\langle \Delta \mathcal{F}^M \right\rangle + kT \left\langle I \right\rangle \\
= \;& kTS(\rho_{\rm ini}^M) + kTS(\rho_{\rm ini}^S) - kTS(\rho_{\rm meas}^M) - \sum_k p_k S(\rho^S(k)) \\
= \;& kTS\left(U^{SM}(\rho_{\rm ini}^S \otimes \rho_{\rm ini}^M) U^{\dagger SM}\right) + kT \sum_k p_k {\rm Tr}_{SM}\left[\rho^{SM}(k)\left(\ln \rho^M(k) + \ln \rho^S(k)\right) + \ln p_k\right] \\
= \;& -kT\Delta S + kTS(\rho_{\rm meas}^{SM}) + kT{\rm Tr}_{SM}\left[\rho_{\rm meas}^{SM} \ln\left(\sum_k p_k \rho^S(k) \otimes \rho^M(k)\right)\right] \\
= \;& -kT\Delta S - kTS\left(\rho_{\rm meas}^{SM} || \sum_k p_k \rho^S(k) \otimes \rho^M(k)\right) \leq 0. \qquad (2.62)
\end{aligned}
$$

Here,

$$\Delta S := S(\rho_{\rm meas}^{SM}) - S\left(U^{SM}(\rho_{\rm ini}^S \otimes \rho_{\rm ini}^M) U^{\dagger SM}\right) \geq 0 \qquad (2.63)$$

gives an increase in the von Neumann entropy of SM via the projection P_k^M [40]. We use (2.63) and the nonnegativity of the quantum relative entropy and obtain (2.62). From (2.62), the second law of thermodynamics which quantifies the extractable work via measurement is obtained [15]:

$$\left\langle W_{\rm ext}^M \right\rangle \leq -\left\langle \Delta \mathcal{F}^M \right\rangle - kT \left\langle I \right\rangle . \qquad (2.64)$$

From (2.58) and (2.64), we find that the information content which enters into the second law when we consider measurement and feedback control is given by the information gain $\langle I \rangle$. By utilizing the obtained information, work can be extracted from the system by the amount of nonequilibrium free-energy difference plus the information gain, whereas the cost $-\left\langle W_{\rm ext}^M \right\rangle$ of performing a measurement to obtain that information is quantified by the nonequilibrium free-energy difference of the measurement device plus the information gain.

2.7 Some comments on experiments

In this section, we review some experiments on information heat engines and information erasure protocols.

The first demonstration of an information heat engine (Maxwell's demon) has been carried out in Ref. [23], by performing a feedback control on a Brownian particle in a tilted periodic potential. The Szilard engine has been experimentally demonstrated by a single-electron box [28] and a colloidal particle using two optical traps [27]. We briefly explain the setup of the experiment done in Ref. [28] in Sec. 6.7.4, and we also discuss a possible experimental test of the main results in this thesis by using a single-electron box. The experimental verification of Landauer's principle has been carried out in Ref. [24] by using a single colloidal particle in a bistable potential (double well potential).

Chapter 3

Nonequilibrium equalities

In this chapter, we review the nonequilibrium equalities which give exact equality relations on thermodynamic quantities, and also give us insights on the irreversibility of the thermodynamic processes even when the dynamics are out of equilibrium. These equality relations offer strong theoretical tools to analyze the nonequilibrium dynamics of the system.

3.1 Detailed fluctuation theorem

The detailed fluctuation theorem relates the ratio of the path probabilities of the forward and backward paths to the total entropy production as

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} = e^{\sigma[\Gamma]}, \tag{3.1}$$

where Γ and $\Gamma^\dagger$ denote the trajectories of the forward and backward (time-reversed) processes, and $P[\Gamma]$ and $\tilde{P}[\Gamma^\dagger]$ are the corresponding path probability distributions. Here,

$$\sigma[\Gamma] := \Delta s[\Gamma] - \beta Q[\Gamma] \tag{3.2}$$

is the total entropy production, where

$$\Delta s[\Gamma] := \ln p_{\rm ini}(x) - \ln p_{\rm fin}(y) \tag{3.3}$$

is a change in the Shannon entropy of the system, $Q[\Gamma]$ is the heat absorbed by the system and β is the inverse temperature of the heat bath. We also denote $p_{\rm ini}(x)$ and $p_{\rm fin}(y)$ as the initial and final probability distributions of the system, respectively. If we assume that the heat bath is very large, the state of the heat bath does not change during the dynamics. Then, the entropy produced by the heat bath is equal to $-\beta Q$. Thus, Eq. (3.2) can be interpreted as the total amount of entropy that is increased during the process.

The detailed fluctuation theorem gives a severe constraint on the path probabilities as shown in Eq. (3.1). The right-hand side of Eq. (3.1) is equal to the exponential of the entropy production, and we find that the probability of observing the negative entropy production during the forward process is exponentially suppressed compared with the corresponding backward probability. We also note that the left-hand side of Eq. (3.1) gives the irreversibility of each stochastic path Γ in terms of the ratio of the forward and backward probability distributions. Because the left-hand side of Eq. (3.1) depends on the protocols (i.e., how to drive the system Hamiltonian from H_{λ_0} to H_{λ_1}), we can use Eq. (3.1) to search for the right protocol which gives the desired form of the entropy production $\sigma[\Gamma]$. We exploit this observation when we derive the main results in the following chapters.

For multiple heat baths, Eq. (3.2) becomes

$$\sigma[\Gamma] = \Delta s[\Gamma] - \sum_i \beta_i Q_i[\Gamma], \tag{3.4}$$

where β_i and $Q_i[\Gamma]$ are the inverse temperature and the heat absorbed by the system for the i-th bath. We note that the detailed fluctuation theorem (3.1) for multiple heat baths has been experimentally verified by using the double quantum dot [46]. For the later purpose, we focus on Eq. (3.2) and review the derivation of the detailed fluctuation theorem (3.1) in this section. We note that we need to specify the microscopic dynamics (e.g. the Langevin equation) in order to find the explicit form of the forward probability distribution $P[\Gamma]$ and the heat $Q[\Gamma]$. In the following subsections, we derive Eq. (3.1) for some typical examples of the microscopic dynamics.

3.1.1 Hamiltonian dynamics

We will first consider the Hamiltonian dynamics for the composite system of the system and the heat bath. Let the initial and final Hamiltonians of the system be $H^S_{\lambda_0}$ and $H^S_{\lambda_1}$, respectively, and the Hamiltonian of the heat bath be H^B. We also denote the interaction as V^{SB}. Let the initial and the final states of the system be $p_{\rm ini}(x)$ and $p_{\rm fin}(y)$, respectively. We assume that the initial state of the heat bath is given by the canonical distribution $p^B_{\rm can}(a) = \exp(-\beta(E^B(a) - F^B))$, where $E^B(a)$ is the eigenenergy of the heat bath and F^B is the equilibrium free energy. We denote $p(y, b|x, a)$ as the transition probability calculated from the Hamiltonian dynamics of the composite system, and

$$P[\Gamma] := p(y, b|x, a)p^S_{\rm ini}(x)p^B_{\rm can}(a) \tag{3.5}$$

gives the forward probability distribution. Since the Hamiltonian dynamics preserves the phase space volume, the transition probability must satisfy the doubly stochastic property

$$\sum_{x,a} p(y, b|x, a) = \sum_{b,y} p(y, b|x, a) = 1. \tag{3.6}$$

The final state of the system is given by

$$p^S_{\rm fin}(y) = \sum_{x,a,b} P[\Gamma]. \tag{3.7}$$

Next, we introduce the backward probability distribution. We start the backward process by a joint distribution

$$p^S_{\rm fin}(y)p^B_{\rm can}(b). \tag{3.8}$$

We denote the transition probability of the backward process as $\tilde{p}(x, a|y, b)$, which is generated by the time-reversal of the forward dynamics. The symmetry in time of the Hamiltonian dynamics implies

$$\tilde{p}(x, a|b, y) = p(b, y|x, a). \tag{3.9}$$

Then, the backward probability distribution is given by

$$\begin{aligned} \tilde{P}[\Gamma^\dagger] &:= p^S_{\rm fin}(y)p^B_{\rm can}(b)\tilde{p}(x, a|y, b) \\ &= p^S_{\rm fin}(y)p^B_{\rm can}(b)p(y, b|x, a). \end{aligned} \tag{3.10}$$

Now the right-hand side of Eq. (3.1) is given by

$$\begin{aligned} \frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} &= \frac{p^S_{\rm ini}(x)p^B_{\rm can}(a)}{p^S_{\rm fin}(y)p^B_{\rm can}(b)} \\ &= \exp(\ln p^S_{\rm ini}(x) - \ln p^S_{\rm fin}(y) - \beta Q[\Gamma]), \end{aligned} \tag{3.11}$$

where we used the relation

$$\frac{p^B_{\rm can}(a)}{p^B_{\rm can}(b)} = \exp(-\beta(E^B(a) - E^B(b))), \tag{3.12}$$

and the definition of the heat absorbed by the system as

$$Q[\Gamma] := E^B(a) - E^B(b). \tag{3.13}$$

Since the right-hand side of Eq. (3.11) is equal to $\exp(\sigma[\Gamma])$, Eq. (3.1) is obtained.

3.1.2 Case of a quantum system

Next, let us consider the case of a quantum system. The derivation is similar to that of the classical system. However, both the underlying dynamics and the explicit form of $p(y, b|x, a)$ are different. Let the initial and final density matrices of the system be

$$\rho^S_{\rm ini} = \sum_x p^S_{\rm ini}(x)|\psi^S_x\rangle\langle\psi^S_x|, \tag{3.14}$$

$$\rho^S_{\rm fin} = \sum_y p^S_{\rm fin}(y)|\phi^S_y\rangle\langle\phi^S_y|. \tag{3.15}$$

We also denote the energy eigenvector of the heat bath as $|\varphi^B_a\rangle$. Then, the transition probability is given by

$$p(y, b|x, a) := \left|\langle\phi^S_y| \otimes \langle\varphi^B_b|U|\psi^S_x\rangle \otimes |\varphi^B_a\rangle\right|^2, \tag{3.16}$$

where U is the unitary operator describing the time evolution calculated by the Schrödinger equation. We can also consider the case in which U is an arbitrary energy-conserving unitary operator. In this case, Eq. (3.16) gives the forward transition probability of the thermal operation (5.1). The backward transition probability is given by using the time-reversal of the time evolution operator $U^\dagger$:

$$\begin{aligned}\tilde{p}(x, a|y, b) &:= \left|\langle\psi^S_x| \otimes \langle\varphi^B_a|U^\dagger|\phi^S_y\rangle \otimes |\varphi^B_b\rangle\right|^2 \\ &= p(y, b|x, a),\end{aligned} \tag{3.17}$$

where the last equality results from

$$\langle\psi^S_x| \otimes \langle\varphi^B_a|U^\dagger|\phi^S_y\rangle \otimes |\varphi^B_b\rangle = \left(\langle\phi^S_y| \otimes \langle\varphi^B_b|U|\psi^S_x\rangle \otimes |\varphi^B_a\rangle\right)^*. \tag{3.18}$$

The forward and the backward probability distributions are given by Eqs. (3.5) and (3.10), respectively. Then, the detailed fluctuation theorem can be derived in a manner to what is done in the classical system (3.11). Here, we note that the average value of the total entropy production takes the form

$$\begin{aligned}\langle\sigma\rangle &= \sum_{x,a,b,y} P[\Gamma]\sigma[\Gamma] \\ &= \sum_x p_{\rm ini}(x)\ln p_{\rm ini}(x) - \sum_y p_{\rm fin}(y)\ln p_{\rm fin}(y) - \beta\sum_a p^B_{\rm can}(a)E^B(a) + \beta\sum_{x,a,b,y} P[\Gamma]E^B(b) \\ &= S(\rho^S_{\rm fin}) - S(\rho^S_{\rm ini}) - \beta\langle Q\rangle,\end{aligned} \tag{3.19}$$

where $S(\rho) = -{\rm Tr}[\rho\ln\rho]$ is the von Neumann entropy and

$$\langle Q\rangle := {\rm Tr}[\rho^B_{\rm can}H^B] - {\rm Tr}[\rho^B_{\rm fin}H^B]. \tag{3.20}$$

Here,

$$\rho^B_{\rm fin} = {\rm Tr}_S[U(\rho^S_{\rm ini} \otimes \rho^B_{\rm can})U^\dagger] \tag{3.21}$$

is the final density matrix of the heat bath.

If we introduce the conditional transition probabilities of the system as

$$P[\Gamma|x] \ := \ p(y,b|x,a)p^B_{\rm can}(a), \tag{3.22}$$
$$\tilde{P}[\Gamma^\dagger|y] \ := \ \tilde{p}(x,a|y,b)p^B_{\rm can}(b), \tag{3.23}$$

the ratio of those probabilities follows the detailed balance relation

$$\frac{P[\Gamma|x]}{\tilde{P}[\Gamma^\dagger|y]} = \exp(-\beta Q[\Gamma]), \tag{3.24}$$

which holds for the thermal operation (5.1). From this relation (3.24), we can formulate the detailed fluctuation theorem (3.1) even when the protocol consists of the thermal operation as a part of the whole protocol.

3.1.3 Stochastic dynamics

Here, let us consider a simple model of a system with a single degrees of freedom (e.g. a colloidal particle) whose trajectory is denoted by $x(t)$. We follow Ref. [53] in this subsection. We consider the following Langevin equation which describes the overdamped motion of a system, given as

$$\dot{x} = \mu F(x,\lambda) + \xi, \tag{3.25}$$

where $F(x,\lambda) = -\partial_x V + f$ is the combination of the force coming from the potential V and the direct force f, depending on the control parameter $\lambda(t)$ and μ is the mobility. Here, $\xi(t)$ is the Gaussian thermal noise satisfying

$$\langle \xi(t)\xi(t')\rangle = 2D\delta(t-t'), \tag{3.26}$$

and we assume the Einstein relation

$$D = \beta^{-1}\mu. \tag{3.27}$$

The conditional probability of observing a trajectory $x(t)$ starting from $x(0) = x_{\rm ini}$ is given by the form

$$p[x(t)|x_{\rm ini}] = \mathcal{N}\exp(-\mathcal{A}[x(t),\lambda(t)]), \tag{3.28}$$

where

$$\mathcal{A}[x(t),\lambda(t)] := \int_0^t {\rm d}\tau \left[\frac{(\dot{x}-\mu F)^2}{4D} + \frac{\mu}{2}\partial_x F\right], \tag{3.29}$$

is a quantity similar to the action.

We define the work and the heat as

$$w[x(t)] \ := \ \int_0^t \left(\frac{\partial V}{\partial \lambda}\dot{\lambda} + f\dot{x}\right){\rm d}\tau, \tag{3.30}$$
$$q[x(t)] \ := \ \int_0^t F\dot{x}{\rm d}\tau. \tag{3.31}$$

Note that the above definition is consistent with the first law of thermodynamics:

$$w[x(t)] \ = \ q[x(t)] + \int_0^t \left(\frac{\partial V}{\partial \lambda}\dot{\lambda} + \frac{\partial V}{\partial x}\dot{x}\right){\rm d}\tau \tag{3.32}$$
$$= \ q[x(t)] + V(x_t,\lambda_t) - V(x_{\rm ini},\lambda_0), \tag{3.33}$$

where $V(x_t,\lambda_t)-V(x_{\rm ini},\lambda_0)$ is the internal energy change of the system. Note that the work done on the system in Eq. (3.30) has two contributions: a change in the potential under fixed position of the particle and an external force applied to the particle. Now the heat can be expressed in terms of the forward and backward transition probabilities. We first rewrite the heat as

$$\begin{aligned}\beta q[x(t)] &= \int_0^t \frac{\mu}{D} F\dot{x}{\rm d}\tau &(3.34)\\ &= -\int_0^t \left[\frac{(\dot{x}-\mu F)^2}{4D}-\frac{(-\dot{x}-\mu F)^2}{4D}\right]{\rm d}\tau. &(3.35)\end{aligned}$$

Next, let us introduce the time-reversed quantities as follows. The time-reversed trajectory is denoted by $\tilde{x}(t)=x(t_{\rm fin}-t)$, which is generated by starting from $x_{\rm fin}=x(t_{\rm fin})$ under the time-reversed protocol $\tilde{\lambda}(t)=\lambda(t_{\rm fin}-t)$. Then, the following relation holds

$$\begin{aligned}&-\int_0^t \left[\frac{(-\dot{x}-\mu F)^2}{4D}+\frac{\mu}{2}\partial_x F\right]{\rm d}\tau\\ &= \int_{t_{\rm fin}}^{t_{\rm fin}-t}\left[\frac{(\dot{x}(t_{\rm fin}-\tau')-\mu F(\lambda(t_{\rm fin}-\tau')))^2}{4D}+\frac{\mu}{2}\partial_x F(\lambda(t_{\rm fin}-\tau'))\right]{\rm d}\tau'\\ &= \int_{t_{\rm fin}}^{t_{\rm fin}-t}\left[\frac{(\dot{\tilde{x}}-\mu F(\tilde{\lambda}(\tau'))^2}{4D}+\frac{\mu}{2}\partial_{\tilde{x}} F(\tilde{\lambda}(\tau'))\right]{\rm d}\tau'\\ &= \mathcal{A}[\tilde{x}(t),\tilde{\lambda}(t)]. &(3.36)\end{aligned}$$

Combining Eqs. (3.36), (3.35) and (3.29), we obtain

$$\begin{aligned}\beta q[x(t)] &= \mathcal{A}[\tilde{x}(t),\tilde{\lambda}(t)]-\mathcal{A}[x(t),\lambda(t)]\\ &= \ln\frac{p[x(t)|x_{\rm ini}]}{\tilde{p}[\tilde{x}(t)|x_{\rm fin}]}, &(3.37)\end{aligned}$$

where we use Eq. (3.28) in deriving the last equality in Eq. (3.37). This expression shows that the ratio of the transition probabilities between the forward dynamics and the backward dynamics is related to the stochastic heat $q[x(t)]$.

The total entropy production in this setup is given by

$$\sigma[x(t)] := \ln\frac{p(x_{\rm ini})}{p(x_{\rm fin})}+\beta q[x(t)], \tag{3.38}$$

where the first term on the right-hand side of Eq. (3.38) is the Shannon entropy difference of the system. Note the sign convention $Q[\Gamma]=-q[x(t)]$ in Eq. (3.38). Then, from the definition of the forward and backward probability distributions

$$\begin{aligned}P[\Gamma] &:= p[x(t)|x_{\rm ini}]p(x_{\rm ini}), &(3.39)\\ \tilde{P}[\Gamma^\dagger] &:= \tilde{p}[\tilde{x}(t)|x_{\rm fin}]p(x_{\rm fin}), &(3.40)\end{aligned}$$

we can derive the detailed fluctuation theorem (3.1) by using Eqs. (3.37) and (3.38) as

$$\begin{aligned}\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} &= \frac{p[x(t)|x_{\rm ini}]}{\tilde{p}[\tilde{x}(t)|x_{\rm fin}]}\frac{p(x_{\rm ini})}{p(x_{\rm fin})}\\ &= \exp(\sigma[x(t)]). &(3.41)\end{aligned}$$

3.2 Nonequilibrium work relations

Next, we relate the total entropy production to the nonequilibrium work and derive general equality relations. We denote the canonical distribution as

$$p^{\rm can}_{\lambda_0}(x) := \exp(-\beta(E_{\lambda_0}(x) - F_{\lambda_0})), \tag{3.42}$$

where $E_{\lambda_0}(x)$ is the initial energy of the system and F_{λ_0} is the initial equilibrium free energy. We also denote the canonical distribution by using the final energy $E_{\lambda_1}(y)$ and the final equilibrium free energy F_{λ_1} as

$$p^{\rm can}_{\lambda_1}(y) := \exp(-\beta(E_{\lambda_1}(y) - F_{\lambda_1})). \tag{3.43}$$

We introduce the following nonequilibrium free energies which generalize the equilibrium free energies into the nonequilibrium situations as

$$\mathcal{F}_{\lambda_0}(x) \;:=\; E_{\lambda_0}(x) - \beta^{-1} s[p_{\rm ini}(x)], \tag{3.44}$$

$$\mathcal{F}_{\lambda_1}(y) \;:=\; E_{\lambda_1}(y) - \beta^{-1} s[p_{\rm fin}(y)], \tag{3.45}$$

where $s[q(x)] = -\ln q(x)$ is the unaveraged Shannon entropy of the system, $p_{\rm ini}(x)$ and $p_{\rm fin}(y)$ are the initial and the final distributions of the system, respectively. Let us also introduce the unaveraged relative entropies

$$\mathcal{D}_{\lambda_0}(x) \;:=\; \ln p_{\rm ini}(x) - \ln p^{\rm can}_{\lambda_0}, \tag{3.46}$$

$$\mathcal{D}_{\lambda_1}(y) \;:=\; \ln p_{\rm fin}(y) - \ln p^{\rm can}_{\lambda_1}, \tag{3.47}$$

where each of them measures the information distance between the state of the system and the corresponding canonical distribution. By using these unaveraged relative entropies, the nonequilibrium free energies can be expressed in terms of the information distances:

$$\mathcal{F}_{\lambda_0}(x) \;:=\; F_{\lambda_0} + \beta^{-1}\mathcal{D}_{\lambda_0}(x), \tag{3.48}$$

$$\mathcal{F}_{\lambda_1}(y) \;:=\; F_{\lambda_1} + \beta^{-1}\mathcal{D}_{\lambda_1}(y). \tag{3.49}$$

The above expressions show that the nonequilibrium free energy is equal to the sum of the equilibrium free energy and the information distance between the state of the system and the canonical distribution. If we take the average of the nonequilibrium free energy, we obtain

$$\begin{aligned}\langle \mathcal{F}_{\lambda_0} \rangle &= \sum_x p_{\rm ini}(x) E_{\lambda_0}(x) - \beta^{-1} S(p_{\rm ini}) \\ &= F_{\lambda_0} + \beta^{-1} D(p_{\rm ini}||p^{\rm can}_{\lambda_0}),\end{aligned} \tag{3.50}$$

where $S(p) = -\sum_x p(x)\ln p(x)$ is the Shannon entropy and $D(p||q) = \sum_x p(x) \ln \frac{p(x)}{q(x)}$ is the relative entropy.

By using the nonequilibrium free energies, we can relate the total entropy production to the extractable work from the system as

$$\sigma[\Gamma] = -\beta W[\Gamma] - \beta \Delta \mathcal{F}(x,y), \tag{3.51}$$

where

$$\Delta\mathcal{F}(x,y) := \mathcal{F}_{\lambda_1}(y) - \mathcal{F}_{\lambda_0}(x) \tag{3.52}$$

is the nonequilibrium free-energy difference and

$$W[\Gamma] := E_{\lambda_0}(x) - E_{\lambda_1}(y) + Q[\Gamma] \tag{3.53}$$

is the extractable work from the system, which is consistent with the first law of thermodynamics. Then, the detailed fluctuation theorem (3.1) can be expressed in terms of the nonequilibrium work

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} = \exp\left[-\beta(W[\Gamma] + \Delta\mathcal{F}(x,y))\right]. \tag{3.54}$$

We note that when the initial and the final distributions are given by the canonical distributions, the nonequilibrium free energies reduce to the equilibrium free energies and Eq. (3.54) reproduces the relation

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} = \exp\left[-\beta(W[\Gamma] + \Delta F)\right]. \tag{3.55}$$

We also note that by introducing the work probability distributions

$$\begin{aligned} p(w) &:= \int d\Gamma \delta(w - W[\Gamma])P[\Gamma], &(3.56)\\ \tilde{p}(-w) &:= \int d\Gamma^\dagger \delta(-w - \tilde{W}[\Gamma^\dagger])\tilde{P}[\Gamma^\dagger], &(3.57)\end{aligned}$$

we can derive the Crook's fluctuation theorem

$$\frac{p(w)}{\tilde{p}(-w)} = \exp(-\beta(w + \Delta F)), \tag{3.58}$$

by using the relation

$$\begin{aligned} p(w) &= \int d\Gamma \delta(w - W[\Gamma])\tilde{P}[\Gamma^\dagger] \exp\left[-\beta(W[\Gamma] + \Delta F)\right] \\ &= \int d\Gamma \delta(w - W[\Gamma])\tilde{P}[\Gamma^\dagger] \exp\left[-\beta(w + \Delta F)\right] \\ &= \tilde{p}(-w) \exp(-\beta(w + \Delta F)). \end{aligned} \tag{3.59}$$

Here, we used the relation $d\Gamma = d\Gamma^\dagger$ and

$$\begin{aligned} \tilde{W}[\Gamma^\dagger] &= E_{\lambda_1}(y) - E_{\lambda_0}(x) - Q[\Gamma] \\ &= -W[\Gamma], \end{aligned} \tag{3.60}$$

in deriving Eq. (3.59). By measuring the work probability distribution of the forward process and the backward process, we can estimate the equilibrium free-energy difference by using Eq. (3.58). The authors in Ref. [44] used this relation and estimated the equilibrium free-energy difference of the RNA molecule by measuring the nonequilibrium work during the folding/unfolding of that molecule.

3.3 Integral fluctuation theorems

Next, we consider the integrated version of the fluctuation theorem. From the detailed fluctuation theorem (3.1), we obtain the integral fluctuation theorem as

$$\begin{aligned} \langle e^{-\sigma} \rangle &= \int d\Gamma P[\Gamma] e^{-\sigma[\Gamma]} \\ &= \int d\Gamma \tilde{P}[\Gamma^\dagger] = 1. \end{aligned} \tag{3.61}$$

Similarly, from Eq. (3.54), we obtain

$$\left\langle e^{\beta(W+\Delta\mathcal{F})} \right\rangle = 1, \tag{3.62}$$

and for the equilibrium initial and final states, Eq. (3.62) reproduces the Jazynski equality

$$\left\langle e^{\beta(W+\Delta F)} \right\rangle = 1. \tag{3.63}$$

However, it should be noted that Eq. (3.63) can be derived without assuming that the final state is given by the canonical distribution as in the original derivation of Jarzynski [50, 51].

These integrated forms of the fluctuation theorems can systematically reproduce the second law of thermodynamics and the fluctuation-dissipation theorem as follows. We use the convexity of the exponential function (the Jensen inequality) given as

$$\left\langle e^{X} \right\rangle \geq e^{\langle X \rangle}, \tag{3.64}$$

where $X[\Gamma]$ is a stochastic quantity. Then, we combine Eq. (3.61) and (3.64) to obtain

$$0 = \ln 1 = \ln \left\langle e^{-\sigma} \right\rangle \geq \ln e^{-\langle \sigma \rangle} = - \langle \sigma \rangle \,, \tag{3.65}$$

which is equivalent to the second-law-like inequality

$$\langle \sigma \rangle \geq 0. \tag{3.66}$$

This relation shows that the average total entropy production is nonnegative for all thermodynamic processes. Equivalently, we can obtain the inequality in which the extractable work is bounded from above as

$$\langle W \rangle \leq - \langle \Delta\mathcal{F} \rangle \,. \tag{3.67}$$

Note that if the initial and the final states are given by the canonical distributions, Eq. (3.67) reproduces the conventional second law of thermodynamics

$$\langle W \rangle \leq -\Delta F. \tag{3.68}$$

Although the inequalities (3.66) and (3.68) are equivalent to the second law of thermodynamics, they do not give the microscopic derivation of the irreversibility of the macroscopic thermodynamics because in deriving those inequalities (3.66) and (3.68), we put the irreversibility by hand in either the microscopic dynamics or the initial state of the backward process. Still, these inequalities (3.66) and (3.68) are useful to analyze the irreversibility of the thermodynamic processes. For example, the thermodynamically reversible protocols can be found by studying the equality conditions of (3.66) and (3.68). The equality condition in (3.68) is equivalent to the condition $W[\Gamma] = -\Delta F$ for all Γ and from Eq. (3.63) we obtain

$$P[\Gamma] = \tilde{P}[\Gamma^{\dagger}]. \tag{3.69}$$

This is possible if we quasi-statically change the control parameter λ_t such that the state of the system is always at equilibrium during the driving. In this situation, the state of the system at both the forward and the backward processes take the same canonical distribution and Eq. (3.69) is satisfied. Then, the upper bound given by the second law (3.68) is achieved and we can extract a maximal amount of work (i.e., the reversible work $-\Delta F$) from the system, which is consistent with the macroscopic thermodynamics. Next, we would like to consider the equality condition of (3.67) or equivalently, that of (3.66). The equality condition is the same as Eq. (3.69):

$$P[\Gamma] = P[\Gamma|x]p_{\mathrm{ini}}(x) = \tilde{P}[\gamma^{\dagger}|y]p_{\mathrm{fin}}(y) = \tilde{P}[\Gamma^{\dagger}]. \tag{3.70}$$

However, we should note that in this case the initial and final states are different from the canonical distributions. Thus, if we quasi-statically change the control parameter λ_t from λ_0 to λ_1, the nonequilibrium distribution $p_{\rm ini}(x)$ changes into the canonical distribution due to the interaction between the system and the heat bath. This thermalization produces dissipation and the entropy production cannot reach the equality condition of (3.66). To avoid this situation, we quench the Hamiltonian at the initial and the final stage of the protocol. Let us introduce the quenched Hamiltonians H'_0 and H'_1 which satisfy the following relations:

$$\begin{aligned} p_{\rm ini} &= \exp(-\beta(H'_0 - F'_0)), \\ p_{\rm fin} &= \exp(-\beta(H'_1 - F'_1)), \end{aligned} \tag{3.71}$$

where F'_0 and F'_1 are the free energies. If we quench the Hamiltonian from H_{λ_0} to H'_0, quasi-statically change the Hamiltonian to H'_1 and quench the Hamiltonian back to the final one H_{λ_1}, the relation (3.69) is satisfied and the equality conditions in (3.67) and (3.66) are achieved. Now, let us focus on the work and the entropy production for each step which achieve the equality conditions in (3.67) and (3.66). The work extracted from the system during each quench process is given by

$$\langle W_0^{\rm q} \rangle = \sum_x p_{\rm ini}(x)(E_{\lambda_0}(x) - E'_0(x)), \tag{3.72}$$

$$\langle W_1^{\rm q} \rangle = \sum_y p_{\rm fin}(y)(E'_1(y) - E_{\lambda_1}(y)), \tag{3.73}$$

and the work extracted during the quasi-static process is given by the equilibrium free-energy difference

$$\langle W^{\text{q-s}} \rangle = F'_0 - F'_1. \tag{3.74}$$

The total extractable work is then given by

$$\begin{aligned} \langle W_0^{\rm q} \rangle + \langle W^{\text{q-s}} \rangle + \langle W_1^{\rm q} \rangle &= \beta^{-1} \sum_x p_{\rm ini}(x) \ln \frac{p_{\rm ini}(x)}{p_{\lambda_0}^{\rm can}(x)} - \beta^{-1} \sum_y p_{\rm fin}(y) \ln \frac{p_{\rm fin}(y)}{p_{\lambda_1}^{\rm can}(y)} - \Delta F \\ &= -\Delta F + \beta^{-1} D(p_{\rm ini} || p_{\lambda_0}^{\rm can}) - \beta^{-1} D(p_{\rm fin} || p_{\lambda_1}^{\rm can}) \\ &= - \langle \Delta \mathcal{F} \rangle, \end{aligned} \tag{3.75}$$

which achieves the equality condition of the second law (3.67). Because the quench process and the quasi-static process are both thermodynamically reversible, the entropy production vanishes during those processes and

$$\langle \sigma \rangle = 0 \tag{3.76}$$

also holds. We note that the nonequilibrium free energy (3.45) acquires a thermodynamic meaning when we have a single heat bath. In this case, the inequality (3.67) shows that the nonequilibrium free energy is equal to the reversible work (i.e., the extractable work for a thermodynamically reversible process) for nonequilibrium initial and/or final states.

Next, let us expand the integral fluctuation theorem (3.61) for small $\sigma[\Gamma]$ as

$$\begin{aligned} 0 &= \ln \left\langle {\rm e}^{-\sigma} \right\rangle = \ln \left(1 - \langle \sigma \rangle + \frac{1}{2} \left\langle \sigma^2 \right\rangle \right) + O(\sigma^3) \\ &= - \langle \sigma \rangle + \frac{1}{2} \left(\left\langle \sigma^2 \right\rangle - \langle \sigma \rangle^2 \right) + O(\sigma^3). \end{aligned} \tag{3.77}$$

We then obtain the fluctuation-dissipation theorem:

$$\left\langle \sigma^2 \right\rangle - \langle \sigma \rangle^2 = 2 \langle \sigma \rangle + O(\sigma^3). \tag{3.78}$$

In particular, for thermalized initial and final states, Eq. (3.78) reduces to

$$\langle W^2 \rangle - \langle W \rangle^2 = 2 \langle \sigma \rangle + O(\sigma^3), \tag{3.79}$$

where the left-hand side of Eq. (3.79) is equal to the work fluctuation and the right-hand side is equal to the dissipation. These relations (3.78) and (3.79) are applicable for small σ, which is equivalent to the condition that the deviation of the process from the thermodynamically reversible process is small. Starting from a equilibrium state, the vanishing work fluctuation protocol and the vanishing dissipation protocol are the same. However, this is not true for nonequilibrium initial and final states because the work generally fluctuates during the quench process (3.69), which is needed to construct a thermodynamically reversible protocol with vanishing dissipation.

3.4 Fluctuation theorems under information processing

In this section, we review how the fluctuation theorems can be generalized to the situations involving information processing. We introduce a measurement device M which records the measurement outcome. Let the joint distribution of the system S and the measurement device be $p^{SM}(x,a)$, where the reduced distribution $p^S(x) = \sum_a p^{SM}(x,a)$ is equal to the premeasurement of S and $p^M(a) = \sum_x p^{SM}(x,a)$ is the probability of obtaining the measurement outcome a. The information acquired by this measurement process is quantified by the classical mutual information

$$I^{SM}(x,a) := \ln \frac{p^{SM}(x,a)}{p^S(x)p^M(a)}. \tag{3.80}$$

The total entropy production of the composite system SM is given by

$$\sigma[\Gamma] := \Delta s^{SM}[\Gamma] - \beta Q[\Gamma], \tag{3.81}$$

where

$$\Delta s^{SM}[\Gamma] := \ln \frac{p^{SM}(x,a)}{p^{SM}_{\text{fin}}(y,a)} \tag{3.82}$$

is the Shannon entropy difference and $p^{SM}_{\text{fin}}(y,a)$ is the joint distribution of SM after the feedback control, which is implemented by changing the control parameter $\lambda_t(a)$ depending on the measurement outcome a. By using the classical mutual information (3.80), we can rewrite Eq. (3.81) as

$$\sigma[\Gamma] = \Delta s[\Gamma] - \beta Q[\Gamma] + I^{SM}(x,a) - I^{SM}_{\text{fin}}(y,a), \tag{3.83}$$

where $\Delta s[\Gamma] := \ln p^S(x) - \ln p^S_{\text{fin}}(y)$ is the Shannon entropy difference of the system and

$$I^{SM}_{\text{fin}}(y,a) := \ln \frac{p^{SM}_{\text{fin}}(y,a)}{p^S_{\text{fin}}(y)p^M(a)} \tag{3.84}$$

quantifies the correlation which is left between S and M when the feedback is completed. Note that if we consider an optimal feedback control protocol, Eq. (3.84) vanishes and the entire correlation (3.80) is utilized to extract work from the system. We apply the detailed fluctuation theorem to the joint system SM as

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} = \exp(\sigma[\Gamma]) = \exp\left(\Delta s[\Gamma] - \beta Q[\Gamma] + I^{SM}(x,a) - I^{SM}_{\text{fin}}(y,a)\right), \tag{3.85}$$

and the integral fluctuation theorem as

$$\left\langle e^{-\sigma} \right\rangle = \left\langle e^{-\Delta s + \beta Q - I^{SM} + I^{SM}_{\text{fin}}} \right\rangle = 1. \tag{3.86}$$

The relation between the total entropy production and the Shannon entropy difference of the system (3.83) allows us to express the nonequilibrium equalities of the feedback-controlled system by introducing the information contents as in Eqs. (3.85) and (3.86). If we introduce the postmeasurement state of the system conditioned on the measurement outcome a as

$$p^{S|M}(x|a) := \frac{p^{SM}(x,a)}{p^{M}(a)}, \tag{3.87}$$

the forward probability distribution can be expressed as

$$P[\Gamma] = P[\Gamma|x,a]p^{SM}(x,a) = p^{M}(a)P[\Gamma|x,a]p^{S|M}(x|a). \tag{3.88}$$

Then, Eq. (3.86) can also be expressed as the ensemble average of the system when the measurement outcome is fixed as

$$1 = \left\langle \mathrm{e}^{-\sigma} \right\rangle = \sum_{a} p^{M}(a) \left\langle \mathrm{e}^{-\Delta s + \beta Q - I^{SM} + I^{SM}_{\mathrm{fin}}} \right\rangle_{a}, \tag{3.89}$$

with

$$\langle X \rangle_{a} := \int \mathrm{d}\Gamma P[\Gamma|x,a]p^{S|M}(x|a)X[\Gamma]. \tag{3.90}$$

By introducing the nonequilibrium free energy (3.38) and use the relation

$$\Delta s[\Gamma] - \beta Q[\Gamma] = -\beta(W[\Gamma] + \Delta\mathcal{F}(x,y)), \tag{3.91}$$

the integral fluctuation theorem (3.86) can be expressed as the generalized form of the Jarzynski equality as

$$\left\langle \mathrm{e}^{\beta(W+\Delta\mathcal{F}) - I^{SM} + I^{SM}_{\mathrm{fin}}} \right\rangle = 1. \tag{3.92}$$

Similarly to Eq. (3.61), we can derive the generalized second law of thermodynamics including the effect of the measurement and the feedback control:

$$\langle W \rangle \leq -\langle \Delta\mathcal{F} \rangle + \beta^{-1}(\langle I^{SM} \rangle - \langle I^{SM}_{\mathrm{fin}} \rangle) \tag{3.93}$$

where

$$\langle I^{SM} \rangle := \sum_{x,a} p^{SM}(x,a) \ln \frac{p^{SM}(x,a)}{p^{S}(x)p^{M}(a)}, \tag{3.94}$$

$$\langle I^{SM}_{\mathrm{fin}} \rangle := \sum_{y,a} p^{SM}_{\mathrm{fin}}(y,a) \ln \frac{p^{SM}_{\mathrm{fin}}(y,a)}{p^{S}_{\mathrm{fin}}(y)p^{M}(a)}, \tag{3.95}$$

are the classical mutual information measuring the initial and final correlations between S and M.

Chapter 4

Quantum fluctuation theorems under measurement and feedback control

In this chapter, we derive fluctuation theorems for a quantum system with measurement and feedback control, which is the first main result of this thesis [72,73]. The obtained equality relations offer a strong theoretical tool to analyze the nonequilibrium dynamics of the general quantum measurement process as well as the feedback control process. In particular, they can be used to search for measurement and feedback protocols with low dissipation. We expect that the obtained quantum Jarzynski equalities can be experimentally tested, and also can be utilized to extract useful information about the thermodynamic quantities in the quantum regime.

4.1 Motivation and the setup of the main results

There have been recent developments in quantum systems which can be used to possess information processing, as in trapped ions [78], superconducting qubits [79] and ultracold atomic gases [80]. Performing measurement and feedback control on those systems have attracted great interest from the viewpoint of state preparation, entanglement generation and quantum computing. Also, in the future, these techniques may be used to realize quantum information heat engines which extract work from the system with the help of feedback control. Here, we would like to analyze thermodynamics of the quantum measurement and feedback control processes, which generally go out of equilibrium. As we reviewed in Sec. 2.6, there are previous studies that analyze the extractable work during the feedback control and the measurement cost by formulating the second law of thermodynamics [15,16]. On the other hand, the fluctuation theorems allow us to consider not only the average value of the thermodynamic quantities but also the unaveraged version of those quantities, which depend on the microscopic trajectory of the system. The detailed fluctuation theorem relates the unaveraged version of the thermodynamic quantities to the forward and backward probability distributions which realize the microscopic trajectory of the measurement and the feedback control processes. The detailed fluctuation theorems provides detailed pieces of information of the quantum measurement and feedback control processes and can be used to find protocols which suppress dissipation during those processes.

Following the above motivation, we would like to derive detailed fluctuation theorems and Jarzynski equalities which are applicable to quantum measurement and feedback control processes. Previous studies have generalized the fluctuation theorems to the case of measurement and feedback control, and also to the general information processing processes in classical systems [14]. As we discussed in Sec. 2.6, the effect of information exchange between the system and the measurement device is quantified by the classical mutual information and this information content has been taken into account in the obtained fluctuation theorems. A generalization of the Jarzynski equality

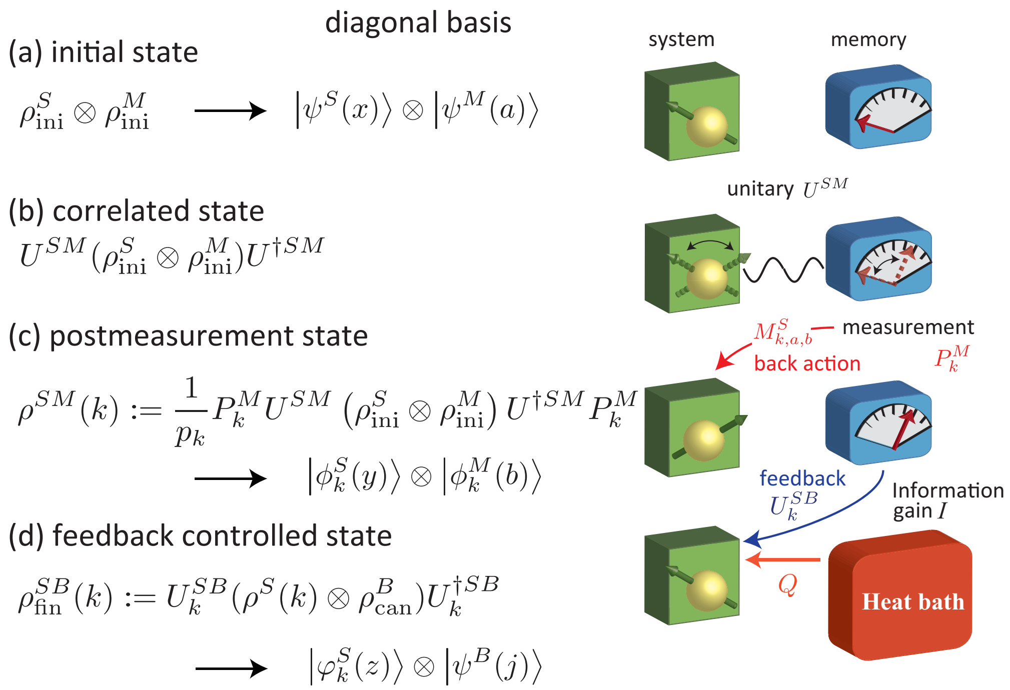

Figure 4.1: Schematic illustration of the quantum measurement and feedback control processes. We show how the density matrices of the composite system change during the measurement and the feedback control. The basis sets which diagonalize the density matrices of the initial state, the postmeasurement state and the feedback controlled state are also presented.

to a feedback controlled system in the quantum regime has been carried out in Ref. [70]. The authors of Ref. [70] have considered a measurement process as a projection measurement on the system followed by a process with classical error on that measurement outcome. Here, the classical mutual information between the two outcomes (i.e., the measurement outcome of the projection and the outcome with classical error) appears in the obtained Jarzynski equality. Apart from the previous studies, we would like to consider detailed fluctuation theorems under general quantum measurement and feedback control by using the general framework of the operational quantum measurement theory [40]. As we discussed in Sec. 2.6, we use the indirect measurement model and perform a general quantum measurement on the system. Below, we summarize the protocol used in Sec. 2.6 which is also the setup we consider when we derive the fluctuation theorems.

We consider a joint system composed of the system S and the memory M. The following process (a)-(c) describes a quantum measurement process which we consider in this chapter. After the measurement, we attach a heat bath B to the system and consider a feedback control which is described by (d). See also Fig. 4.1.

(a) Initial state:

$$\rho^S_{\rm ini} \otimes \rho^M_{\rm ini}. \tag{4.1}$$

The Hamiltonian of the composite system during the measurement process is given by $H^S_{\lambda_0} + H^M + V^{SM}(t)$, where the interaction $V^{SM}(t)$ vanishes at the initial stage and the final stage of the measurement process.

(b) Correlated state:

$$U^{SM}(\rho^S_{\rm ini} \otimes \rho^M_{\rm ini})U^{\dagger SM}, \tag{4.2}$$

where the time evolution operator $U^{SM} = \mathrm{T}\exp(-\mathrm{i}\int \mathrm{d}t' H^S_{\lambda_0} + H^M + V^{SM}(t'))$ generates a correlation between S and M.

(c) Postmeasurement state:

$$\rho^{SM}(k) := \frac{1}{p_k}\left(P^M_k \otimes \mathbb{1}^S\right) U^{SM}\left(\rho^S_{\mathrm{ini}} \otimes \rho^M_{\mathrm{ini}}\right) U^{\dagger SM}\left(P^M_k \otimes \mathbb{1}^S\right), \tag{4.3}$$

where $P^M_k = \sum_b \left|\phi^M_k(b)\right\rangle\left\langle\phi^M_k(b)\right|$ is a projection operator and $\{\left|\phi^M_k(b)\right\rangle\}_{k,b}$ is an orthonormal basis set of M. The postmeasurement state of the system conditioned on the measurement outcome k is given by

$$\rho^S(k) = \frac{1}{p_k}\sum_{a,b} M^S_{k,a,b}\rho^S_{\mathrm{ini}} M^{\dagger S}_{k,a,b}, \tag{4.4}$$

where $M^S_{k,a,b} := \left\langle\phi^M_k(b)\right| U^{SM}\left|\psi^M(a)\right\rangle$ is a measurement operator which acts on the system satisfying the normalization condition $\sum_{k,a,b} M^{\dagger S}_{k,a,b} M^S_{k,a,b} = \mathbb{1}^S$. Here, $\{\left|\psi^M(a)\right\rangle\}_a$ is a basis set which diagonalizes the initial density matrix of M: $\rho^M_{\mathrm{ini}} = \sum_a p^M_{\mathrm{ini}}\left|\psi^M(a)\right\rangle\left\langle\psi^M(a)\right|$.

(d) Feedback controlled state:

$$\rho^{SB}_{\mathrm{fin}}(k) := U^{SB}_k(\rho^S(k) \otimes \rho^B_{\mathrm{can}})U^{\dagger SB}_k. \tag{4.5}$$

At the beginning of the feedback control, we attach the heat bath B to S and consider a density matrix $\rho^S(k) \otimes \rho^B_{\mathrm{can}}$. We change the Hamiltonian of the system depending on the measurement outcome k through the control parameter $\lambda_t(k)$. The total Hamiltonian during the feedback control is defined as $H_{\lambda_t(k)} + H^B + V^{SB}$. Then, the time evolution operator $U^{SB}_k = \mathrm{T}\exp(-\mathrm{i}\int \mathrm{d}t'(H_{\lambda_{t'}(k)} + H^B + V^{SB}))$ acts as a feedback control on the system.

Using the setup of the quantum measurement and the feedback control described above, we derive two main results in this chapter. The first one is the detailed fluctuation theorem for the quantum measurement process (4.47):

$$\frac{\tilde{p}^{\mathrm{meas}}(x,a,k,y,b)}{p^{\mathrm{meas}}(x,a,k,y,b)} = \mathrm{e}^{\beta(W^M_{\mathrm{ext}}(a,k,b)+\Delta\mathcal{F}^M(a,k,b))+I(x,k,y)}, \tag{4.6}$$

and the second one is the detailed fluctuation theorem for the feedback controlled process (4.33):

$$\frac{\tilde{p}^{\mathrm{fb}}(h,k,y,j,z)}{p^{\mathrm{fb}}(h,k,y,j,z)} = \mathrm{e}^{\beta(W^S_{\mathrm{ext}}(x,h,k,j,z)+\Delta\mathcal{F}^S(x,k,z))-I(x,k,y)}. \tag{4.7}$$

Both detailed fluctuation theorems relate the ratio of the forward and backward probabilities of the process to the unaveraged values of the extractable work, nonequilibrium free-energy difference and the information content (the information gain $I(x,k,y)$ defined in Eq. (4.22)). We also use this result to derive Jarzynski equalities and also discuss a method to reduce dissipation during measurement and feedback protocol processes. When we derive Eqs. (4.47) and (4.33), we assume that the initial and the final density matrices of S and M are diagonalized by their corresponding energy eigenbases. For the initial and final states which have quantum coherence in the energy eigenbases, we can derive the detailed fluctuation theorems (4.27) and (4.42) instead of Eqs. (4.6) and (4.7) by using the dissipated work σ^{SB} and σ^M.

Note that the detailed fluctuation theorems depend on the microscopic trajectories of the system as in Eqs. (4.6) and (4.7). We define those trajectories by introducing the orthonormal basis sets which diagonalize the density matrices of the initial state, the postmeasurement state, and the feedback controlled state (See also Fig. 4.1). Then, we define microscopic trajectories as a transition between those bases via the measurement process and the feedback control process and also define probability distributions which realize those trajectories.

Finally, we would like to discuss some possible generalization of the measurement and feedback control that we consider in this chapter. The type of measurement and feedback control that we use are called the measurement-based feedback control in which we apply the projection P_k^M on M followed by a feedback control U_k^{SB} on SB. We can also consider the coherent feedback control in which we apply the joint unitary transformation $U_{\rm fb}^{SBM} = \sum_k U_k^{SB} \otimes P_k^M$ on the composite system. Note that in this way, we do not discard the information concerning the coherence between S and M which has some advantage as discussed in Ref. [81]. At this moment, there is a technical difficulty to treat this situation because the joint density matrix of SM and its marginal density matrices of S and M do not commute with each other in general and thus a consistent definition of the microscopic path is hard to introduce. We also note that the feedback control is implemented by a unitary transformation on SB. We can also consider different operations for the feedback control process such as the CPTP maps. However, we note that under the state change of the system by applying the CPTP map, the amount of work that can be extracted from the system may be larger than the free-energy difference and this process may apparently violate the second law of thermodynamics. Here, we would like to consider the additional amount of work that can be extracted from the system with the help of the feedback control. Thus, we implement a feedback control as a unitary transformation since we can fairly compare the extractable work via this type of feedback control with the case of the extractable work without feedback control. If we consider a unital map instead of the unitary time evolution, we can derive a quantum fluctuation theorem as discussed in Ref. [82] and a similar relation can also be obtained. In Ref. [15], a quantum second law under feedback control was derived by coupling the system and the heat bath before the measurement process. A generalization of the main result to this kind of situation may also be interesting.

4.2 Detailed fluctuation theorems

4.2.1 Detailed fluctuation theorems for the feedback control process

As we derive quantum fluctuation theorems without feedback control in Sec. (3.1.2), we also introduce orthonormal basis sets $\{|\psi^S(x)\rangle\}_x$ and $\{|\varphi_k^S(z)\rangle\}_z$ which diagonalize the initial and the final density matrices:

$$\rho_{\rm ini}^S = \sum_x p_{\rm ini}^S(x) \left|\psi^S(x)\right\rangle \left\langle\psi^S(x)\right|, \tag{4.8}$$

$$\rho_{\rm fin}^S(k) = \sum_z p_{\rm fin}^S(z|k) \left|\varphi_k^S(z)\right\rangle \left\langle\varphi_k^S(z)\right|, \tag{4.9}$$

where $\rho_{\rm fin}^S(k) := {\rm Tr}_B[\rho_{\rm fin}^{SB}(k)]$ and $\rho_{\rm fin}^{SB}(k)$ is given in Eq. (4.5). Let us also introduce a set of energy eigenbasis $\{|\psi^B(h)\rangle\}_h$ which diagonalizes the initial density matrix of B:

$$\rho_{\rm can}^B = \sum_h p_{\rm can}^B(h) \left|\psi^B(h)\right\rangle \left\langle\psi^B(h)\right|. \tag{4.10}$$

Next, let us define define a quantity

$$\left\langle\sigma^{SB}\right\rangle := -\beta(\left\langle W_{\rm ext}^S\right\rangle + \left\langle\Delta\mathcal{F}^S\right\rangle) = \sum_k p_k S(\rho_{\rm fin}^S(k)) - S(\rho_{\rm ini}^S) - \beta\left\langle Q\right\rangle, \tag{4.11}$$

where $\langle W_{\rm ext}^S\rangle$ and $\langle\Delta\mathcal{F}^S\rangle$ are given in Eqs. (2.50) and (2.56), respectively. Note that $\langle\sigma^{SB}\rangle$ quantifies the entropy produced by SB (or the dissipated work of the system normalized by β), and the combination of $\langle \sigma^{SB} \rangle + \langle I \rangle \geq 0$ quantifies the irreversiblity of the system during the feedback control as can be seen from the second law of thermodynamics (2.58). Using the relation

$$\langle \sigma^{SB} \rangle = -S(\rho^S_{\rm ini}) - S(\rho^B_{\rm can}) - \sum_k p_k {\rm Tr}\left[\rho^{SB}_{\rm fin}(k) \ln\left(\rho^S_{\rm fin}(k) \otimes \rho^B_{\rm can}\right)\right], \tag{4.12}$$

we can define a quantity σ^{SB} from the following argument:

$$\begin{aligned} \langle \sigma^{SB} \rangle &= \sum_x p^S_{\rm ini}(x) \ln p^S_{\rm ini}(x) + \sum_h p^B_{\rm can}(h) \ln p^B_{\rm can}(h) \\ &\quad + \sum_k p_k \langle \varphi^S_k(z)| \otimes \langle \psi^B(j)| \rho^{SM}_{\rm fin}(k) |\varphi^S_k(z)\rangle \otimes |\psi^B(j)\rangle \ln[p^S_{\rm fin}(z|k) p^B_{\rm can}(j)] \\ &= \sum_{x,a,h,k,y,b,j,z} p(x,a,h,k,y,b,j,z) \ln \frac{p^S_{\rm ini}(x) p^B_{\rm can}(h)}{p^S_{\rm fin}(z|k) p^B_{\rm can}(j)}. \end{aligned} \tag{4.13}$$

In Eq. (4.13), we introduce the forward probability distribution of the total system:

$$p(x,a,h,k,y,b,j,z) := p(z,j|k,y,h) p(k,y,b|x,a) p^S_{\rm ini}(x) p^M_{\rm ini}(a) p^B_{\rm can}(h), \tag{4.14}$$

where

$$\begin{aligned} p(k,y,b|x,a) &:= \left|\langle \phi^S_k(y)| \otimes \langle \phi^M_k(b)| U^{SM} |\psi^S(x)\rangle \otimes |\psi^M(a)\rangle\right|^2 \\ &= \frac{1}{p^M_{\rm ini}(a)} \left|\langle \phi^S_k(y)| M^S_{k,a,b} |\psi^S(x)\rangle\right|^2 \end{aligned} \tag{4.15}$$

is the transition probability from the state labeled by x, a to the state labeled by k, y, b during the measurement process, and

$$p(z,j|k,y,h) := \left|\langle \varphi^S_k(z)| \otimes \langle \psi^B(j)| U^{SB}_k |\phi^S_k(y)\rangle \otimes |\psi^B(h)\rangle\right|^2 \tag{4.16}$$

is the transition probability from the state labeled by k, y, h to the state labeled by z, j during the feedback process. In Eq. (4.15), we use the orthonormal basis set $|\phi^S_k(y)\rangle$ which diagonalizes the postmeasurement state of the system

$$\rho^S(k) = \sum_y p^S(k|y) |\phi^S_k(y)\rangle \langle \phi^S_k(y)|. \tag{4.17}$$

We also note that in deriving the last equality of Eq. (4.13), we used the relation

$$\begin{aligned} &\sum_k p_k \langle \varphi^S_k(z)| \otimes \langle \psi^B(j)| \rho^{SM}_{\rm fin}(k) |\varphi^S_k(z)\rangle \otimes |\psi^B(j)\rangle \\ &= \sum_{h,k,y,j,z} p(z,j|k,y,h) p_k p^B_{\rm can}(h) p^S(k|y) \\ &= \sum_{x,a,h,k,y,b,j,z} p(z,j|k,y,h) p(k,y,b|x,a) p^B_{\rm ini}(x) p^M_{\rm ini}(a) p^B_{\rm can}(h), \end{aligned} \tag{4.18}$$

by noting that

$$\begin{aligned} \sum_{x,a,b} p(k,y,b|x,a) p^M_{\rm ini}(a) p^B_{\rm can}(h) &= \sum_{x,a,b} {\rm Tr}\left[\langle \phi^S_k(y)| M^S_{k,a,b} \rho^S_{\rm ini} M^{\dagger S}_{k,a,b} |\phi^S_k(y)\rangle\right] \\ &= p_k p^S(k|y). \end{aligned} \tag{4.19}$$

From Eq. (4.13), we define the stochastic quantity σ^{SB} as

$$\sigma^{SB}(x,h,j,z) := \ln \frac{p^S_{\rm ini}(x)p^B_{\rm can}(h)}{p^S_{\rm fin}(z|k)p^B_{\rm can}(j)} = \ln p^S_{\rm ini}(x) - \ln p^S_{\rm fin}(z|k) - \beta Q(h,j), \tag{4.20}$$

where $Q(h,j) := E^B(h) - E^B(j)$ is the heat absorbed by the system.

From Eq (4.17), the information gain can be rewritten as

$$\begin{aligned} \langle I \rangle &= -\sum_x p^S_{\rm ini}(x) \ln p^S_{\rm ini}(x) + \sum_{k,y} p_k p^S(y|k) \ln p^S(y|k) \\ &= \sum_{x,a,h,k,y,b,j,z} p(x,a,h,k,y,b,j,z) \ln \frac{p^S(y|k)}{p^S_{\rm ini}(x)}, \end{aligned} \tag{4.21}$$

and we define a stochastic quantity I as

$$I(x,k,y) := \ln \frac{p^S(y|k)}{p^S_{\rm ini}(x)}. \tag{4.22}$$

Next, let us consider the combination of $\sigma^{SB} + I$, which measures the irreversibility of the feedback control process as follows:

$$\begin{aligned} \left\langle \sigma^{SB} \right\rangle + \langle I \rangle &= \sum_{h,k,y,j,z} p^{\rm fb}(h,k,y,j,z) \ln \frac{p^S(y|k)^B_{\rm can}(h)}{p^S_{\rm fin}(z|k)p^B_{\rm can}(j)} \\ &= \sum_{h,k,y,j,z} p^{\rm fb}(h,k,y,j,z) \ln \frac{p^{\rm fb}(h,k,y,j,z)}{\tilde{p}^{\rm fb}(h,k,y,j,z)} \geq 0, \end{aligned} \tag{4.23}$$

where we introduce the forward and the backward probability distributions corresponding to the feedback control process:

$$p^{\rm fb}(h,k,y,j,z) := p(z,j|k,y,h)p_k p^S(y|k)^B_{\rm can}(h), \tag{4.24}$$
$$\tilde{p}^{\rm fb}(h,k,y,j,z) := \tilde{p}(k,y,h|z,j)p_k p^S_{\rm fin}(z|k)p^B_{\rm can}(j), \tag{4.25}$$

with

$$\begin{aligned} \tilde{p}(k,y,h|z,j) &:= \left| \left\langle \phi^S_k(y) \right| \otimes \left\langle \psi^B(h) \right| U^{\dagger SB}_k \left| \varphi^S_k(z) \right\rangle \otimes \left| \psi^B(j) \right\rangle \right|^2 \\ &= p(z,j|k,y,h). \end{aligned} \tag{4.26}$$

Note that Eq. (4.24) gives a forward probability distribution starting from the postmeasurement state of the system and the canonical distribution of the heat bath ($\rho^S(k) \otimes \rho^B_{\rm can}$), and Eq. (4.25) gives a backward probability distribution starting from the final state of the system and the canonical distribution of the heat bath ($\rho^S_{\rm fin}(k) \otimes \rho^B_{\rm can}$), where the time evolution of the backward process is given by $U^{\dagger SB}_k$. From (4.23), we obtain a detailed fluctuation theorem for the feedback control process given by

$$\frac{\tilde{p}^{\rm fb}(h,k,y,j,z)}{p^{\rm fb}(h,k,y,j,z)} = {\rm e}^{\sigma^{SB}(x,h,j,z) + I(x,k,y)}. \tag{4.27}$$

Note that the x dependence on the right-hand side of Eq. (4.27) is actually canceled out and the combination of $\Sigma(h,k,y,j,z) := \sigma^{SB}(x,h,j,z) + I(x,k,y)$ quantifies the irreversibility of the feedback control process. However, we can divide this term Σ into σ^{SB} and I as in (4.23), where both terms σ^{SB} and I have physically important meanings.

Next, let us derive a detailed fluctuation theorem which relates the ratio of the forward and the backward probability distributions and the work. Let us assume that the initial and the final density matrices of the system are diagonalized by their corresponding energy eigenbases of the system. Then, the basis sets $\{|\psi^S(x)\rangle\}_x$ and $\{|\phi_k^S(z)\rangle\}_z$ satisfy

$$H_{\lambda_0}^S = \sum_x E_{\lambda_0}^S(x) \left|\psi^S(x)\right\rangle \left\langle\psi^S(x)\right|, \tag{4.28}$$

$$H_{\lambda_1}^S(k) = \sum_z E_{\lambda_1}^S(z|k) \left|\phi_k^S(z)\right\rangle \left\langle\phi_k^S(z)\right|. \tag{4.29}$$

Using Eqs. (4.28) and (4.29), we can define W_{ext}^S, $\mathcal{F}_{\lambda_0}$ and $\mathcal{F}_{\lambda_1}$ as follows:

$$W_{\mathrm{ext}}^S(x,h,k,j,z) := E_{\lambda_0}^S(x) - E_{\lambda_1}^S(z|k) + Q(h,j), \tag{4.30}$$

$$\mathcal{F}_{\lambda_0}(x) := E_{\lambda_0}^S(x) + kT \ln p_{\mathrm{ini}}^S(x), \tag{4.31}$$

$$\mathcal{F}_{\lambda_1}(z|k) := E_{\lambda_1}^S(z|k) + kT \ln p_{\mathrm{fin}}^S(z|k). \tag{4.32}$$

Note that if the initial density matrix is not diagonalized by the energy eigenbasis, we cannot use the same label x and the basis set $\{|\psi^S(x)\rangle\}_x$ to define the stochastic quantities $E_{\lambda_0}^S(x)$ and $\ln p_{\mathrm{ini}}^S(x)$ on the right-hand side of Eq. (4.31). We also note that a similar argument also holds for Eq. (4.32).

From Eq. (4.27), we obtain a detailed fluctuation theorem under feedback control process:

$$\frac{\tilde{p}^{\mathrm{fb}}(h,k,y,j,z)}{p^{\mathrm{fb}}(h,k,y,j,z)} = \mathrm{e}^{\beta(W_{\mathrm{ext}}^S(x,h,k,j,z)+\Delta\mathcal{F}(x,k,z))-I(x,k,y)}, \tag{4.33}$$

where $\Delta\mathcal{F}(x,k,z) = \mathcal{F}_{\lambda_1}(z|k) - \mathcal{F}_{\lambda_0}(x)$.

We note that Eqs. (4.27) and (4.33) holds for the labels y satisfying $p(y|k) \neq 0$, and we denote this set of labels y as Y:

$$Y := \{y | p(y|k) \neq 0\}. \tag{4.34}$$

If $p(y|k) = 0$, we do not observe a forward trajectory starting from a state labeled by y and thus we cannot assign a stochastic quantity of the physical variables to that trajectory. Moreover, the left-hand sides of Eqs. (4.27) and (4.33) formally diverge to ∞ and they are not well defined as has been pointed out in the classical case in Refs. [54, 55]. Using the probability measure, we say $\tilde{p}^{\mathrm{fb}}(h,k,y \notin Y,j,z)$ is singular with respect to $p^{\mathrm{fb}}(h,k,y,j,z)$ and $\tilde{p}^{\mathrm{fb}}(h,k,y \in Y,j,z)$ is absolutely continuous with respect to $p^{\mathrm{fb}}(h,k,y,j,z)$ following the general argument done in the classical case in Ref. [56].

4.2.2 Detailed fluctuation theorems for the measurement process

Next, we derive a detailed fluctuation theorem for the measurement process in a similar manner as we derive Eq. (4.27). We begin by defining a quantity $\langle\sigma^M\rangle$ which quantifies the amount of entropy that is increased in M (or the dissipated work of M normalized by β) as

$$\langle\sigma^M\rangle := -\beta(\langle W_{\mathrm{ext}}^M\rangle + \langle\Delta\mathcal{F}^M\rangle) = S(\rho_{\mathrm{meas}}^M) - S(\rho_{\mathrm{ini}}^M). \tag{4.35}$$

Let us denote $p_k p_{\mathrm{meas}}^M(b|k) = \langle\phi_k^M(b)|\, \rho_{\mathrm{meas}}^M \,|\phi_k^M(b)\rangle$. Then, Eq. (4.35) can be further transformed as

$$\begin{aligned}\langle\sigma^M\rangle &= \sum_a p_{\mathrm{ini}}^M(a) \ln p_{\mathrm{ini}}^M(a) - \sum_{k,b} p_k p_{\mathrm{meas}}^M(b|k) \ln[p_k p_{\mathrm{meas}}^M(b|k)] \\ &= \sum_{x,a,k,y,b} p^{\mathrm{meas}}(x,a,k,y,b) \ln \frac{p_{\mathrm{ini}}^M(a)}{p_k p_{\mathrm{meas}}^M(b|k)},\end{aligned} \tag{4.36}$$

where

$$p^{\rm meas}(x,a,k,y,b) := p(k,y,b|x,a)p_{\rm ini}^S(x)p_{\rm ini}^M(a) \tag{4.37}$$

is the forward probability distribution corresponding to the measurement process. From Eq. (4.36), we define the quantity σ^M as

$$\sigma^M(a,k,b) := \ln \frac{p_{\rm ini}^M(a)}{p_k p_{\rm meas}^M(b|k)}. \tag{4.38}$$

Next, let us consider the combination $\langle\sigma^M\rangle - \langle I\rangle$:

$$\begin{aligned}\langle\sigma^M\rangle - \langle I\rangle &= \sum_{x,a,k,y,b} p^{\rm meas}(x,a,k,y,b)\ln\frac{p_{\rm ini}^S(x)p_{\rm ini}^M(a)}{p_k p^S(y|k)p_{\rm meas}^M(b|k)} \\ &= \sum_{x,a,k,y,b} p^{\rm meas}(x,a,k,y,b)\ln\frac{p^{\rm meas}(x,a,k,y,b)}{\tilde{p}^{\rm meas}(x,a,k,y,b)} \geq 0, \end{aligned} \tag{4.39}$$

where the last inequality results from the nonnegativity of the relative entropy and we also note that (4.39) is equivalent to the second law (2.64). In (4.39), we introduce the backward probability distribution of the measurement process as

$$\begin{aligned}\tilde{p}^{\rm meas}(x,a,k,y,b) &:= \tilde{p}(x,a|k,y,b)p_k p^S(y|k)p_{\rm meas}^M(b|k) \\ &= p(k,y,b|x,a)p_k p^S(y|k)p_{\rm meas}^M(b|k),\end{aligned} \tag{4.40}$$

where

$$\tilde{p}(x,a|k,y,b) := \left|\langle\psi^S(x)|\otimes\langle\psi^M(a)|\,U^{\dagger SM}\,|\varphi_k^S(y)\rangle\otimes|\phi_k^M(b)\rangle\right|^2 \tag{4.41}$$

and Eq. (4.40) gives a joint probability distribution starting from $\sum_k p_k\rho^S(k)\otimes\rho^M(k)$ and perform a reversal of establishing a correlation between S and M by applying $U^{\dagger SM}$. Using (4.39), we obtain the detailed fluctuation theorem for the measurement process given by

$$\frac{\tilde{p}^{\rm meas}(x,a,k,y,b)}{p^{\rm meas}(x,a,k,y,b)} = \mathrm{e}^{-\sigma^M(a,k,b)+I(x,k,y)}. \tag{4.42}$$

Let us assume that the initial and the final density matrices of the measurement device are diagonalized by the energy eigenbasis of M and derive a detailed fluctuation theorem which relates the ratio of the forward and backward probability distributions and the work extracted from M during the measurement process. Then, the basis sets $\{|\psi^M(a)\rangle\}_a$ and $\{|\phi_k^M(b)\rangle\}_{k,b}$ satisfy

$$H^M = \sum_a E^M(a)\,|\psi^M(a)\rangle\langle\psi^M(a)|, \tag{4.43}$$

$$H^M = \sum_{b,k} E^M(b,k)\,|\phi_k^M(b)\rangle\langle\phi_k^M(b)|, \tag{4.44}$$

which means $a=(b,k)$ and $|\psi^M(a)\rangle = |\phi_k^M(b)\rangle$. For example, we can start from a initial state of M in the "standard state" $k=0$, such that the initial density matrix of M is localized in the subspace $\mathbf{H_0^M}$ spanned by the basis set $\{|\phi_0^M(b)\rangle\}_b$ (i.e., $p_{\rm ini}^M(b,k)=0$ for $k\neq 0$) [16]. Using Eqs. (4.43) and (4.44), we can define the stochastic quantities $W_{\rm ext}^M$ and $\Delta\mathcal{F}^M$ as follows:

$$W_{\rm ext}^M(a,k,b) := E^M(a) - E^M(b,k), \tag{4.45}$$

$$\Delta\mathcal{F}^M(a,k,b) := \left(E^M(b,k)+kT\ln[p_k p_{\rm meas}^M(b|k)]\right) - \left(E^M(a)+kT\ln p_{\rm ini}^M(a)\right). \tag{4.46}$$

By substituting $\sigma^M = -\beta(W_{\rm ext}^M + \Delta\mathcal{F}^M)$ in Eq. (4.42), we obtain the detailed fluctuation theorem which relates the ratio of the forward and the backward probability distributions to the extractable work from the measurement device during the measurement process:

$$\frac{\tilde{p}^{\rm meas}(x,a,k,y,b)}{p^{\rm meas}(x,a,k,y,b)} = {\rm e}^{\beta(W_{\rm ext}^M(a,k,b)+\Delta\mathcal{F}^M(a,k,b))+I(x,k,y)}. \tag{4.47}$$

We note that Eqs. (4.42) and (4.47) hold for the labels x, a satisfying $p_{\rm ini}^S(x)p_{\rm ini}^M(a) \neq 0$, and we denote this set of labels (x,a) as A:

$$A := \{(x,a)|p_{\rm ini}^S(x)p_{\rm ini}^M(a) \neq 0\}. \tag{4.48}$$

Because of the same reason as for the feedback control case, if $p_{\rm ini}^S(x)p_{\rm ini}^M(a) = 0$, we do not observe a forward trajectory starting from a state labeled by (x,a) and thus we cannot assign a stochastic quantity of the physical variables to that trajectory.

4.3 Integral fluctuation theorems

We derive the integrated version of the quantum fluctuation theorem and the quantum Jarzynski equality under measurement and feedback control, by using the detailed fluctuation theorems derived in the previous section. We start from the normalization condition of the backward probability corresponding to the feedback control process (4.25):

$$1 = \sum_{h,k,y,j,z} \tilde{p}^{\rm fb}(h,k,y,j,z). \tag{4.49}$$

Note that we can use the detailed fluctuation theorem (4.27) for $y \in Y$. By decomposing the backward probability distribution into two parts $\tilde{p}^{\rm fb}(h,k,y\in Y,j,z)$ and $\tilde{p}^{\rm fb}(h,k,y\notin Y,j,z)$, we have

$$\begin{aligned} 1 &= \sum_{h,k,y\notin Y,j,z} \tilde{p}^{\rm fb}(h,k,y,j,z) + \sum_{h,k,y\in Y,j,z} \tilde{p}^{\rm fb}(h,k,y,j,z) \\ &= \lambda^{\rm fb} + \sum_{h,k,y\in Y,j,z} p^{\rm fb}(h,k,y,j,z)\frac{\tilde{p}^{\rm fb}(h,k,y,j,z)}{p^{\rm fb}(h,k,y,j,z)} \\ &= \lambda^{\rm fb} + \sum_{x,a,h,k,y\in Y,b,j,z} p(x,a,h,k,y,b,j,z){\rm e}^{-\sigma^{SB}(x,h,k,j,z)-I(x,k,y)} \\ &= \lambda^{\rm fb} + \left\langle {\rm e}^{-\sigma^{SB}-I} \right\rangle, \end{aligned} \tag{4.50}$$

where

$$\begin{aligned} \lambda^{\rm fb} &:= \sum_{h,k,y\notin Y,j,z} \tilde{p}^{\rm fb}(h,k,y,j,z) \\ &= \sum_{h,k,y\notin Y,j,z} p_k \left\langle \varphi_k^S(y)\right| {\rm Tr}_B[U_k^{\dagger SB}(\rho_{\rm fin}^S(k)\otimes\rho_{\rm can}^B)U_k^{SB}] \left|\varphi_k^S(y)\right\rangle \end{aligned} \tag{4.51}$$

is the sum of the singular paths of the backward probabilities, where the overlap between the post-measurement state $\{|\varphi_k^S(y)\rangle\}_{y\in Y}$ and the final density matrix of the backward process ${\rm Tr}_B[U_k^{\dagger SB}(\rho_{\rm fin}^S(k)\otimes \rho_{\rm can}^B)U_k^{SB}]$ is zero. Let us introduce the Renyi divergence of order 0 as

$$D_0(p||q) := -\ln\sum_{x|p(x)>0} q(x). \tag{4.52}$$

Then, $\lambda^{\rm fb}$ satisfies the following relation:

$$1-\lambda^{\rm fb}={\rm e}^{-D_0(p^{\rm fb}||\tilde{p}^{\rm fb})}. \tag{4.53}$$

We note that q is absolutely continuous with respect to p if and only if $D_0(p||q)=0$ [77].

We finally obtain the integral quantum fluctuation theorem under feedback control as

$$\left\langle {\rm e}^{-\sigma^{SB}-I}\right\rangle = 1-\lambda^{\rm fb}. \tag{4.54}$$

Equivalently, using the Renyi divergence of order 0, we have

$$\left\langle {\rm e}^{-\sigma^{SB}-I+D_0(p^{\rm fb}||\tilde{p}^{\rm fb})}\right\rangle = 1. \tag{4.55}$$

If the initial and the final density matrices are diagonalized in their corresponding energy eigenbasis sets, we can use Eq. (4.33) and obtain the quantum Jarzynski equality under feedback control:

$$\left\langle {\rm e}^{\beta(W^S_{\rm ext}+\Delta\mathcal{F}^S)-I+D_0(p^{\rm fb}||\tilde{p}^{\rm fb})}\right\rangle = 1. \tag{4.56}$$

Applying Jensen's inequality to Eq. (4.55), we obtain an inequality that lower bounds the measure of irreversibility of the feedback control process as

$$\left\langle\sigma^{SB}\right\rangle + \langle I\rangle \geq D_0(p^{\rm fb}||\tilde{p}^{\rm fb}). \tag{4.57}$$

Note that the left-hand side of (4.57) is related to the relative entropy:

$$\left\langle\sigma^{SB}\right\rangle + \langle I\rangle = D(p^{\rm fb}||\tilde{p}^{\rm fb}). \tag{4.58}$$

This quantity vanishes if and only if $p^{\rm fb}=\tilde{p}^{\rm fb}$, which means that the feedback control process is thermodynamically reversible (D_0 also vanishes in this case). We also note that (4.57) can be directly obtained by using the relation $D(p||q)\geq D_0(p||q)$ [77].

We can also apply Jensen's inequality to Eq. (4.56) and obtain the second law of thermodynamics under feedback control:

$$\left\langle W^S_{\rm ext}\right\rangle \leq -\left\langle\Delta\mathcal{F}^S\right\rangle + kT\langle I\rangle - kTD_0(p^{\rm fb}||\tilde{p}^{\rm fb}). \tag{4.59}$$

Due to the term D_0 on the right-hand side of (4.59), the inequality (4.59) gives a stronger constraint on the extractable work compared with the second law given in (2.58). By combining Eq. (4.11) and (4.57), we can obtain (4.59) without assuming that the initial and the final density matrices are diagonalized in the corresponding energy eigenbasis sets. However, if we want to obtain information about higher moments of the extractable work, we need to use Eq. (4.56) rather than Eq. (4.55) since $H^S_{\lambda_0}$ and $\rho^S_{\rm ini}$ cannot be diagonalized simultaneously in general.

From (2.55), we find that $\left\langle W^S_{\rm ext}\right\rangle$ takes the maximum value if and only if

$$\rho^S(k)\otimes\rho^B_{\rm can}=\tilde{\rho}^{SB}_{\rm fin}(k), \tag{4.60}$$

which means that the final density matrix of the backward process exactly returns to the post-measurement state of the system. Let us compare this fact with the term $D_0(p^{\rm fb}||\tilde{p}^{\rm fb})$ on the right hand side of (4.59) by introducing the quantum Renyi zero divergence as

$$D(\rho||\sigma):=-\ln{\rm Tr}[\Pi_\rho\sigma], \tag{4.61}$$

where Π_ρ is a projection onto the support of ρ. Then, $\lambda^{\rm fb}$ is given by

$$\begin{aligned} e^{-D_0(p^{\rm fb}||\tilde{p}^{\rm fb})} = 1-\lambda^{\rm fb} &= \sum_k p_k {\rm Tr}[\Pi_{\rho^S(k)\otimes\rho^B_{\rm can}}\tilde{\rho}^{SB}(k)] \\ &= \sum_k p_k e^{-D_0(\rho^S(k)\otimes\rho^B||\tilde{\rho}^{SB}(k))}. \end{aligned} \tag{4.62}$$

Then, we find that to increase the extractable work $\langle W^S_{\rm ext}\rangle$, we need to choose a feedback control U^{SB}_k such that the overlap between $\rho^S(k)\otimes\rho^B_{\rm can}$ and $\tilde{\rho}^{SB}_{\rm fin}(k)$ is large, thereby making $D_0(\rho^S(k)\otimes\rho^B||\tilde{\rho}^{SB}(k))$ small. In this sense, $D_0(\rho^S(k)\otimes\rho^B||\tilde{\rho}^{SB}(k))$ and $D_0(p^{\rm fb}||\tilde{p}^{\rm fb})$ quantifies the inefficiency of the feedback control. Note that using Jensen's inequality, we also obtain

$$\langle W^S_{\rm ext}\rangle \le -\langle\Delta\mathcal{F}^S\rangle + kT\langle I\rangle - kT\sum_k p_k D_0(\rho^S(k)\otimes\rho^B||\tilde{\rho}^{SB}(k)). \tag{4.63}$$

Next, we use the normalization condition of the backward probability distribution for the measurement process (4.37) and derive integral quantum fluctuation theorems:

$$1 = \sum_{x,a,y,k,b}\tilde{p}^{\rm meas}(x,a,k,y,b). \tag{4.64}$$

Let us decompose the backward probability distribution into two parts $\tilde{p}^{\rm meas}((x,a)\in A,h,k,y,b,j,z)$ and $\tilde{p}^{\rm meas}((x,a)\notin A,h,k,y,b,j,z)$ and use the detailed fluctuation theorem (4.42) to obtain

$$\begin{aligned} 1 &= \sum_{(x,a)\notin A,y,k,b}\tilde{p}^{\rm meas}(x,a,k,y,b) + \sum_{(x,a)\in A,y,k,b}\tilde{p}^{\rm meas}(x,a,k,y,b) \\ &= \lambda^{\rm meas} + \sum_{(x,a)\in A,y,k,b} p^{\rm meas}(x,a,k,y,b)\frac{\tilde{p}^{\rm meas}(x,a,k,y,b)}{p^{\rm meas}(x,a,k,y,b)} \\ &= \lambda^{\rm meas} + \sum_{(x,a)\in A,h,k,y,b,j,z} p(x,a,h,k,y,b,j,z)e^{-\sigma^M(a,k,b)+I(x,k,y)} \\ &= \lambda^{\rm meas} + \left\langle e^{-\sigma^M+I}\right\rangle, \end{aligned} \tag{4.65}$$

where

$$\begin{aligned} \lambda^{\rm meas} &:= \sum_{(x,a)\notin A,y,k,b}\tilde{p}^{\rm meas}(x,a,k,y,b) \\ &= \sum_{(x,a)\notin A,y,k,b} p_k\langle\psi^S(x)|\otimes\langle\psi^M(a)|[U^{\dagger SM}(\rho^S(k)\otimes\rho^M(k))U^{SM}]|\psi^S(s)\rangle\otimes|\psi^M(a)\rangle \\ &= 1-e^{-D_0(\rho^S_{\rm ini}\otimes\rho^M_{\rm ini}||\tilde{\rho}^{SM}_{\rm fin})} \end{aligned} \tag{4.66}$$

is the sum of the singular paths of the backward probabilities such that the overlap between the initial density matrix $\rho^S_{\rm ini}\otimes\rho^M_{\rm ini}$ and the final density matrix of the backward process

$$\tilde{\rho}^{SM}_{\rm fin} := U^{\dagger SM}\left(\sum_k p_k\rho^S(k)\otimes\rho^M(k)\right)U^{SM} \tag{4.67}$$

is zero. Using the Renyi divergence of order 0 (4.53), we also have

$$D_0(p^{\rm meas}||\tilde{p}^{\rm meas}) = D_0(\rho^S_{\rm ini}\otimes\rho^M_{\rm ini}||\tilde{\rho}^{SM}_{\rm fin}). \tag{4.68}$$

From Eq. (4.65), we obtain the integral quantum fluctuation theorem for the measurement process as

$$\left\langle e^{-\sigma^M + I} \right\rangle = 1 - \lambda^{\text{meas}}, \tag{4.69}$$

and equivalently,

$$\left\langle e^{-\sigma^M + I + D_0} \right\rangle = 1. \tag{4.70}$$

If the initial and the final density matrices of M are diagonalized by the energy eigenbasis, we can use Eq. (4.47) and obtain the quantum Jarzynski equality for the measurement process:

$$\left\langle e^{\beta(W_{\text{ext}}^M + \Delta \mathcal{F}^M) + I + D_0} \right\rangle = 1. \tag{4.71}$$

Applying Jensen's inequality to Eq. (4.70), we obtain an inequality that lower bounds the measure of irreversibility of the measurement process as

$$\left\langle \sigma^{SM} \right\rangle - \langle I \rangle = D(p^{\text{meas}} || \tilde{p}^{\text{meas}}) \geq D_0(p^{\text{meas}} || \tilde{p}^{\text{meas}}). \tag{4.72}$$

We can also apply Jensen's inequality to Eq. (4.71) or substitute Eq. (4.35) in (4.72) to obtain the second law of thermodynamics for the measurement process:

$$\left\langle W_{\text{ext}}^M \right\rangle \leq - \left\langle \Delta \mathcal{F}^M \right\rangle - kT \langle I \rangle - kT D_0(\rho_{\text{ini}}^S \otimes \rho_{\text{ini}}^M || \tilde{\rho}_{\text{fin}}^{SM}), \tag{4.73}$$

which gives a stronger constraint on the extractable work compared with the second law given in (2.64). Note that the density matrix $\rho^S(k) \otimes \rho^M(k)$ is obtained by destroying all the correlation between S and M from the postmeasurement state $\rho^{SM}(k)$. Then, Eq. (4.66) quantifies how well the decohered state returns back to the initial density matrix by undoing the correlation by applying the unitary operator $U^{\dagger SM}$.

4.4 Some discussions on the main results

4.4.1 Effect of measurement back action and the information quantity

The stochastic quantity $I(x, k, y) := \ln p_{\text{ini}}^S(x) - \ln p(y|k)$ differs from the classical mutual information $I^{SM}(x, k) := \ln p_{\text{ini}}^S(x) - \ln p(x|k)$, where $p(x|k) := p^{SM}(x, k)/p^M(k)$ is the postmeasurement state of the system conditioned on the measurement outcome k. Because of the back action of the quantum measurement, the postmeasurement state depends on the label y, which is different from the label x that quantifies the microscopic state of the initial density matrix. Usually we assume that the classical measurement does not alter the state of the system and this kind of the measurement back action does not occur. Then, the corresponding information quantity is given by the classical mutual information and it always takes a nonnegative value because the postmeasurement state conditioned on the measurement outcome is always "less mixed" compared to the initial state. However, the density matrix of the system conditioned on the measurement outcome may be "more mixed" than the initial density matrix due to the effect of the measurement back action and thus $\langle I \rangle$ may take a negative value. We note that we can also consider a measurement back action in a classical system. In that case, we can use the information gain $I(x, k, y)$ to quantify the obtained information via the classical measurement and derive fluctuation theorems by using this information quantity $I(x, k, y)$ instead of the classical mutual information.

Let us also analyze how we can estimate the postmeasurement state of the system from the knowledge of the measurement device by dividing the information gain into two parts, i.e., $\langle I \rangle :=$

$I_{\rm gain} - I_{\rm loss}$. Here, $I_{\rm gain}$ quantifies the amount of information gain about the system if we know the microscopic trajectory of the measurement device which is labeled by a, k, b:

$$I_{\rm gain} := S(\rho^S_{\rm ini}) - \sum_{a,k,b} p_{a,k,b} S(\rho^S(a,k,b)) \geq 0, \tag{4.74}$$

where

$$\rho^S(a,k,b) := \frac{1}{p_{a,k,b}} M^S_{k,a,b} \rho^S_{\rm ini} M^{\dagger S}_{k,a,b} \tag{4.75}$$

and $p_{a,k,b} := {\rm Tr}[M^S_{k,a,b}\rho^S_{\rm ini} M^{\dagger S}_{k,a,b}]$. Here $I_{\rm loss}$ quantifies the amount of information loss due to the measurement back action given by

$$I_{\rm loss} := \sum_k p_k S(\rho^S(k)) - \sum_{a,k,b} p_{a,k,b} S(\rho^S(a,k,b)) \geq 0. \tag{4.76}$$

If we have an additional knowledge about the measurement device and know that the initial state and the final state of M are given by $|\psi^M(a)\rangle$ and $|\phi^M_k(b)\rangle$ rather than the statistical mixtures $\rho^M_{\rm ini}$ and $\rho^M(k)$, we can estimate the postmeasurement state of the system given by $\rho^S(a,k,b)$. Then, the extractable work is characterized by $I_{\rm meas}$. However, we usually do not know the microscopic trajectory of the measurement device and thus from the obtained measurement outcome k, we can only estimate that the postmeasurement state is given by $\rho^S(k) = \sum_{a,b} p_{a,k,b}\rho^S(a,k,b)/p_k$. The lack of the knowledge of the full microscopic trajectory of M is quantified by $I_{\rm loss}$.

In Ref. [56], the sum of the singular path probabilities λ was introduced and the fluctuation theorems and Jarzynski equalities for classical systems were obtained: $\langle {\rm e}^{-\sigma}\rangle = 1-\lambda$. Furthermore, the authors of Ref. [56] called the process with nonzero singular paths as an absolutely irreversible process, and discussed its physical meanings. Here, we use the same technique to treat the singular paths that naturally arise in the case of the measurement and the feedback control processes.

4.4.2 Relation with the quantum fluctuation theorem for the total system

In this subsection, we consider a composite system of S, M and B, and derive the fluctuation theorems for the entire process involving measurement and feedback control. Let us first define the backward probability distribution of the entire process as

$$\tilde{p}(x,a,h,k,b,j,z) := \tilde{p}(x,a|k,y,b)\tilde{p}(k,y,h|z,j) p_k p^S_{\rm fin}(z|k) p^M_{\rm meas}(b|k) p^B_{\rm can}(j). \tag{4.77}$$

Then, the logarithm of the ratio between the forward and backward probability distributions is related to the total entropy production during the entire process:

$$\begin{aligned} \sigma^{\rm tot} &:= \ln \frac{p(x,a,h,k,b,j,z)}{\tilde{p}(x,a,h,k,b,j,z)} \\ &= (\sigma^{SB}(x,h,k,j,z) + I(x,k,y)) + (\sigma^M(a,k,b) - I(x,k,y)). \end{aligned} \tag{4.78}$$

The integral fluctuation theorem can be obtained as

$$\left\langle {\rm e}^{-\sigma_{\rm tot} + D_0(p||\tilde{p})} \right\rangle = 1, \tag{4.79}$$

and the second law, which is applicable to the entire system, is given by

$$\langle \sigma_{\rm tot} \rangle \geq D_0(p||\tilde{p}). \tag{4.80}$$

Note that the effect of the information exchange between S and M is canceled out, and the information gain does not appear in Eq. (4.79). On the other hand, we can divide the total entropy production into $\sigma^{SB} + I$ and $\sigma^M - I$ [71] as in Eq. (4.78) and derive fluctuation theorems for the measurement process (4.70) and the feedback control process (4.55) separately.

4.4.3 Two-point measurement scheme

In the quantum regime, the detailed fluctuation theorem is usually obtained by using the so-called two-point measurement scheme, i.e., we perform energy projective measurements at the beginning and end of the protocol and define the stochastic energy difference of the system during a single run of the protocol. In this way, we can assign physical meanings to path probabilities and quantities for each path and formulate the detailed fluctuation theorem in the same form as that in the classical system, except that the underlying dynamics is quantum. The work statistics of an isolated quantum system has been obtained by using this method on a trapped-ion system and the quantum Jarzynski equality has been verified [47]. We also use this approach to define path probabilities and physical quantities such as the extractable work and the information gain to each trajectory. In particular, we need to perform a projection measurement $\left|\varphi_k^S(y)\right\rangle\left\langle\varphi_k^S(y)\right|$ on the system after the general quantum measurement $M_{k,a,b}^S$ in order to define the quantity $I(x,k,y)$ and also the path probability $p^{\mathrm{fb}}(h,k,y,j,z)$ during the feedback control process. Note that the time-reversal of the projection measurement (using the basis $\left|\varphi_k^S(y)\right\rangle$) at the beginning of the feedback control implies that we consider a projective measurement using the basis $\left|\varphi_k^S(y)\right\rangle \otimes \left|\psi^B(h)\right\rangle$ at the end of the backward process of the feedback control. Then, the backward probability distribution of the feedback control is given by $\tilde{p}^{\mathrm{fb}}(h,k,y,j,z)$ and does not depend on x which labels the initial state of the system.

4.4.4 Definition of work in the quantum regime

There are some different approaches of defining work in the quantum regime [30, 84–88]. We note that the two-point measurement scheme does not tell us where the extracted work goes and hence some studies concerning about this problem has been carried out by several authors. Let us consider a semi-classical model where the system is coupled to a classical field which acts like a control parameter. Then, energy must be consumed by the classical field to drive the Hamiltonian and extract the free-energy difference as work. From the energy conservation, it is natural to consider that the extracted work goes to the classical field as a back-action of the driving. At this moment, we have a few models which analyze this back-action and relate it to the extracted work.

The authors in Ref. [84] considered a system and an auxiliary system with a time-independent composite Hamiltonian. They showed that the effective Hamiltonian of the system changes in time, and the extractable work W_{ext} via an isothermal-like process from a qubit ($W_{\mathrm{ext}} = kT\ln 2$ is the maximum value) is stored in the average energy of the auxiliary system at the end of the protocol This means that the driving field stores the extracted work as the back-action of the driving of the system for a specific model.

The authors in Ref. [85] formulated a general framework of the work extraction protocol by using a measurement based formalism. It allows us to overcome the negative point of the two-point measurement scheme, and also it is physically reasonable to model a quantum composite system including the measurement device which detects the extracted work.

Chapter 5

Single-shot statistical mechanics

In this chapter, we review the recent developments in quantum thermodynamics, using the idea of resource theory. This framework is used to construct the deterministic work extraction protocol (and often refereed to as the single-shot statistical mechanics).

5.1 Thermal operation

We begin by introducing the thermal operation, which plays a key role in considering the deterministic work extraction protocol. The thermal operation is expressed as

$$\mathcal{E}_{\mathrm{thermal}}(\rho) := \mathrm{Tr}_B \left[U \left(\rho \otimes \frac{\mathrm{e}^{-\beta H^B}}{Z^B} \right) U^\dagger \right] = \sigma, \tag{5.1}$$

where ρ is an input state of the system and σ is the output state of the system after the operation. We also note that in Eq. (5.1), H^B and $Z^B = \mathrm{Tr}[\exp(-\beta H^B)]$ are the Hamiltonian and the partition function of the heat bath, respectively. Here, U in Eq. (5.1) is an arbitrary energy-conserving unitary operator:

$$[U, H + H^B] = 0, \tag{5.2}$$

where H is the Hamiltonian of the system. Note that the unitary operator U is not the unitary operator $\exp(-\mathrm{i}(H + H^B)t)$ which describes the time evolution of the system. In this framework, the canonical distribution of the heat bath $\mathrm{e}^{-\beta H^B}/Z^B$ is regarded as a free resource, i.e., we assume that we can freely utilize heat baths at the same inverse temperatures β during the process. Then, the thermal operation (5.1) gives a transformation of the density matrix of the system by letting the system interact with the heat bath via the unitary operator U. We note that in this formalism, we assume that the interaction Hamiltonian V vanishes identically. Instead, we effectively generate the interaction between the system and the heat bath by mixing those states via the unitary operator U.

The important property of this thermal operation is that the canonical distribution of the system is conserved:

$$\mathcal{E}_{\mathrm{thermal}} \left(\frac{\mathrm{e}^{-\beta H}}{Z} \right) = \frac{\mathrm{e}^{-\beta H}}{Z}. \tag{5.3}$$

This can be shown by noting that the unitary operator and the canonical distribution of the entire

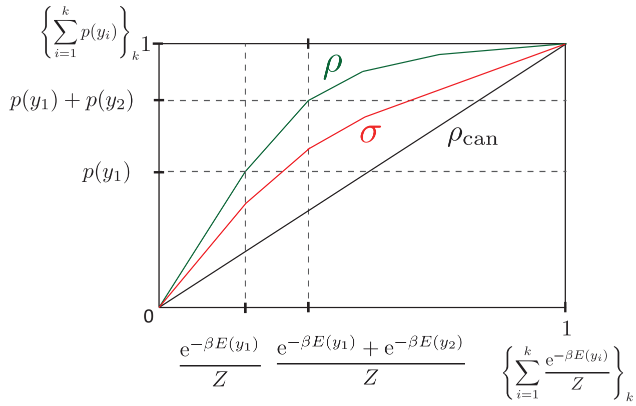

Figure 5.1: Lorenz curve and the thermo-majorization criteria. We plot $\{\sum_{i=1}^k e^{-\beta E(y_i)}/Z\}_k$ and $\{\sum_{i=1}^k p(y_i)\}_k$ in the (X, Y) plane, where the curve $p = \rho$ is represented by the blue curve, $p = \sigma$ is represented by the red curve and $p = \rho_{\rm can} = e^{-\beta H}/Z$ is represented by the black line. Then, the thermo-majorization criterion tells us that the state ρ can be transformed into σ via a thermal operation if the Lorenz curve $(\sigma, \frac{e^{-\beta H}}{Z})$ is a subset of the curve $(\rho, \frac{e^{-\beta H}}{Z})$.

system plus bath commute with each other because of Eq. (5.2). Then,

$$\begin{aligned}\mathcal{E}_{\rm thermal}\left(\frac{e^{-\beta H}}{Z}\right) &= {\rm Tr}_B\left[U\left(\frac{e^{-\beta H}}{Z}\otimes\frac{e^{-\beta H^B}}{Z^B}\right)U^\dagger\right]\\ &= {\rm Tr}_B\left[\left(\frac{e^{-\beta H}}{Z}\otimes\frac{e^{-\beta H^B}}{Z^B}\right)UU^\dagger\right]\\ &= \frac{e^{-\beta H}}{Z},\end{aligned} \tag{5.4}$$

and Eq. (5.3) is obtained. We note that maps which conserves the canonical distribution is called Gibbs-preserving maps. From Eq. (5.3), the thermal operation is a Gibbs-preserving map, but the converse is not true in general [37].

5.2 Thermo-majorization

Next, we describe a simple method which can decide whether a transition

$$\rho \to \sigma \tag{5.5}$$

is possible or not by the thermal operation (5.1). In the following, we assume that ρ and σ commute with the Hamiltonian H. We also denote the diagonal element of ρ and that of σ as $p(x)$ and $q(x)$, respectively.

Let us arrange the probability distribution $p(x)$ according to the following order:

$$\frac{p(x_1)}{e^{-\beta E(x_1)}/Z} \geq \frac{p(x_2)}{e^{-\beta E(x_2)}/Z} \geq \frac{p(x_3)}{e^{-\beta E(x_3)}/Z} \geq \cdots. \tag{5.6}$$

Then, we plot the Lorenz curve denoted as $(\rho, e^{-\beta H}/Z)$ in the (X, Y) plane as in Fig. 5.1 in which each point is given by

$$\left(\frac{\mathrm{e}^{-\beta H}}{Z},\rho\right) = \left(\frac{\mathrm{e}^{-\beta E(x_1)}}{Z},p(x_1)\right),\ \left(\sum_{i=1}^{2}\frac{\mathrm{e}^{-\beta E(x_i)}}{Z},\sum_{i=1}^{2}p(x_i)\right),$$
$$\left(\sum_{i=1}^{3}\frac{\mathrm{e}^{-\beta E(x_i)}}{Z},\sum_{i=1}^{3}p(x_i)\right),\ \cdots,\ (1,1)), \tag{5.7}$$

Note that the ordering (5.6) ensures that the curve (5.7) is convex. The thermo-majorization criteria [30] allow that the transition (5.5) is possible by the thermal operation if and only if the curve $(\frac{\mathrm{e}^{-\beta H}}{Z},\sigma)$ is a subset of the curve $(\frac{\mathrm{e}^{-\beta H}}{Z},\rho)$ (see also Fig. 5.1). Note that the Lorenz curve $(\mathrm{e}^{-\beta H}/Z,\mathrm{e}^{-\beta H}/Z)$ gives a straight line, and any state can be transformed into the canonical distribution by the thermal operation. We also find that the thermal operation always brings the state of the system closer to the canonical distribution i.e., the input state majorizes to a thermal distribution by the thermal operation.

Let us denote $F(t)$, $t\in[0,1]$ as the curve that upper bounds the Lorenz curve (p,q). Then, the Renyi divergence is given by:

$$D_\alpha(p||q) = \frac{1}{\alpha-1}\ln\int_0^1 (F'(t))^\alpha \mathrm{d}t = \frac{1}{\alpha-1}\ln\left[\left(\frac{p(x_1)}{q(x_1)}\right)^\alpha q(x_1) + \left(\frac{p(x_2)}{q(x_2)}\right)^\alpha q(x_2)+\cdots\right], \tag{5.8}$$

which can be seen by noting that the slope of the Lorenz curve is given by

$$F'(t) := \frac{p(x_i)}{q(x_i)} \quad \text{if} \quad t\in\left[\sum_{k=1}^{i}q(x_k),\sum_{k=1}^{i+1}q(x_k)\right]. \tag{5.9}$$

5.3 Deterministic work extraction protocol

Next let us introduce a work qubit, whose energy difference is given by w. We denote the ground state and the excited state of the qubit as $|0\rangle_W$ and $|w\rangle_W$. The Hamiltonian of the work qubit is then given by $H^W = w\,|w\rangle\langle w|_W$. Let us consider a joint state of the system and the qubit given by $\rho\otimes|0\rangle\langle 0|_W$. Then, we examine whether the following transformation is possible by the thermal operation:

$$\mathcal{E}_{\mathrm{thermal}}(\rho\otimes|0\rangle\langle 0|_W) = \mathrm{Tr}_B\left[U\left(\rho\otimes|0\rangle\langle 0|_W\otimes\frac{\mathrm{e}^{-\beta H^B}}{Z^B}\right)U^\dagger\right] = \sigma\otimes|w\rangle\langle w|_W\,. \tag{5.10}$$

If this is possible, we can extract work w by transforming the system from ρ to σ, because energy is stored in the work qubit. We note that in general, we need to a priori know the amount of extractable work w and tune the energy level of the work qubit before the operation. Since the extracted work w does not fluctuates, this work extraction protocol is sometimes called the deterministic work extraction protocol. We also note that the extracted work w can be determined by a single measurement of the extracted work. In this sense, we also call this work extraction protocol (5.10) as the single-shot statistical mechanics.

In the following subsections, we review the results given in Ref. [30], which quantify the amount of extractable work starting from a nonequilibrium distribution and ending at a canonical distribution, and the amount of work cost needed to create a nonequilibrium distribution starting from the canonical distribution.

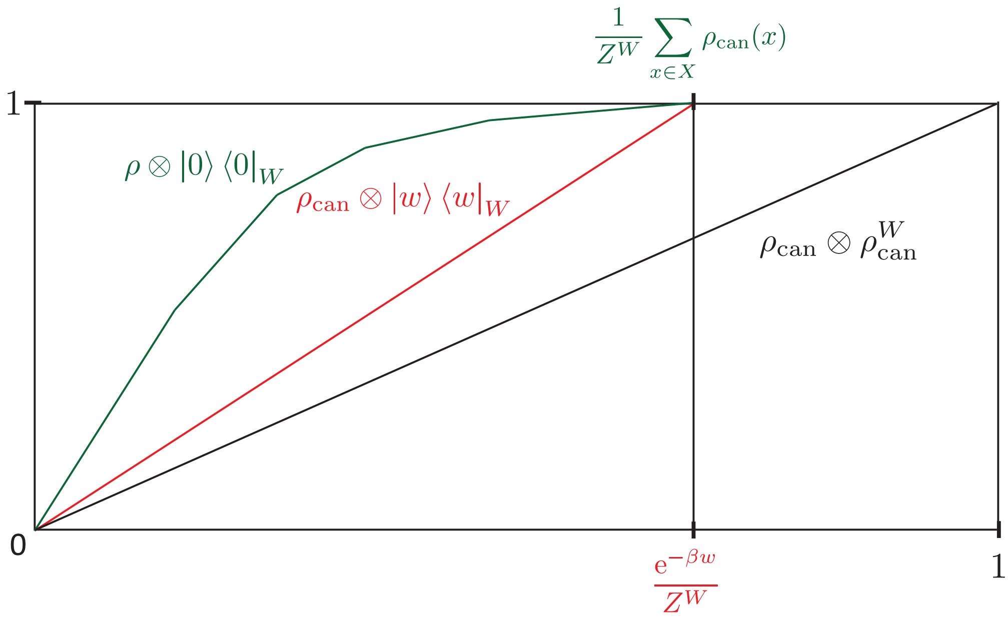

Figure 5.2: Lorenz curves of Eqs. (5.12) and (5.13). Here, X denotes the support which is equal to that of ρ. Then, $(\sum_{x\in X}\rho_{\rm can}(x)/Z^W, 1)$ gives a point in which the Lorenz curve of Eq. (5.12) first touches the $Y = 1$ point. The Lorenz curves of Eq. (5.13) is shown by the red solid line. Using the thermo-majorization criteria, we find that the transformation (5.11) is possible if and only if $\sum_{x\in X}\rho_{\rm can}(x) \leq {\rm e}^{-\beta w}$. Note that we illustrate the case of w satisfying the equality condition $\sum_{x\in X}\rho_{\rm can}(x) = {\rm e}^{-\beta w}$.

5.3.1 Extractable work starting from a nonequilibrium state

Let us consider the amount of extractable work starting from a nonequilibrium distribution ρ and ending at the canonical distribution $\rho_{\rm can}$. In this setup, Eq. (5.10) becomes:

$$\mathcal{E}_{\rm thermal}(\rho \otimes |0\rangle\langle 0|_W) = \rho_{\rm can} \otimes |w\rangle\langle w|_W . \tag{5.11}$$

We plot two Lorenz curves (defined by Eq. (5.7))

$$(\rho_{\rm can} \otimes \rho_{\rm can}^W, \rho \otimes |0\rangle\langle 0|_W) \tag{5.12}$$

and

$$(\rho_{\rm can} \otimes \rho_{\rm can}^W, \rho_{\rm can} \otimes |w\rangle\langle w|_W) \tag{5.13}$$

in Fig. 5.2, where $\rho_{\rm can}^W$ denotes the canonical distribution of the work qubit. Then, from the thermo-majorization criteria, the transformation (5.11) is possible if and only if

$$\sum_{x\in X}\rho_{\rm can}(x) \leq {\rm e}^{-\beta w}. \tag{5.14}$$

Using the Renyi divergence of order 0 (4.52): $D_0(p||q) := -\ln\sum_{x|p(x)>0} q(x)$, the condition (5.14) can be rewritten as

$$\beta w \leq D_0(\rho||\rho_{\rm can}). \tag{5.15}$$

We can conclude that the maximum extractable work starting from a nonequilibrium distribution ρ and ending at the canonical distribution $\rho_{\rm can}$ is given by

$$w_{\rm ext}^{\rm max} = kTD_0(\rho||\rho_{\rm can}). \tag{5.16}$$

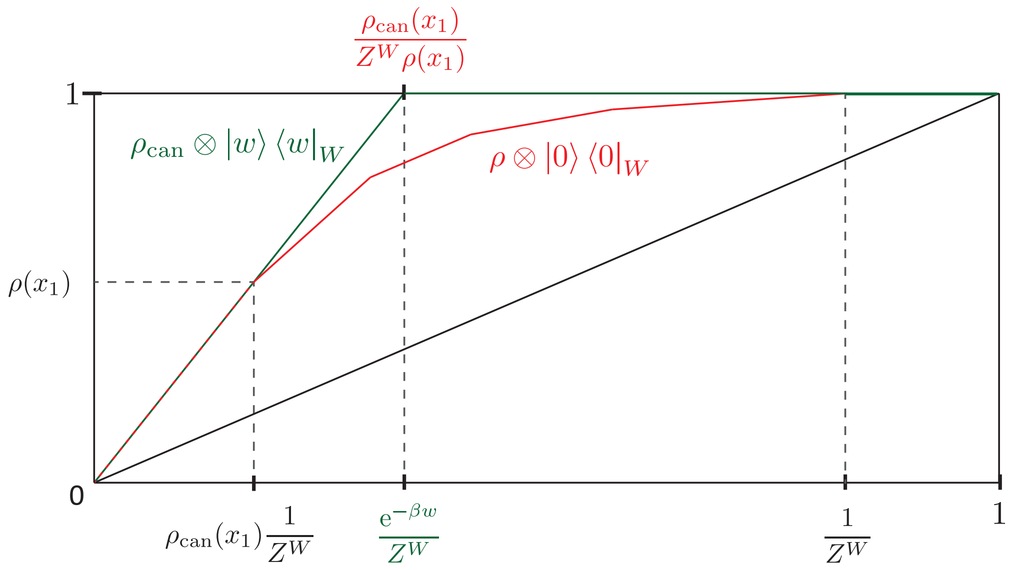

Figure 5.3: Lorenz curves of Eqs. (5.18) and (5.19). Here, the slope connecting two points $(0, 0)$ and $(\rho_{\mathrm{can}}(x_1)/Z^W, \sigma(x_1))$ is equal to $\rho_{\mathrm{can}}(x_1)/(Z^W \rho(x_1))$. Using the thermo-majorization criteria, we find that the transformation (5.17) is possible if and only if $\rho_{\mathrm{can}}(x_1)/(Z^W \rho(x_1)) \geq \mathrm{e}^{-\beta w}/Z^W$.

5.3.2 Work cost of creating a nonequilibrium state from a thermalized state

Let us next consider the work cost of creating a nonequilibrium state ρ starting from the canonical distribution ρ_{can}. In this setup, Eq. (5.10) becomes

$$\mathcal{E}_{\mathrm{thermal}}(\rho_{\mathrm{can}} \otimes |w\rangle \langle w|_W) = \rho \otimes |0\rangle \langle 0|_W . \tag{5.17}$$

We plot two Lorenz curves

$$(\rho_{\mathrm{can}} \otimes \rho^W_{\mathrm{can}}, \rho_{\mathrm{can}} \otimes |w\rangle \langle w|_W) \tag{5.18}$$

and

$$(\rho_{\mathrm{can}} \otimes \rho^W_{\mathrm{can}}, \rho \otimes |0\rangle \langle 0|_W) \tag{5.19}$$

in Fig. 5.3. Then, from the thermo-majoriztion criteria, the transformation (5.17) is possible if and only if

$$\frac{\rho_{\mathrm{can}}(x_1)}{Z^W \rho(x_1)} \geq \frac{\mathrm{e}^{-\beta w}}{Z^W}. \tag{5.20}$$

By introducing the Renyi divergence of order ∞:

$$D_\infty(p||q) := \ln \max_x \frac{p(x)}{q(x)}, \tag{5.21}$$

we can rewrite the condition (5.20) as

$$\beta w \geq D_\infty(\rho||\rho_{\mathrm{can}}). \tag{5.22}$$

We can conclude that the minimum work cost of creating a nonequilibrium distribution ρ from the canonical distribution ρ_{can} is given by

$$w^{\mathrm{min}}_{\mathrm{cost}} = kTD_\infty(\rho||\rho_{\mathrm{can}}). \tag{5.23}$$

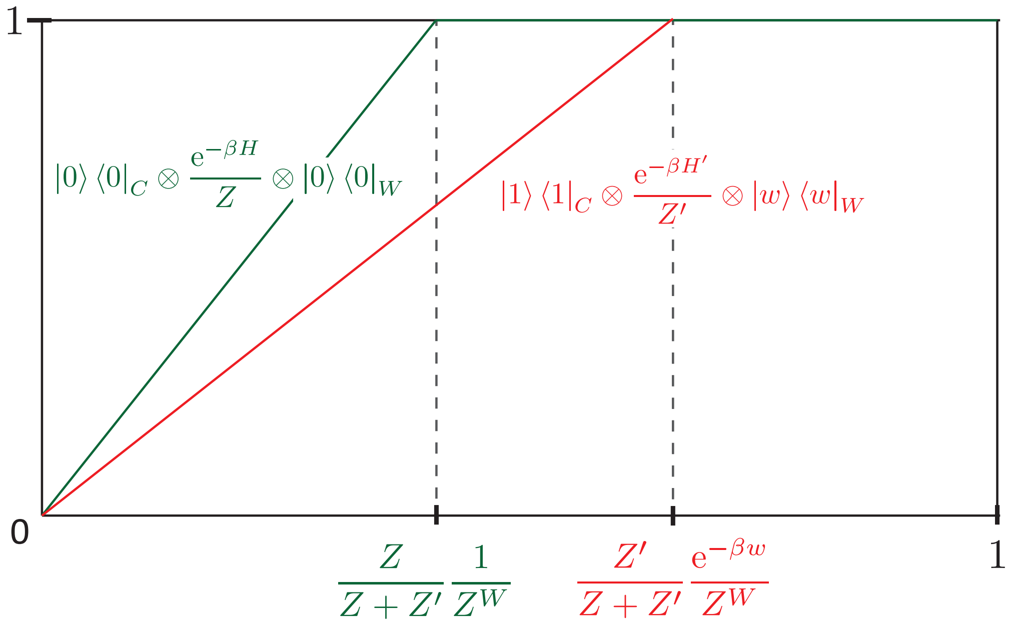

Figure 5.4: Two Lorenz curves $|0\rangle\langle 0|_C \otimes e^{-\beta H}/Z \otimes |0\rangle\langle 0|_W$ and $|1\rangle\langle 1|_C \otimes e^{-\beta H'}/Z' \otimes |w\rangle\langle w|_W$. The former can be transformed into the latter by the thermal operation if and only if $\frac{Z}{Z+Z'}\frac{1}{Z^W} \leq \frac{Z'}{Z+Z'}\frac{e^{-\beta w}}{Z^W}$.

5.4 Changing Hamiltonians

We have considered the thermal operation for a fixed Hamiltonian of the system. However, we often consider thermodynamic processes by changing the Hamiltonian of the system. The authors in Ref. [30] introduced a way to treat the change of the Hamiltonian as an effective dynamics of a fixed Hamiltonian of the larger system. Note that this is different from the time-dependent Hamiltonian approach in which we drive the control parameter of the Hamiltonian depending on time.

We introduce a qubit system C which switches the Hamiltonian of the system from H to H' by taking the total Hamiltonian to be

$$H_{\text{tot}} := |0\rangle\langle 0|_C \otimes H + |1\rangle\langle 1|_C \otimes H' + w\,|w\rangle\langle w|_W\,, \tag{5.24}$$

where we also introduce the work qubit W. If the initial state is

$$|0\rangle\langle 0|_C \otimes \rho \otimes |0\rangle\langle 0|_W\,, \tag{5.25}$$

and the final state is

$$|1\rangle\langle 1|_C \otimes \sigma \otimes |w\rangle\langle w|_W\,, \tag{5.26}$$

the Hamiltonian of the system changes from H to H' and the state of the system changes from ρ to σ by extracting work w.

For example, a transformation between two canonical distributions $\rho = e^{-\beta H}/Z$ and $\sigma = e^{-\beta H'}/Z'$ is possible if and only if

$$\frac{Z}{Z+Z'}\frac{1}{Z^W} \leq \frac{Z'}{Z+Z'}\frac{e^{-\beta w}}{Z^W}, \tag{5.27}$$

from the thermo-majorization criteria using the Lorenz diagram shown in Fig. (5.4). Then, we can reproduce the second law of thermodynamics

$$w \leq F - F', \tag{5.28}$$

by using the thermal operation. Here, F and F' are the initial and final equilibrium free energies. Note that the reverse operation $\mathrm{e}^{-\beta H'}/Z' \to \mathrm{e}^{-\beta H}/Z$ can be done if and only if

$$\frac{Z}{Z+Z'}\frac{1}{Z^W} \geq \frac{Z'}{Z+Z'}\frac{\mathrm{e}^{-\beta w}}{Z^W}. \tag{5.29}$$

Combining (5.29) and (5.27), we obtain the maximum amount of extractable work from the system as

$$w = F - F' = -\Delta F^{\mathrm{eq}}. \tag{5.30}$$

5.5 Thermal operation and the detailed fluctuation theorem

In this section, we consider the relation between the thermal operation and the detailed fluctuation theorem for a general Hamiltonian (5.24). We assume that the initial and final states are given by Eqs. (5.25) and (5.26). We denote the diagonal elements of the initial and final states of the system as $p(x)$ and $q(y)$, respectively. The forward transition probability via the thermal operation is given by

$$P[\Gamma|x] := \frac{\mathrm{e}^{-\beta E_a^B}}{Z^B}\left|\langle 1|_C \otimes \langle\,\phi_y| \otimes \langle w|_W \otimes \langle\,\varphi_b^B|U\,|0\rangle_C \otimes |\psi_x\rangle \otimes |0\rangle_W \otimes |\varphi_a^B\rangle\right|^2, \tag{5.31}$$

and the backward transition probability is given by

$$\tilde{P}[\Gamma^\dagger|y] := \frac{\mathrm{e}^{-\beta E_b^B}}{Z^B}\left|\langle 0|_C \otimes \langle\,\psi_x| \otimes \langle 0|_W \otimes \langle\,\varphi_a^B|U^\dagger\,|1\rangle_C \otimes |\phi_y\rangle \otimes |w\rangle_W \otimes |\varphi_b^B\rangle\right|^2. \tag{5.32}$$

If we take the ratio of the transition probabilities (5.31) and (5.32), we obtain the detailed fluctuation theorem

$$\frac{P[\Gamma|x]}{\tilde{P}[\Gamma^\dagger|y]} = \exp(-\beta Q[\Gamma]), \tag{5.33}$$

where $Q[\Gamma] = E^B(a) - E^B(b)$ is the heat emitted from the heat bath. Let us denote $E(x)$ and $E'(y)$ as the initial and final eigenenergies. Since U is the total-energy-conserving unitary operator, we obtain the following relation for the forward process with the trajectory Γ:

$$E^B(a) + E(x) = E^B(b) + E(y) + w. \tag{5.34}$$

If we use the definition of the extractable work (3.53) given as

$$W[\Gamma] := E(x) - E'(y) + Q[\Gamma], \tag{5.35}$$

we find that

$$w = W[\Gamma]. \tag{5.36}$$

Thus, the extractable work is equal to the excited energy of the work qubit. Again, we note that we have to *a priori* tune w before the operation which is exactly equal to Eq. (5.36). If we consider a general work storage system W whose Hamiltonian is given by $H^W = \sum_i E_i^W |i\rangle\langle i|_W$, we obtain

$$\Delta E^W[\Gamma] = W[\Gamma], \tag{5.37}$$

where the energy change of the work storage is found to be equal to the extractable work from the system. If we multiply $p(x)/q(y)$ on both hand-sides of Eq. (5.33), we obtain the detailed fluctuation theorem

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} = \exp(-\beta Q[\Gamma] + \Delta s[\Gamma]), \tag{5.38}$$

where $\Delta s[\Gamma] = \ln p(x) - \ln q(y)$ is the Shannon entropy difference of the system. Let us denote

$$\begin{aligned} \mathcal{F}_\rho(x) &:= E(x) + \beta^{-1}\ln p(x), & (5.39)\\ \mathcal{F}_\sigma(y) &:= E'(y) + \beta^{-1}\ln q(y), & (5.40) \end{aligned}$$

as the initial and final nonequilibrium free energies. Then, we combine Eqs. (5.34) and (5.38) to obtain the detailed fluctuation theorem:

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} = \exp(-\beta w - \beta\Delta\mathcal{F}[\Gamma]), \tag{5.41}$$

where

$$\Delta\mathcal{F}[\Gamma] := \mathcal{F}_\sigma(y) - \mathcal{F}_\rho(x) \tag{5.42}$$

is the nonequilibrium free-energy difference. We also note that by defining

$$\begin{aligned} P[\Gamma]_{\mathrm{can}} &:= P[\Gamma|x]\frac{\mathrm{e}^{-\beta E(x)}}{\sum_{x:p(x)>0}\mathrm{e}^{-\beta E(x)}}, & (5.43)\\ P[\Gamma^\dagger]_{\mathrm{can}} &:= \tilde{P}[\Gamma^\dagger|y]\frac{\mathrm{e}^{-\beta E'(y)}}{Z'}, & (5.44) \end{aligned}$$

the detailed fluctuation theorem (5.41) can be rewritten as

$$\frac{P[\Gamma]_{\mathrm{can}}}{\tilde{P}[\Gamma^\dagger]_{\mathrm{can}}} = \exp(-\beta w - \beta\Delta F^{\mathrm{eq}} + D_0(p||\mathrm{e}^{-\beta H}/Z)). \tag{5.45}$$

Here, we note that $P[\Gamma|x]$ is defined only when the support of x is equal to that of $p(x)$. We use this fact and the normalization factor of the canonical distribution is taken into account in Eq. (5.45). From this relation, we find that the extractable work w via the transition $\rho \to \sigma$ for a unitary operator U can be obtained by using Eq. (5.45), where the transition is given by $\mathrm{e}^{-\beta H + D_0}/Z \to \mathrm{e}^{-\beta H'}/Z'$ by using the same unitary operator U.

In this section, we have shown how we can derive the detailed fluctuation theorem when we consider the thermal operation. In the next chapter, we will show how we can reproduce the single-shot results from the detailed fluctuation theorem as a special case of the main results.

5.6 Relations between majorization and thermo-majorization

In this section, we discuss how thermo-majorization generalizes the idea of majorization.

To explain the idea of majorization, we first introduce the noisy operation given as

$$\mathcal{E}_{\mathrm{noisy}}(\rho) := \mathrm{Tr}_A\left[U\left(\rho\otimes\frac{\mathbb{1}_A}{d_A}\right)U^\dagger\right], \tag{5.46}$$

where A is an auxiliary system whose dimension is d_A and U is an arbitrary unitary operator on the total system. Note that we consider a special case in which the dimensions of the input state ρ

and the output state $\mathcal{E}_{\text{noisy}}(\rho)$ are the same as in Eq. (5.46). The noisy operation (5.46) conserves the maximally mixed state of the system $\mathbb{1}/d$ as

$$\mathcal{E}_{\text{noisy}}\left(\frac{\mathbb{1}}{d}\right) = \frac{\mathbb{1}}{d}, \tag{5.47}$$

where d is the total dimension of the system. We can interpret the maximally mixed state $\mathbb{1}_A/d_A$ as the noise source, and the noisy operation gives a change in the state of the system when it interacts with that noise source. The thermal operation (5.1) can be viewed as a generalization of the noisy operation by replacing the maximally mixed state $\mathbb{1}_A/d_A$ of the auxiliary system by the canonical distribution $\mathrm{e}^{-\beta H^B}/Z^B$ of the heat bath.

Now the notion of majorization is introduced as follows. Let us first denote the diagonal element of ρ and that of σ as $p(x)$ and $q(x)$, respectively. We say p majorizes q if and only if the noisy operation $\mathcal{E}_{\text{noisy}}(\rho) = \sigma$ is possible. We label x and arrange $p(x)$ in a decreasing order:

$$p(x_1) \geq p(x_2) \geq \cdots \geq p(x_d), \tag{5.48}$$

where d is the total dimension of the system. Then, p majorizes q (which is sometimes denoted as $p \succ q$) if and only if

$$\sum_{i=1}^{k} p(x_i) \geq \sum_{i=1}^{k} q(y_i) \quad \text{for } k = 1, 2, \cdots, d. \tag{5.49}$$

If the dimension of the input state ρ and that of the output state σ are different, we use the Lorenz curve (5.7) given by

$$\left(\frac{\mathbb{1}}{d}, \rho\right) = \left(\frac{1}{d}, p(x_1)\right), \left(\frac{2}{d}, \sum_{i=1}^{2} p(x_i)\right), \cdots, (1,1). \tag{5.50}$$

The noisy operation $\mathcal{E}_{\text{noisy}}(\rho) = \sigma$ is possible if and only if the Lorenz curve $(\mathbb{1}/d, \sigma)$ is a subset of $(\mathbb{1}/d, \rho)$. This condition is equivalent to that given by Eq. (5.49) when the dimension of ρ and σ are the same, because the condition (5.49) ensures that each point of the Lorenz curve $(\mathbb{1}/d, \rho)$ is larger than that of the Lorenz curve $(\mathbb{1}/d, \sigma)$ in the vertical direction (in the Y plane).

Thus, majorization gives an ordering relation between distributions, which essentially gives an ordering that brings the distribution closer to a uniform distribution. Thermo-majorization is a generalization of majorization by extending the uniform distribution to a canonical distribution.

5.7 Proof of (2.41)

To show (2.41), we use the monotonicity of the quantum relative entropy

$$S(\rho||\sigma) \geq S(\mathcal{E}(\rho)||\mathcal{E}(\sigma)), \tag{5.51}$$

where $S(\rho||\sigma)$ is a quantum relative entropy and $\mathcal{E}(\bullet)$ is a completely positive, trace-preserving (CPTP) map [40]. It is known that a CPTP map can be expressed as

$$\mathcal{E}(\rho) = \mathrm{Tr}_A\left[U(\rho \otimes \rho^A)U^\dagger\right], \tag{5.52}$$

where ρ^A is a density matrix of an auxiliary system A and U is a unitary operator that acts on the composite system. Note that the thermal operation is a CPTP map because it takes the form of Eq. (5.52). Then, let us take $\sigma = \mathrm{e}^{-\beta H}/Z$ and apply (5.51) to the thermal operation, which takes the form

$$S(\rho||\frac{\mathrm{e}^{-\beta H}}{Z}) \geq S(\mathcal{E}_{\text{thermal}}(\rho)||\frac{\mathrm{e}^{-\beta H}}{Z}), \tag{5.53}$$

where we use Eq. (5.3). Now let us consider the case in which ρ and $\mathcal{E}_{\mathrm{thermal}}(\rho)$ are diagonal in the energy eigenbasis, and denote their diagonal components by p and $\mathcal{E}_{\mathrm{thermal}}(p)$. Because ρ and $\mathrm{e}^{-\beta H}/Z$ are diagonalized by the same basis, the quantum relative entropy is equal to the classical relative entropy in (5.53). Then (5.53) takes the following form:

$$D(p||\frac{\mathrm{e}^{-\beta H}}{Z}) \geq D(\mathcal{E}_{\mathrm{thermal}}(p)||\frac{\mathrm{e}^{-\beta H}}{Z}), \tag{5.54}$$

By combining this result with the definition of the nonequilibrium entropy (2.36), we obtain (2.41).

Chapter 6

Trade-off relation in heat engines

In the previous chapters, we focused on two approaches (single-shot statistical mechanics [29,30] and the second law of thermodynamics) which search for protocols that minimize either work fluctuation or dissipation under nonequilibrium situations. We first show a trade-off relation between work fluctuation and fluctuation in dissipation for arbitrary initial and final states. Then, we give a method to construct explicit protocols that achieve the lower bound of the trade-off relation, which reproduces the deterministic work extraction protocol in the limit of vanishing work fluctuation and the thermodynamically reversible protocol in the limit of vanishing dissipation [74].

6.1 Setup of the main result

Our setup is basically the same as the one discussed in Chap. 3. We consider a system interacting with a single heat bath with inverse temperature β^{-1}, where the coupling between the heat bath and the system is constant (we also consider the case in which the coupling can be turned on and off arbitrary). The dynamics of the system interacting with the heat bath can be arbitrary as far as the detailed fluctuation theorem (3.1) holds, e.g., the Hamiltonian dynamics (in the quantum case, the dynamics generated by the unitary time evolution) and the stochastic dynamics. We assume that we can change the energy level of the Hamiltonian freely during the process, and in particular, a sudden quench process is assumed. We also assume that we can perform the thermal operation during the process, while it generally requires a detailed control of the system-bath coupling. We denote the (given) initial and final states of the system as p_{ini} and p_{fin}. The initial and final Hamiltonians are denoted by H_{λ_0} and H_{λ_1}. The extractable work is defined in Eq. (3.53):

$$W[\Gamma] := E_{\lambda_0}(x) - E_{\lambda_1}(y) + Q[\Gamma], \tag{6.1}$$

and dissipation (the total entropy production) is defined in Eq. (3.51):

$$\sigma[\Gamma] := -\beta(W[\Gamma] + \Delta\mathcal{F}(x,y)). \tag{6.2}$$

Here,

$$\Delta\mathcal{F}(x,y) := \mathcal{F}_{\lambda_1}(y) - \mathcal{F}_{\lambda_0}(x) \tag{6.3}$$

is the nonequilibrium free-energy difference and the nonequilibrium free energy is defined in Eq. (3.49):

$$\begin{aligned} \mathcal{F}_{\lambda_0}(x) &:= E_{\lambda_0}(x) + \beta^{-1}\ln p_{\mathrm{ini}}(x) \\ &= F^{\mathrm{eq}} + \beta^{-1}\ln\frac{p_{\mathrm{ini}}(x)}{p^{\mathrm{can}}_{\lambda_0}(x)}, \end{aligned} \tag{6.4}$$

and a similar relation holds for $\mathcal{F}_{\lambda_1}(y)$. In the following, we derive a fundamental bound on work and dissipation for arbitrary processes which connect $(p_{\mathrm{ini}}, H_{\lambda_0})$ and $(p_{\mathrm{fin}}, H_{\lambda_1})$ in Sec. 6.2.

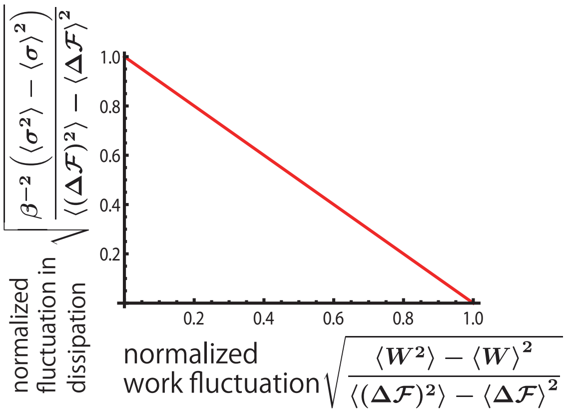

Figure 6.1: **Trade-off relation between work fluctuation and fluctuation in dissipation.** The ordinate shows the normalized standard deviation of dissipation σ and the abscissa shows the normalized standard deviation of work W. The solid line shows the lower bound of the trade-off relation (6.5). For thermalized initial and final states, the lower bound shrinks to a single point in the origin and work fluctuation and fluctuation in dissipation can take arbitrary values. This figure is taken from Ref. [74].

6.2 Trade-off relation between work fluctuation and fluctuation in dissipation

The first main result of our work is the trade-off relation between work and dissipation in terms of their standard deviations (see also Fig. 6.1):

$$\sqrt{\langle W^2\rangle - \langle W\rangle^2} + \beta^{-1}\sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2} \geq \sqrt{\langle(\Delta\mathcal{F})^2\rangle - \langle\Delta\mathcal{F}\rangle^2}. \tag{6.5}$$

The right-hand side of (6.5) is equal to the standard deviation of the nonequilibrium free-energy difference (6.3), which gives the information distance between the initial (final) distribution and the canonical distribution as discussed in Sec. 2.5. If the initial and final states are far from equilibrium, the right-hand side of (6.5) takes on a large value. If we want to reduce work fluctuation, fluctuation in dissipation inevitably takes a large value and vice versa, which implies that work and dissipation cannot take definite values simultaneously. If the initial and final states are given by the canonical distributions, the right-hand side of (6.5) vanish and σ and W can take on arbitrary values. In particular, dissipation and work fluctuation vanishes for a thermodynamically reversible process starting from equilibrium, and Eq. (6.5) is consistent with the conventional second law of thermodynamics.

We note that the right-hand side of (6.5) depends on the path probability $P[\Gamma]$ in general. Thus, there is a possibility that the left-hand side of (6.5) for a process which satisfies the equality condition of (6.5) does not give the minimum value of the sum of the standard deviation of work and that of the total entropy production. If we take $p_{\mathrm{fin}}(y) = p^{\mathrm{can}}_{\lambda_1}(y)$, the right-hand side of (6.5) depends only on $p_{\mathrm{fin}}(y)$ and $p^{\mathrm{can}}_{\lambda_1}(y)$, and if we take $p_{\mathrm{ini}}(x) = p^{\mathrm{can}}_{\lambda_0}(x)$, the right-hand side of (6.5) depends only on $p_{\mathrm{ini}}(x)$ and $p^{\mathrm{can}}_{\lambda_0}(x)$. Thus, if we consider a thermalized final state or an initially thermalized state, the protocol that achieves the equality condition of (6.5) minimizes both work fluctuation and the fluctuation in dissipation as much as possible.

6.3 Proof of the main result

To show the first main result (6.5), we first calculate the variance of $\sigma + \beta W$. By using Eq. (6.2), we obtain

$$\beta^2(\langle W^2\rangle - \langle W\rangle^2) + \langle \sigma^2\rangle - \langle\sigma\rangle^2 + 2\beta(\langle\sigma W\rangle - \langle\sigma\rangle\langle W\rangle) = \langle(\Delta\mathcal{F})^2\rangle - \langle\Delta\mathcal{F}\rangle^2 . \tag{6.6}$$

Next, we consider the property of the variance-covariance matrix defined by

$$V_{ij} := \langle (X_i - \langle X_i\rangle)(X_j - \langle X_j\rangle)\rangle . \tag{6.7}$$

The eigenvalues of V_{ij} are positive semi-definite and thus the determinant of V_{ij} is nonnegative. By taking $X_1 = \beta W$ and $X_2 = \sigma$, and from $\mathrm{Det}[V_{ij}] \geq 0$, we obtain

$$(\langle W^2\rangle - \langle W\rangle^2)(\langle\sigma^2\rangle - \langle\sigma\rangle^2) \geq (\langle\sigma W\rangle - \langle\sigma\rangle\langle W\rangle)^2. \tag{6.8}$$

Let us take the square root of Eq. (6.8) and obtain

$$\beta\sqrt{\langle W^2\rangle - \langle W\rangle^2}\sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2} \geq \langle\sigma W\rangle - \langle\sigma\rangle\langle W\rangle , \tag{6.9}$$

where the above inequality holds trivially in the case of $\langle\sigma W\rangle - \langle\sigma\rangle\langle W\rangle \leq 0$. Combining Eqs. (6.6) and (6.9), we obtain

$$\begin{aligned}\langle(\Delta\mathcal{F})^2\rangle - \langle\Delta\mathcal{F}\rangle^2 &\leq \beta^2(\langle W^2\rangle - \langle W\rangle^2) + \langle\sigma^2\rangle - \langle\sigma\rangle^2 + 2\beta\sqrt{\langle W^2\rangle - \langle W\rangle^2}\sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2} \\ &= \left(\sqrt{\langle W^2\rangle - \langle W\rangle^2} + \beta^{-1}\sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2}\right)^2 . \end{aligned} \tag{6.10}$$

By taking the square root of both hand sides of (6.10), we obtain the trade-off relation (6.5).

6.4 Equality condition of the trade-off relation

The equality condition in (6.5) is satisfied if and only if one of the eigenvalues of the matrix V_{ij} is zero. Let us introduce some constants a and b, and define new variables $X_1' := a\sigma + b\beta W$ and $X_2' := a\sigma - b\beta W$, such that the matrix V_{ij} is diagonalized. The equality condition in (6.5) is re-expressed as $\mathrm{Var}[a\sigma - b\beta W] = 0$, where $\mathrm{Var}[X] := \langle X^2\rangle - \langle X\rangle^2$. Without the loss of generality, we can take $a \geq 0$.

Let us rewrite Eq. (6.2) as

$$\sigma[\Gamma] = -\beta W[\Gamma] - \beta\Delta F^{\mathrm{eq}} + \ln\frac{p_{\mathrm{ini}}(x)}{p^{\mathrm{can}}_{\lambda_0}(x)} - \ln\frac{p_{\mathrm{fin}}(y)}{p^{\mathrm{can}}_{\lambda_1}(y)}. \tag{6.11}$$

Combining the detailed fluctuation theorem (3.1) and Eq. (6.11), we have

$$a\sigma[\Gamma] - b\beta W[\Gamma] = b\beta\Delta F^{\mathrm{eq}} + (a+b)\ln\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]} - b\ln\frac{p_{\mathrm{ini}}(x)}{p^{\mathrm{can}}_{\lambda_0}(x)} + b\ln\frac{p_{\mathrm{fin}}(y)}{p^{\mathrm{can}}_{\lambda_1}(y)}. \tag{6.12}$$

We set the condition

$$a\sigma[\Gamma] - b\beta W[\Gamma] = b\beta\Delta F^{\mathrm{eq}} + c \tag{6.13}$$

to make $\mathrm{Var}[a\sigma - b\beta W] = 0$. Here, we introduced a constant c. Then, Eqs. (6.12) and (6.13) lead to

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]}\left(\frac{p_{\mathrm{ini}}(x)}{p^{\mathrm{can}}_{\lambda_0}(x)}\right)^{\frac{-b}{a+b}}\left(\frac{p_{\mathrm{fin}}(y)}{p^{\mathrm{can}}_{\lambda_1}(y)}\right)^{\frac{b}{a+b}} = \mathrm{e}^{\frac{c}{a+b}}. \tag{6.14}$$

If we fix a protocol λ_t, the forward and backward probability distributions $P[\Gamma]$ and $\tilde{P}[\Gamma^\dagger]$ are uniquely fixed. Thus, the lower bound of the trade-off relation can be achieved if and only if there exists a protocol λ_t that satisfies Eq. (6.14).

6.4.1 Case of an equilibrium final state

We first assume $p_{\rm fin}(y) = p^{\rm can}_{\lambda_1}(y)$ and show the equality condition of the trade-off relation. We discuss the case of a nonequilibrium final state in Sec. 6.4.3.

Under this assumption, the condition (6.14) takes the form

$$\begin{aligned}\tilde{P}[\Gamma^\dagger] &= P[\Gamma]\left(\frac{p_{\rm ini}(x)}{p^{\rm can}_{\lambda_0}(x)}\right)^{\frac{-b}{a+b}} \mathrm{e}^{-\frac{c}{a+b}} \\ &= P[\Gamma|x]p_{\frac{a}{a+b}}(x), \end{aligned} \tag{6.15}$$

where $P[\Gamma|x] := P[\Gamma]/p_{\rm ini}(x)$ is the conditional forward probability conditioned on x and we define the distribution

$$p_{\frac{a}{a+b}}(x) := (p^{\rm can}_{\lambda_0}(x))^{\frac{b}{a+b}}(p_{\rm ini}(x))^{\frac{a}{a+b}}\mathrm{e}^{-\frac{c}{a+b}}. \tag{6.16}$$

This distribution $p_{\frac{a}{a+b}}$ should be normalized, and the constant c is fixed:

$$\mathrm{e}^{\frac{c}{a+b}} = \sum_x (p^{\rm can}_{\lambda_0}(x))^{\frac{b}{a+b}}(p_{\rm ini}(x))^{\frac{a}{a+b}}. \tag{6.17}$$

By introducing the Renyi divergence [75]

$$D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0}) := \frac{1}{\alpha-1}\ln\left[\sum_x (p_{\rm ini}(x))^\alpha (p^{\rm can}_{\lambda_0}(x))^{1-\alpha}\right], \tag{6.18}$$

the constant c takes the form

$$\frac{c}{a+b} = -(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0}), \tag{6.19}$$

with $\alpha = a/(a+b)$. Using Eq. (6.17), σ and W take the following forms:

$$\sigma[\Gamma] = (1-\alpha)\left(\mathcal{D}_{\lambda_0}(x) - D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})\right), \tag{6.20}$$

$$\beta W[\Gamma] = -\beta\Delta F^{\rm eq} + \alpha\mathcal{D}_{\lambda_0}(x) + (1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0}), \tag{6.21}$$

where $\mathcal{D}_{\lambda_0}(x) = \ln p_{\rm ini}(x) - \ln p^{\rm can}_{\lambda_0}(x)$. Note that $\mathcal{F}_{\lambda_0}(x) = F^{\rm eq}_{\lambda_0} + \mathcal{D}_{\lambda_0}(x)$. In particular, the first and the second terms on the left-hand side of Eq. (6.26) are given by

$$\sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2} = \beta^{-1}|1-\alpha|\sqrt{\langle(\mathcal{F}_{\lambda_0})^2\rangle - \langle\mathcal{F}_{\lambda_0}\rangle^2}, \tag{6.22}$$

$$\sqrt{\langle W^2\rangle - \langle W\rangle^2} = \alpha\sqrt{\langle(\mathcal{F}_{\lambda_0})^2\rangle - \langle\mathcal{F}_{\lambda_0}\rangle^2}. \tag{6.23}$$

To satisfy the equality condition in (6.5), α should take the value in the range $0 \le \alpha \le 1$. Note that the existence of an explicit protocol λ_t, which satisfies the condition Eq. (6.15), is needed to achieve the lower bound of the trade-off relation. Substituting Eq. (6.19) into Eq. (6.17), the condition (6.15) is now described as

$$\tilde{P}[\Gamma^\dagger] = P[\Gamma|x]p_\alpha(x), \tag{6.24}$$

with

$$p_\alpha(x) := (p^{\rm can}_{\lambda_0}(x))^{1-\alpha}(p_{\rm ini}(x))^\alpha \mathrm{e}^{(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})}. \tag{6.25}$$

We discuss in the next section that the condition (6.24) is satisfied by implementing λ_t as a thermodynamically reversible protocol starting from $p_\alpha(x)$ and ending in $p_{\rm fin}(y)$. We conclude that the equality condition of the trade-off relation for a thermalized final state

$$\sqrt{\langle W^2\rangle - \langle W\rangle^2} + \beta^{-1}\sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2} \ge \sqrt{\langle(\mathcal{F}_{\lambda_0})^2\rangle - \left\langle\mathcal{F}^2_{\lambda_0}\right\rangle} \tag{6.26}$$

is satisfied if and only if Eqs. (6.20) and (6.21) hold and α takes the value in the range $0 \le \alpha \le 1$.

6.4.2 Explicit protocols that achieve the lower bound of the trade-off relation

Let us discuss the explicit protocols which satisfy Eq. (6.24). We introduce a Hamiltonian $H(p_{\rm ini})$ such that the canonical distribution with respect to $H(p_{\rm ini})$ is equal to the initial distribution:

$$p_{\rm ini}(x) = \frac{{\rm e}^{-\beta E_{p_{\rm ini}}(x)}}{Z(p_{\rm ini})}, \tag{6.27}$$

where $E_{p_{\rm ini}}(x)$ is the eigenenergy of $H(p_{\rm ini})$ and $Z(p_{\rm ini}) = \sum_x \exp(-\beta E_{p_{\rm ini}}(x))$. Then, the distribution p_α is found to be equal to the canonical distribution with respect to the Hamiltonian $H_\alpha(p_{\rm ini})$:

$$\begin{aligned} p_\alpha(x) &= {\rm e}^{-\beta(\alpha E_{p_{\rm ini}}(x)+(1-\alpha)E_{\lambda_0}(x))}\frac{\exp\left((1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})\right)}{(Z(p_{\rm ini}))^\alpha (Z_{\lambda_0})^{1-\alpha}} \\ &= \frac{{\rm e}^{-\beta E^{p_{\rm ini}}_\alpha(x)}}{Z_\alpha}, \end{aligned} \tag{6.28}$$

where $H_\alpha(p_{\rm ini}) = \sum_{x\in X}\{\alpha E_{p_{\rm ini}}(x) + (1-\alpha)E_{\lambda_0}(x)\}$, which has the same support with that of the initial distribution, i.e., $X = \{x|p_{\rm ini}(x) > 0\}$. We note that

$$\begin{aligned} \exp\left(-(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})\right) &= \sum_x (p^{\rm can}_{\lambda_0}(x))^{1-\alpha}(p_{\rm ini}(x))^\alpha \\ &= \sum_x \frac{{\rm e}^{-\beta(\alpha E_{p_{\rm ini}}(x)+(1-\alpha)E_{\lambda_0}(x))}}{(Z(p_{\rm ini}))^\alpha (Z_{\lambda_0})^{1-\alpha}} \\ &= \frac{Z_\alpha}{(Z(p_{\rm ini}))^\alpha (Z_{\lambda_0})^{1-\alpha}}. \end{aligned} \tag{6.29}$$

As the condition (6.24) shows, we need to construct a thermodynamically reversible protocol starting from $(p_\alpha, H_\alpha(p_{\rm ini}))$ and ending at $(p^{\rm can}_{\lambda_1}, H_{\lambda_1})$, where (p, H) denotes a pair of the state p and the Hamiltonian H. Noting that p_α is the canonical distribution with respect to $H_\alpha(p_{\rm ini})$, we should change the Hamiltonian from $H_\alpha(p_{\rm ini})$ to H_{λ_1} in a quasi-static manner and let the system interact with the heat bath:

$$(p_\alpha, H_\alpha(p_{\rm ini})) \overset{\text{quasi-static}}{\longrightarrow} (p^{\rm can}_{\lambda_1}, H_{\lambda_1}). \tag{6.30}$$

Then, the condition (6.24) is satisfied. Noting that the protocol λ_t starts from $(p_{\rm ini}, H_{\lambda_0})$, one should consider a protocol that gives a transition

$$(p_{\rm ini}, H_{\lambda_0}) \to (p_\alpha, H_\alpha(p_{\rm ini})). \tag{6.31}$$

Since the transition Eq. (6.30) is associated with vanishing dissipation and work fluctuation, dissipation and work fluctuation during the process (6.31) are given by Eqs. (6.20) and (6.23). Let us rewrite Eq. (6.20) by using p_α:

$$\begin{aligned} \sigma[\Gamma] &= \ln\frac{p_{\rm ini}(x)}{(p^{\rm can}_{\lambda_0}(x))^{1-\alpha}(p_{\rm ini}(x))^\alpha {\rm e}^{(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})}} \\ &= \ln\frac{p_{\rm ini}(x)}{p_\alpha(x)}, \end{aligned} \tag{6.32}$$

and also rewrite Eq. (6.23) by using $E^{p_{\rm ini}}_\alpha(x)$:

$$\begin{aligned} \sqrt{\langle W^2\rangle - \langle W\rangle^2} &= \sqrt{\langle(\mathcal{D}_{\lambda_0}(x)-\sigma[\Gamma])^2\rangle - \langle \mathcal{D}_{\lambda_0}(x)-\sigma[\Gamma]\rangle^2} \\ &= \sqrt{\left\langle\left(\ln p_\alpha(x) - \ln p^{\rm can}_{\lambda_0}\right)^2\right\rangle - \left\langle \ln p_\alpha(x) - \ln p^{\rm can}_{\lambda_0}\right\rangle^2} \\ &= \sqrt{\left\langle(E_{\lambda_0}(x) - E^{p_{\rm ini}}_\alpha(x))^2\right\rangle - \langle E_{\lambda_0}(x) - E^{p_{\rm ini}}_\alpha(x)\rangle^2}. \end{aligned} \tag{6.33}$$

From Eq. (6.33), we find that the transition (6.31) is partially implemented by quenching the Hamiltonian from H_{λ_0} to $H_\alpha(p_{\rm ini})$ by keeping the initial distribution $p_{\rm ini}(x)$ fixed. We find from Eq. (6.32) that we need to change the distribution from $p_{\rm ini}(x)$ to $p_\alpha(x)$ while the Hamiltonian $H_\alpha(p_{\rm ini})$ is fixed. This can be achieved by letting the system interact with the heat bath and the distribution $p_{\rm ini}(x)$ thermalize to the canonical distribution $p_\alpha(x)$. The complete protocol is determined by combining the above protocols. The explicit protocol which satisfies the condition (6.24) and realizes the lower bound of the trade-off relation (6.26) is given by

$$(p_{\rm ini}, H_{\lambda_0}) \overset{\rm quench}{\longrightarrow} (p_{\rm ini}, H_\alpha(p_{\rm ini})) \overset{\rm thermalize}{\longrightarrow} (p_\alpha, H_\alpha(p_{\rm ini})) \overset{\text{quasi-static}}{\longrightarrow} (p^{\rm can}_{\lambda_1}, H_{\lambda_1}). \tag{6.34}$$

If we can turn off the interaction between the system and the heat bath during the protocols, the quench process can be replaced by an ordinary adiabatic process in which the system is detached from the heat bath.

In the regime of vanishing dissipation ($\alpha = 1$), we only need to perform a quench process. By quenching the Hamiltonian into $H(p_{\rm ini})$, the initial distribution is equal to a canonical distribution with respect to the Hamiltonian $H(p_{\rm ini})$. Thus, dissipation does not occur when we let the system interact with the heat bath. This protocol is thermodynamically reversible, i.e., if we perform the time-reversal of the protocol, the final state $p^{\rm can}_{\lambda_1}$ is transformed into $p_{\rm ini}$.

In the regime of vanishing work fluctuation ($\alpha = 0$), we first need to quench the Hamiltonian from H_{λ_0} to $H_{\alpha=0}$, which can be done by rising the energy levels which are initially unoccupied ($\{E_{\lambda_0}(x)\}_{x\notin X}$) to infinity. Then, we let the system interact with the heat bath, and then the system thermalizes. Note that the nonequilibrium free-energy

$$\begin{aligned}\mathcal{F}_0(p_{\rm ini}) &= F^{\rm eq}_{\lambda_0} - \beta^{-1}\ln\left[\sum_{x\in X} p^{\rm can}_{\lambda_0}(x)\right] \\ &= -\beta^{-1}\ln\left[\sum_{x\in X}\exp(-\beta E_{\lambda_0}(x))\right],\end{aligned} \tag{6.35}$$

is equal to the equilibrium local free energy whose support X is the same as the initial distribution $p_{\rm ini}(x)$. One can increase the free energy of the system $\mathcal{F}_0(p_{\rm ini}) \geq F^{\rm eq}_{\lambda_0}$ and keep the amount of work fluctuation vanishing.

For general α, we quench the Hamiltonian by changing the energy levels by mixing the thermodynamically reversible ($\alpha = 1$) protocol and the single-shot ($\alpha = 0$) protocol by the ratio $\alpha : 1-\alpha$.

6.4.3 The case of a nonequilibrium final state

In this section, we consider the case in which the final state is out of equilibrium. Let us derive the condition which achieves the lower bound of the trade-off relation by rewriting Eq. (6.14) as

$$P[\Gamma|x]p_\alpha(x) = \tilde{P}[\Gamma^\dagger|y]p'_\alpha(y), \tag{6.36}$$

where we use Eq. (6.25). Here, we define the distribution p'_α by

$$p'_\alpha(y) := (p_{\rm fin}(y))^\alpha (p^{\rm can}_{\lambda_1}(y))^{1-\alpha} {\rm e}^{c-(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})}. \tag{6.37}$$

Let us introduce a Hamiltonian $H(p_{\rm fin})$ such that the canonical distribution with respect to $H(p_{\rm fin})$ is equal to the final distribution:

$$p_{\rm fin}(y) = \frac{{\rm e}^{-\beta E_{p_{\rm fin}}(y)}}{Z(p_{\rm fin})}, \tag{6.38}$$

where $E_{p_{\rm fin}}(y)$ is the eigenenergy of $H(p_{\rm fin})$ and $Z(p_{\rm fin}) = \sum_y \exp(-\beta E_{p_{\rm fin}}(y))$. Let us also define a Hamiltonian $H_\alpha(p_{\rm fin}) = \sum_{y\in Y'}\{\alpha E_{p_{\rm fin}}(y) + (1-\alpha)E_{\lambda_1}(y)\}$, which has the same support as that of the final distribution, i.e., $Y' = \{y | p_{\rm fin}(y) > 0\}$. The canonical distribution $\mathrm{e}^{-\beta H_\alpha(p_{\rm fin})}/Z_\alpha(p_{\rm fin})$ is given by

$$\frac{\mathrm{e}^{-\beta E_\alpha^{p_{\rm fin}}(y)}}{Z_\alpha(p_{\rm fin})} = (p_{\rm fin}(y))^\alpha (p_{\lambda_1}^{\rm can}(y))^{1-\alpha} \mathrm{e}^{(1-\alpha)D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can})}. \tag{6.39}$$

Then, $p'_\alpha(y)$ is equal to the canonical distribution with respect to $H_\alpha(p_{\rm fin})$ if the support of the distribution $p'_\alpha(y)$ is the same as that of the final distribution $p_{\rm fin}(y)$ by comparing Eqs. (6.37) and (6.39) and the normalization of p'_α:

$$p'_\alpha(y) = \frac{\mathrm{e}^{-\beta E_\alpha^{p_{\rm fin}}(y)}}{Z_\alpha(p_{\rm fin})} \quad \text{if the support of } p'_\alpha \text{ is equal to } Y'. \tag{6.40}$$

Then, a similar argument in deriving (6.34) suggests us that the explicit protocol to achieve the lower bound of the trade-off relation is given by

$$\begin{aligned}
&(p_{\rm ini}, H_{\lambda_0}) \xrightarrow{\text{quench}} (p_{\rm ini}, H_\alpha(p_{\rm ini})) \xrightarrow{\text{thermalize}} (p_\alpha, H_\alpha(p_{\rm ini})) \xrightarrow{\text{quasi-static}} (\frac{\mathrm{e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}, H_\alpha(p_{\rm fin})) \\
&\xrightarrow{?} (p_{\rm fin}, H_\alpha(p_{\rm fin})) \xrightarrow{\text{quench}} (p_{\rm fin}, H_{\lambda_1}),
\end{aligned} \tag{6.41}$$

where the transition

$$(\frac{\mathrm{e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}, H_\alpha(p_{\rm fin})) \xrightarrow{?} (p_{\rm fin}, H_\alpha(p_{\rm fin})) \tag{6.42}$$

needs the time-reversal of the thermalization process. Since this is not possible by a natural time-evolution of the system which interacts with a single heat bath, we denote this transition as $\xrightarrow{?}$. Thus, we need to consider an alternative protocol that replaces Eq. (6.42). From the conditions (6.13) and (6.14), the work fluctuation is given by

$$\begin{aligned}
\langle W^2 \rangle - \langle W \rangle^2 &= \left\langle (E_{\lambda_0}(x) - E_\alpha^{p_{\rm ini}}(x))^2 \right\rangle - \langle E_{\lambda_0}(x) - E_\alpha^{p_{\rm ini}}(x) \rangle^2 \\
&\quad + \left\langle (E_{\lambda_1}(y) - E_\alpha^{p_{\rm fin}}(y))^2 \right\rangle - \langle E_{\lambda_1}(y) - E_\alpha^{p_{\rm fin}}(y) \rangle^2,
\end{aligned} \tag{6.43}$$

which comes from the two quench processes in Eq. (6.41). This observation requires us to construct an alternative protocol with vanishing work fluctuation. Thus, we consider a thermal operation as the alternative protocol. We determine a distribution q which can be transformed to $p_{\rm fin}$ by using the thermo-majorization criteria [30] introduced in Sec. 5.2. As discussed in the following sections, we have two choices of q: (1) a local canonical distribution $p'_\alpha(y)$ with respect to the Hamiltonian $H_\alpha(p_{\rm fin})$ whose support is restricted to Y and (2) the joint distribution $\mathrm{e}^{-\beta H'_\alpha}/Z_\alpha(p_{\rm fin}) \otimes p_\epsilon^A$ by introducing an auxiliary system A whose initial distribution is given by p_ϵ^A.

(a) The case of preparing a local canonical distribution

Let us consider the first case: prepare a local canonical distribution $p'_\alpha(y) := \mathrm{e}^{-\beta H_\alpha^Y(p_{\rm fin})}/Z_\alpha^Y(p_{\rm fin})$ whose support is restricted to Y. We plot two Lorentz curves

$$\left(p_{\rm fin}, \frac{\mathrm{e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}\right) \tag{6.44}$$

and

$$\left(\frac{\mathrm{e}^{-\beta H_\alpha^Y(p_{\rm fin})}}{Z_\alpha^Y(p_{\rm fin})}, \frac{\mathrm{e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}\right) \tag{6.45}$$

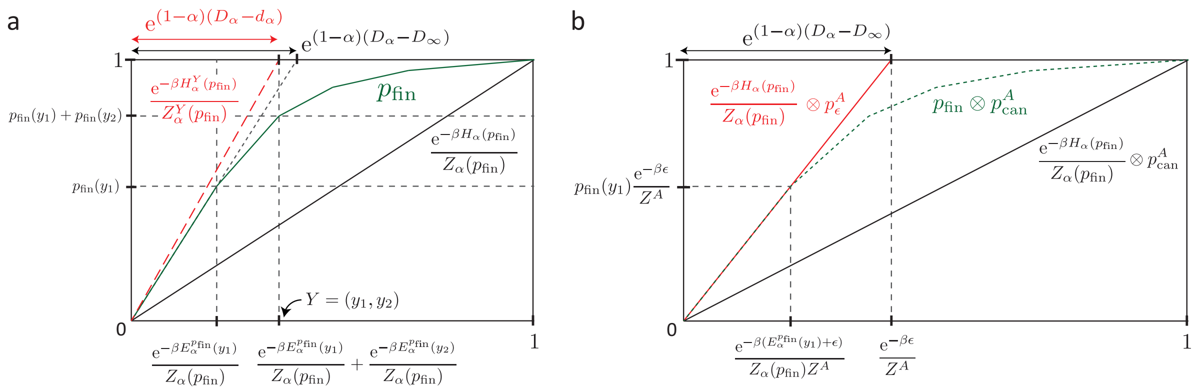

Figure 6.2: **Lorentz curve.** (a) We plot two Lorentz curves $(p_{\rm fin}, \frac{e^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})})$ and $(\frac{e^{-\beta H^Y_\alpha(p_{\rm fin})}}{Z^Y_\alpha(p_{\rm fin})}, \frac{e^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})})$, where the curve $\frac{e^{-\beta H^Y_\alpha(p_{\rm fin})}}{Z^Y_\alpha(p_{\rm fin})}$ is represented by the dashed line and $p_{\rm fin}$ is represented by the solid curve. The slope of the dotted line is given by $\exp\left[(1-\alpha)(D_\infty(p_{\rm fin}||p^{\rm can}_{\lambda_1}) - D_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1}))\right]$. The support Y is defined as the largest support such that the slope of the line p'_α is equal to or larger than that of the dotted line. Then, the thermo-majorization criterion tells us that the local equilibrium state $\frac{e^{-\beta H^Y_\alpha(p_{\rm fin})}}{Z^Y_\alpha(p_{\rm fin})}$ can be transformed into $p_{\rm fin}$ via a thermal operation. (b) The case of introducing an auxiliary system. By introducing an auxiliary system, the joint probability distribution $e^{-\beta H_\alpha(p_{\rm fin})}/Z_\alpha(p_{\rm fin}) \otimes p^A_\epsilon$ (red line) can be transformed into $p_{\rm fin}\otimes p^A_{\rm can}$ (blue dotted curve) if the slope of the former line is equal to or larger than the largest slope of the latter curve, i.e., $p_{\rm fin}(y_1)Z_\alpha(p_{\rm fin})/e^{-\beta E^{p_{\rm fin}}_\alpha(y_1)}$. This figure is taken from Ref. [74].

as in Fig. 6.2. From the thermo-majorization criteria, a transition

$$\left(\frac{e^{-\beta H^Y_\alpha(p_{\rm fin})}}{Z^Y_\alpha(p_{\rm fin})}, H_\alpha(p_{\rm fin})\right) \overset{\text{thermal operation}}{\longrightarrow} (p_{\rm fin}, H_\alpha(p_{\rm fin})) \tag{6.46}$$

is possible if the former curve is a subset of the latter curve. This is equivalent to the condition that the slope of the curve $(\frac{e^{-\beta H^Y_\alpha(p_{\rm fin})}}{Z^Y_\alpha(p_{\rm fin})}, \frac{e^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})})$ is larger than that of the curve connecting two points: $(0,0)$ and $(e^{-\beta E^{p_{\rm fin}}_\alpha(y_1)}/Z_\alpha(p_{\rm fin}), p_{\rm fin}(y_1))$ satisfying

$$\frac{1}{\sum_{y\in Y}(Z_\alpha(p_{\rm fin}))^{-1}e^{-\beta E^{p_{\rm fin}}_\alpha(y)}} \geq \frac{p_{\rm fin}(y_1)}{(Z_\alpha(p_{\rm fin}))^{-1}e^{-\beta E^{p_{\rm fin}}_\alpha(y_1)}}. \tag{6.47}$$

We choose the largest support of Y that satisfies the condition (6.47). Now the normalization condition of $\frac{e^{-\beta H^Y_\alpha(p_{\rm fin})}}{Z^Y_\alpha(p_{\rm fin})}$ given by Eq. (6.37) determines the constant c:

$$c = (1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0}) + (1-\alpha)d_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1}), \tag{6.48}$$

where

$$(1-\alpha)d_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1}) := -\ln\sum_{y\in Y}(p_{\rm fin}(y))^\alpha(p^{\rm can}_{\lambda_1}(y))^{1-\alpha}. \tag{6.49}$$

Using Eq. (6.39), the right-hand side of (6.47) can be expressed as

$$\begin{aligned}\frac{Z_\alpha(p_{\rm fin})p_{\rm fin}(y_1)}{e^{-\beta E^{p_{\rm fin}}_\alpha(y_1)}} &= \frac{p_{\rm fin}(y_1)}{(p_{\rm fin}(y_1))^\alpha(p^{\rm can}(y_1))^{1-\alpha}}e^{-(1-\alpha)D_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1})} \\ &= \exp\left[(1-\alpha)(D_\infty(p_{\rm fin}||p^{\rm can}_{\lambda_1}) - D_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1}))\right], \end{aligned} \tag{6.50}$$

where we used the Renyi divergence of order ∞. The left-hand side of (6.47) can be expressed as

$$\begin{aligned}\frac{Z_\alpha(p_{\rm fin})}{\sum_{y\in Y}{\rm e}^{-\beta E_\alpha^{p_{\rm fin}}(y)}} &= \frac{1}{\sum_{y\in Y}(p_{\rm fin}(y_1))^\alpha(p^{\rm can}(y_1))^{1-\alpha}}{\rm e}^{-(1-\alpha)D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can})}\\ &= \exp\left[(1-\alpha)(d_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can})-D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can}))\right]. \end{aligned}\tag{6.51}$$

The condition (6.47) to determine Y is rewritten as

$$\exp\left[(1-\alpha)(d_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can})-D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can}))\right]\geq\exp\left[(1-\alpha)(D_\infty(p_{\rm fin}||p_{\lambda_1}^{\rm can})-D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can}))\right],\tag{6.52}$$

which can be further simplified as

$$d_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can})\geq D_\infty(p_{\rm fin}||p_{\lambda_1}^{\rm can}).\tag{6.53}$$

The explicit protocol to achieve the lower bound of the trade-off relation is given by

$$\begin{aligned}&(p_{\rm ini},H_{\lambda_0})\xrightarrow{\text{quench}}(p_{\rm ini},H_\alpha(p_{\rm ini}))\xrightarrow{\text{thermalize}}\left(\frac{{\rm e}^{-\beta H_\alpha^Y(p_{\rm fin})}}{Z_\alpha^Y(p_{\rm fin})},H_\alpha(p_{\rm ini})\right)\\ &\xrightarrow{\text{quasi-static}}\left(\frac{{\rm e}^{-\beta H_\alpha^Y(p_{\rm fin})}}{Z_\alpha^Y(p_{\rm fin})},H_\alpha(p_{\rm fin})\right)\xrightarrow{\text{thermal operation}}(p_{\rm fin},H_\alpha(p_{\rm fin}))\xrightarrow{\text{quench}}(p_{\rm fin},H_{\lambda_1}).\end{aligned}\tag{6.54}$$

Combining Eqs. (6.13), (6.14) and (6.48), the functional forms of the dissipation and the work that achieve the lower bound of the trade-off relation (6.5) are determined:

$$\sigma[\Gamma] = (1-\alpha)(\mathcal{D}_{\lambda_0}(x)-D_\alpha(p_{\rm ini}||p_{\lambda_0}^{\rm can})-\mathcal{D}_{\lambda_1}(y)+d_\alpha),\tag{6.55}$$

$$\beta W[\Gamma] = -\beta\Delta F^{\rm eq}+(1-\alpha)(D_\alpha(p_{\rm ini}||p_{\lambda_0}^{\rm can})-d_\alpha)+\alpha(\mathcal{D}_{\lambda_0}(x)-\mathcal{D}_{\lambda_1}(y)).\tag{6.56}$$

We also note that the averaged values of Eqs. (6.55) and (6.56) are given as

$$\langle\sigma\rangle = (1-\alpha)(D(p_{\rm ini}||p_{\lambda_0}^{\rm can})-D_\alpha(p_{\rm ini}||p_{\lambda_0}^{\rm can})-D(p_{\rm fin}||p_{\lambda_1}^{\rm can})+d_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can})),\tag{6.57}$$

$$\begin{aligned}\beta\langle W\rangle &= -\beta\Delta F^{\rm eq}+(1-\alpha)(D_\alpha(p_{\rm ini}||p_{\lambda_0}^{\rm can})-d_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can}))\\ &\quad+\alpha(D(p_{\rm ini}||p_{\lambda_0}^{\rm can})-D(p_{\rm fin}||p_{\lambda_1}^{\rm can})).\end{aligned}\tag{6.58}$$

(b) The case with an auxiliary system

Let us consider the second case: introduce a two-level auxiliary system A, whose eigenenergies are given by $E^A(0)=0,\ E^A(1)=\epsilon$. The excited state of A is denoted by p_ϵ^A, i.e., $p_\epsilon^A(0)=0$ and $p_\epsilon^A(1)=1$. We consider a transition

$$\left(\frac{{\rm e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}\otimes p_\epsilon^A,H_\alpha(p_{\rm fin})\otimes H^A\right)\xrightarrow{\text{thermal operation}}(p_{\rm fin}\otimes p_{\rm can}^A,H_\alpha(p_{\rm fin})\otimes H^A),\tag{6.59}$$

where H^A is the Hamiltonian of A, $p_{\rm can}^A={\rm e}^{-\beta H^A}/Z^A$ is the canonical distribution and $Z^A=1+{\rm e}^{-\beta\epsilon}$ is the partition function. Using the thermo-majorization criteria, this operation is possible if ϵ satisfies (see Fig. 6.2. (b))

$$\exp\left[(1-\alpha)(D_\infty(p_{\rm fin}||p_{\lambda_1}^{\rm can})-D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can}))\right]=\frac{p_{\rm fin}(y_1)p_{\rm can}^A(0)}{p_{\rm can}^A(0){\rm e}^{-\beta E_\alpha^{p_{\rm fin}}(y_1)}/Z_\alpha(p_{\rm fin})}=\frac{1}{{\rm e}^{-\beta\epsilon}/Z^A}.\tag{6.60}$$

The introduction of an auxiliary system changes Eq. (6.36) into

$$P[\Gamma|x,z]p_\alpha(x)p_\epsilon^A(z)=\tilde{P}[\Gamma^\dagger|y,z]\frac{{\rm e}^{-\beta E_\alpha^{p_{\rm fin}}(y)}}{Z_\alpha(p_{\rm fin})}p_{\rm can}(z){\rm e}^{c-(1-\alpha)(D_\alpha(p_{\rm ini}||p_{\lambda_0}^{\rm can})+D_\alpha(p_{\rm fin}||p_{\lambda_1}^{\rm can}))}.\tag{6.61}$$

Since the probability distribution

$$p_{\rm can}(z){\rm e}^{c-(1-\alpha)(D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})+D_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1}))} \tag{6.62}$$

which appears in the right-hand side of Eq. (6.61) should be normalized, the constant c is determined by

$$c=(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})+(1-\alpha)D_\alpha(p_{\rm fin}||p^{\rm can}_{\lambda_1})-\ln p^A_{\rm can}(1). \tag{6.63}$$

In deriving Eq. (6.63), we note that the support of the distribution in Eq. (6.62) is equal to $Z=\{z=1\}$ by comparing both had sides of Eq. (6.61). Combining Eqs. (6.60) and (6.63), we obtain

$$c=(1-\alpha)D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})+(1-\alpha)D_\infty(p_{\rm fin}||p^{\rm can}_{\lambda_1}). \tag{6.64}$$

The explicit protocol to achieve the lower bound of the trade-off relation is given by

$$\begin{aligned}
&(p_{\rm ini},H_{\lambda_0})\overset{\rm quench}{\longrightarrow}(p_{\rm ini},H_\alpha(p_{\rm ini}))\overset{\rm thermalize}{\longrightarrow}(p_\alpha,H_\alpha(p_{\rm ini}))\overset{\text{quasi-static}}{\longrightarrow}\left(\frac{{\rm e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})},H_\alpha(p_{\rm fin})\right)\\
&\overset{{\rm add}\ A}{\longrightarrow}\left(\frac{{\rm e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}\otimes p^A_\epsilon,H_\alpha(p_{\rm fin})\otimes H^A\right)\overset{\text{thermal operation}}{\longrightarrow}(p_{\rm fin}\otimes p^A_{\rm can},H_\alpha(p_{\rm fin})\otimes H^A)\\
&\overset{\rm quench}{\longrightarrow}(p_{\rm fin}\otimes p^A_{\rm can},H_{\lambda_1}\otimes H^A).
\end{aligned} \tag{6.65}$$

Combining Eqs. (6.13), (6.14) and (6.64), the functional forms of the dissipation and the work that achieve the lower bound of the trade-off relation (6.5) are determined as follows

$$\sigma[\Gamma] = (1-\alpha)(\mathcal{D}_{\lambda_0}(x)-D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})-\mathcal{D}_{\lambda_1}(y)+D_\infty), \tag{6.66}$$

$$\beta W[\Gamma] = -\beta\Delta F^{\rm eq}+(1-\alpha)(D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})-D_\infty)+\alpha(\mathcal{D}_{\lambda_0}(x)-\mathcal{D}_{\lambda_1}(y)). \tag{6.67}$$

We note that the averaged values of Eqs. (6.66) and (6.67) are given by

$$\langle\sigma\rangle = (1-\alpha)(D(p_{\rm ini}||p^{\rm can}_{\lambda_0})-D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})-D(p_{\rm fin}||p^{\rm can}_{\lambda_1})+D_\infty(p_{\rm fin}||p^{\rm can}_{\lambda_1})), \tag{6.68}$$

$$\begin{aligned}\beta\langle W\rangle &= -\beta\Delta F^{\rm eq}+(1-\alpha)(D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})-D_\infty(p_{\rm fin}||p^{\rm can}_{\lambda_1}))\\ &\quad+\alpha(D(p_{\rm ini}||p^{\rm can}_{\lambda_0})-D(p_{\rm fin}||p^{\rm can}_{\lambda_1})).\end{aligned} \tag{6.69}$$

By comparing Eqs. (6.57) and (6.68), we find that the minimum amount of dissipation takes a smaller value if we introduce an auxiliary system, compared to the case without an auxiliary system by using Eq. (6.53).

6.5 Explicit protocol and information-theoretic quantities

In this section, we analyze the extractable work $\langle W\rangle$ and the dissipated work $\langle W_{\rm dis}\rangle=\beta^{-1}\langle\sigma\rangle$ during each step of the protocol given in Eq. (6.65) which achieves the lower bound of the trade-off relation (6.5) (see also Fig. 6.3).

1. Quench process $(p_{\rm ini},H_{\lambda_0})\to(p_{\rm ini},H_\alpha(p_{\rm ini}))$.

   The extractable work during this quench process is given by

$$\begin{aligned}\langle W\rangle_1 &= \sum_x p_{\rm ini}(x)(E_{\lambda_0}(x)-E^{p_{\rm ini}}_\alpha(x))\\ &= \alpha\sum_x p_{\rm ini}(x)(E_{\lambda_0}(x)-E_{p_{\rm ini}}(x))\\ &= \alpha\langle\mathcal{F}_{\lambda_0}\rangle-\alpha F^{\rm eq}(p_{\rm ini}).\end{aligned} \tag{6.70}$$

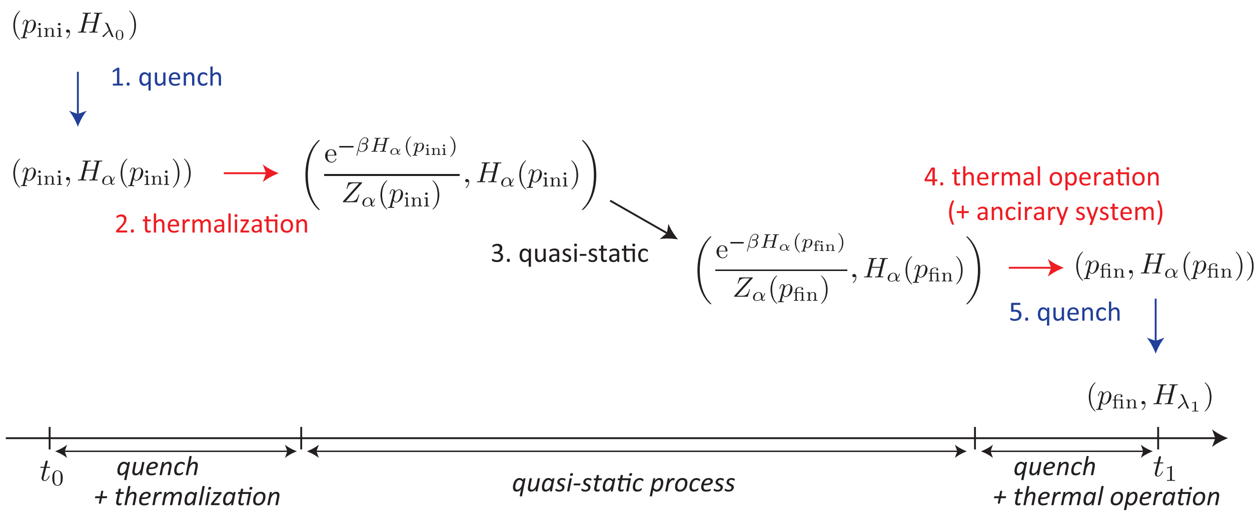

Figure 6.3: Protocol achieving the lower bound of the trade-off relation. We denote (p, H) as a pair of the state p and the Hamiltonian H, and a change in the Hamiltonian is shown in a vertical direction and a change in the state is shown in a horizontal direction.

Note that the nonequilibrium free energy $\langle \mathcal{F}_{\lambda_0} \rangle$ is the reversible work (the maximum extractable work) from a nonequilibrium state. From Eq. (6.70), we find that only a part of this nonequilibrium free energy $\alpha \langle \mathcal{F}_{\lambda_0} \rangle$ can be extracted by the quench process.

2. Thermalization process $(p_{\rm ini}, H_\alpha(p_{\rm ini})) \to {\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini}), H_\alpha(p_{\rm ini}))$.

The extractable work vanishes and the dissipation is given by the nonequilibrium free-energy difference:

$$\begin{aligned}\langle W_{\rm dis}\rangle_2 &= \mathcal{F}(p_{\rm ini}, H_\alpha(p_{\rm ini})) - \mathcal{F}({\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini}), H_\alpha(p_{\rm ini})) \\ &= (1-\alpha)(\langle \mathcal{F}_{\lambda_0}\rangle - \mathcal{F}_\alpha(p_{\rm ini})) \\ &= \beta^{-1} D(p_{\rm ini}||{\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini})). \end{aligned} \quad (6.71)$$

Note that before the thermalization process, the unexpended nonequilibrium free energy is given by $(1-\alpha)\langle \mathcal{F}_{\lambda_0}\rangle$. From Eq. (6.71), we find that this unexpended nonequilibrium free energy is partly lost during the thermalization and the remaining free energy is given by $(1-\alpha)\mathcal{F}_\alpha(p_{\rm ini})$, where

$$\mathcal{F}_\alpha(p_{\rm ini}) := \beta^{-1} D_\alpha(p_{\rm ini}||p_0^{\rm can}) + F^{\rm eq}(p_{\lambda_0}^{\rm can}) \quad (6.72)$$

is the α-generalization of the nonequilibrium free energy. This amount of free energy can be extracted as work by the following quasi-static process. We also note that $\beta^{-1}(1-\alpha)D_\alpha$ quantifies the amount of the increased free energy due to the mixture of two Hamiltonians H_{λ_0} and $H(p_{\rm ini})$ by the ratio $1-\alpha : \alpha$ (which means $H_\alpha = (1-\alpha)H_{\lambda_0} + \alpha H(p_{\rm ini})$) as follows:

$$F^{\rm eq}({\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini})) = (1-\alpha)F^{\rm eq}(p_{\lambda_0}^{\rm can}) + \alpha F^{\rm eq}(p_{\rm ini}) + \beta^{-1}(1-\alpha)D_\alpha(p_{\rm ini}||p_{\lambda_0}^{\rm can}). \quad (6.73)$$

This amount of the "increased free-energy" $\beta^{-1}(1-\alpha)D_\alpha$ corresponds to the extractable work from the system after the thermalization process ($(1-\alpha)\mathcal{F}_\alpha(p_{\rm ini})$). Also, we note that

$$\langle \sigma \rangle_2 = D(p_{\rm ini}||{\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini})) \quad (6.74)$$

gives the information distance between the initial state and the auxiliary intermediate canonical distribution ${\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini})$.

3. Quasi-static process $(\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{ini}})}/Z_\alpha(p_{\mathrm{ini}}), H_\alpha(p_{\mathrm{ini}})) \to (\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}/Z_\alpha(p_{\mathrm{fin}}), H_\alpha(p_{\mathrm{fin}}))$.

   The extractable work is equal to the equilibrium free-energy difference:

$$\begin{aligned}\langle W\rangle_3 &= F^{\mathrm{eq}}\left(\frac{\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{ini}})}}{Z_\alpha(p_{\mathrm{ini}})}\right) - F^{\mathrm{eq}}\left(\frac{\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}}{Z_\alpha(p_{\mathrm{fin}})}\right) \\ &= (1-\alpha)\mathcal{F}_\alpha(p_{\mathrm{ini}}) + \alpha F^{\mathrm{eq}}(p_{\mathrm{ini}}) - (1-\alpha)\mathcal{F}_\alpha(p_{\mathrm{fin}}) - \alpha F^{\mathrm{eq}}(p_{\mathrm{fin}}). \end{aligned} \tag{6.75}$$

   Here, we use the α-generalization of the nonequilibrium free energy corresponding to the final state:

$$\mathcal{F}_\alpha(p_{\mathrm{fin}}) = \beta^{-1} D_\alpha(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1}) + F^{\mathrm{eq}}(p^{\mathrm{can}}_{\lambda_1}). \tag{6.76}$$

4. Thermal operation $\left(\frac{\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}}{Z_\alpha(p_{\mathrm{fin}})} \otimes p^A_\epsilon, H_\alpha(p_{\mathrm{fin}}) \otimes H^A\right) \longrightarrow (p_{\mathrm{fin}} \otimes p^A_{\mathrm{can}}, H_\alpha(p_{\mathrm{fin}}) \otimes H^A)$.

   The extractable work vanishes and the dissipation is given by the nonequilibrium free-energy difference:

$$\begin{aligned}\langle W_{\mathrm{dis}}\rangle_4 &= \mathcal{F}\left(\frac{\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}}{Z_\alpha(p_{\mathrm{fin}})} \otimes p^A_\epsilon, H_\alpha(p_{\mathrm{fin}}) \otimes H^A\right) - \mathcal{F}(p_{\mathrm{fin}} \otimes p^A_{\mathrm{can}}, H_\alpha(p_{\mathrm{fin}}) \otimes H^A) \\ &= \beta^{-1} D(p^A_\epsilon||p^A_{\mathrm{can}}) - \beta^{-1} D(p_{\mathrm{fin}}||\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}/Z_\alpha(p_{\mathrm{fin}})) \\ &= -\beta^{-1}\ln\left(\frac{\mathrm{e}^{-\beta\epsilon}}{Z^A}\right) - \beta^{-1}(1-\alpha)(D(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1}) - D_\alpha(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1})) \\ &= \beta^{-1}(1-\alpha)(D_\infty(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1}) - D(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1})), \end{aligned} \tag{6.77}$$

   where we used Eq. (6.60) in deriving the third equality. We note that

$$\begin{aligned}\langle \sigma\rangle_4 &= (1-\alpha)(D_\infty(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1}) - D(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1})) \\ &= D_\infty(p_{\mathrm{fin}}||\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}/Z_\alpha(p_{\mathrm{fin}})) \end{aligned} \tag{6.78}$$

   gives the minimum amount of dissipation if we create the final state p_{fin} starting from the auxiliary intermediate canonical distribution $\mathrm{e}^{-\beta H_\alpha(p_{\mathrm{fin}})}/Z_\alpha(p_{\mathrm{fin}})$ via the thermal operation.

5. Quench process $(p_{\mathrm{fin}}, H_\alpha(p_{\mathrm{fin}})) \to (p_{\mathrm{fin}}, H_{\lambda_1})$.

   The extractable work is given by

$$\langle W\rangle_5 = \sum_y p_{\mathrm{fin}}(y)(E^{p_{\mathrm{fin}}}_\alpha(y) - E_{\lambda_1}(y)) = -\alpha\,\langle \mathcal{F}_{\lambda_1}\rangle + \alpha F^{\mathrm{eq}}(p_{\mathrm{fin}}). \tag{6.79}$$

We conclude that the total amount of dissipation is given by

$$\langle\sigma\rangle = (1-\alpha)(D(p_{\mathrm{ini}}||p^{\mathrm{can}}_{\lambda_0}) - D_\alpha(p_{\mathrm{ini}}||p^{\mathrm{can}}_{\lambda_0}) + D_\infty(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1}) - D(p_{\mathrm{fin}}||p^{\mathrm{can}}_{\lambda_1}), \tag{6.80}$$

and it quantifies the amount of dissipation that occurs during the thermalization and the thermal operation. Equivalently, the average value of the extractable work is given by

$$\langle W\rangle = -\alpha\,\langle \Delta\mathcal{F}\rangle + (1-\alpha)\mathcal{F}_\alpha(p_{\mathrm{ini}}) - (1-\alpha)\mathcal{F}_\infty(p_{\mathrm{fin}}), \tag{6.81}$$

where the first, the second and the third term on the right-hand side of Eq. (6.81) is the extractable work from the two quench processes, the free energy which remains in the system after the thermalization and the free energy needed to complete the thermal operation, respectively.

As we discuss in Sec. 6.2, either by taking the initial state or the final state as the canonical distribution, the left-hand side of (6.5) is independent of path probabilities and takes a constant value.

Let us consider the case in which the final state is given by the canonical distribution, i.e., $p_{\rm fin}(y) = p^{\rm can}_{\lambda_1}(y)$. In this case, the explicit protocol that achieves the lower bound of the trade-off relation (6.5) is given by (6.34):

$$(p_{\rm ini}, H_{\lambda_0}) \stackrel{\rm quench}{\longrightarrow} (p_{\rm ini}, H_\alpha(p_{\rm ini})) \stackrel{\rm thermalize}{\longrightarrow} ({\rm e}^{-\beta H_\alpha(p_{\rm ini})}/Z_\alpha(p_{\rm ini}), H_\alpha(p_{\rm ini})) \stackrel{\text{quasi-static}}{\longrightarrow} (p^{\rm can}_{\lambda_1}, H_{\lambda_1}). \tag{6.82}$$

We note this protocol (6.82) minimizes both work fluctuation and fluctuation in dissipation simultaneously and they take the values

$$\begin{aligned} {\rm Var}[W] &= {\rm Var}[\alpha \mathcal{F}_{\lambda_0}], & (6.83)\\ {\rm Var}[\sigma] &= {\rm Var}[(1-\alpha)\mathcal{F}_{\lambda_0}], & (6.84)\end{aligned}$$

for all $\alpha \in [0,1]$. We also note that the amount of dissipation and the extractable work satisfies

$$\begin{aligned} \langle \sigma \rangle &= (1-\alpha)(D(p_{\rm ini}||p^{\rm can}_{\lambda_0}) - D_\alpha(p_{\rm ini}||p^{\rm can}_{\lambda_0})), & (6.85)\\ \langle W \rangle &= \alpha \langle \mathcal{F}_{\lambda_0}\rangle + (1-\alpha)\mathcal{F}_\alpha(p_{\rm ini}) - F^{\rm eq}(p^{\rm can}_{\lambda_1}). & (6.86)\end{aligned}$$

We can also consider the case of a nonequilibrium final distribution starting from a thermalized initial state $p_{\rm ini}(x) = p^{\rm can}_{\lambda_0}(x)$. In this case, the explicit protocol that achieves the lower bound of the trade-off relation (6.5) is given by (6.34):

$$\begin{aligned} &(p^{\rm can}_{\lambda 0}, H_{\lambda_0}) \stackrel{\text{quasi-static}}{\longrightarrow} \left(\frac{{\rm e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})}, H_\alpha(p_{\rm fin})\right) \stackrel{{\rm add}\ A}{\longrightarrow} \left(\frac{{\rm e}^{-\beta H_\alpha(p_{\rm fin})}}{Z_\alpha(p_{\rm fin})} \otimes p^A_\epsilon, H_\alpha(p_{\rm fin}) \otimes H^A\right)\\ &\stackrel{\text{thermal operation}}{\longrightarrow} (p_{\rm fin} \otimes p^A_{\rm can}, H_\alpha(p_{\rm fin}) \otimes H^A) \stackrel{\rm quench}{\longrightarrow} (p_{\rm fin} \otimes p^A_{\rm can}, H_{\lambda_1} \otimes H^A). \end{aligned} \tag{6.87}$$

We note this protocol (6.87) minimizes both work fluctuation and fluctuation in dissipation simultaneously and they take the values

$$\begin{aligned} {\rm Var}[W] &= {\rm Var}[\alpha \mathcal{F}_{\lambda_1}], & (6.88)\\ {\rm Var}[\sigma] &= {\rm Var}[(1-\alpha)\mathcal{F}_{\lambda_1}], & (6.89)\end{aligned}$$

for all $\alpha \in [0,1]$. We also note that the amount of dissipation and the extractable work satisfy

$$\begin{aligned} \langle \sigma \rangle &= (1-\alpha)(D_\infty(p_{\rm fin}||p^{\rm can}_{\lambda_1}) - D(p_{\rm fin}||p^{\rm can}_{\lambda_1})), & (6.90)\\ \langle W \rangle &= F^{\rm eq}(p^{\rm can}_{\lambda_0}) - \alpha \langle \mathcal{F}_{\lambda_1}\rangle - (1-\alpha)\mathcal{F}_\infty(p_{\rm fin}). & (6.91)\end{aligned}$$

From (6.82), we find that a combination of a quench process, thermalization and a quasi-static process minimize the fluctuation in dissipation and the work fluctuation simultaneously for a protocol that starts from a nonequilibrium state and ends at a thermalized state. A similar relation holds in the case of a thermalized initial state and a nonequilibrium final state. In this case, a combination of a quasi-static process, a thermal operation and a quench process (6.87) minimizes dissipation and work fluctuation simultaneously. Note that ${\rm Var}[W] = 0$ holds for thermalization and a thermal operation and ${\rm Var}[\sigma] = 0$ holds for a quench process. We also note that both ${\rm Var}[W] = 0$ and ${\rm Var}[\sigma] = 0$ holds for a quasi-static process.

6.6 Comparison with the previous studies

In this section, we consider the protocol (6.54) that achieves the lower bound of the trade-off relation (6.5) and compare it with the previous studies (the deterministic work extraction protocol [29, 30] and the thermodynamically reversible protocol [1, 3, 4]).

For $\alpha = 1$, the protocol (6.54) consists of two quench processes and a quasi-static process, which reproduces the thermodynamically protocol with vanishing dissipation. From Eqs. (6.80) and (6.81), the dissipation and the extractable work are given by $\langle\sigma\rangle = 0$ and $\langle W\rangle = -\langle\Delta\mathcal{F}\rangle$, respectively. Note that these values are equal to the lower bound of the second law of thermodynamics $\langle\sigma\rangle \geq 0$ and $\langle W\rangle \leq -\Delta\langle\mathcal{F}\rangle$. In this case, the amount of work fluctuation takes a large value $\sqrt{\langle W^2\rangle - \langle W\rangle^2} = \sqrt{\langle(\Delta\mathcal{F})^2\rangle - \langle\Delta\mathcal{F}\rangle^2}$.

Next, let us consider the case of $\alpha = 0$. In this case, the protocols (6.82) and (6.87) reproduce the single-shot protocols [29,30]. Also, Eqs. (6.86) and (6.91) take the form $\langle W\rangle = \mathcal{F}_0(p_{\rm ini}) - F^{\rm eq}(p^{\rm can}_{\lambda_1})$ and $\langle W\rangle = F^{\rm eq}(p^{\rm can}_{\lambda_0}) - \mathcal{F}_\infty(p_{\rm fin})$, respectively and reproduce the single-shot results. Note that we can increase the free energy of the system $F^{\rm eq}_{\lambda_0} \leq \mathcal{F}_0(p_{\rm ini})$ by raising the energy levels which are initially unoccupied. This amount of free energy can be extracted as the "fluctuation-free" work. Let the initial and final Hamiltonians to be the same and consider a protocol starting from a nonequilibrium state $p_{\rm ini}$ and ending at a thermalized state. Then, the maximal extractable work is given by $\mathcal{F}_0$. If we consider the time-reversal of this process, i.e., start from a thermalized state and end with $p_{\rm ini}$, the minimal work cost is given by $\mathcal{F}_\infty$. Since $\mathcal{F}_\infty \geq \mathcal{F}_0$, we have a clear irreversibility in this setup [30] because the amount of dissipation cannot vanish during the thermalization process and also during the thermal operation. Note the method to obtain the protocol (6.82) can be used to derive the deterministic work extraction protocol from the detailed fluctuation theorem. This observation may offer us a key relation between two different methods (the fluctuation theorems and the single-shot statistical mechanics) which are used to analyze thermodynamics of small systems.

Because we consider a single heat bath in our setup, the obtained results cannot be applied to situations involving multiple heat baths. One crucial reason is that the nonequilibrium free-energy acquires a thermodynamic meaning only when we let the nonequilibrium system interact with a single heat bath. We do not know whether we can assign a thermodynamic meaning to the nonequilibrium free energy for multiple heat baths or we can define a generalized nonequilibrium free energy for multiple heat baths.

The method we used to derive the trade-off relations can be used to more general settings and also for different stochastic quantities. We only need an equality relation between two stochastic quantities to derive an inequality relation which is similar to the first main result (6.5). However, we note that the existence of the equality condition is nontrivial and that it should depend on the relations which are satisfied by those stochastic quantities (in the present case, dissipation follows the detailed fluctuation theorem).

6.7 Application to information heat engines

In this section, we apply the obtained trade-off relations to information heat engines.

6.7.1 Trade-off relation

For simplicity, we consider a classical system and assume that the premeasurement state $p^S(x)$ is given by the canonical distribution. A measurement is done on the system by recording the measurement outcome a on the memory M, where the joint probability distribution of the postmeasurement state is given by $p^{SM}(x, a)$. We also denote the probability of obtaining the measurement outcome a as $p^M(a)$. Then, the information acquired by this measurement process is quantified by the classical mutual information

$$I^{SM}(x,a) := \ln \frac{p^{SM}(x,a)}{p^S(x)p^M(a)}. \tag{6.92}$$

By performing a feedback control depending on the measurement outcome, one can extract work from the system up to [18]

$$\langle W_{\max}\rangle := -\Delta F^S_{\rm eq} + \beta^{-1}\left\langle I^{SM}\right\rangle, \tag{6.93}$$

where $\Delta F^S_{\rm eq}$ is the equilibrium free-energy difference of the system. We define dissipation as the difference between the maximum amount of extractable work and the actually extracted work:

$$\sigma[\Gamma] := -\beta(W^S[\Gamma] + \Delta F^S_{\rm eq}) + I^{SM}(x,a). \tag{6.94}$$

If we introduce the nonequilibrium postmeasurement state conditioned on the measurement outcome $p^{S|M}(x|a) := p^{SM}(x,a)/p^M(a)$, We can define the nonequilibrium free energy

$$\mathcal{F}(p^{S|M}) := F^S_{\lambda_0} + \beta^{-1} I^{SM}(x,a) = E^S_{\lambda_0}(x) - \beta^{-1} S[p^{S|M}(x|a)], \tag{6.95}$$

which gives the distance between the postmeasurement state and the canonical distribution $p^S(x) = \exp(-\beta(E^S_{\lambda_0}(x) - F^S_{\lambda_0}))$.

By replacing $\Delta\mathcal{F}$ by I^{SM} in Sec. 6.3, we can derive the following trade-off relation

$$\beta\sqrt{\langle W^2\rangle - \langle W\rangle^2} + \sqrt{\langle \sigma^2\rangle - \langle \sigma\rangle^2} \geq \sqrt{\langle (I^{SM})^2\rangle - \langle I^{SM}\rangle^2}, \tag{6.96}$$

showing that the fluctuation in work and dissipation are bounded from below by the degree of fluctuation in the obtained information. The right-hand side of (6.96) does not depend on the trajectory of the system, i.e. it only depends on $p^{SM}(x,a)$, $p^S(x)$ and $p^M(a)$. Note that Eq. (6.96) is a generalization of the first result (6.5) to information heat engines. If the initial and final states are out of equilibrium, the quantity $\Delta\mathcal{F}$ also enters in the right-hand side of Eq. (6.96).

6.7.2 Equality condition of the trade-off relation

Let us consider the equality condition of the trade-off relation for information heat engines. We use a similar method carried out in Sec. 6.4. We set the condition

$$\begin{aligned} \alpha\sigma[\Gamma] - (1-\alpha)\beta W[\Gamma] &= \sigma[\Gamma] + (1-\alpha)\beta\Delta F^{\rm eq} - (1-\alpha) I^{SM}(x,a) \\ &= (1-\alpha)\beta\Delta F^{\rm eq} + c \end{aligned} \tag{6.97}$$

to make $\mathrm{Var}[\alpha\sigma - (1-\alpha)\beta W] = 0$. Combined with the detailed fluctuation theorem (6.5), we obtain

$$\frac{P[\Gamma]}{\tilde{P}[\Gamma^\dagger]}\left(\frac{p^{SM}(x,a)}{p^S(x)p^M(a)}\right)^{\alpha-1} = \mathrm{e}^c. \tag{6.98}$$

Using the relation $P[\Gamma] = P[\Gamma|x,a]p^{SM}(x,a)$, Eq. (6.98) is rewritten as

$$P[\Gamma|x,a]p^{SM}_\alpha(x,a) = \tilde{P}[\Gamma^\dagger], \tag{6.99}$$

where

$$p^{SM}_\alpha(x,a) := (p^{SM}(x,a))^\alpha (p^S(x)p^M(a))^{1-\alpha}\mathrm{e}^{(1-\alpha)I^{SM}_\alpha}, \tag{6.100}$$

and c is determined as

$$c = -(1-\alpha)I^{SM}_\alpha. \tag{6.101}$$

Then, the functional forms of σ and W which achieve the lower bound of the trade-off relations are given by

$$\sigma_\alpha[\Gamma] := (1-\alpha)(I^{SM}(x,a) - I^{SM}_\alpha), \tag{6.102}$$

$$\beta W_\alpha[\Gamma] := \alpha I^{SM}(x,a) + (1-\alpha)I^{SM}_\alpha - \beta\Delta F^S_{\rm eq}. \tag{6.103}$$

6.7.3 Explicit protocol

Next, let us consider the explicit protocol which achieves the lower bound of the trade-off relation. First, let us define a Hamiltonian H^S_a depending on the measurement outcome a which satisfies

$$p^{S|M}(x|a) = \frac{\mathrm{e}^{-\beta E^S_a(x)}}{Z_a}. \tag{6.104}$$

Then, the conditional distribution $p_\alpha(x,a)/p^M(a)$ can be written as

$$\begin{aligned} p^{S|M}_\alpha(x|a) := \frac{p_\alpha(x,a)}{p^M(a)} &= (p^{S|M}(x|a))^\alpha (p^S(x))^{1-\alpha} \mathrm{e}^{(1-\alpha)I^{SM}_\alpha} \\ &= \frac{\exp(-\beta(\alpha E^S_a(x) + (1-\alpha)E^S_{\lambda_0}(x)))}{Z_{\alpha,a}}, \end{aligned} \tag{6.105}$$

which is equal to the canonical distribution with respect to the Hamiltonian $H_{\alpha,a} = \alpha H^S_a + (1-\alpha)H^S_{\lambda_0}$. Thus, the explicit protocol can be expressed as

$$\begin{aligned} &(p^S(x), H^S_{\lambda_0}(x)) \xrightarrow{\text{measurement}} (p^{S|M}(x|a), H^S_{\lambda_0}) \xrightarrow{\text{quench}} (p^{S|M}(x|a), H^S_{\alpha,a}) \\ &\xrightarrow{\text{thermalization}} (p^{S|M}_\alpha(x|a), H^S_{\alpha,a}) \xrightarrow{\text{quasi-static}} (p^{\text{can}}_{\lambda_1}(y), H^S_{\lambda_1}). \end{aligned} \tag{6.106}$$

Note the processes starting from $(p^{S|M}(x|a), H^S_{\lambda_0})$ in (6.106) is equivalent to the one given in Eq. (6.82), where the postmeasurement state $p^{S|M}(x|a)$ corresponds to the initial state in Eq. (6.82).

6.7.4 Example: the Szilard engine

Here, we consider the Szilard engine as an example, following the setup used in the experiment done in the single electron box to realize the Szilard engine [25,26]. Due to the Coulomb blockade, one can prepare a system such that only up to a single electron can enter into the quantum dot at once. Then, we can describe the system as a two-level system with $n=0$ (no electron in the quantum dot) and $n=1$ (a single electron in the quantum dot). The internal energy of the system is given by

$$E(n) := E_{\mathrm{c}}(n - n_{\mathrm{g}})^2, \tag{6.107}$$

where E_{c} is the total charging energy, n_{g} is a control parameter which can tune the energy level of the system by changing the gate voltage. Here, we consider the case in which the initial and final Hamiltonians are degenerate i.e., $n_{\mathrm{g}} = \frac{1}{2}$. Then, the initial state of the system is given by the canonical distribution

$$p^S(n=0) = p^S(n=1) = \frac{1}{2}. \tag{6.108}$$

We next perform a measurement to obtain information about the internal state of the system. Let the measurement error rate be ϵ. Then, the joint probability distribution of the system being n and the measurement outcome being m is given by

$$\begin{aligned} p^{SM}(n,m) &= \frac{1-\epsilon}{2} \qquad \text{for } m = n; \\ p^{SM}(n,m) &= \frac{\epsilon}{2} \qquad \text{for } m \neq n. \end{aligned} \tag{6.109}$$

We perform a feedback control by lowering the energy level of the state $n=m$ relative to that of the other state by changing n_{g}. This change of the energy level is done quickly (ideally, we quench the Hamiltonian). Then, we slowly (ideally, quasi-statically) change the energy levels back to the degenerate point.

Note that the postmeasurement state for the measurement outcome $m = 0$ is given by

$$p^S(n = 0|m = 0) = 1 - \epsilon, \quad p^S(n = 1|m = 0) = \epsilon. \tag{6.110}$$

Let us denote the Hamiltonian $H^S_{m=0}$ that gives a canonical distribution which is equal to the postmeasurement state:

$$\frac{\mathrm{e}^{-\beta E^S_{m=0}(n)}}{Z_{m=0}} = p^S(n|m = 0), \tag{6.111}$$

Also, let us define the Hamiltonian

$$H^S_{\alpha,m=0} := \alpha H^S_{m=0} + (1 - \alpha)H^S(n_{\mathrm{g}} = \frac{1}{2}). \tag{6.112}$$

If we quench the initial Hamiltonian into $H^S_{\alpha,m=0}$, let the system thermalize and quasi-statically drive the Hamiltonian back to $H^S(n_{\mathrm{g}} = \frac{1}{2})$, we can achieve the lower bound of the trade-off relation for a given α. If we denote the control parameter $n_{\alpha,m=0}$ which gives the Hamiltonian $H^S_{\alpha,m=0}$, the explicit protocol is given by

$$\begin{aligned}&(p^S(n), n_{\mathrm{g}} = \frac{1}{2}) \overset{\text{measurement}}{\longrightarrow} (p^S(n|m = 0), n_{\mathrm{g}} = \frac{1}{2}) \overset{\text{quench}}{\longrightarrow} (p^S(n|m = 0), n_{\mathrm{g}} = n_{\alpha,m=0})\\ &\overset{\text{quasi-static}}{\longrightarrow} (p^S(n), n_{\mathrm{g}} = \frac{1}{2}).\end{aligned} \tag{6.113}$$

where we denote (p, n_{g}) as the pair of the state p and the control parameter n_{g}. For the measurement outcome $m = 1$, we can construct an explicit protocol in a similar manner.

We note that the fundamental lower bound of the trade-off relation is independent of the choice of the feedback control protocol and in this case, the information-theoretic quantities take the form

$$\langle I^{SM} \rangle = ln2 + (1 - \epsilon)\ln(1 - \epsilon) + \epsilon \ln \epsilon, \tag{6.114}$$

$$\langle (I^{SM})^2 \rangle - \langle I^{SM} \rangle^2 = \epsilon(1 - \epsilon) \left[(\ln 2\epsilon)^2 + (\ln 2(1 - \epsilon))^2\right], \tag{6.115}$$

$$I^{SM}_\alpha = \frac{1}{\alpha - 1} \ln \left[\frac{1}{2}(2\epsilon)^\alpha + \frac{1}{2}(2(1 - \epsilon))^\alpha\right]. \tag{6.116}$$

Chapter 7

Summary and discussion

We have studied the quantum fluctuation theorems under measurement and feedback control and the fundamental bound on work fluctuation and fluctuation in dissipation by using the information distances that measure the degrees of nonequilibriumness of the initial and final states. The obtained results in this thesis provide us with methods to control thermodynamic systems under nonequilibrium situations and perform quantum measurement and feedback control in an efficient way as well as to suppress work fluctuation and fluctuation in dissipation as much as possible. In the following, we summarize these results and discuss some outstanding issues concerning them.

After a brief introduction of the subjects treated in this thesis in chapter 1, we have reviewed the second law of thermodynamics starting from and/or ending at arbitrary nonequilibrium states in chapter 2. We have also reviewed a few selected topics of information processing protocols and analyzed them by using the second law.

In chapter 3, we have reviewed the derivation of the nonequilibrium equalities which are the basic tools for studying nonequilibrium statistical mechanics.

In chapter 4, we have derived the quantum versions of the detailed fluctuation theorems and the integral-type fluctuation theorems for both the measurement process and the feedback control process. From the obtained equality relations, we can analyze the nonequilibrium dynamics of the general quantum measurement process as well as the feedback control process and search for protocols which suppress dissipation as much as possible. We have clarified the difference between the quantum system and the classical system by focusing on the obtained information via the measurement process. An extension of the fluctuation theorems to the case of initial states which possess coherence in the energy eigenbasis and also an extension to states that are initially entangled both remain outstanding issues.

In chapter 5, we have reviewed the single-shot statistical mechanics. The deterministic work extraction protocol is reviewed, and the extractable work under this protocol has been characterized by the information-theoretic quantities.

In chapter 6, we have derived a fundamental trade-off relation between work fluctuation and fluctuation in dissipation for arbitrary initial and final states. Furthermore, we have proposed a method to construct explicit protocols that achieve the lower bound of the trade-off relation. In the regime of vanishing work fluctuation, we reproduce the deterministic work extraction protocol [29, 30]. In the regime of vanishing fluctuation in dissipation, we reproduce the thermodynamically reversible protocol [3]. If we either take the initial state or the final state as a canonical distribution, the obtained explicit protocol is found to minimize both work fluctuation and fluctuation in dissipation as much as possible, which provide us with methods to control thermodynamic systems under nonequilibrium situations. We have generalized the trade-off relation to the case of information heat engines, and show that the obtained information enters into the trade-off relation which measures the nonequilibriumness of the postmeasurement state of the system. If we fix the amount of work fluctuation, the obtained protocol minimizes fluctuation in dissipation, and we expect that the average value of the dissipation is also minimized by the same protocol. I intend to derive this relation (the work-fluctuation dissipation trade-off relation) in a future study. We have assumed that the system interacts with a single heat bath when we derive the main result. For multiple heat baths, we cannot operationally give a thermodynamic meaning to the nonequilibrium free energy, and a straightforward generalization of the trade-off relation to the case of multiple heat baths is not possible. However, the detailed fluctuation theorem is obtained for a system interacting with two heat baths at different temperatures, which is applicable to the Carnot engine-like processes [89,90]. It is challenging to start from the detailed fluctuation theorem obtained in Ref. [89, 90] and derive a generalized trade-off relation, and also to identify the information-theoretic quantity which takes place of the nonequilibrium free energy in this multiple heat bath situation. Finding the relation between nonequilibrium equalities and the single-shot statistical mechanics is interesting, where we partly show their relation by obtaining a method to derive the single-shot result from a detailed fluctuatoin theorem in Sec. 6.5.

In the future, I intend to derive trade-off relations for different thermodynamic quantities. Note that we have only used the property of the variance-covariance matrix to prove the trade-off relation between work fluctuation and fluctuation in dissipation. This suggests that similar inequalities can be obtained for two stochastic quantities which are related by some equality relations (in the present case, the two stochastic quantities are the extractable work and the total entropy production, which are related by Eq. (3.51)). Since the obtained trade-off relation between work fluctuation and dissipation generalizes the fluctuation-dissipation relation to a nonequilibrium situation, we expect that the generalized method of the trade-off relation allows us to analyze the relation between thermodynamic quantities beyond the linear-response regime.

Bibliography

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

[2] S. Deffner and E. Lutz, Preprint at http://arXiv.org/abs/1201.3888 (2012).

[3] M. Esposito and C. Van den Broeck, Euro. Phys. Lett. **95**, 40004 (2011).

[4] H.-H. Hasegawa, J. Ishikawa, K. Takara and D. J. Driebe, Phys. Lett. A **374**, 1001-1004 (2010).

[5] K. Takara, H.-H. Hasegawa and K. J. Driebe, Phys. Lett. A **375**, 88-92 (2010).

[6] H. S. Leff and A. F. Rex, *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing* (Institute of Physics Publishing, 2003).

[7] K. Maruyama, F. Nori and V. Vedral, Rev. Mod. Phys. **81**, 1-23 (2009).

[8] J. C. Maxwell, *Theory of Heat* (Appleton, London, 1871).

[9] L. Szilard, Z. Phys. **53**, 840 (1929).

[10] L. Brillouin, J. Appl. Phys. **22**, 334-337 (1951).

[11] R. Landauer, IBM J. Res. Dev. **5**, 183-191 (1961).

[12] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).

[13] T. Sagawa, in *Lectures on Quantum Computing, Thermodynamics, and Statistical Physics*, edited by M. Nakahara and S. Tanaka (World Scientific, Singapore, 2012).

[14] Thermodynamics of Information Processing in Small Systems, T. Sagawa (Springer, 2013).

[15] T. Sagawa and M. Ueda, Phys. Rev. Lett. **100**, 080403 (2008).

[16] T. Sagawa and M. Ueda, Phys. Rev. Lett. **102**, 250602 (2009).

[17] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010).

[18] T. Sagawa and M. Ueda, Phys. Rev. Lett. **109**, 180602 (2012).

[19] J. M. Horowitz and J. M. R. Parrondo, EPL. **95**, 1 (2011).

[20] J. M. Horowitz, T. Sagawa and J. M. R. Parrondo, Phys. Rev. Lett. **111**, 010602 (2013).

[21] L. del Rio, J. Aberg, R. Renner, O. Dahlsten and V. Vedral, Nature **474**, 61-63 (2011).

[22] P. Sartori, L. Granger, C. F. Lee, and J. M. Horowitz, PLoS Comput. Biol. **10**(12): e1003974 (2014).

[23] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki and M. Sano, Nat. Phys. **6**, 988 (2010).

[24] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider and E. Lutz, Nature **483**, 187-189 (2012).

[25] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola, Phys. Rev. Lett. **113**, 030601 (2014).

[26] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin, PNAS **111**, 13786 (2014).

[27] E. Roldán, I. A. Martinez, J. M. R. Parrondo and D. Petrov, Nature Phys. **10**, 457-461 (2014).

[28] J. P. Pekola, *Nat. Phys.* **11**, 118 (2015).

[29] Aberg, J. Nat. Commun. **4**, 1925 (2013).

[30] Horodecki, M. and Oppenheim, J. Nat. Commun. **4**, 2059 (2013).

[31] Brandão, F. G. S. L., Horodecki, M., Oppenheim, J., Renes, J. M. and Spekkens, R. W. Phys. Rev. Lett. **111**, 250404 (2013).

[32] Brandão, F. G. S. L., Horodecki, M., Ng, N. H. Y., Oppenheim, J. and Wehner, S. PNAS **112**, 3275 (2015).

[33] Lostaglio, M., Jennings, D. and Rudolph, T. Nat. Commun. **6**, 6383 (2015).

[34] D. Egloff, O. C. O. Dahlsten, R. Renner and V. Vedral, arXiv:1207.0434.

[35] L. Wang, R. Colbeck and R. Renner, IEEE Int. Symp. Inf. Theory 2009, 1804 (2009).

[36] S. Salek and K. Wiesner, arXiv:1504.05111.

[37] P. Faist, J. Oppenheim and R. Renner, New J. Phys. **17** 043003 (2015).

[38] H. Y. Halpern, A. J. Garner, O. C. Dahlsten and V. Vedral, arXiv:1409.3878.

[39] C. Perry, P. Cwiklinski, J. Anders, M. Horodecki and J. Oppenheim, arXiv:1511.06553.

[40] Nielsen, M. A. and Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[41] A. W. Marshall, I. Olkin, B. C. Arnold, Inequalities: Theory of Majorization and its Applications, Springer (2010).

[42] G. Gour, M. P. Muller, V. Narasimhachar, R. W. Spekkens and N. Y. Halpern, Phys. Rep. **583**, 1-58 (2015).

[43] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, Jr and C. Bustamante, Nature **437**, 231-234 (2005).

[44] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr. and C. Bustamante, Sience **296**, 1832 (2002).

[45] Y. Utsumi, D. S. Golubev, M. Marthaler, K. Saito, T. Fujisawa, and G. Schön, Phys. Rev. B **81** 125331 (2010).

[46] J. V. Koski, T. Sagawa, O-P. Saira, Y. Yoon, A. Kutovonen, P. Solinas, M. Möttönen, T. Ala-Nissila and J. P. Pekola, Nature Phys. **9**, 644-648 (2013).

[47] Shuoming An, Jing-Ning Zhang, Mark Um, Dingshun Lv, Yao Lu, Junhua Zhang, Zhang-Qi Yin, H. T. Quan and Kihwan Kim, Nat. Phys. **11**, 193-199 (2015).

[48] D. J. Evans, G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).

[49] G. Gallavotti, and E. G. D. Cohen, Phys. Rev. Lett. **74** 2694 (1995).

[50] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).

[51] C. Jarzynski, Phys. Rev. E.**56**, 5018 (1997).

[52] G. E. Crooks, Phys. Rev. E **60**, 2721-2726 (1999).

[53] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[54] J. M. Horowitz and S. Vaikuntanathan, Phys. Rev. E **82**, 061120 (2010).

[55] J. M. Horowitz, Phys. Rev. E **85**, 031110 (2012).

[56] Y. Murashita K. Funo, and M. Ueda, Phys. Rev. E **90**, 042110 (2014).

[57] Y. Ashida, K. Funo Y. Murashita, and M. Ueda, Phys. Rev. E **90**, 052125 (2014).

[58] P. Talkner, E. Lutz and P. Hänggi, Phys. Rev. E **75** 050102 (2007).

[59] P. Talkner and P. Hänggi, J. Phys. A: Math. Theor. **40** F569 (2007).

[60] H. Tasaki, Preprint cond-mat/0009244 (1999).

[61] J. Kurchan, Preprint cond-mat/0007360 (2000).

[62] Crooks, G. E. J. Stat. Mech.: Theor. Exp. P10023 (2008).

[63] Campisi, M., Hänggi, P. and Talkner, P. Rev. Mod. Phys. **83** 771 (2011).

[64] J. M. Horowitz and J. M. R. Parrondo, New J. Phys. **15** 085028 (2013).

[65] R. Chetrite, K. Malick, J. Stat. Phys. **148**, 480-501 (2012).

[66] Y. Subasi and B. L. Hu, Phys. Rev. E **85**, 011112 (2012).

[67] M. Esposito, U. Harbola, S. Mukamel, Rev. Mod. Phys. **81**, 1665 (2009).

[68] T. Monnai, Phys. Rev. E **72**, 027102 (2005).

[69] M. Campisi, P. Talkner, and P. Hänggi, Phys. Rev. Lett. **105**, 140601 (2010).

[70] M. Morikuni and H. Tasaki, J. Stat. Phys. **143**, 1 (2011).

[71] H. Tasaki, Preprint cond-mat/1308.3776 (2013).

[72] K. Funo, Y. Watanabe, and M. Ueda, Phys. Rev. E **88**, 052121 (2013).

[73] K. Funo, Y. Murashita, and M. Ueda, New J. Phys. **17**, 075005 (2015).

[74] K. Funo and M. Ueda, Phys. Rev. Lett. **115**, 260601 (2015).

[75] Rényi, A., Rev. Int. Stat. Inst. **33**, 1 (1965).

[76] van Erven, T. and Harremoës, P. IEEE Int. Symp. Inf. Theory 2010, 1335 (2010).

[77] van Erven, T. and Harremoës, P. IEEE Int. Symp. Inf. Theory, vol.60, no.7, 3797-3820 (2014).

[78] D. J. Wineland, M. Barrett, J. Britton, J. Chiaverini, B. DeMarco, W. M. Itano, B. Jelenković, C. Langer, D. Leibfried, V. Meyer, T. Rosenband, T. Schätz, Phil. Trans. R. Soc. Lond. A **361**, 1349-1361 (2003).

[79] M. H. Devorte and R. J. Schoelkopf, Science, **339**, 6124, 1169-1174 (2013).

[80] I. Bloch Nature **453**, 1016-1022 (2008).

[81] J. M. Horowitz and K. Jacobs, Phys. Rev. E. **89**, 042134 (2014).

[82] T. Albash, D. A. Lidar, M. Marvian, and P. Zanardi, Phys. Rev. E. **88**, 032146 (2013).

[83] J. P. Pekola, V. F. Maisi, S. Kafanov, N. Chekurov, A. Kemppinen, Yu. A. Pashkin, O.-P. Saira, M. Mottonen and J. S. Tsai, Phys. Rev. Lett. **105**, 026803 (2010).

[84] M. Frenzel, D. Jennings and T. Rudolph, Phys. Rev. E **90**, 052136 (2014).

[85] M. Hayashi and H. Tajima, arXiv:1504.06150.

[86] P. Talkner and P. Hanggi, arXiv:1512.02516.

[87] A. J. Roncaglia, F. Cerisola and J. P. Paz, Phys. Rev. Lett. **113**, 250601 (2014).

[88] G. de Chiara, A. J. Roncaglia and J. P. Paz, New J. Phys. **17**, 035004 (2015).

[89] N. A. Sinitsyn, J. Phys. A: Math. Theor. **44** 405001 (2011).

[90] M. Campisi, J. Phys. A: Math. Theor. **47** 245001 (2014).