

Master Thesis

*The exchange fluctuation theorem
in quantum systems*

Shiho Akagawa

January 8, 2009

*Department of Physics, Graduate School of Science,
The University of Tokyo,
4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan*

Abstract

We study the heat transfer between two finite quantum systems initially at different temperatures. We find that a recently proposed fluctuation theorem for heat exchange, namely the exchange fluctuation theorem [C. Jarzynski and D. K. Wójcik, *Phys. Rev. Lett.* **92**, 230602 (2004)], does not hold in the presence of a finite heat transfer. The deviation from the exchange fluctuation theorem is explicitly calculated for simple models. We confirm that the deviation has a finite value as far as the coupling between the two systems is finite. We discover a condition for the exchange fluctuation theorem to hold in the presence of a finite heat transfer. We check the condition analytically and numerically.

Contents

1	Introduction	5
2	Fluctuation theorems and Jarzynski's equality	7
2.1	Transient fluctuation theorem	7
2.2	Steady-state fluctuation theorem	8
2.3	Jarzynski's equality	8
3	Two-time measurement scheme	11
3.1	Procedure of a two-time measurement	11
3.2	Joint probability of the measurement	13
3.3	Time-evolution operator	15
4	Examination of the Exchange Fluctuation Theorem	19
4.1	Original Derivation of the XFT	19
4.1.1	Time-reversal operator in quantum mechanics	20
4.1.2	Derivation of the XFT by Jarzynski and Wójcik	21
4.2	Deviation from the XFT and the IXFT	25
4.2.1	Deviation from the XFT and the IXFT	26
4.2.2	Example: a two-spin 1/2 system	28
4.2.3	Example: coupled harmonic oscillators	32
4.3	Commutable-coupling condition	33
4.3.1	XFT and IXFT under the commutable-coupling condition	33
4.3.2	Example: a two-spin 1/2 system	37
4.3.3	Example: coupled harmonic oscillators	37
5	Conclusions	41
A	Time-reversed evolution	43

1 Introduction

In the present thesis, we study the heat transfer between two finite quantum systems initially at different temperatures. Our main interest is to understand statistical-physical properties of the heat transfer between finite systems out of equilibrium. In this situation, Jarzynski and Wójcik [1] have presented a symmetry relation about the heat transfer. They referred to the relation as the exchange fluctuation theorem (XFT). We show that the exchange fluctuation theorem does not hold generally in the presence of a finite heat transfer for quantum systems. We also find a specific condition with which the exchange fluctuation theorem holds rigorously in the presence of a finite heat transfer.

The development of the modern techniques of microscopic manipulation enables us to treat small systems, for example, nano-devices and molecular motors. In such small systems, classical thermodynamics is not well applicable to quantification of the heat flow or the work. At the nano- and micro-scales, the available thermal energy per degree of freedom is comparable to the energy of the small systems. This thermal energy enhances fluctuations, whose effects measurably appear in such small systems. Thus, we cannot apply classical thermodynamics to these small systems. We need a substitute framework if we try to design or control nano-devices and molecular motors as macroscopic heat engines.

Because of this motivation, we will analyze thermal fluctuations in small quantum-mechanical systems in the present thesis, but we first briefly review facts that have been discovered for fluctuations in small classical-mechanical systems. In 1993, the first quantitative description of the entropy production in finite systems was provided by the discovery of a fluctuation theorem [2]. In its most general form, the theorem gives an analytical expression of the probability that a dissipative heat flux flows in the direction opposite to the one required by the second law of thermodynamics. For thermostated dissipative systems, the theorem relates the probability $p_\tau(\Omega)$ of observing over the time duration τ the entropy increase Ω , to the probability $p_\tau(-\Omega)$ of observing the entropy decrease of the same magnitude:

$$\frac{p_\tau(\Omega)}{p_\tau(-\Omega)} = e^\Omega. \quad (1)$$

Since the entropy production is an extensive quantity and the total entropy production increases in time, the fluctuation theorem shows that a positive entropy production is overwhelmingly likely as either the system size or the observation time increases. In this sense, we can view the fluctuation theorem as a generalization of the second law of thermodynamics. When the fluctuation theorem is applied to the transient response of a system, the theorem is referred to as the transient fluctuation theorem [3].

Since the above works, many fluctuation theorems have been presented for a variety of classical-mechanical situations: thermostated systems [2, 3, 4], stochastic systems [5, 6], and the externally driven systems [7, 8, 9]. These classical fluctuation theorems have been reviewed in Refs [10, 11]. Some of these theorems were verified experimentally [12, 13, 14].

When the system becomes further small, quantum effects may become significant. It is, however, not straightforward to extend the afore-mentioned fluctuation theorems to

quantum-mechanical systems. The crucial difference of the quantum systems from the classical systems is an essential role of measurements. In order to generalize the fluctuation theorems to quantum systems, we need to identify the entropy, the work, and the heat that are measured in the quantum-mechanical context. There have been two attempts to do this: first, defining operators to represent the heat and the work; second, measuring the system and using the measurement outcomes to represent the heat and the work. In general, the former attempt has led to quantum corrections to the classical results [15, 16, 17, 18]. In the latter attempt, on the other hand, several fluctuation theorems have been shown without quantum corrections [19, 20, 21, 22, 23]. Both the heat and the work are defined as the difference between the results of two measurements, a two-point quantity. We refer to this attempt as a two-time measurement scheme. The exchange fluctuation theorem (XFT) [1] was presented with this scheme.

The situation in which the exchange fluctuation theorem was proved [1] is quite simple; the two finite quantum systems are prepared in equilibrium at different temperatures then placed in thermal contact with one another. There is no work resource such as an external field nor an external force. In this situation, we simply identify the energy increase of each system as a heat flowing into the system.

To summarize the present thesis, we find the following:

- (i) The exchange fluctuation theorem does not hold in the presence of a finite heat transfer.
- (ii) If the Hamiltonian that couples the two systems commutes with the total Hamiltonian, the exchange fluctuation theorem becomes an exact relation.

The present thesis is organized as follows. In Chapter 2, we review previous results of quantum fluctuation theorems. In Chapter 3, we explain in details the measurement procedure referred to as the two-time measurement and show a general expression for the probability at which we observe a measurement outcome. In Chapter 4, we first follow the original derivation of the exchange fluctuation theorem [1], and then show the deviation from the theorem. We also show an explicit form of the deviation for two simple models. After that, we introduce a condition on the coupling Hamiltonian for the exchange fluctuation theorem to hold exactly and demonstrate that the deviation from the theorem vanishes. This is also confirmed in two specific models. Conclusions will be presented in Chapter 5.

2 Fluctuation theorems and Jarzynski's equality

In this section, we briefly review the transient fluctuation theorem, the steady-state fluctuation theorem and Jarzynski's equality. These relations hold in finite and far-from-equilibrium systems. They have been experimentally confirmed and have provided us an insight into statistical-physical characteristics of non-equilibrium systems.

2.1 Transient fluctuation theorem

The transient fluctuation theorem describes how irreversibility emerges from a completely time-reversible dynamics in the time-evolution. Here, we explain the Crooks fluctuation theorem [9, 20] as an example.

We use an externally driven system [20] initially prepared in the Gibbs state at an inverse temperature β . We first consider a time-dependent non-equilibrium “forward” process for which the system Hamiltonian changes from $H(0)$ to $H(\tau)$ over the finite time duration $0 \leq t \leq \tau$ and denote the forward process by $x_F(t)$. The forward process starts from the equilibrium state of $H(0)$ and ends at a non-equilibrium state of $H(\tau)$. In the “reverse” process, an initial equilibrium state of $H(\tau)$ evolves to a non-equilibrium state of $H(0)$. The non-equilibrium protocol for the reverse process, denoted by $x_R(t)$, is a time-reversed process of the forward one: $x_R(t) = x_F(\tau - t)$. Let $p'_\tau(\Omega)$ and $p'^R_\tau(\Omega)$ denote the probability distribution of the entropy production Ω in the forward and the reverse processes, respectively. Then the Crooks fluctuation theorem is expressed as

$$\frac{p'_\tau(\Omega)}{p'^R_\tau(-\Omega)} = e^\Omega. \quad (2)$$

To identify the entropy production Ω , we consider an isolated quantum system as an example. Let $H(t)$ and $|m_t\rangle$ denote the Hamiltonian of the system and an instantaneous eigenstate of the Hamiltonian $H(t)$, respectively. The joint probabilities $P(m_0, n_\tau|\tau)$ for the forward process and $P^R(n_\tau, m_0|\tau)$ for the reverse process (defined in Appendix A) are

$$P(m_0, n_\tau|\tau) = \frac{e^{-\beta E_{m_0}}}{Z_0} |\langle n_\tau|U|m_0\rangle|^2, \quad (3)$$

$$P^R(n_\tau, m_0|\tau) = \frac{e^{-\beta E_{n_\tau}}}{Z_\tau} |\langle m_0|U^\dagger|n_\tau\rangle|^2, \quad (4)$$

where U is the unitary time-evolution operator and Z_0 and Z_τ are the partition function of the equilibrium states at $t = 0$ and $t = \tau$, respectively. Using the joint probabilities, we define the probability distribution of the work W performed on the system as

$$p_\tau(W) = \sum_{m_0, n_\tau} P(m_0, n_\tau|\tau) \delta(W - (E_{n_\tau} - E_{m_0})), \quad (5)$$

$$p^R_\tau(W) = \sum_{m_0, n_\tau} P^R(n_\tau, m_0|\tau) \delta(W - (E_{m_0} - E_{n_\tau})), \quad (6)$$

where $\delta(W)$ is the delta function. Substituting Eq. (3) into Eq. (5), we have

$$\begin{aligned}
p_\tau(W) &= \sum_{m_0, n_\tau} \frac{e^{-\beta E_{m_0}}}{Z_0} |\langle n_\tau | U | m_0 \rangle|^2 \delta(W - (E_{n_\tau} - E_{m_0})) \\
&= \sum_{m_0, n_\tau} \frac{e^{-\beta(E_{n_\tau} - W)}}{Z_0} |\langle n_\tau | U | m_0 \rangle|^2 \delta(W - (E_{n_\tau} - E_{m_0})) \\
&= e^{\beta W} \frac{Z_\tau}{Z_0} \sum_{m_0, n_\tau} \frac{e^{-\beta E_{n_\tau}}}{Z_\tau} |\langle m_0 | U^\dagger | n_\tau \rangle|^2 \delta(-W - (E_{m_0} - E_{n_\tau})) \\
&= e^{\beta[W - (F_\tau - F_0)]} \sum_{m_0, n_\tau} \frac{e^{-\beta E_{n_\tau}}}{Z_\tau} |\langle m_0 | U^\dagger | n_\tau \rangle|^2 \delta(-W - (E_{m_0} - E_{n_\tau})) \\
&= e^{\beta(W - \Delta F)} p_\tau^R(-W), \tag{7}
\end{aligned}$$

where F_t is the free energy of the equilibrium state at time t and we used $Z_t = e^{-\beta F_t}$. The free energy difference between the equilibrium state at $t = 0$ and $t = \tau$ is $\Delta F = F_\tau - F_0$. Regarding $p_\tau(W)$ and $p_\tau^R(-W)$ in Eqs. (5) and (6) as $p'_\tau(\Omega)$ and $p_\tau^R(-\Omega)$ in Eq. (7), we identify the entropy production Ω as the irreversible work $W_{\text{irr}} = W - \Delta F$:

$$p'_\tau(\Omega) = e^{\beta W_{\text{irr}}} p_\tau^R(-\Omega). \tag{8}$$

This type of fluctuation theorem was proved by Crooks [9] in the classical context, and its quantum version was proved by Tasaki [20].

2.2 Steady-state fluctuation theorem

The steady-state fluctuation theorem is considered in externally driven systems. The external driving force can be made by applying an external field or boundary conditions. The steady state fluctuation theorem is written in the form

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \ln \frac{\bar{p}_\tau(\bar{\Omega})}{\bar{p}_\tau(-\bar{\Omega})} = \bar{\Omega}, \tag{9}$$

where $\bar{\Omega}$ is the time average of the entropy production Ω , $\bar{\Omega} = \Omega/\tau$, and $\bar{p}_\tau(\bar{\Omega})$ is the probability at which we observe the time-averaged value $\bar{\Omega}$. This expression was first established by Gallabotti and Cohen [4]. Equation (9) indicates that a steady-state system is more likely to produce the entropy rather than to consume the entropy. The steady-state fluctuation theorem has been proved recently in quantum systems [24, 25, 26].

2.3 Jarzynski's equality

Jarzynski's equality [7] relates the work W done on a system to the free-energy difference of the system ΔF between the initial and final equilibrium states. This equality is expressed as

$$\langle e^{-\beta W} \rangle_\tau = e^{-\beta \Delta F}, \tag{10}$$

where β is the inverse temperature of the initial state and the brackets $\langle \cdot \rangle_\tau$ denote the ensemble average to be defined in Sec. 3.2. The remarkable point of Jarzynski's equality is that we can determine the free-energy difference from a non-equilibrium irreversible process. Jarzynski's equality is easily obtained from the Crooks-Tasaki fluctuation theorem (7) as follows:

$$\begin{aligned}
\langle e^{-\beta W} \rangle_\tau &= \int dW e^{-\beta W} p_\tau(W) \\
&= \int dW e^{-\beta W} e^{\beta(W-\Delta F)} p_\tau^{\text{R}}(-W) \\
&= \int dW p_\tau^{\text{R}}(-W) e^{-\beta \Delta F} \\
&= e^{-\beta \Delta F}, \tag{11}
\end{aligned}$$

where we used the normalization of $p_\tau^{\text{R}}(W)$, $\int dW p_\tau^{\text{R}}(W) = 1$. Note that Jarzynski [7] first presented this expression in 1997, before the Crooks fluctuation theorem [9] was derived.

3 Two-time measurement scheme

In quantum systems, measurement affects the dynamics of the system, and thus we need to consider the effect of the measurement in general. However, modeling the effect of the system-detector interaction or the detector itself is complicated. Here, we consider a projection measurement which can be viewed as a fundamental way of modeling quantum measurements.

In this section, we explain a procedure of the two-time measurement which we will consider hereafter. The measurements in this procedure are modeled by the projection measurement. Then we introduce the corresponding joint probability and define an ensemble average in the two-time measurement procedure. In order to discuss a quantity related to the change of the system, we measure the system twice at least.

3.1 Procedure of a two-time measurement

We consider two quantum systems A and B with their respective Hilbert spaces $\mathcal{H}^{(A)}$ and $\mathcal{H}^{(B)}$. The state space \mathcal{H} of the combined system $A + B$ is given by the tensor product of the Hilbert spaces pertaining to the subsystems $\mathcal{H}^{(A)}$ and $\mathcal{H}^{(B)}$,

$$\mathcal{H} = \mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}. \quad (12)$$

The system A is described by the Hamiltonian H_A on $\mathcal{H}^{(A)}$ and the system B is described by H_B on $\mathcal{H}^{(B)}$. The connection between the two systems is described by the coupling Hamiltonian H_c on \mathcal{H} . The total Hamiltonian is

$$H = H_0 + \gamma H_c, \quad (13)$$

$$H_0 = H_A \otimes 1_B + 1_A \otimes H_B, \quad (14)$$

where 1_A and 1_B are the identity operators on $\mathcal{H}^{(A)}$ and $\mathcal{H}^{(B)}$, respectively, and γ is a parameter controlling the coupling strength between the two systems. Under the assumption that the total system $A + B$ is isolated, the density operator of the total system evolves in time according to the von Neumann equation:

$$\frac{d}{dt}\rho(t) = \frac{1}{i\hbar} [H, \rho(t)]. \quad (15)$$

The density operator at time t is written as

$$\rho(t) = U(t)\rho_0 U^\dagger(t), \quad (16)$$

where ρ_0 is the density operator at $t = 0$ and $U(t)$ is the solution of the Schrödinger equation

$$i\hbar \frac{d}{dt}U(t) = HU(t) \quad (17)$$

with the initial condition $U(0) = 1$.

Let X denote an observable, a Hermitian operator on \mathcal{H} , whose eigenvalues and eigenstates are $\{x\}$ and $\{|x, \lambda_x\rangle\}$, respectively. We distinguish the degenerate eigenstates of an eigenvalue x with a quantum number λ_x . Using the eigenstates $\{|x, \lambda_x\rangle\}$, we can write the projection operator onto the eigenspace of X with the eigenvalue x as

$$\Pi_x = \sum_{\lambda_x} |x, \lambda_x\rangle\langle x, \lambda_x|. \quad (18)$$

In the rest of the present thesis, we consider projective measurements (ideal measurements) only. Therefore, the measurement of X is expressed by a set of the projection operators $\{\Pi_x\}$. The state space on which the projection operators are defined depends on what we measure. If we measure the particle number of the system B , for example, the corresponding projection operator is defined as

$$\Pi_{n_B} = 1_A \otimes \sum_{\lambda_{n_B}} |n_B, \lambda_{n_B}\rangle\langle n_B, \lambda_{n_B}|, \quad (19)$$

where $|n_B, \lambda_{n_B}\rangle$ is an eigenstate of the number operator of the system B and λ_{n_B} is a quantum number which labels states with a particle number n_B .

We can divide the measuring procedure of X in the two-time measurement into the following four stages:

- Stage 1

For time $t < 0$, we prepare the system in the state described by ρ_{init} :

$$\rho(t < 0) = \rho_{\text{init}}. \quad (20)$$

The coupling between the two systems is off in this stage, and thus the two systems are initially decoupled.

- Stage 2

At $t = 0$, we perform the first measurement of the observable X and obtain an outcome x_1 . The state of the system is projected onto the eigenspace of X with the eigenvalue x_1 . The density operator after the first measurement with the measurement outcome is given by

$$\begin{aligned} \rho(t = 0) &= \rho_{x_1} \\ &:= \frac{\Pi_{x_1} \rho_{\text{init}} \Pi_{x_1}}{\text{Tr}(\Pi_{x_1} \rho_{\text{init}})}, \end{aligned} \quad (21)$$

where the denominator in the right-hand of the second line guarantees the normalization of the density operator after the measurement.

- Stage 3

After the first measurement, we turn on the coupling between the systems. Then the total system evolves from $t = 0$ to $t = \tau$ according to the von Neumann equation:

$$i\hbar \frac{d}{dt} \rho(t) = [H, \rho(t)]. \quad (22)$$

Therefore, the density operator at time t ($0 < t < \tau$) is given by

$$\rho(t) = U(t) \rho_{x_1} U^\dagger(t). \quad (23)$$

- Stage 4

At $t = \tau$, we separate the two systems and perform the second measurement of X . The state of the system is projected onto the corresponding eigenspace of X . The density operator after the measurement with an outcome x_2 is given by

$$\begin{aligned} \rho(\tau) &= \frac{\Pi_{x_2} U(\tau) \rho_{x_1} U^\dagger(\tau) \Pi_{x_2}}{\text{Tr}(\Pi_{x_2} U(\tau) \rho_{x_1} U^\dagger(\tau))} \\ &= \frac{\Pi_{x_2} U(\tau) \Pi_{x_1} \rho_{\text{init}} \Pi_{x_1} U^\dagger(\tau) \Pi_{x_2}}{\text{Tr}(\Pi_{x_2} U(\tau) \Pi_{x_1} \rho_{\text{init}} \Pi_{x_1} U^\dagger(\tau))}, \end{aligned} \quad (24)$$

where the denominator in the second line guarantees the normalization of the density operator $\rho(\tau)$.

In the two-time measurement procedure, we can discuss quantities indicating the state change: for example, the heat flow, the particle flow, and the work done in the intervals of the two measurements.

3.2 Joint probability of the measurement

Here, we introduce the joint probability that the result of the measurement is x_1 at time $t = 0$ and x_2 at time $t = \tau$ as

$$P(x_1, x_2 | \tau) = \text{Tr}(\Pi_{x_2} U(\tau) \Pi_{x_1} \rho_{\text{init}} \Pi_{x_1} U^\dagger(\tau)). \quad (25)$$

We can calculate the ensemble average in the two-time measurement with the joint probability. Summing over all possible measurement results, we can confirm the normalization of the joint probability $P(x_1, x_2 | \tau)$ as

$$\begin{aligned} \sum_{x_1, x_2} P(x_1, x_2 | \tau) &= \sum_{x_1, x_2} \text{Tr}(\Pi_{x_2} U(\tau) \Pi_{x_1} \rho_{\text{init}} \Pi_{x_1} U^\dagger(\tau)) \\ &= \sum_{x_1} \text{Tr}(U(\tau) \Pi_{x_1} \rho_{\text{init}} \Pi_{x_1} U^\dagger(\tau)) \\ &= \sum_{x_1} \text{Tr}(\Pi_{x_1} \rho_{\text{init}} \Pi_{x_1}) \\ &= \sum_{x_1} \text{Tr}(\rho_{\text{init}} \Pi_{x_1}) \\ &= \text{Tr}(\rho_{\text{init}}) = 1, \end{aligned} \quad (26)$$

where we used the properties of projection operators,

$$\begin{aligned}\Pi_x^2 &= \Pi_x, \\ \sum_x \Pi_x &= 1.\end{aligned}$$

The joint probability can be reduced to a product form by taking the trace in Eq. (25):

$$\begin{aligned}P(x_1, x_2|\tau) &= \text{Tr}(\Pi_{x_2} U(\tau) \Pi_{x_1} \rho_{\text{init}} \Pi_{x_1} U^\dagger(\tau)) \\ &= \sum_{\lambda_{x_1}, \lambda_{x_2}} \langle x_2, \lambda_{x_2} | U(\tau) | x_1, \lambda_{x_1} \rangle \langle x_1, \lambda_{x_1} | \rho_{\text{init}} | x_1, \lambda_{x_1} \rangle \langle x_1, \lambda_{x_1} | U^\dagger(\tau) | x_2, \lambda_{x_2} \rangle \\ &= \sum_{\lambda_{x_1}, \lambda_{x_2}} \langle x_1, \lambda_{x_1} | \rho_{\text{init}} | x_1, \lambda_{x_1} \rangle |\langle x_2, \lambda_{x_2} | U(\tau) | x_1, \lambda_{x_1} \rangle|^2 \\ &= \sum_{\lambda_{x_1}, \lambda_{x_2}} p_{\text{init}}(x_1, \lambda_{x_1}) T_{(x_1, \lambda_{x_1}) \rightarrow (x_2, \lambda_{x_2})}(\tau),\end{aligned}\tag{27}$$

where $p_{\text{init}}(x_1, \lambda_{x_1})$ denotes the probability of sampling the state $|x_1, \lambda_{x_1}\rangle$ from the initial state ρ_{init} :

$$p_{\text{init}}(x_1, \lambda_{x_1}) = \langle x_1, \lambda_{x_1} | \rho_{\text{init}} | x_1, \lambda_{x_1} \rangle\tag{28}$$

and $T_{(x_1, \lambda_{x_1}) \rightarrow (x_2, \lambda_{x_2})}(\tau)$ denotes the transition probability from $|x_1, \lambda_{x_1}\rangle$ to $|x_2, \lambda_{x_2}\rangle$:

$$T_{(x_1, \lambda_{x_1}) \rightarrow (x_2, \lambda_{x_2})}(\tau) = |\langle x_2, \lambda_{x_2} | U(\tau) | x_1, \lambda_{x_1} \rangle|^2.\tag{29}$$

If the observable X has no degeneracy, the joint probability is simply written as

$$\begin{aligned}P(x_1, x_2|\tau) &= \langle x_1 | \rho_{\text{init}} | x_1 \rangle |\langle x_2 | U(\tau) | x_1 \rangle|^2 \\ &= p_{\text{init}}(x_1) T_{x_1 \rightarrow x_2}(\tau).\end{aligned}\tag{30}$$

We will use the two expressions of the joint probability, Eqs. (25) and (27), in the present thesis.

In order to discuss a quantity concerning the change of the measurement outcome between the two measurements, we introduce the probability distribution defined as

$$p_\tau(\Delta x) = \sum_{x_1, x_2} P(x_1, x_2|\tau) \delta(\Delta x - (x_2 - x_1)),\tag{31}$$

where x_1 and x_2 are the results at the first and the second measurements, respectively, and $\Delta x = x_2 - x_1$ is the difference of the outcome between the two measurements. Here $\delta(\cdot)$ is the delta function. We can easily check the normalization of $p_\tau(\Delta x)$ by integrating it over all possible values of Δx and using Eq. (26):

$$\begin{aligned}\int d(\Delta x) p_\tau(\Delta x) &= \sum_{x_1, x_2} P(x_1, x_2|\tau) \int d(\Delta x) \delta(\Delta x - (x_2 - x_1)) \\ &= \sum_{x_1, x_2} P(x_1, x_2|\tau) \\ &= 1.\end{aligned}\tag{32}$$

We can calculate the ensemble average of the measurement difference Δx by either of $p_\tau(\Delta x)$ and $P(x_1, x_2|\tau)$. Using the probability distribution $p_\tau(\Delta x)$, we define the ensemble average in the two-time measurement as follows:

$$\langle f(\Delta x) \rangle_\tau := \int d(\Delta x) p_\tau(\Delta x) f(\Delta x), \quad (33)$$

where $f(\cdot)$ is an arbitrary function. Integrating over all possible values, we can represent the ensemble average with the joint probability $P(x_1, x_2|\tau)$ as

$$\langle f(\Delta x) \rangle_\tau = \sum_{x_1, x_2} P(x_1, x_2|\tau) f(x_2 - x_1). \quad (34)$$

In the following discussions, we will use the brackets $\langle \cdot \rangle_\tau$ to denote the ensemble average defined above.

3.3 Time-evolution operator

In this subsection, we represent the time-evolution operator with the commutation relation between the Hamiltonian of the decoupled system $H_0 = H_A \otimes 1_B + 1_A \otimes H_B$ and the Hamiltonian connecting the two systems H_c . For convenience, we introduce an operator $C(t)$ which is of the first order of the coupling Hamiltonian and contains all the commutation relations of H_0 and H_c . The coupling Hamiltonian H_c appears mainly through the operator $C(t)$ in the following discussions.

In the Schrödinger picture, the unitary time-evolution operator $U(t)$ is expressed for the static Hamiltonian as

$$U(t) = e^{-i\frac{H_0 + \gamma H_c}{\hbar}t}, \quad (35)$$

In the interaction picture, the unitary time-evolution operator $\tilde{U}(t)$ is expressed as follows:

$$\tilde{U}(t) = \mathcal{T} \exp \left[-\gamma \frac{i}{\hbar} \int_0^t \tilde{H}_c(s) ds \right] \quad (36)$$

$$= 1 + \sum_{k=1}^{\infty} \left(\frac{\gamma}{i\hbar} \right)^k \int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n \\ \times \tilde{H}_c(s_1) \tilde{H}_c(s_2) \cdots \tilde{H}_c(s_n), \quad (37)$$

where

$$\tilde{H}_c(t) = e^{i\frac{H_0}{\hbar}t} H_c e^{-i\frac{H_0}{\hbar}t} \quad (38)$$

is the coupling Hamiltonian in the interaction picture and \mathcal{T} is an anti-chronological time ordering from left to right. The operator $\tilde{U}(t)$ is the solution of the Schrödinger equation in the interaction picture

$$i\hbar \frac{\partial}{\partial t} \tilde{U}(t) = \gamma \tilde{H}_c(t) \tilde{U}(t) \quad (39)$$

with the initial condition $\tilde{U}(0) = 1$. The relation between the two pictures of the time-evolution operators is

$$U(t) = e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t). \quad (40)$$

We can confirm Eq. (40) by differentiating the right-hand side with respect to t and seeing that it obeys the Schrödinger equation and fulfills the same initial condition as $U(t)$, $U(t) = 1$:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\left(e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t)\right) &= H_0e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t) + \gamma e^{-\frac{i}{\hbar}H_0t}\tilde{H}_c\tilde{U}(t) \\ &= H_0e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t) + \gamma H_c e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t) \\ &= (H_0 + \gamma H_c)\left(e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t)\right) \\ &= H\left(e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t)\right), \\ e^{-\frac{i}{\hbar}H_0t}\tilde{U}(t)\Big|_{t=0} &= \tilde{U}(0) = 1. \end{aligned} \quad (41)$$

Defining δ_A as the inner-derivation operator

$$\delta_A = [A, \], \quad (42)$$

and using the following relation

$$\begin{aligned} e^A B e^{-A} &= B + \delta_A B + \frac{1}{2!}\delta_A\delta_A B + \frac{1}{3!}\delta_A\delta_A\delta_A B + \dots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (\delta_A)^n B, \end{aligned} \quad (43)$$

we can represent the coupling Hamiltonian in the interaction picture as follows:

$$\tilde{H}_c(t) = e^{i\frac{H_0}{\hbar}t}H_c e^{-i\frac{H_0}{\hbar}t} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{it}{\hbar}\right)^n (\delta_{H_0})^n H_c. \quad (44)$$

Since the Hamiltonian H_0 and H_c are independent of time, the time dependence of $\tilde{H}_c(t)$ appears only in the form t^n , and hence we can easily integrate it as

$$iC(t) := \frac{i}{\hbar} \int_0^t \tilde{H}_c(s) ds = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \left(\frac{it}{\hbar}\right)^{n+1} (\delta_{H_0})^n H_c, \quad (45)$$

where $C(t)$ defined above is a Hermitian operator. We mainly use $C(t)$ instead of H_c in the following discussions. Note that if the coupling Hamiltonian commutes with the total Hamiltonian, $[H, H_c] = 0$, only the term $n = 0$ remains in the summation in Eq. (45), and thus we have $C(t) = H_c t/\hbar$. We will consider this situation in Sec. 4.3.

The time-evolution operator in the interaction picture $\tilde{U}(t)$ is expressed with $C(t)$ as

$$\begin{aligned}
\tilde{U}(t) &= 1 - \gamma \frac{i}{\hbar} \int_0^t \tilde{H}_c(s_1) ds_1 + \gamma^2 \left(\frac{i}{\hbar} \right)^2 \int_0^t \tilde{H}_c(s_1) ds_1 \int_0^{s_1} \tilde{H}_c(s_2) ds_2 \\
&\quad - \gamma^3 \left(\frac{i}{\hbar} \right)^3 \int_0^t \tilde{H}_c(s_1) ds_1 \int_0^{s_1} \tilde{H}_c(s_2) ds_2 \int_0^{s_2} \tilde{H}_c(s_3) ds_3 + \cdots \\
&= 1 - i\gamma C(t) + (i\gamma)^2 \int_0^t ds_1 \dot{C}(s_1) C(s_1) \\
&\quad - (i\gamma)^3 \int_0^t ds_1 \dot{C}(s_1) \int_0^{s_1} ds_2 \dot{C}(s_2) C(s_2) + \cdots \\
&= 1 - i\gamma C(t) + \frac{1}{2!} (i\gamma C(t))^2 - \frac{1}{3!} (i\gamma C(t))^3 + \cdots \\
&= e^{-i\gamma C(t)}.
\end{aligned} \tag{46}$$

From the above discussion, the time-evolution operator in the Schrödinger picture is written as

$$U(t) = e^{-i\frac{H_0}{\hbar}t} e^{-i\gamma C(t)}. \tag{47}$$

Using this representation of the time-evolution operator, we can represent the transition probability (29) and the corresponding joint probability (25) in the following forms

$$\begin{aligned}
T_{m \rightarrow n}(\tau) &= \sum_{\lambda_m, \lambda_n} |\langle n, \lambda_n | U(\tau) | m, \lambda_m \rangle|^2 \\
&= \sum_{\lambda_m, \lambda_n} |e^{-i\frac{E_n}{\hbar}\tau} \langle n, \lambda_n | e^{-i\gamma C(\tau)} | m, \lambda_m \rangle|^2, \\
&= \sum_{\lambda_m, \lambda_n} |\langle n, \lambda_n | e^{-i\gamma C(\tau)} | m, \lambda_m \rangle|^2,
\end{aligned} \tag{48}$$

$$P(m, n|\tau) = p_{\text{init}}(m) \sum_{\lambda_m, \lambda_n} |\langle n, \lambda_n | e^{-i\gamma C(\tau)} | m, \lambda_m \rangle|^2, \tag{49}$$

where E_m and $|m, \lambda_m\rangle$ are an eigenvalue and eigenstate of the Hamiltonian H_0 , respectively, $H_0|m, \lambda_m\rangle = E_m|m, \lambda_m\rangle$, and λ_m is a quantum number labeling the eigenstates with the same eigenvalue. In these expressions, we can easily see the dependence of the coupling strength γ on the transition probability and the joint probability, and thus these expressions suit the perturbative expansion with respect to γ .

Before going to the next section, we represent the joint probability in the trace form:

$$P(m, n|\tau) = \text{Tr} \left(\Pi_n e^{-i\gamma C(\tau)} \Pi_m \rho_{\text{init}} \Pi_m e^{i\gamma C(\tau)} \right).$$

If we choose the initial state as a stationary state in the decoupled system $[H_0, \rho_{\text{init}}] = 0$, the initial state ρ_{init} can be diagonalized with the eigenstates of H_0 and, we have

$$P(m, n|\tau) = \text{Tr} \left(\Pi_n e^{-i\gamma C(\tau)} \rho_{\text{init}} \Pi_m e^{i\gamma C(\tau)} \right). \tag{50}$$

In the next section, we chose the initial state as a product state of the Gibbs state [Eq. (62)] and this initial state is commutable with the Hamiltonian of the decoupled system H_0 . Thus, we use Eq. (50) when we represent the joint probabilities in the trace form.

4 Examination of the Exchange Fluctuation Theorem

C. Jarzynski and D. K. Wójcik [1] derived a fascinating symmetry relation regarding the statistics of heat exchange between two finite systems initially prepared at different temperatures in Hamiltonian dynamics. Let β_A and β_B denote the inverse temperatures at which the system A and the system B are prepared, respectively. The symmetry relation is expressed with the probability distribution $p_\tau(Q)$ of the net heat transfer Q as follows:

$$p_\tau(Q) = e^{\Delta\beta Q} p_\tau^R(-Q), \quad (51)$$

where $\Delta\beta = \beta_B - \beta_A$ is the difference between the inverse temperatures and τ is the time duration between the two measurements as in the previous section. They referred to this symmetry relation as the exchange fluctuation theorem (XFT) because the exchanged energy Q between the two systems is regarded as a heat transfer. They derived the exchange fluctuation theorem in both classical and quantum systems. We will discuss the exchange fluctuation theorem only for quantum systems in the present thesis.

To derive the exchange fluctuation theorem, Jarzynski and Wójcik assumed the following three conditions. First, the Hamiltonian of the system is time-reversal invariant. Second, both of the two systems are initially in the Gibbs state described as follows:

$$\begin{aligned} \rho_{\text{init}} &= \rho_A \otimes \rho_B \\ \rho_A &= \frac{e^{-\beta_A H_A}}{Z_A}, & Z_A &= \text{Tr} e^{-\beta_A H_A}, \\ \rho_B &= \frac{e^{-\beta_B H_B}}{Z_B}, & Z_B &= \text{Tr} e^{-\beta_B H_B}, \end{aligned}$$

where H_α and Z_α are the Hamiltonians and the partition function of the system α ($\alpha = A, B$), respectively. Finally, the interaction is weak enough for the total energy to be preserved:

$$E_{m_A}^A + E_{m_B}^B \simeq E_{n_A}^A + E_{n_B}^B, \quad (52)$$

where $(E_{m_A}^A, E_{m_B}^B)$ and $(E_{n_A}^A, E_{n_B}^B)$ are the measurement outcomes in the first and second measurements, respectively, and here we regard the sum $E_{m_A}^A + E_{m_B}^B$ and $E_{n_A}^A + E_{n_B}^B$ as the total energy of the systems since the coupling between the systems is off at $t \leq 0$ and $t \geq \tau$.

We inspect the exchange fluctuation theorem under these three conditions analytically and numerically in Sec. 4. 2. To make the point clear, we review the original derivation of the exchange fluctuation theorem [1] in Sec. 4. 1.

4.1 Original Derivation of the XFT

In this section, we further assume the eigenvalues of H_A and H_B to be nondegenerate for simplicity. Jarzynski and Wójcik did not mention the degeneracy of the Hamiltonians in their original paper [1]. The degeneracy indeed causes no corrections to their discussions.

4.1.1 Time-reversal operator in quantum mechanics

Before following the original derivation of the exchange fluctuation theorem, we note some properties of the time-reversal operation in our setup. The time-reversal operation in quantum mechanics is described by an antilinear operator Θ [27]. The antilinearity is represented as

$$\Theta(\alpha_1|\Psi\rangle + \alpha_2|\Phi\rangle) = \alpha_1^*\Theta|\Psi\rangle + \alpha_2^*\Theta|\Phi\rangle, \quad (53)$$

where α^* is complex conjugate of a scalar number α , and $|\Psi\rangle$ and $|\Phi\rangle$ are arbitrary states. The time-reversal invariance of the system is expressed for the Hamiltonian as

$$\Theta H = H \Theta. \quad (54)$$

Under the time-reversal invariance, every eigenstate of H , $|n\rangle$, has the corresponding time-reversed eigenstate $\Theta|n\rangle$ with the same energy:

$$\begin{aligned} H(\Theta|n\rangle) &= \Theta H|n\rangle \\ &= \Theta E_n|n\rangle \\ &= E_n(\Theta|n\rangle). \end{aligned} \quad (55)$$

These two states, $|n\rangle$ and $\Theta|n\rangle$, are either linear independent or identical apart from an overall phase. From Eqs. (53) and (54), we have the relation between Θ and $U(t)$:

$$\begin{aligned} \Theta U(t) &= \Theta \left(1 + \frac{H}{i\hbar}t + \frac{1}{2!} \left(\frac{H}{i\hbar}t \right)^2 + \dots \right) \\ &= \left(1 - \frac{H}{i\hbar}t + \frac{1}{2!} \left(-\frac{H}{i\hbar}t \right)^2 + \dots \right) \Theta \\ &= U^\dagger(t)\Theta. \end{aligned} \quad (56)$$

Finally, we introduce a notation of the inner product when we use the time-reversal operator. The Dirac bra-ket notation may be confusing when we treat the time-reversal operator; the expression $\langle\Psi|\Theta|\Phi\rangle$ is ambiguous as to whether Θ is acting on the right state or on the left state. To avoid this confusion, let us denote the inner product as $(|\Psi\rangle, |\Phi\rangle)$. (We use the Dirac bra-ket notation to represent the inner product whenever the time-reversal operator Θ does not cause confusion.) Using this notation, the transition probability from $|\Phi\rangle$ to $|\Psi\rangle$ is written as follows:

$$|(|\Psi\rangle, U(\tau)|\Phi\rangle)|^2. \quad (57)$$

The inner product $(\Theta|\Psi\rangle, \Theta|\Phi\rangle)$ are related to the original inner product $(|\Psi\rangle, |\Phi\rangle)$ as

$$(\Theta|\Psi\rangle, \Theta|\Phi\rangle) = (|\Psi\rangle, |\Phi\rangle)^*$$

because the time-reversal operator Θ preserves the wave function normalization. We can confirm this relation using an orthogonal complete set $\{|\alpha\rangle\}$ as follows:

$$\begin{aligned}
(\Theta|\Psi\rangle, \Theta|\Phi\rangle) &= \left(\Theta \sum_{\alpha_1} |\alpha_1\rangle \langle \alpha_1 | \Psi \rangle, \Theta \sum_{\alpha_2} |\alpha_2\rangle \langle \alpha_2 | \Phi \rangle \right) \\
&= \sum_{\alpha_1, \alpha_2} (\langle \alpha_1 | \Psi \rangle^* \Theta |\alpha_1\rangle, \langle \alpha_2 | \Phi \rangle^* \Theta |\alpha_2\rangle) \\
&= \sum_{\alpha_1, \alpha_2} \langle \alpha_1 | \Psi \rangle \langle \Phi | \alpha_2 \rangle (\Theta |\alpha_1\rangle, \Theta |\alpha_2\rangle) \\
&= \sum_{\alpha_1, \alpha_2} \langle \alpha_1 | \Psi \rangle \langle \Phi | \alpha_2 \rangle \delta_{\alpha_1, \alpha_2} \\
&= \sum_{\alpha_1} \langle \Phi | \alpha_1 \rangle \langle \alpha_1 | \Psi \rangle \\
&= \langle \Phi | \Psi \rangle \\
&= (|\Psi\rangle, |\Phi\rangle)^*. \tag{58}
\end{aligned}$$

Thus, the time-reversal operator Θ is not only antilinear but also antiunitary. These properties and notations were used in their original derivation [1].

4.1.2 Derivation of the XFT by Jarzynski and Wójcik

Jarzynski and Wójcik [1] considered two finite quantum systems given by the Hamiltonian

$$H = H_A \otimes 1_B + 1_A \otimes H_B + \gamma H_c, \tag{59}$$

where H_A and H_B are the Hamiltonians of the systems A and B , respectively. The third term of the right hand side of Eq. (59), H_c , is the coupling Hamiltonian which describes the connection between the two systems and γ is a factor controlling the coupling strength between the two systems as mentioned before. The assumption of the time-reversal invariance is described as follows:

$$\Theta H_A = H_A \Theta, \quad \Theta H_B = H_B \Theta, \quad \Theta H_c = H_c \Theta. \tag{60}$$

Let $|m_\alpha\rangle$ and $E_{m_\alpha}^\alpha$ denote an eigenstate of H_α and the corresponding eigenvalue ($\alpha = A, B$), respectively. We refer to the product states $|m_A\rangle \otimes |m_B\rangle$ as $|m_A, m_B\rangle$, or just as $|m\rangle$ for simplicity. Jarzynski and Wójcik considered energy measurement with the two-time measurement procedure, and thus the energy measurement is described by the projection operator

$$\begin{aligned}
\Pi_m &= (|m_A\rangle \otimes |m_B\rangle) (\langle m_A| \otimes \langle m_B|) \\
&= |m\rangle \langle m|. \tag{61}
\end{aligned}$$

For time $t < 0$, each of the systems A and B is separately connected to a heat reservoir at the inverse temperatures, β_A and β_B , respectively, for sufficiently long time to reach its

equilibrium state. The reservoirs are removed just before the first measurement, and the initial state of the total system is given by the following product state:

$$\rho_{\text{init}} = \rho_A \otimes \rho_B \quad (62)$$

$$\rho_A = \frac{e^{-\beta_A H_A}}{Z_A}, \quad Z_A = \text{Tr} e^{-\beta_A H_A}, \quad (63)$$

$$\rho_B = \frac{e^{-\beta_B H_B}}{Z_B}, \quad Z_B = \text{Tr} e^{-\beta_B H_B}, \quad (64)$$

where Z_A and Z_B are the partition function of each system, respectively. At $t = 0$, we perform the first measurement of the energy of each system. Suppose that we obtained the outcome $(E_{m_A}^A, E_{m_B}^B)$. Then, the state of the systems is projected onto the eigenstate of $H_0 = H_A \otimes 1_B + 1_A \otimes H_B$: $|m\rangle = |m_A, m_B\rangle$:

$$\begin{aligned} \rho(0) &= \frac{\Pi_m \rho_{\text{init}} \Pi_m}{\text{Tr} (\Pi_m \rho_{\text{init}})} \\ &= \frac{(\Pi_{m_A} \otimes \Pi_{m_B}) \rho_{\text{init}} (\Pi_{m_A} \otimes \Pi_{m_B})}{\text{Tr} [(\Pi_{m_A} \otimes \Pi_{m_B}) \rho_{\text{init}}]} \\ &= \frac{e^{-\beta_A E_{m_A}^A} e^{-\beta_B E_{m_B}^B} / Z_A Z_B}{e^{-\beta_A E_{m_A}^A} e^{-\beta_B E_{m_B}^B} / Z_A Z_B} |m_A, m_B\rangle \langle m_A, m_B| \\ &= |m\rangle \langle m|. \end{aligned} \quad (65)$$

From $t = 0$ to $t = \tau$, the total system evolves according to the von Neumann equation, and thus we have

$$\rho(t) = U(t) \rho(0) U^\dagger(t). \quad (66)$$

The coupling between the two systems is turned on after the first measurement. The time evolution is described by

$$U = e^{-i \frac{H_0 + \gamma H_c}{\hbar}}.$$

At $t = \tau$, we separate the two systems again and measure the energy of each system. Suppose that we obtained the outcome $(E_{n_A}^A, E_{n_B}^B)$. Then the density operator becomes after the second measurement

$$\begin{aligned} \rho(\tau) &= \frac{\Pi_n U(\tau) \rho(0) U^\dagger(\tau) \Pi_n}{\text{Tr} (\Pi_n U(\tau) \rho(0) U^\dagger(\tau))} \\ &= \frac{(\Pi_{n_A} \otimes \Pi_{n_B}) U(\tau) (|m_A, m_B\rangle \langle m_A, m_B|) U^\dagger(\tau) (\Pi_{n_A} \otimes \Pi_{n_B})}{\text{Tr} [(\Pi_{n_A} \otimes \Pi_{n_B}) U(\tau) (|m_A, m_B\rangle \langle m_A, m_B|) U^\dagger(\tau)]} \\ &= \frac{\langle n_A, n_B | U(\tau) |m_A, m_B\rangle \langle m_A, m_B| U^\dagger(\tau) |n_A, n_B\rangle}{\langle n_A, n_B | U(\tau) |m_A, m_B\rangle \langle m_A, m_B| U^\dagger(\tau) |n_A, n_B\rangle} |n_A, n_B\rangle \langle n_A, n_B| \\ &= |n\rangle \langle n|. \end{aligned} \quad (67)$$

From Eq. (30), the joint probability in the forward process that we observe $(E_{m_A}^A, E_{m_B}^B)$ at time $t = 0$ and $(E_{n_A}^A, E_{n_B}^B)$ at time $t = \tau$ is written as

$$P(m, n|\tau) = p_{\text{init}}(m) T_{m \rightarrow n}(\tau). \quad (68)$$

Here, we introduce the reverse process corresponding to the forward process. In the reverse process, the time evolves inversely and observation outcomes are $(E_{n_A}^A, E_{n_B}^B)$ at time $t = 0$ and $(E_{m_A}^A, E_{m_B}^B)$ at time $t = \tau$. The joint probability in the reverse process is given by

$$P^{\text{R}}(n, m|\tau) = \text{Tr} (\Pi_m U(\tau)^\dagger \Pi_n \rho_{\text{init}} \Pi_n U(\tau)). \quad (69)$$

We give the verification of this definition for the reverse process in Appendix A. From Eq. (69), we have

$$\begin{aligned} P^{\text{R}}(n, m|\tau) &= \text{Tr} (\Pi_m U(\tau)^\dagger \Pi_n \rho_{\text{init}} \Pi_n U(\tau)) \\ &= \langle m | e^{i\frac{H}{\hbar}\tau} | n \rangle \langle n | \rho_{\text{init}} | n \rangle \langle n | e^{-i\frac{H}{\hbar}\tau} | m \rangle \\ &= p_{\text{init}}(n) \left| \langle n | e^{-i\frac{H}{\hbar}\tau} | m \rangle \right|^2 \\ &= p_{\text{init}}(n) T_{m \rightarrow n}(\tau), \end{aligned} \quad (70)$$

and

$$\begin{aligned} P^{\text{R}}(n, m|\tau) &= p_{\text{init}}(n) \left| \langle m | e^{i\frac{H}{\hbar}(-\tau)} | n \rangle \right|^2 \\ &= p_{\text{init}}(n) T_{n \rightarrow m}(-\tau) \\ &= P(n, m | -\tau). \end{aligned} \quad (71)$$

When the Hamiltonian is independent of time, the joint probability in the reverse process $P^{\text{R}}(n, m|\tau)$ is expressed by the joint probability in the forward process, $P(n, m | -\tau)$.

Using Eqs. (68) and (70), we obtain a symmetry relation between the joint probabilities of the forward and reverse processes as

$$\begin{aligned} \frac{P(m, n|\tau)}{P^{\text{R}}(n, m|\tau)} &= \frac{p_{\text{init}}(m) T_{m \rightarrow n}(\tau)}{p_{\text{init}}(n) T_{m \rightarrow n}(\tau)} \\ &= \frac{p_{\text{init}}(m)}{p_{\text{init}}(n)} \\ &= e^{-\beta_A E_{m_A}^A - \beta_B E_{m_B}^B} e^{\beta_A E_{n_A}^A + \beta_B E_{n_B}^B} \\ &= e^{-\beta_A (E_{m_A}^A - E_{n_A}^A)} e^{-\beta_B (E_{m_B}^B - E_{n_B}^B)} \\ &= e^{(\beta_B - \beta_A) (E_{m_A}^A - E_{n_A}^A)} e^{-\beta_B (E_{m_B}^B - E_{n_B}^B + E_{m_A}^A - E_{n_A}^A)} \\ &= e^{\Delta\beta Q_{m \rightarrow n}^A} e^{\beta_B \Delta E_{m \rightarrow n}}, \end{aligned} \quad (72)$$

where

$$Q_{m \rightarrow n}^A = E_{m_A}^A - E_{n_A}^A \quad (73)$$

is the energy decrease in the system A and

$$\Delta E_{m \rightarrow n} = E_{n_A}^A + E_{n_B}^B - E_{m_A}^A - E_{m_B}^B \quad (74)$$

is the energy change of the total system. We interpret $Q_{m \rightarrow n}^A$ as the heat draining from the system A to the system B .

Jarzynski and Wójcik assumed that the total energy of the systems A and B is almost preserved between the measurements if the coupling between the two systems is sufficiently weak;

$$\Delta E_{m \rightarrow n} = E_{n_A}^A + E_{n_B}^B - E_{m_A}^A - E_{m_B}^B \simeq 0. \quad (75)$$

It follows that the energy changes in the two systems are approximately equal:

$$E_{m_A}^A - E_{n_A}^A \simeq E_{n_B}^B - E_{m_B}^B. \quad (76)$$

Then we can remove the superscript A from $Q_{m \rightarrow n}^A$ in Eq. (73), because

$$Q_{m \rightarrow n} = Q_{m_A \rightarrow n_A}^A \simeq E_{n_B}^B - E_{m_B}^B, \quad (77)$$

where the right-hand side of Eq. (77) is the energy increase in the system B and we interpret it as the heat flowing to the system B . Substituting Eqs. (75) and (77) into Eq. (72), they obtained the following symmetry relation:

$$\frac{P(m, n|\tau)}{P^R(n, m|\tau)} \simeq e^{\Delta\beta Q_{m \rightarrow n}}. \quad (78)$$

Using this relation, we obtain the symmetry relation of the probability distribution of the net heat transfer Q is obtained as follows:

$$p_\tau(Q) = \sum_{n,m} P(m, n|\tau) \delta(Q - Q_{m \rightarrow n}) \quad (79)$$

$$\simeq e^{\Delta\beta Q} \sum_{n,m} P^R(n, m|\tau) \delta(Q + Q_{n \rightarrow m}) \quad (80)$$

$$\begin{aligned} &= e^{\Delta\beta Q} \sum_{n,m} P^R(n, m|\tau) \delta(-Q - Q_{n \rightarrow m}) \\ &= e^{\Delta\beta Q} p_\tau^R(-Q). \end{aligned} \quad (81)$$

This is the exchange fluctuation theorem.

Equation (81) implies that the average of $e^{-\Delta\beta Q}$ over the ensemble of realizations for any time interval τ is unity:

$$\langle e^{-\Delta\beta Q} \rangle_\tau = \int_{-\infty}^{\infty} dQ p_\tau(Q) e^{-\Delta\beta Q} \quad (82)$$

$$\begin{aligned} &\simeq \int_{-\infty}^{\infty} dQ p_\tau^R(-Q) \\ &= 1. \end{aligned} \quad (83)$$

This integral form of equality is a direct consequence of the exchange fluctuation theorem, and thus we refer to this equality as the integral exchange fluctuation theorem (IXFT).

We note that the crucial assumption in the original derivation [1] is that the total energy difference $\Delta E_{m \rightarrow n}$ becomes negligibly small in the weak coupling limit as stated in Eq. (75). Since we are interested in non-equilibrium phenomena, we hope that the exchange fluctuation theorem holds in the presence of a finite heat transfer; a non-equilibrium situation. However, the case $\gamma = 0$ means that the two systems are decoupled and that the heat never transfers between the two systems. Indeed, we are interested in the case of finite γ . In this respect, we examine whether the exchange fluctuation theorem holds or not in the presence of a finite heat transfer in the present thesis. In the next section, we derive deviation from the exchange fluctuation theorem and discuss whether the exchange fluctuation holds or not in the presence of a finite heat transfer.

4.2 Deviation from the XFT and the IXFT

In this section, we examine both the exchange fluctuation theorem and the integral exchange fluctuation theorem and show deviation from these theorems in the same situation as the original derivation [1]. The deviation in general has a finite value for a finite coupling strength γ , which implies that the exchange fluctuation theorem does not hold for a finite γ . In the limit $\gamma \rightarrow 0$, the deviation term vanishes and the exchange fluctuation theorem becomes an exact relation. In such limit, however, the k th moment of the probability distribution of the heat transfer, $p_\tau(Q)$, also vanishes:

$$\lim_{\gamma \rightarrow 0} \langle Q^k \rangle_\tau = \int dQ p_\tau(Q) Q^k = 0.$$

This implies that if the coupling between the two systems is weak enough for the exchange fluctuation theorem to hold, no finite heat flows between the systems. This is not the case we are interested in. If any k th moment of $p_\tau(Q)$ vanishes in the weak coupling limit, the exchange fluctuation theorem becomes just a trivial relation:

$$\begin{aligned} \lim_{\gamma \rightarrow 0} p_\tau(Q) &= \delta(Q), \\ \lim_{\gamma \rightarrow 0} (p_\tau(Q) - e^{\Delta Q} p_\tau^R(-Q)) &= \delta(Q) - e^0 \delta(-Q) = 0. \end{aligned} \tag{84}$$

We derive the deviation from the integral fluctuation theorem as an explicit expression for a specific model and seek a possibility that the exchange fluctuation theorem holds in the presence of the finite heat transfer. In this model, the deviation $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ and the net heat transfer $\langle Q \rangle_\tau$ have the same dependence on the coupling strength. Therefore, if we take the limit $\gamma \rightarrow 0$, both quantities go to 0 in the same manner. We show the deviation from the exchange fluctuation theorem for another model numerically.

4.2.1 Deviation from the XFT and the IXFT

First, we derive the deviation from the exchange fluctuation theorem. Substituting the symmetry relation (72) into Eq. (79), we have

$$\begin{aligned}
p_\tau(Q) &= \sum_{m,n} P(m, n|\tau) \delta(Q - Q_{m \rightarrow n}) \\
&= \sum_{m,n} e^{\Delta\beta Q_{m \rightarrow n}} e^{\beta_B \Delta E_{m \rightarrow n}} P^{\text{R}}(n, m|\tau) \delta(Q - Q_{m \rightarrow n}) \\
&= e^{\Delta\beta Q} \sum_{m,n} e^{\beta_B \Delta E_{m \rightarrow n}} P^{\text{R}}(n, m|\tau) \delta(-Q - Q_{n \rightarrow m}) \\
&= e^{\Delta\beta Q} \sum_{m,n} P^{\text{R}}(n, m|\tau) \delta(-Q - Q_{n \rightarrow m}) \\
&\quad + e^{\Delta\beta Q} \sum_{m,n} P^{\text{R}}(n, m|\tau) (e^{\beta_B \Delta E_{m \rightarrow n}} - 1) \delta(-Q - Q_{n \rightarrow m}) \\
&= e^{\Delta\beta Q} p_\tau^{\text{R}}(-Q) \\
&\quad + e^{\Delta\beta Q} \sum_{m,n} P^{\text{R}}(n, m|\tau) (e^{\beta_B \Delta E_{m \rightarrow n}} - 1) \delta(Q - Q_{m \rightarrow n}), \tag{85}
\end{aligned}$$

and thus the deviation from the exchange fluctuation theorem is

$$p_\tau(Q) - e^{\Delta\beta Q} p_\tau^{\text{R}}(-Q) = e^{\Delta\beta Q} \sum_{m,n} P^{\text{R}}(n, m|\tau) (e^{\beta_B \Delta E_{m \rightarrow n}} - 1) \delta(Q - Q_{m \rightarrow n}). \tag{86}$$

In general, this deviation term (the right-hand side of Eq. (86)) has a finite value. We can check this by examining the deviation from the integral exchange fluctuation theorem since the integral exchange fluctuation theorem is a direct consequence of the exchange fluctuation theorem. To see the deviation from the integral exchange fluctuation theorem, we multiply Eq. (86) by $e^{-\Delta\beta Q}$ and integrate it over the heat transfer Q , having

$$\begin{aligned}
\int dQ (p_\tau(Q) e^{-\Delta\beta Q} - p_\tau^{\text{R}}(-Q)) &= \langle e^{-\Delta\beta Q} \rangle_\tau - 1 \\
&= \sum_{m,n} P^{\text{R}}(n, m|\tau) (e^{\beta_B \Delta E_{m \rightarrow n}} - 1). \tag{87}
\end{aligned}$$

Using Eq. (71) and the trace representation of the joint probability in Eq. (50), we have

$$\begin{aligned}
\langle e^{-\Delta\beta Q} \rangle_\tau - 1 &= \sum_{m,n} \text{Tr} (\Pi_m e^{i\gamma C(\tau)} \rho_{\text{init}} \Pi_n e^{-i\gamma C(\tau)}) \left(e^{\beta_B (E_{n_A}^A + E_{n_B}^B - E_{m_A}^A - E_{m_B}^B)} - 1 \right) \\
&= \sum_{m,n} \text{Tr} (e^{-\beta_B H_0} \Pi_m e^{i\gamma C(\tau)} \rho_{\text{init}} e^{\beta_B H_0} \Pi_n e^{-i\gamma C(\tau)}) \\
&\quad - \sum_{m,n} \text{Tr} (\Pi_m e^{i\gamma C(\tau)} \rho_{\text{init}} \Pi_n e^{-i\gamma C(\tau)}) \\
&= \text{Tr} (e^{-\beta_B H_0} e^{i\gamma C(\tau)} \rho_{\text{init}} e^{\beta_B H_0} e^{-i\gamma C(\tau)}) - \text{Tr} (e^{i\gamma C(\tau)} \rho_{\text{init}} e^{-i\gamma C(\tau)}) \\
&= \text{Tr} (\rho_{\text{init}} e^{\beta_B H_0} e^{-i\gamma C(\tau)} e^{-\beta_B H_0} e^{i\gamma C(\tau)}) - 1
\end{aligned}$$

and then using Eq. (43), we obtain

$$\begin{aligned}
\langle e^{-\Delta\beta Q} \rangle_\tau - 1 &= \sum_{k=0}^{\infty} \frac{1}{k!} (-i\gamma)^k \text{Tr} [\rho_{\text{init}} e^{\beta_B H_0} (\delta_{C(\tau)})^k e^{-\beta_B H_0}] - 1 \\
&= -i\gamma \text{Tr} \{ \rho_{\text{init}} e^{\beta_B H_0} [C(\tau), e^{-\beta_B H_0}] \} + \sum_{k=2}^{\infty} \frac{1}{k!} (-i\gamma)^k \text{Tr} [\rho_{\text{init}} e^{\beta_B H_0} (\delta_{C(\tau)})^k e^{-\beta_B H_0}] \\
&= -i\gamma \text{Tr} \{ [e^{-\beta_B H_0}, \rho_{\text{init}} e^{\beta_B H_0}] C(\tau) \} + \sum_{k=2}^{\infty} \frac{1}{k!} (-i\gamma)^k \text{Tr} [\rho_{\text{init}} e^{\beta_B H_0} (\delta_{C(\tau)})^k e^{-\beta_B H_0}] \\
&= \sum_{k=2}^{\infty} \frac{1}{k!} (-i\gamma)^k \text{Tr} [\rho_{\text{init}} e^{\beta_B H_0} (\delta_{C(\tau)})^k e^{-\beta_B H_0}], \tag{88}
\end{aligned}$$

where we used $[e^{\beta_B H_0}, \rho_{\text{init}}] = 0$ in the last line. This commutability is a consequence of taking the initial state as a Gibbs state. We thus obtain the deviation, which does not vanish for finite γ in general. In Sec. 4.2.2, the deviation is explicitly calculated for a simple model and is shown to have a finite value for finite γ .

We can see that the deviation term (88) vanishes in the limit $\gamma \rightarrow 0$, and the exchange fluctuation theorem holds. This is consistent with the result obtained by taking the limit $\gamma \rightarrow 0$ in Eq. (86):

$$\begin{aligned}
\lim_{\gamma \rightarrow 0} (p_\tau(Q) - e^{\Delta\beta Q} p_\tau^R(-Q)) &= e^{\Delta\beta Q} \sum_{m,n} \lim_{\gamma \rightarrow 0} P(n, m | -\tau) (e^{\beta_B \Delta E_{m \rightarrow n}} - 1) \delta(Q - Q_{m \rightarrow n}) \\
&= e^{\Delta\beta Q} \sum_{m,n} \lim_{\gamma \rightarrow 0} p_{\text{init}}(n) \delta_{m,n} (e^{\beta_B \Delta E_{m \rightarrow n}} - 1) \delta(Q - Q_{m \rightarrow n}). \\
&= 0, \tag{89}
\end{aligned}$$

where we used

$$\begin{aligned}
\lim_{\gamma \rightarrow 0} P(m, n | \tau) &= p_{\text{init}}(m) \lim_{\gamma \rightarrow 0} |\langle n | e^{-i\gamma C(\tau)} | m \rangle|^2 \\
&= p_{\text{init}}(m) \delta_{m,n}.
\end{aligned}$$

Both exchange fluctuation theorem and the integral fluctuation theorem hold together in the limit $\gamma \rightarrow 0$.

However, we must pay attention to the behavior of the net heat transfer $\langle Q \rangle_\tau$ in such a limit because we are interested in a non-equilibrium system, or in the presence of a finite heat flow. The net heat transfer is also represented in the power series of γ :

$$\begin{aligned}
\langle Q \rangle_\tau &= \int_{-\infty}^{\infty} dQ p_\tau(Q) Q \\
&= \sum_{m,n} P(m, n|\tau) \int_{-\infty}^{\infty} dQ Q \delta(Q - Q_{m \rightarrow n}) \\
&= \sum_{m,n} P(m, n|\tau) (E_{m_A}^A - E_{n_A}^A) \\
&= \sum_{m,n} \text{Tr} (\Pi_n e^{-i\gamma C(\tau)} \rho_{\text{init}} \Pi_m e^{i\gamma C(\tau)}) (E_{m_A}^A - E_{n_A}^A) \\
&= \sum_{m,n} \text{Tr} (\Pi_n e^{-i\gamma C(\tau)} \rho_{\text{init}} H_A \Pi_m e^{i\gamma C(\tau)} - H_A \Pi_n e^{-i\gamma C(\tau)} \rho_{\text{init}} \Pi_m e^{i\gamma C(\tau)}) \\
&= \text{Tr} (e^{-i\gamma C(\tau)} \rho_{\text{init}} H_A e^{i\gamma C(\tau)} - H_A e^{-i\gamma C(\tau)} \rho_{\text{init}} e^{i\gamma C(\tau)}) \\
&= \text{Tr} (\rho_{\text{init}} H_A - \rho_{\text{init}} e^{i\gamma C(\tau)} H_A e^{-i\gamma C(\tau)}) \\
&= \text{Tr} \rho_{\text{init}} H_A - \sum_{k=2}^{\infty} \frac{1}{k!} (i\gamma)^k \text{Tr} [\rho_{\text{init}} (\delta_{C(\tau)})^k H_A] \\
&= -i\gamma \text{Tr} \{ \rho_{\text{init}} [C(\tau), H_A] \} - \sum_{k=2}^{\infty} \frac{1}{k!} (i\gamma)^k \text{Tr} [\rho_{\text{init}} (\delta_{C(\tau)})^k H_A] \\
&= -i\gamma \text{Tr} \{ C(\tau) [H_A, \rho_{\text{init}}] \} - \sum_{k=2}^{\infty} \frac{1}{k!} (i\gamma)^k \text{Tr} [\rho_{\text{init}} (\delta_{C(\tau)})^k H_A] \\
&= - \sum_{k=2}^{\infty} \frac{1}{k!} (i\gamma)^k \text{Tr} [\rho_{\text{init}} (\delta_{C(\tau)})^k H_A], \tag{90}
\end{aligned}$$

where we used $[\rho_{\text{init}}, H_A] = 0$ in the last line. The lowest order of γ in $\langle Q \rangle_\tau$ is the second order, which is the same order as the deviation term in Eq. (88). We can show that the higher moments of $p_\tau(Q)$ has the same dependence of γ , and thus the exchange fluctuation theorem becomes just a trivial relation in the limit $\gamma \rightarrow 0$ as shown in Eq. (84).

4.2.2 Example: a two-spin 1/2 system

We consider a quantum system which consists of two spin 1/2s initially prepared at different temperatures. The spins exchange the energy via a coupling between the two spins and this coupling is given by the Heisenberg coupling. Thus, this system is given by the following

Hamiltonian:

$$\begin{aligned} H &= H_0 + \gamma H_c, \\ H_0 &= H_A \otimes 1_B + 1_A \otimes H_B, \end{aligned}$$

where H_A , H_B and H_c are

$$H_A = -\frac{\epsilon_A}{2} \sigma_A^z \quad (91)$$

$$H_B = -\frac{\epsilon_B}{2} \sigma_B^z \quad (92)$$

$$H_c = -\frac{J}{4} \vec{\sigma}_A \cdot \vec{\sigma}_B \quad (93)$$

with σ^x , σ^y and σ^z denoting the Pauli matrices,

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (94)$$

This model is analytically solvable and the deviation from the exchange fluctuation theorem, $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$, and the heat transfer $\langle Q \rangle_\tau$ are calculated explicitly as follows:

$$\begin{aligned} \langle e^{-\Delta\beta Q} \rangle_\tau - 1 &= \frac{\gamma^2 J^2}{(\epsilon_A - \epsilon_B)^2 + \gamma^2 J^2} \operatorname{sech} \left[\frac{\beta_A \epsilon_A}{2} \right] \operatorname{sech} \left[\frac{\beta_B \epsilon_B}{2} \right] \\ &\quad \times \sinh \left[\frac{\epsilon_A}{2} (\beta_B - \beta_A) \right] \sinh \left[\frac{\beta_B}{2} (\epsilon_A - \epsilon_B) \right] \\ &\quad \times \sin^2 \left(\frac{\tau}{2\hbar} \sqrt{(\epsilon_A - \epsilon_B)^2 + \gamma^2 J^2} \right), \end{aligned} \quad (95)$$

$$\begin{aligned} \langle Q \rangle_\tau &= \frac{\epsilon_A}{2} \frac{\gamma^2 J^2}{(\epsilon_A - \epsilon_B)^2 + \gamma^2 J^2} \operatorname{sech} \left[\frac{\beta_A \epsilon_A}{2} \right] \operatorname{sech} \left[\frac{\beta_B \epsilon_B}{2} \right] \\ &\quad \times \sinh \left[\frac{1}{2} (\beta_B \epsilon_B - \beta_A \epsilon_A) \right] \\ &\quad \times \sin^2 \left(\frac{\tau}{2\hbar} \sqrt{(\epsilon_A - \epsilon_B)^2 + \gamma^2 J^2} \right). \end{aligned} \quad (96)$$

The time dependence of the deviation term $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ and the heat transfer $\langle Q \rangle_\tau$ are shown in Fig. 1.

Note that $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ and $\langle Q \rangle_\tau$ have the same dependence on γ . Thus, the ratio of these two quantities is independent of the coupling strength γ :

$$\frac{\langle e^{-\Delta\beta Q} \rangle_\tau - 1}{\langle Q \rangle_\tau / (\frac{\epsilon_A}{2})} = \frac{\sinh \left[\frac{\epsilon_A}{2} (\beta_B - \beta_A) \right] \sinh \left[\frac{\beta_B}{2} (\epsilon_A - \epsilon_B) \right]}{\sinh \left[\frac{1}{2} (\beta_B \epsilon_B - \beta_A \epsilon_A) \right]}. \quad (97)$$

Figure 2 shows parameters' dependence of the ratio that. The ratio has a finite value for almost all range of energy level difference $\epsilon_B - \epsilon_A$. As a consequence, if the coupling

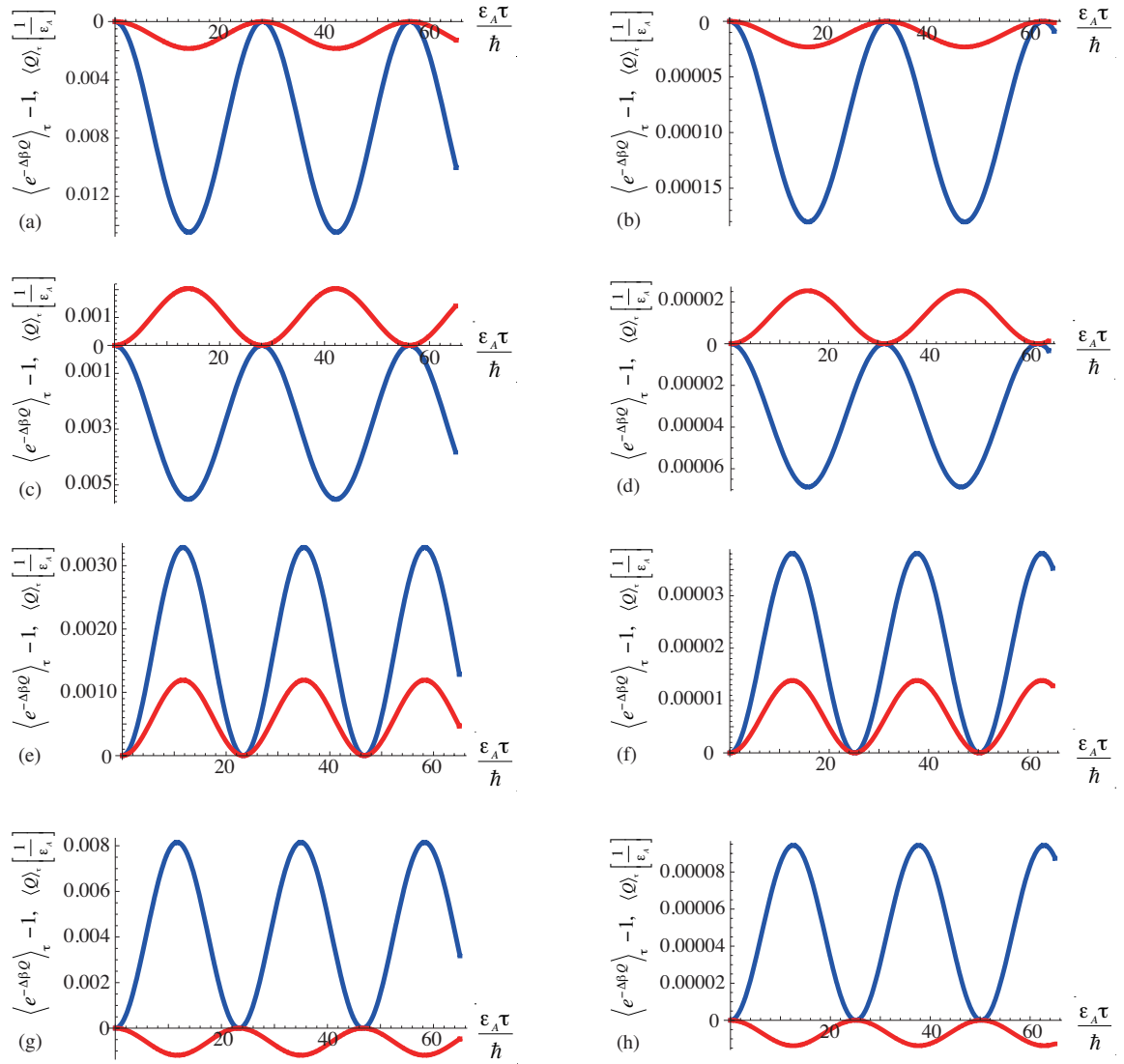


Figure 1: The quantities $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ (red line) and $\langle Q \rangle_\tau$ (blue line) in the two-spin 1/2 system. We fixed $\beta_A \epsilon_A = 2$: for $\epsilon_B / \epsilon_A = 0.8$, $\beta_B / \beta_A = 0.9$ with (a) $\gamma J / \epsilon_A = 0.1$ and (b) $\gamma J / \epsilon_A = 0.01$; for $\epsilon_B / \epsilon_A = 0.8$, $\beta_B / \beta_A = 1.1$ with (c) $\gamma J / \epsilon_A = 0.1$ and (d) $\gamma J / \epsilon_A = 0.01$; for $\epsilon_B / \epsilon_A = 1.25$, $\beta_B / \beta_A = 0.9$ with (e) $\gamma J / \epsilon_A = 0.1$ and (f) $\gamma J / \epsilon_A = 0.01$; for $\epsilon_B / \epsilon_A = 1.25$, $\beta_B / \beta_A = 1.1$ with (g) $\gamma J / \epsilon_A = 0.1$ and (h) $\gamma J / \epsilon_A = 0.01$.

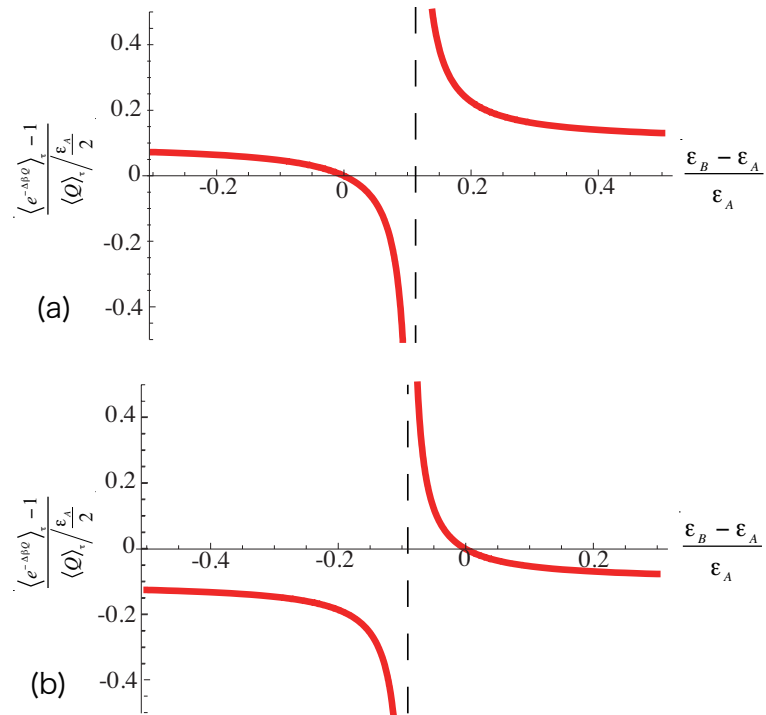


Figure 2: The ratio of the $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ to $\langle Q \rangle_\tau$. We fixed $\beta_A \epsilon_A = 2$ with (a) $\beta_B/\beta_A = 0.9$ and (b) $\beta_B/\beta_A = 1.1$. The commutable-coupling condition is fulfilled at the point $\epsilon_B - \epsilon_A = 0$.

strength γ is weak enough to neglect the deviation from the integral exchange fluctuation theorem, Eq. (95), the net heat transfer $\langle Q \rangle_\tau$ is also negligibly small. This result clearly shows that the exchange fluctuation theorem does not generally hold in the presence of a finite heat transfer.

Note, however, that the ratio in Fig. 2 vanishes for $\epsilon_A = \epsilon_B$; that is, the exchange fluctuation theorem is recovered at this particular point with a finite heat transfer. This is indeed the main target of Sec. 4.3.

4.2.3 Example: coupled harmonic oscillators

The second example is a system which consists of two harmonic oscillators. This system is given by the following Hamiltonian:

$$H = H_0 + \gamma H_c, \quad (98)$$

$$H_0 = H_A \otimes 1_B + 1_A \otimes H_B, \quad (99)$$

where the Hamiltonian H_A , H_B and H_c are

$$\begin{aligned} H_A &= \hbar\omega_A a^\dagger a, \\ H_B &= \hbar\omega_B b^\dagger b, \\ H_c &= \nu (a^\dagger b + b^\dagger a), \end{aligned} \quad (100)$$

where a and a^\dagger are the creation and annihilation operators of the oscillator A , b and b^\dagger are those of B , $[a, a^\dagger] = 1$, $[a, a] = 0$ and $[b, b^\dagger] = 1$, $[b, b] = 0$, and ν is a real number. After the transformation

$$\begin{aligned} X &= a \cos \theta + b \sin \theta, \\ Y &= -a \sin \theta + b \cos \theta, \end{aligned} \quad (101)$$

the total Hamiltonian (98) takes the diagonal form

$$H = \hbar\omega_X X^\dagger X + \hbar\omega_Y Y^\dagger Y, \quad (102)$$

with

$$\omega_X = \omega_A \cos^2 \theta + \omega_B \sin^2 \theta + 2\gamma\nu \sin \theta \cos \theta \quad (103)$$

$$\omega_Y = \omega_A \sin^2 \theta + \omega_B \cos^2 \theta - 2\gamma\nu \sin \theta \cos \theta, \quad (104)$$

and

$$\tan 2\theta = \frac{2\gamma\nu}{\omega_A - \omega_B}, \quad (\omega_A \neq \omega_B), \quad (105)$$

$$\theta = \frac{\pi}{4}, \quad (\omega_A = \omega_B). \quad (106)$$

Since the coupling Hamiltonian conserves the total boson number of the systems, $(n_a + n_b)$, once the boson number is measured, the transition during time τ occurs only between the states with the same boson number. However, if the frequencies of the both systems are not equal, $\omega_a \neq \omega_b$, the total energy does not conserve in a transition between different states. Therefore, the conservation of the total boson number does not mean the conservation of the total energy, and the energy change of the total system results from the coupling between the systems.

The time dependence of the deviation from the integral exchange fluctuation theorem and the net heat transfer during τ are shown in Fig. 3. We calculated the results numerically with the approximation that the boson number is seven at maximum. The probability to have more than seven bosons is of the order of 10^{-12} .

4.3 Commutable-coupling condition

We showed in the previous section that the exchange fluctuation theorem does not hold for arbitrary strength of the coupling. However, we discover an additional condition for which the exchange fluctuation theorem and the integral exchange fluctuation theorem hold under a finite heat transfer. In this section, we show that these theorems become exact relations when we impose an additional condition on the coupling Hamiltonian H_c . Our additional condition is

$$[H_0, H_c] = 0. \quad (107)$$

We refer to this condition as the commutable-coupling condition since the coupling Hamiltonian in the Heisenberg picture is independent of time under this condition. Jarzynski's equality for open quantum systems was discussed under this condition [28]. As we show in the following discussions, the exchange fluctuation theorem and the integral exchange fluctuation theorem hold under the commutable-coupling condition for a finite heat transfer and for a finite coupling strength γ . In Sec. 4.3.2, we confirm the presence of a finite heat transfer under the commutable-coupling condition in specific models.

4.3.1 XFT and IXFT under the commutable-coupling condition

First of all, we prove the integral exchange fluctuation theorem under the commutable-coupling condition. In this condition, the Hermitian operator $C(t)$ is reduce to a simple form

$$\begin{aligned} C(\tau) &= -i \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \left(\frac{i\tau}{\hbar} \right)^{n+1} (\delta_{H_0})^n H_c \\ &= \frac{\tau}{\hbar} H_c, \end{aligned} \quad (108)$$

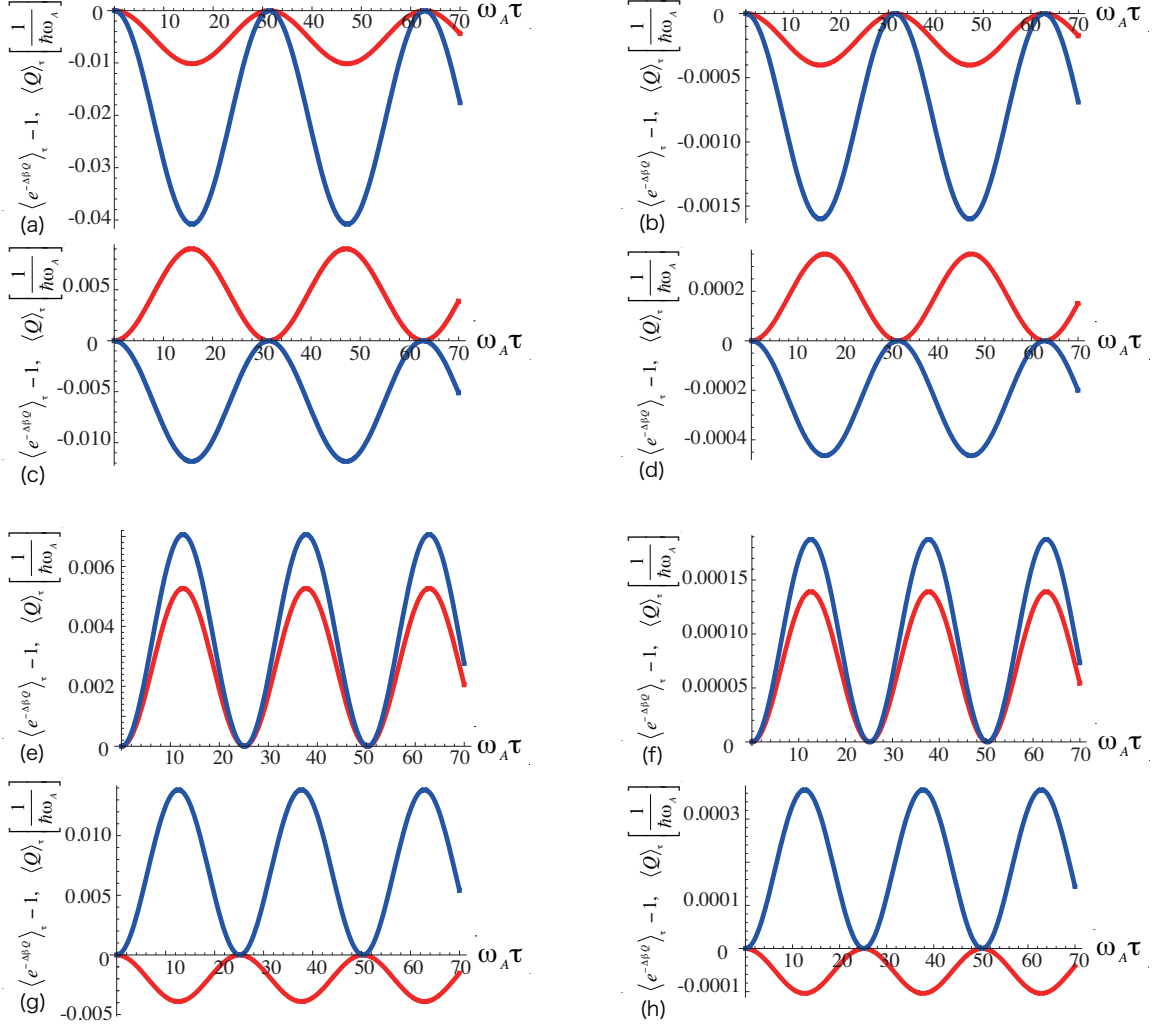


Figure 3: The quantities $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ (red line) and $\langle Q \rangle_\tau$ (blue line) in the coupled harmonic oscillators. We fixed $\beta_A \hbar \omega_A = 4$: for $\omega_B/\omega_A = 0.8$, $\beta_B/\beta_A = 0.9$ with (a) $\gamma\nu/\hbar\omega_A = 0.1$ and (b) $\gamma\nu/\hbar\omega_A = 0.01$; for $\omega_B/\omega_A = 0.8$, $\beta_B/\beta_A = 1.1$ with (c) $\gamma\nu/\hbar\omega_A = 0.1$ and (d) $\gamma\nu/\hbar\omega_A = 0.01$; for $\omega_B/\omega_A = 1.25$, $\beta_B/\beta_A = 0.9$ with (e) $\gamma\nu/\hbar\omega_A = 0.1$ and (f) $\gamma\nu/\hbar\omega_A = 0.01$; for $\omega_B/\omega_A = 1.25$, $\beta_B/\beta_A = 1.1$ with (g) $\gamma\nu/\hbar\omega_A = 0.1$ and (h) $\gamma\nu/\hbar\omega_A = 0.01$.

and we can immediately confirm that the integral exchange fluctuation theorem rigorously holds by substituting Eq. (108) to Eq. (88) and using the commutable-coupling condition:

$$\begin{aligned}
\langle e^{-\Delta\beta Q} \rangle_\tau - 1 &= \sum_{k=2}^{\infty} \frac{1}{k!} (-i\gamma)^k \text{Tr} [\rho_{\text{init}} e^{-\beta_B H_0} (\delta_{C(\tau)})^k e^{\beta_B H_0}] \\
&= \sum_{k=2}^{\infty} \frac{1}{k!} (-i\frac{\gamma}{\hbar} \tau)^k \text{Tr} [\rho_{\text{init}} e^{-\beta_B H_0} (\delta_{H_c})^k e^{\beta_B H_0}] \\
&= 0.
\end{aligned} \tag{109}$$

Note that the derivation of Eq. (109) is independent of the coupling strength γ . This suggests that this relation holds even if the system is in a strongly out-of-equilibrium situation.

Next, we prove the exchange fluctuation theorem under the commutable-coupling condition. From Eq. (47), the time-evolution operator in the Schrödinger picture is

$$U(t) = e^{-i\frac{H_0}{\hbar}t} e^{-i\gamma\frac{H_c}{\hbar}t} = e^{-i\gamma\frac{H_c}{\hbar}t} e^{-i\frac{H_0}{\hbar}t}, \tag{110}$$

and we have

$$\begin{aligned}
\langle n|U(t)|m \rangle &= e^{-i\frac{E_n}{\hbar}t} \langle n|e^{-i\gamma\frac{H_c}{\hbar}t}|m \rangle \\
&= \langle n|e^{-i\gamma\frac{H_c}{\hbar}t}|m \rangle e^{-i\frac{E_m}{\hbar}t},
\end{aligned}$$

which is followed by

$$\left(e^{-i\frac{E_n}{\hbar}t} - e^{-i\frac{E_m}{\hbar}t} \right) \langle n|e^{-i\gamma\frac{H_c}{\hbar}t}|m \rangle = 0. \tag{111}$$

If $\langle n|e^{-i\gamma\frac{H_c}{\hbar}t}|m \rangle$ has a finite value, Eq. (111) implies

$$e^{-i\frac{E_m}{\hbar}\tau} - e^{-i\frac{E_n}{\hbar}\tau} = 0.$$

This equation relates the second measurement result E_n to the first measurement result E_m as follows:

$$E_n = E_m + \frac{2\pi\hbar}{\tau} k, \tag{112}$$

where k is an integer. We further assume that there exists more than two points of time at which $\langle n|e^{-i\gamma\frac{H_c}{\hbar}\tau}|m \rangle$ is finite for fixed (m, n) . For $\tau_2 \neq \tau_1$, this assumption leads to

$$\begin{aligned}
E_n &= E_m + \frac{2\pi\hbar}{\tau_1} k \\
&= E_m + \frac{2\pi\hbar}{\tau_2} k.
\end{aligned}$$

Then we have

$$\left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) k = 0. \quad (113)$$

Therefore we arrive at $k = 0$ in Eq. (112) and obtain from Eq. (111)

$$\langle n | e^{-i\gamma \frac{H_c}{\hbar} \tau} | m \rangle = \langle n | e^{-i\gamma \frac{H_c}{\hbar} \tau} | m \rangle \delta_{E_n, E_m}. \quad (114)$$

From the above discussion, we see that the commutable-coupling condition restricts the realized measurement outcomes because the quantity $\langle n | e^{-i\gamma \frac{H_c}{\hbar} \tau} | m \rangle$ is directly related to the transition probability as follows:

$$\begin{aligned} T_{m \rightarrow n}(\tau) &= |\langle n | e^{-i\gamma \frac{H_c}{\hbar} \tau} | m \rangle|^2 \\ &= |\langle n | e^{-i\gamma \frac{H_c}{\hbar} \tau} | m \rangle|^2 \delta_{E_n, E_m} \end{aligned} \quad (115)$$

with $E_m = E_{m_A}^A + E_{m_B}^B$. Substituting this transition probability into Eq. (79), we conclude that the exchange fluctuation theorem becomes an exact relation under the commutable-coupling condition:

$$\begin{aligned} p_\tau(Q) &= \sum_{m,n} P(m, n | \tau) \delta [Q - Q_{m \rightarrow n}] \times \delta_{E_n, E_m} \\ &= e^{\Delta\beta Q} \sum_{m,n} e^{\beta_B (E_{m_A}^A + E_{m_B}^B - E_{n_A}^A - E_{n_B}^B)} P(n, m | -\tau) \delta [Q + (E_{n_A}^A - E_{m_A}^A)] \times \delta_{E_n, E_m} \\ &= e^{\Delta\beta Q} \sum_{m,n} P(n, m | -\tau) \delta [-Q - Q_{n \rightarrow m}] \times \delta_{E_n, E_m} \\ &= e^{\Delta\beta Q} p_\tau^R(-Q). \end{aligned} \quad (116)$$

We thus proved the both exchange fluctuation theorem and integral exchange fluctuation theorem under the commutable-coupling condition.

Next, we show that a finite heat transfer between two systems does exist under the commutable-coupling condition. The ensemble average of the heat transfer is written from Eq. (90) as

$$\langle Q \rangle_\tau = - \sum_{k=2}^{\infty} \frac{1}{k!} \left(i \frac{\gamma \tau}{\hbar} \right)^k \text{Tr} \left[\rho_{\text{init}} (\delta_{H_c})^k H_A \right].$$

Note that $\delta_{H_c} H_A$ is generally finite although we assume $\delta_{H_c} H_0 = \delta_{H_c} (H_A + H_B) = 0$. Therefore the heat can flow between the two systems and the exchange fluctuation theorem is a nontrivial relation. In the next section, we show the presence of a finite heat transfer for the specific models that we considered in Sec. 4.2.2.

4.3.2 Example: a two-spin 1/2 system

We consider again the quantum system which consists of two spin 1/2s. In order to identify the commutable-coupling condition in this model, we calculate the commutation relations

$$\begin{aligned} [H_A, H_c] &= \frac{J}{8} \epsilon_A \sum_{j=x,y,z} [\sigma_A^z, \sigma_A^j] \sigma_B^j \\ &= i \frac{J}{4} \epsilon_A (\sigma_A^y \sigma_B^x - \sigma_A^x \sigma_B^y), \end{aligned} \quad (117)$$

$$\begin{aligned} [H_B, H_c] &= \frac{J}{8} \epsilon_B \sum_{j=x,y,z} [\sigma_B^z, \sigma_B^j] \sigma_A^j \\ &= i \frac{J}{4} \epsilon_B (\sigma_A^x \sigma_B^y - \sigma_A^y \sigma_B^x) \\ &= -i \frac{J}{4} \epsilon_B (\sigma_A^y \sigma_B^x - \sigma_A^x \sigma_B^y). \end{aligned} \quad (118)$$

Then the commutable-coupling condition is written as

$$[H_0, H_c] = i \frac{J}{4} (\epsilon_A - \epsilon_B) (\sigma_A^y \sigma_B^x - \sigma_A^x \sigma_B^y) = 0. \quad (119)$$

From Eq. (119), we find that the relation $\epsilon_A = \epsilon_B$ corresponds to the commutable-coupling condition. Indeed, Fig. 2 shows that the exchange fluctuation theorem is recovered with a finite heat transfer for $\epsilon_A = \epsilon_B$.

Under the commutable-coupling condition $\epsilon_A = \epsilon_B$, Eqs. (95) and (96) become

$$\langle e^{-\Delta\beta Q} \rangle_\tau - 1 = 0, \quad (120)$$

$$\langle Q \rangle_\tau = \frac{\epsilon}{2} \operatorname{sech} \left[\frac{\epsilon}{2} \beta_A \right] \operatorname{sech} \left[\frac{\epsilon}{2} \beta_B \right] \sinh \left[\frac{\epsilon}{2} (\beta_B - \beta_A) \right] \sin^2 \left[\frac{\gamma J}{2\hbar} \tau \right]. \quad (121)$$

The time dependence of the net heat transfer $\langle Q \rangle_\tau$ and the deviation from the integral fluctuation theorem $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ are shown in Fig. 4. The net heat transfer has finite values, while the integral exchange fluctuation theorem holds.

4.3.3 Example: coupled harmonic oscillators

We show how the commutable-coupling condition is expressed in the model system which consists of coupled harmonic oscillators. In order to identify the commutable-coupling condition in this model, we calculate the commutation relations

$$\begin{aligned} [H_A, H_c] &= \hbar\nu\omega_A [a^\dagger a, a^\dagger b + b^\dagger a] \\ &= \hbar\nu\omega_A (a^\dagger b - b^\dagger a), \end{aligned} \quad (122)$$

$$\begin{aligned} [H_B, H_c] &= \hbar\nu\omega_B [b^\dagger b, a^\dagger b + b^\dagger a] \\ &= -\hbar\nu\omega_B (a^\dagger b - b^\dagger a). \end{aligned} \quad (123)$$

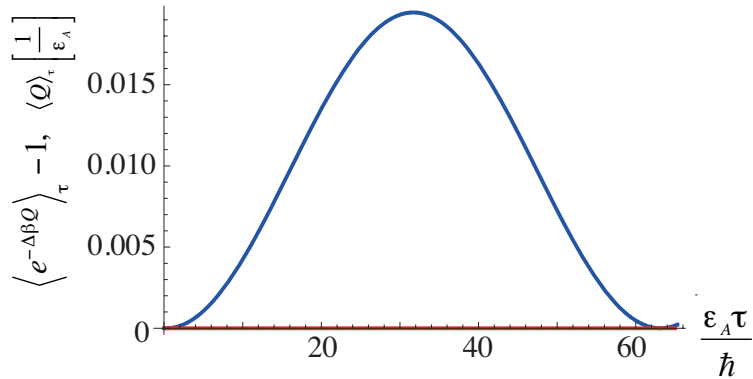


Figure 4: The quantities $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ (red line) and $\langle Q \rangle_\tau$ (blue line) under the commutable-coupling condition in two-spin 1/2 system for $\epsilon_B/\epsilon_A = 1$, $\beta_A\epsilon_A = 2$, $\beta_B/\beta_A = 1.1$, and $\gamma J/\epsilon_A = 0.01$.

Then the commutable-coupling condition is written as

$$[H_0, H_c] = \hbar\nu(\omega_A - \omega_B)(a^\dagger b - b^\dagger a) = 0. \quad (124)$$

From Eq. (124), we find that the relation $\omega_A = \omega_B$ corresponds to the commutable-coupling condition. The time dependence of the net heat transfer $\langle Q \rangle_\tau$ and the deviation from the integral fluctuation theorem $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ are shown in Fig. 5. We calculated the results numerically with the approximation that the boson number is seven at maximum. The probability to have more than seven bosons is of the order of 10^{-12} .

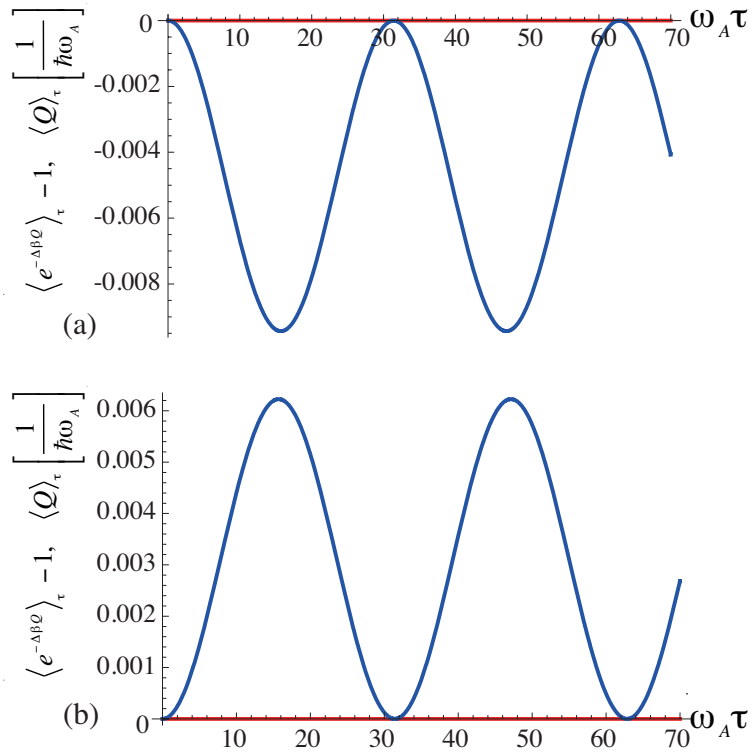


Figure 5: The quantities $\langle e^{-\Delta\beta Q} \rangle_\tau - 1$ (red line) and $\langle Q \rangle_\tau$ (blue line) under the commutable-coupling condition in coupled harmonic oscillators for $\omega_B/\omega_A = 1$, $\beta_A \hbar \omega_A = 4$ and $\gamma\nu/\hbar\omega_A = 0.1$ with (a) $\beta_B/\beta_A = 0.9$ and (b) $\beta_B/\beta_A = 1.1$.

5 Conclusions

To summarize, we showed that the exchange fluctuation theorem in its original form does not generally hold in the presence of a finite heat transfer. In the limit $\gamma \rightarrow 0$, the k th moments of $p_\tau(Q)$ also vanish. We also showed that the deviation from the exchange fluctuation theorem has generally the same dependence on the coupling strength γ as the ensemble average of the heat transfer between the systems and both analytically and numerically confirmed this with specific models. This means that there is no heat transfer when the coupling strength γ is small enough to neglect the deviation from the exchange fluctuation theorem. In this case, the exchange fluctuation theorem reduces to a trivial relation and has no information about the heat transfer.

However, we found a condition for the exchange fluctuation theorem to hold exactly and we referred to it as the commutable-coupling condition. Under this condition, the exchange fluctuation theorem becomes an exact relation independently of the coupling strength γ under the existence of a finite heat transfer. We confirmed this in specific models.

The deviation from the exchange fluctuation theorem consists of the commutation relation between the Hamiltonian of the total system and the coupling Hamiltonian. Therefore, the non-commutativity of the observable in quantum mechanics plays an important role in the deviation.

Acknowledgments

I would like to express my gratitude to Professor Naomichi Hatano for his detailed comments, valuable suggestions, and patient support. I am also grateful to Professor Hisao Hayakawa for his valuable comments and suggestions. Finally I express my sincere thanks and appreciation for all the members of the Hatano group for their tender supports.

A Time-reversed evolution

We explain why the joint probability in the reverse process can be written as Eqs. (6) and (69). As in Eq. (60), the time-reversal invariance for the time-dependent Hamiltonian $H(t)$, represented as

$$\Theta H(t) = H(t)\Theta, \quad (125)$$

is equivalent to

$$\Theta \Pi_{m_t} = \Pi_{m_t} \Theta, \quad (126)$$

where Π_{m_t} is the projection operator onto an instantaneous eigenstate m_t of the Hamiltonian at time t , $H(t)$, with the eigenvalue E_{m_t} .

The density operator in the time-reversed dynamics evolves according to the von Neumann equation

$$i\hbar \frac{d}{dt} \rho^R(t) = [H^R(t), \rho^R(t)]$$

with the initial condition

$$\rho^R(0) = \rho_{\text{init}}^R. \quad (127)$$

The Hamiltonian in the time-reversed dynamics is defined as

$$H^R(t) = \Theta H(\tau - t) \Theta^{-1}, \quad (128)$$

where τ is the time duration of the dynamics that we consider. The solution of the von Neumann equation is given by

$$\rho^R(t) = U^R(t, 0) \rho^R(0) U^R(t, 0)^\dagger,$$

where $U^R(t, 0)$ is the time-evolution operator in the time-reversed dynamics and is a solution of the Schrödinger equation,

$$i\hbar \frac{d}{dt} U^R(t, 0) = H^R(t, 0) U^R(t, 0) \quad (129)$$

with the initial condition $U^R(0, 0) = 1$. The solution $U^R(t, 0)$ is given by [29, 30]

$$U^R(t, 0) = \Theta U(\tau - t, \tau) \Theta^{-1}. \quad (130)$$

We can confirm this by substituting Eq. (130) into the Schrödinger equation (129):

$$\begin{aligned} i\hbar \frac{d}{dt} (\Theta U(\tau - t, \tau) \Theta^{-1}) &= \Theta \left(-i\hbar \frac{d}{dt} U(\tau - t, \tau) \Theta^{-1} \right) \\ &= \Theta (H(\tau - t) U(\tau - t, \tau) \Theta^{-1}) \\ &= \Theta H(\tau - t) \Theta^{-1} \Theta U(\tau - t, \tau) \Theta^{-1} \\ &= H^R(t) (\Theta U(\tau - t, \tau) \Theta^{-1}), \end{aligned} \quad (131)$$

and

$$\begin{aligned}
\Theta U(\tau - t, \tau) \Theta^{-1} |_{t=0} &= \Theta U(\tau, \tau) \Theta^{-1} \\
&= \Theta \Theta^{-1} \\
&= 1.
\end{aligned} \tag{132}$$

We used the antilinearity of Θ in the first line in Eq. (131).

Using these quantities, we define the joint probability in the reverse process as follows:

$$P^{\text{R}}(n, m|\tau) = \text{Tr} \left[\Pi_{m_0} U_{\text{R}}(\tau, 0) \Pi_{n_\tau} \rho_{\text{init}}^{\text{R}} \Pi_{n_\tau} U_{\text{R}}^\dagger(\tau, 0) \right] \tag{133}$$

with

$$\rho_{\text{init}}^{\text{R}} = \Theta \rho_\tau \Theta^{-1}. \tag{134}$$

If the initial state of the forward process is the Gibbs state $\rho_{\text{init}} = e^{-\beta H(0)}/Z_0$ at an inverse temperature β , for example, we choose ρ_τ as the Gibbs state of the system described by the Hamiltonian $H(\tau)$ at the inverse temperature β :

$$\rho_\tau = \frac{e^{-\beta H(\tau)}}{Z_\tau}, \tag{135}$$

where $Z_\tau = \text{Tr} e^{-\beta H(\tau)}$ is the partition function.

If the system is time-reversal invariant, $\Theta H(t) = H(t) \Theta$, the Hamiltonian in the time-reversed dynamics at time t coincides with the Hamiltonian in the forward dynamics at time $\tau - t$:

$$\begin{aligned}
H^{\text{R}}(t) &= \Theta H(\tau - t) \Theta^{-1} \\
&= H(\tau - t) \Theta \Theta^{-1} \\
&= H(\tau - t),
\end{aligned} \tag{136}$$

where we used Eq. (125). In this case, we can represent the joint probability in the reverse process with the quantities in the forward dynamics by substituting Eqs. (127) and (130) into Eq. (133):

$$\begin{aligned}
P^{\text{R}}(n, m|\tau) &= \text{Tr} \left[\Pi_{m_0} (\Theta U(0, \tau) \Theta^{-1}) \Pi_{n_\tau} (\Theta \rho_\tau \Theta^{-1}) \Pi_{n_\tau} (\Theta U^\dagger(0, \tau) \Theta^{-1}) \right] \\
&= \text{Tr} \left[\Theta \Pi_{m_0} U^\dagger(\tau, 0) \Theta^{-1} \Theta \Pi_{n_\tau} \rho_\tau \Theta^{-1} \Theta \Pi_{n_\tau} U(\tau, 0) \Theta^{-1} \right] \\
&= \text{Tr} \left[\Pi_{m_0} U^\dagger(\tau, 0) \Pi_{n_\tau} \rho_\tau \Pi_{n_\tau} U(\tau, 0) \right],
\end{aligned} \tag{137}$$

where we used Eq. (126).

We can generalize the representation of the joint probability in the time-reversed dynamics to the case where a static magnetic field is present.

References

- [1] C. Jarzynski and D. K. Wójcik, Phys. Rev. Lett. **92**, 230602 (2004).
- [2] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
- [3] D. J. Evans, D. J. Searles, Phys. Rev. E. **50**, 1645 (1994).
- [4] G. Gallavotti and E. G. D. Cohen, Phys. Rev. E. **74**, 2694 (1995).
- [5] J. Kurchan, J. Phys. A: Math. Gen. **31**, 3719 (1998).
- [6] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95** 333 (1999).
- [7] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
- [8] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
- [9] G. E. Crooks, Phys. Rev. E. **60**, 2721 (1999).
- [10] D. J. Evans and D. J. Searles, Adv. Phys. **51**, 1529 (2002).
- [11] E. M. Sevick, R. Prabhakar, S. R. Williams, and D. J. Searles, Annu. Rev. Phys. Chem. **59**, 603 (2008).
- [12] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **92**, 140601 (2004).
- [13] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco (Jr), and C. Bustamante, Science **296**, 1832 (2002).
- [14] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco (Jr), and C. Bustamante, Nature **437**, 231 (2005).
- [15] S. Yukawa, J. Phys. Soc. Jpn. **69**, 2367 (2000).
- [16] T. Monnai and S. Tasaki, cond-mat/0308337.
- [17] A. E. Allahverdyan and Th. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).
- [18] M. F. Gelin and D. S. Kosov, Phys. Rev. E **78**, 011116 (2008).
- [19] J. Kurchan, cond-mat/0007360.
- [20] H. Tasaki, cond-mat/0009244.
- [21] S. Mukamel, Phys. Rev. Lett. **90**, 170604 (2003).
- [22] T. Monnai, Phys. Rev. E. **72**, 027102 (2005).

- [23] P. Talkner, E. Lutz, and P. Hanggi, Phys. Rev. E. (R)**75**, 050102 (2007).
- [24] K. Saito and A. Dhar, Phys. Rev. Lett. **99**, 180601 (2007).
- [25] K. Saito and Y. Utsumi, Phys. Rev. B. **78**, 115429 (2008).
- [26] D. Andrieux, P. Gaspard, T. Monnai, S. Tasaki cond-mat/08113687.
- [27] J. J. Sakurai, *Modern Quantum Mechanics*, (Benjamin, Menlo Park, California, 1985).
- [28] J. Teifel and G. Mahler, Phys. Rev. E **76**, 051126 (2007).
- [29] M. Esposito, U. Harbola, and S. Mukamel, cond-mat/08113717.
- [30] D. Andrieux and P. Gaspard Phys. Rev. Lett. **100**, 230404 (2008).