Fluctuation theorems for quantum master equations

Massimiliano Esposito* and Shaul Mukamel
Department of Chemistry, University of California, Irvine, California 92697, USA
(Received 17 November 2005; published 24 April 2006)

A quantum fluctuation theorem for a driven quantum subsystem interacting with its environment is derived based solely on the assumption that its reduced density matrix obeys a closed evolution equation—i.e., a quantum master equation (QME). Quantum trajectories and their associated entropy, heat, and work appear naturally by transforming the QME to a time-dependent Liouville space basis that diagonalizes the instantaneous reduced density matrix of the subsystem. A quantum integral fluctuation theorem, a steady-state fluctuation theorem, and the Jarzynski relation are derived in a similar way as for classical stochastic dynamics.

DOI: 10.1103/PhysRevE.73.046129 PACS number(s): 05.30.Ch, 05.70.Ln, 03.65.Yz

I. INTRODUCTION

The fluctuation theorems and the Jarzynski relation are some of a handful of powerful results of nonequilibrium statistical mechanics that hold far from thermodynamic equilibrium. Originally derived in the context of classical mechanics [1], the Jarzynski relation has been subsequently extended to stochastic dynamics [2]. It relates the distribution of the work done by a driving force of arbitrary speed on a system initially at equilibrium (nonequilibrium property) to the free energy difference between the initial and final equilibrium states of the system (equilibrium property). This remarkable relation has recently been shown to hold for arbitrary coupling strength between the system and environment (see Jarzynski's reply [3] to criticism from Cohen and Mauzerall [4]). The fluctuation theorems are based on a fundamental relation connecting the entropy production of a single system trajectory to the logarithm of the ratio of the probability of forward and backward trajectories [5]. The ensemble average of the trajectory entropy production is the macroscopic entropy production of the system whereas its distribution gives rise to various kinds of fluctuation theorems. The first has been derived for classical mechanics and initially for deterministic (but non-Hamiltonian) thermostatted systems [6–8]. Some interesting studies of fluctuation relations valid for far from equilibrium classical Hamiltonian systems were made even earlier [9–11]. Fluctuation theorems for systems with stochastic dynamics have also been developed [12–19]. For classical stochastic dynamics, the connection between the fluctuation theorem and the Jarzynski relation has been established by Crooks [17]. Seifert has recently provided a unified description of the different fluctuation relations and of the Jarzynski relation for classical stochastic processes described by master equations [19].

The understanding of these two fundamental relations in quantum mechanics is still not fully established. Quantum Jarzynski relations have been investigated in [20–23]. Quantum fluctuation theorems have been developed only in a few restricted situations [24–27]. A quantum exchange fluctuation theorem has also been considered in [28]. Some interesting considerations of the quantum definition of work in the previous studies have been made in [29].

It should be noted that the dynamics of an isolated (whether driven or not) quantum system is unitary and its von Neumann entropy is time independent. Therefore, fluctuation theorems for such closed systems are useful only provided one defines some reduced macrovariable dynamics or some measurement process on the system [30].

The purpose of this paper is to provide a unified derivation for the different quantum fluctuation relations (an integral fluctuation theorem, a steady-state fluctuation theorem, and the Jarzynski relation). We build upon the unification of the different fluctuation relations recently accomplished by Seifert [19] for classical stochastic dynamics described by a birth and death master equation (BDME). Quantum evolution involves coherences which make its interpretation in term of trajectories not obvious. Nevertheless, we show that it is possible to formally develop a trajectory picture of quantum dynamics which allows one to uniquely represent entropy, heat, and work distributions. This relies on the single assumption that the reduced dynamics of a driven quantum subsystem interacting with its environment is described by a closed evolution equation for the density matrix of the subsystem—i.e., a quantum master equation (QME) [31–34]. However, while the physical quantities defined along classical trajectories are conceptually clear and experimentally measurable, how to measure the physical quantities associated with quantum trajectories remains an open issue intimately connected to quantum measurement.

The plan of the paper is as follows: We start in Sec. II by defining quantum heat and quantum work for a driven subsystem interacting with its environment, consistent with thermodynamics. We then discuss the consequences of defining heat and work in terms of the time-dependent basis which diagonalizes the subsystem density matrix in Sec. III. In Sec. IV, we show that by assuming a QME for the subsystem reduced density matrix we can recast its solution in a representation which takes the form of a BDME with time-dependent rates. In Sec. V, we show that the BDME representation allows us to split the entropy evolution into two parts: the entropy flow associated with exchange processes with the environment and the entropy production associated

*Also at Center for Nonlinear Phenomena and Complex Systems, Université Libre de Bruxelles, Code Postal 231, Campus Plaine, B-1050 Brussels, Belgium.

with subsystem internal irreversible processes. In Sec. VI, we show that the BDME representation naturally allows one to define quantum trajectories as well as their associated entropy flow and production. We then derive the fundamental relation of this paper [Eq. (67)] which will allow us to derive, in Sec. VII, a quantum integral fluctuation theorem and, in Sec. VIII, a quantum steady-state fluctuation theorem. Having identified in Sec. IX the heat and work associated with the quantum trajectories, we show in Sec. X that the fundamental relation of Sec. VI also allows one to derive a quantum Jarzynski relation. We finally draw conclusions in Sec. XI.

II. AVERAGE HEAT AND WORK

We start by defining the average quantum heat and work for a driven subsystem interacting with its environment and show the consistency of these definitions with thermodynamics. Heat and work can be rigorously expressed in terms of the reduced density matrix of the subsystem without having to refer explicitly to the environment.

We consider a driven subsystem with Hamiltonian $\hat{H}_S(t)$. Everywhere in this paper we denote operators with a caret (and superoperators with two carets) and use the Schrödinger picture where the time dependence of the observables is explicit and comes exclusively from external driving. We could also have written $\hat{H}_S(\lambda(t))$, where $\lambda(t)$ is the external-time-dependent driving. This subsystem is interacting with its environment whose Hamiltonian is $\hat{H}_B$. The interaction energy between the subsystem and environment is described by $\hat{H}_I$. The Hamiltonian of the total system reads therefore

$$\hat{H}_T(t) = \hat{H}_S(t) + \hat{H}_B + \hat{H}_I. \tag{1}$$

We have assumed that the driving acts exclusively on the subsystem and does not affect $\hat{H}_B$ and $\hat{H}_I$.

The state of the total system is described by the density matrix $\hat{\rho}(t)$ which obeys the von Neumann equation

$$\dot{\hat{\rho}}(t) = -i[\hat{H}_T(t), \hat{\rho}(t)] = \hat{\hat{\mathcal{L}}}(t)\hat{\rho}(t). \tag{2}$$

The energy of the total system is given by

$$\langle \hat{H}_T \rangle_t \equiv \mathrm{Tr}\hat{H}_T(t)\hat{\rho}(t). \tag{3}$$

The change in the total energy between time 0 and t due to the time-dependent driving is therefore given by

$$\Delta E_T(t) \equiv \int_0^t d\tau \frac{d\langle \hat{H}_T \rangle_\tau}{d\tau} = W_T(t) + Q_T(t), \tag{4}$$

where the work and heat have, respectively, been defined as

$$W_T(t) \equiv \int_0^t d\tau \mathrm{Tr}\dot{\hat{H}}_T(t)\hat{\rho}(t), \tag{5}$$

$$Q_T(t) \equiv \int_0^t d\tau \mathrm{Tr}\hat{H}_T(t)\dot{\hat{\rho}}(t). \tag{6}$$

Using the von Neumann equation (2) and the invariance of the trace under cyclic permutation [35], we find that no heat is generated in the isolated total system

$$Q_T(t) = -i\int_0^t d\tau \mathrm{Tr}\hat{H}_T(t)[\hat{H}_T(t), \hat{\rho}(t)] = 0. \tag{7}$$

We next turn to the subsystem. Its reduced density matrix is defined as $\hat{\sigma}(t) \equiv \mathrm{Tr}_B\, \hat{\rho}(t)$, and its energy is given by

$$\langle \hat{H}_S \rangle_t \equiv \mathrm{Tr}\hat{H}_S(t)\hat{\rho}(t) = \mathrm{Tr}_S\hat{H}_S(t)\hat{\sigma}(t). \tag{8}$$

The change in this energy between time 0 and t is given by

$$\Delta E_S(t) \equiv \int_0^t d\tau \frac{d\langle \hat{H}_S \rangle_\tau}{d\tau} = W_S(t) + Q_S(t), \tag{9}$$

where the work and heat are defined as

$$W_S(t) \equiv \int_0^t d\tau \mathrm{Tr}\dot{\hat{H}}_S(\tau)\hat{\rho}_T(\tau) = \int_0^t d\tau \mathrm{Tr}_S\dot{\hat{H}}_S(\tau)\hat{\sigma}(\tau), \tag{10}$$

$$Q_S(t) \equiv \int_0^t d\tau \mathrm{Tr}\hat{H}_S(\tau)\dot{\hat{\rho}}_T(\tau) = \int_0^t d\tau \mathrm{Tr}_S\hat{H}_S(\tau)\dot{\hat{\sigma}}(\tau). \tag{11}$$

Since the time dependence of the total system Hamiltonian comes solely from the subsystem Hamiltonian, $\dot{\hat{H}}_B = \dot{\hat{H}}_I = 0$, $\dot{\hat{H}}_T = \dot{\hat{H}}_S$, and the work done by the driving force on the subsystem is the same as the work done by this force on the total system,

$$W_T(t) = W_S(t) \equiv W(t). \tag{12}$$

This also means that the energy increase in the subsystem minus the amount of heat which went to the environment is equal to the energy increase in the total system:

$$W(t) = \Delta E_T(t) = \Delta E_S(t) - Q_S(t). \tag{13}$$

It should be noticed that due to the absence of heat flux in the total system $Q_T(t) = 0$. Using Eq. (6) with Eqs. (1) and (11), we can also express the heat going from the subsystem to the environment as

$$Q_S(t) = -\int_0^t d\tau \frac{d\langle \hat{H}_B \rangle_\tau}{d\tau} - \int_0^t d\tau \frac{d\langle \hat{H}_I \rangle_\tau}{d\tau}. \tag{14}$$

III. CALCULATING HEAT AND WORK IN A TIME-DEPENDENT BASIS

As will become clear in Sec. IV, in order to associate trajectories with the quantum dynamics, one needs to represent the dynamics in a time-dependent basis. This is a fundamental difference from classical thermodynamics where the basis set (coordinate system) is fixed. In order to associate heat and work with single trajectories they must be de-

fined with respect to a time-dependent basis set. The ensemble average of the quantities defined for the trajectories will therefore also depend on this basis set. For this reason, we introduce a modified definition of heat and work. The effect of the basis time dependence on heat and work is given in Appendix A.

The energy of the total system (3) can also be written as

$$\langle \hat{H}_T \rangle_t = \sum_\alpha P_t^T(\alpha) \langle \alpha_t | \hat{H}_T(t) | \alpha_t \rangle, \tag{15}$$

where we have introduced the time-dependent basis $\{|\alpha_t\rangle\}$ which diagonalizes the instantaneous density matrix at all times:

$$\langle \alpha_t | \hat{\rho}(t) | \alpha_t' \rangle = \langle \alpha_t | \hat{\rho}(t) | \alpha_t \rangle \delta_{\alpha\alpha'} \equiv P_t^T(\alpha) \delta_{\alpha\alpha'}. \tag{16}$$

The change in this energy between time 0 and t due to the time-dependent driving can therefore be rewritten as

$$\Delta E_T(t) = \tilde{W}_T(t) + \tilde{Q}_T(t), \tag{17}$$

where the modified work and heat have, respectively, been defined as

$$\tilde{W}_T(t) \equiv \int_0^t d\tau \sum_\alpha P_\tau^T(\alpha) \frac{d}{d\tau} [\langle \alpha_\tau | \hat{H}_T(\tau) | \alpha_\tau \rangle], \tag{18}$$

$$\tilde{Q}_T(t) \equiv \int_0^t d\tau \sum_\alpha \dot{P}_\tau^T(\alpha) \langle \alpha_\tau | \hat{H}_T(\tau) | \alpha_\tau \rangle. \tag{19}$$

Because the total system is driven but otherwise isolated, its evolution is unitary and we have (see Appendix A)

$$\tilde{W}_T(t) = W_T(t) = \Delta E_T(t), \tag{20}$$

$$\tilde{Q}_T(t) = Q_T(t) = 0. \tag{21}$$

This means that defining heat and work on the time-dependent basis which diagonalizes the instantaneous density matrix for unitary evolution is equivalent to the original definition of heat and work in a time-independent basis.

The energy of the subsystem (9) can also be written in analogy to Eq. (15) as

$$\langle \hat{H}_S \rangle_t = \sum_m P_t(m) \langle m_t | \hat{H}_S(t) | m_t \rangle, \tag{22}$$

where we have introduced the time-dependent basis $\{|m_t\rangle\}$ which diagonalizes the instantaneous subsystem reduced density matrix:

$$\langle m_t | \hat{\sigma}(t) | m_t' \rangle = \langle m_t | \hat{\sigma}(t) | m_t \rangle \delta_{mm'} \equiv P_t(m) \delta_{mm'}. \tag{23}$$

Let us note for future reference that

$$\frac{d}{dt}(\langle m_t | \hat{\sigma}(t) | m_t' \rangle) = \langle m_t | \dot{\hat{\sigma}}(t) | m_t' \rangle + \langle \dot{m}_t | \hat{\sigma}(t) | m_t' \rangle + \langle m_t | \hat{\sigma}(t) | \dot{m}_t' \rangle. \tag{24}$$

Equations (23) and (24) give

$$\langle m_t | \dot{\hat{\sigma}}(t) | m_t' \rangle = \dot{P}_t(m) \delta_{mm'} - \langle \dot{m}_t | m_t' \rangle P_t(m') - \langle m_t | \dot{m}_t' \rangle P_t(m). \tag{25}$$

Notice also that for $m=m'$, we have

$$\langle m_t | \dot{\hat{\sigma}}(t) | m_t \rangle = \dot{P}_t(m) \tag{26}$$

because $\langle \dot{m}_t | m_t \rangle + \langle m_t | \dot{m}_t \rangle = \frac{d}{dt}(\langle m_t | m_t \rangle) = 0$.

Using Eq. (22), the change in the subsystem energy between time 0 and t can be rewritten as

$$\Delta E_S(t) = \tilde{W}_S(t) + \tilde{Q}_S(t), \tag{27}$$

where the work and heat are defined in analogy to Eqs. (18) and (19) as

$$\tilde{W}_S(t) \equiv \int_0^t d\tau \sum_m P_\tau(m) \frac{d}{d\tau} [\langle m_\tau | \hat{H}_S(\tau) | m_\tau \rangle], \tag{28}$$

$$\tilde{Q}_S(t) \equiv \int_0^t d\tau \sum_m \dot{P}_\tau(m) \langle m_\tau | \hat{H}_S(\tau) | m_\tau \rangle. \tag{29}$$

It is shown in Appendix A that the work and heat defined in the time-dependent basis $\{|m_t\rangle\}$ are related to the original work and heat defined in any time-independent basis by

$$\tilde{W}_S(t) = W_S(t) + A_S(t), \tag{30}$$

$$\tilde{Q}_S(t) = Q_S(t) - A_S(t), \tag{31}$$

where

$$A_S(t) \equiv \int_0^t d\tau \sum_m P_\tau(m) [\langle \dot{m}_\tau | \hat{H}_S(\tau) | m_\tau \rangle + \langle m_\tau | \hat{H}_S(\tau) | \dot{m}_\tau \rangle]. \tag{32}$$

It should be emphasized that both the original and modified work and heat of the subsystem can be defined exclusively in terms of subsystem quantities, without referring explicitly to the environment.

Using Eqs. (12) and (13) with Eqs. (30) and (31), we get

$$\Delta E_T(t) = W(t) = \tilde{W}_S(t) - A_S(t) = \Delta E_S(t) - \tilde{Q}_S(t) - A_S(t). \tag{33}$$

IV. BDME REPRESENTATION OF THE QME SOLUTION

In this section we show that if we assume a closed evolution equation for the subsystem reduced density matrix, we can recast its solution in a BDME form with time-dependent rates.

We assume that the reduced subsystem density matrix $\hat{\sigma}(t)$ obeys a closed QME [31–34]. This QME can be derived microscopically by perturbation theory like in the Redfield theory or using a quantum dynamical semigroup approach leading to Lindblad-type master equations. In Liouville space [36], the QME of the externally driven subsystem interacting with its environment reads

$$|\dot{\hat{\sigma}}(t)\rangle\rangle = \hat{\hat{\mathcal{K}}}(t)|\hat{\sigma}(t)\rangle\rangle. \tag{34}$$

If the interaction with the environment vanishes, the generator $\hat{\hat{\mathcal{K}}}(t)$ becomes the anti-Hermitian superoperator $\hat{\hat{\mathcal{K}}}(t) = \hat{\hat{\mathcal{L}}}_S(t) = -i[\hat{H}_S(t), \cdot]$ and the evolution superoperator $\hat{\hat{\mathcal{M}}}_t$ defined by $|\hat{\sigma}(t)\rangle\rangle = \hat{\hat{\mathcal{M}}}_t|\hat{\sigma}(0)\rangle\rangle$ becomes the unitary superoperator $\hat{\hat{\mathcal{M}}}_t = \exp_+\{\int_0^t d\tau \hat{\hat{\mathcal{L}}}_S(\tau)\}$. However, for nonvanishing coupling this generator is not anti-Hermitian and leads to a nonunitary evolution.

The QME in some given (possibly time-dependent) basis reads

$$\langle\langle ii'|\dot{\hat{\sigma}}(t)\rangle\rangle = \sum_{jj'} \langle\langle ii'|\hat{\hat{\mathcal{K}}}(t)|jj'\rangle\rangle\langle\langle jj'|\hat{\sigma}(t)\rangle\rangle, \tag{35}$$

where $\langle\langle jj'|\hat{\sigma}(t)\rangle\rangle$ is the superoperator representation of $\langle j|\hat{\sigma}(t)|j'\rangle$. Let us now use the time-dependent basis $\{|m_t\rangle\}$ introduced in Eq. (23). Since the QME keeps $\hat{\sigma}(t)$ Hermitian, this diagonalization is always possible:

$$\langle\langle m_t m_t'|\hat{\sigma}(t)\rangle\rangle = P_t(m)\delta_{m,m'}. \tag{36}$$

A crucial property of this basis is that [see Eq. (26)]

$$\dot{P}_t(m) = \langle\langle m_t m_t|\dot{\hat{\sigma}}(t)\rangle\rangle. \tag{37}$$

The consequence of this property is that by defining

$$W_t(m',m) \equiv \langle\langle m_t m_t|\hat{\hat{\mathcal{K}}}(t)|m_t' m_t'\rangle\rangle \tag{38}$$

and by projecting the QME (34) onto the time-dependent superbra $\langle\langle m(t)m(t)|$ we get

$$\dot{P}_t(m) = \sum_{m'} W_t(m',m)P_t(m'). \tag{39}$$

Since the QME (34) preserves probability, we have $\Sigma_m W_t(m',m)=0$ and $W_t(m',m)$ real. Therefore, we can rewrite Eq. (39) as

$$\dot{P}_t(m) = \sum_{m'\neq m} \{W_t(m',m)P_t(m') - W_t(m,m')P_t(m)\}. \tag{40}$$

Even though this equation appears like a BDME, it should not be viewed as an equation of motion. It is merely a way of recasting the solution of the QME (34) in a diagonal basis. In fact, in order to get the $P_t(m)$'s and $W_t(m',m)$'s, we need to solve the QME first and find the time-dependent unitary transformation diagonalizing the solution $\hat{\sigma}(t)$ at any time. Equation (39) should therefore be viewed as a formal definition of the rate matrix $W_t(m',m)$. We will show that $W_t(m',m)$ defined in this way can be used to derive quantum fluctuation relations. Note that $W_t(m',m)$ depends on the subsystem initial condition $\hat{\sigma}(0)$.

If the subsystem (driven or not) does not interact with the environment, the generator is anti-Hermitian and the evolution superoperator unitary. In this case $|\hat{\sigma}(t)\rangle\rangle = \hat{\hat{\mathcal{M}}}_t|\hat{\sigma}(0)\rangle\rangle$ and $\langle\langle m_t m_t| = \langle\langle m_0 m_0|\hat{\hat{\mathcal{M}}}_t^{-1}$, so that

$$P_t(m) = \langle m_0|\hat{\sigma}(0)|m_0\rangle = P_0(m). \tag{41}$$

This shows that the $P_t(m)$'s evolve only if the dynamics is nonunitary.

When there is no driving and the subsystem does interact with its environment, the dynamics is nonunitary and the subsystem will reach equilibrium $\hat{\sigma}^{\mathrm{eq}}$ on long time scales. For an infinite isothermal environment this equilibrium state will correspond to the canonical subsystem reduced density matrix $\hat{\sigma}^{\mathrm{eq}} = e^{-\beta \hat{H}_S}/Z_S$ where $Z_S = \mathrm{Tr}\, e^{-\beta \hat{H}_S}$ and $\beta = 1/T$ ($k_B \equiv 1$). In this case the basis diagonalizing $\hat{\sigma}^{\mathrm{eq}}$ becomes time independent and will also diagonalize the subsystem Hamiltonian so that $P^{\mathrm{eq}}(m) = e^{-\beta E_m}/Z_S$ where E_m are the eigenvalues of the subsystem Hamiltonian.

For a subsystem with nonequilibrium boundary conditions and interacting with its environment, the subsystem can reach a steady state $\hat{\sigma}^{\mathrm{st}}$ at long times. In this case the matrix diagonalizing the density matrix is again time independent and both the probabilities $P_t(m) = P^{\mathrm{st}}(m)$ and rates $W_t(m',m) = W^{\mathrm{st}}(m',m)$ become time independent.

V. ENTROPY FOR QUANTUM ENSEMBLES

In this section we define the von Neumann entropy associated with the subsystem and separate its evolution into two parts: the entropy flow associated with the heat going from the subsystem to the environment and the always positive entropy production associated with the internal entropy growth of the subsystem.

The von Neumann entropy of the subsystem is defined by

$$S(t) \equiv -\,\mathrm{Tr}\hat{\sigma}(t)\ln\hat{\sigma}(t) = -\sum_m P_t(m)\ln P_t(m). \tag{42}$$

Using Eq. (40), we can write its time derivative as

$$\dot{S}(t) = -\sum_m \dot{P}_t(m)\ln P_t(m) \tag{43}$$

$$= -\sum_{m,m'} P_t(m)W_t(m,m')\ln\frac{P_t(m')}{P_t(m)}. \tag{44}$$

In analogy with [13,15] for classical systems, this can be partitioned as

$$\dot{S}(t) = \dot{S}_e(t) + \dot{S}_i(t), \tag{45}$$

where

$$\dot{S}_e(t) \equiv -\sum_{m,m'} P_t(m)W_t(m,m')\ln\frac{W_t(m,m')}{W_t(m',m)} \tag{46}$$

and

$$\dot{S}_i(t) \equiv \sum_{m,m'} P_t(m) W_t(m,m') \ln \frac{P_t(m) W_t(m,m')}{P_t(m') W_t(m',m)}. \quad (47)$$

As a consequence of the inequality $(R_1 - R_2)\ln(R_1/R_2) \geqslant 0$, we notice that $\dot{S}_i(t) \geqslant 0$ is always a positive quantity. We will therefore identify it with the entropy production. The remaining part of the entropy $\dot{S}_e(t)$ is thus associated with the entropy flow to the environment since in thermodynamics the entropy evolution is partitioned in the (reversible) entropy flow to the environment and the (irreversible) entropy production [37,38]. To further rationalize this identification, let us assume that $W_t(m,m')$ satisfy the detailed balance condition [17,38,39]. For isothermal environments at temperature T, the detailed balance condition with respect to $\hat{H}_s(t)$, which means that the nondriven subsystem tends to thermal equilibrium at long time, reads

$$\frac{W_t(m,m')}{W_t(m',m)} = e^{\beta(\langle m_t|\hat{H}_S(t)|m_t\rangle - \langle m_t'|\hat{H}_S(t)|m_t'\rangle)}. \quad (48)$$

Noticing that the heat (29) can be rewritten as

$$\begin{aligned} \dot{\tilde{Q}}(t) &= \sum_m \dot{P}_t(m) \langle m_t|\hat{H}_S(t)|m_t\rangle \\ &= -\sum_{m,m'} P_t(m) W_t(m,m') [\langle m_t|\hat{H}_S(t)|m_t\rangle - \langle m_t'|\hat{H}_S(t)|m_t'\rangle] \\ &= -T \sum_{m,m'} P_t(m) W_t(m,m') \ln \frac{e^{\beta\langle m_t|\hat{H}_S(t)|m_t\rangle}}{e^{\beta\langle m_t'|\hat{H}_S(t)|m_t'\rangle}} \end{aligned} \quad (49)$$

and using Eq. (46), the immediate consequence of Eq. (48) is that the entropy flow is equal to the modified heat going from the subsystem to the environment divided by the environment temperature as expected from thermodynamics:

$$\dot{S}_e(t) = \frac{\dot{\tilde{Q}}_S}{T}. \quad (50)$$

This motivates our partition of the entropy (45) and the definition of the modified heat in Sec. III.

We can further show that the entropy flow is associated with reversible entropy variations. In a thermodynamic sense, a reversible transformation is a one during which the entropy production is zero, $\dot{S}_i(t) = 0$. This property holds provided the following condition is satisfied [see Eq. (47)]:

$$P_t(m) W_t(m,m') = P_t(m') W_t(m',m). \quad (51)$$

Using Eq. (48), we find that for a reversible transformation the subsystem has to be in the time-dependent state

$$P_t(m) = \frac{e^{-\beta\langle m_t|\hat{H}_S(t)|m_t\rangle}}{\sum_m e^{-\beta\langle m_t|\hat{H}_S(t)|m_t\rangle}}. \quad (52)$$

This state correspond to the instantaneous Gibbs state of the subsystem $\hat{\sigma}(t) = e^{-\beta\hat{H}_S(t)}/Z_S$. In this case $\{|m\rangle_t\}$ in Eq. (51) becomes the adiabatic basis (basis diagonalizing the subsystem Hamiltonian). We thus show that for reversible transformations the probability distribution remains Gibbsian along the adiabatic levels. Because $\dot{S}_i(t) = 0$, we also have $\dot{S}(t) = \dot{S}_e(t)$. Using Eq. (48), this means that for a reversible transformation the change in the entropy of the subsystem results exclusively from the heat flow to the environment, $\dot{S}(t) = \dot{\tilde{Q}}_S/T$, consistent with thermodynamics.

When there is no driving, Eq. (51) with Eq. (48) defines equilibrium. At equilibrium we have $\dot{S}(t) = \dot{S}_i(t) = \dot{S}_e(t) = 0$.

VI. ENTROPY FOR QUANTUM TRAJECTORIES

In this section we introduce quantum trajectories and distributions. We will associate an entropy with these trajectories and identify the entropy flow and production of these trajectories, whose ensemble averages recover the entropies discussed in Sec. V. This will allow us to derive a fundamental quantum relation similar to the classical relation obtained by Crooks [17] and Seifert [19] connecting the ratio of the probability of a forward trajectory and the backward one with the trajectory entropy production.

From Eq. (40) it seems natural to unravel the evolution equation for the probability $P_t(m)$ in the same way as is done for classical stochastic processes [19]. Let us consider a stochastic trajectory of duration t which contains N jumps. Different trajectories can of course have a different number of jumps N. $\tau = [0,t]$ labels time during the process. $j = 1, \ldots, N$ labels the jumps. The trajectory $n_{(\tau)}$ (see Fig. 1) is made by the successive states taken by the system in time:

$$n_{(\tau)} = n_0 \to n_1 \to n_2 \to \cdots \to n_N. \quad (53)$$

The system starts in n_0, jumps at time τ_j from n_{j-1} to n_j, and ends up at time t in n_N. We will denote $\tau_0 = 0$ and $\tau_{N+1} = t$.

The entropy associated with the trajectory $n_{(\tau)}$ reads

$$s(\tau) \equiv -\ln P_\tau(n_{(\tau)}), \quad (54)$$

where $P_\tau(n_{(\tau)})$ is the solution of Eq. (40) for an initial condition $P_0(n_0)$, evaluated along the trajectory $n_{(\tau)}$.

The time derivative of this trajectory entropy,

$$\dot{s}(\tau) = -\left.\frac{\partial_\tau P_\tau(n)}{P_\tau(n)}\right|_{n_{(\tau)}} - \sum_{j=1}^{N} \delta(\tau - \tau_j) \ln \frac{P_\tau(n_j)}{P_\tau(n_{j-1})}, \quad (55)$$

will be partitioned as

$$\dot{s}(\tau) = \dot{s}_e(\tau) + \dot{s}_i(\tau), \quad (56)$$

where the trajectory entropy flux reads

$$\dot{s}_e(\tau) \equiv -\sum_{j=1}^{N} \delta(\tau - \tau_j) \ln \frac{W_\tau(n_{j-1}, n_j)}{W_\tau(n_j, n_{j-1})} \quad (57)$$

and the trajectory entropy production reads

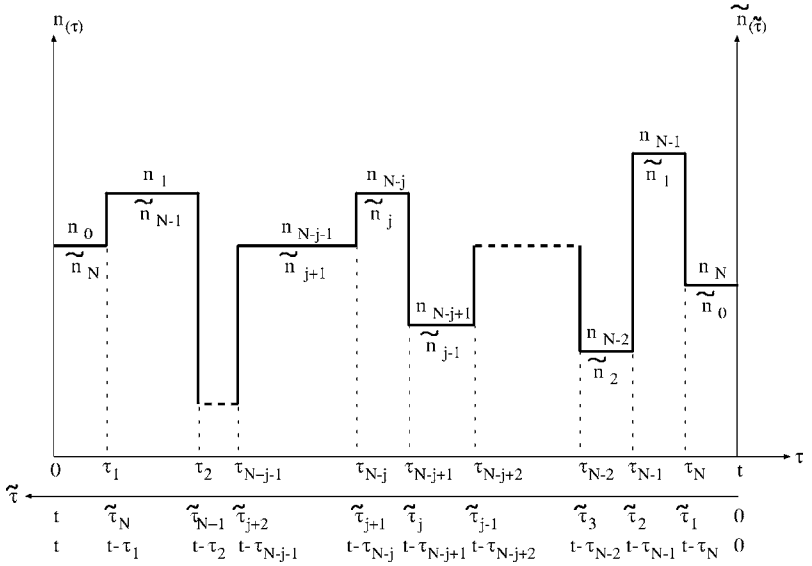

FIG. 1. Representation of a quantum forward trajectory $n_{(\tau)}$ and of the associated backward trajectory $\tilde{n}_{(\tilde{\tau})}$.

$$\dot{s}_i(\tau) \equiv -\left.\frac{\partial_t P_\tau(n)}{P_\tau(n)}\right|_{n_{(\tau)}} - \sum_{j=1}^{N} \delta(\tau-\tau_j)\ln\frac{P_\tau(n_j)W_\tau(n_j,n_{j-1})}{P_\tau(n_{j-1})W_\tau(n_{j-1},n_j)}. \quad (58)$$

The ensemble average over trajectories is carried out by using the probability $P_\tau(n_{j-1})W_\tau(n_{j-1},n_j)$ that a transition occurs at time τ_j between n_{j-1} and n_j. We get

$$\dot{S}(\tau) = \langle \dot{s}(\tau)\rangle, \quad (59)$$

$$\dot{S}_e(\tau) = \langle \dot{s}_e(\tau)\rangle, \quad (60)$$

$$\dot{S}_i(\tau) = \langle \dot{s}_i(\tau)\rangle. \quad (61)$$

The probability of a forward trajectory $n_{(\tau)}$ starting at time 0 and ending at time t is given by

$$\mu_F[n_{(\tau)}] = P_0(n_0)\left[\prod_{j=1}^{N}\exp\left(-\int_{\tau_{j-1}}^{\tau_j} d\tau' \sum_m W_{\tau'}(n_{j-1},m)\right) \times W_{\tau_j}(n_{j-1},n_j)\right]\exp\left(-\int_{\tau_N}^{t} d\tau' \sum_m W_{\tau'}(n_N,m)\right). \quad (62)$$

The exponentials represent the probabilities to stay in a given state during the time interval between two successive jumps, and the transition rates evaluated at the jump times give the probability for the jumps to occur at these times.

Defining the backward process as done for classical stochastic dynamics [19] is not possible. In the classical case, it is sufficient after the forward process to revert the driving protocol $\tilde{\lambda}(\tau)=\lambda(t-\tau)$ and to ask for the probability of a backward trajectory (system taking the sequence of states of the forward trajectory but in the reversed order) to occur. The reversal of the driving protocol has the consequence of reversing the time dependence of the transition matrix $\tilde{W}_\tau(m,m')=W_{t-\tau}(m,m')$. The time dependence of the quantum transition matrix $W_t(m,m')$ does not come exclusively from the external driving force and is different for different initial conditions of the subsystem $\hat{\sigma}(0)$. Therefore, reversing the driving protocol does not simply reverse the time dependence of the transition matrix. In the quantum case, we define the backward process by taking $\tau\to t-\tau$ everywhere in the QME (the protocol is reversed and a minus sign appears in the left hand side of Eq. (34) due to the time derivative). This has the effect of reversing the time dependence of the quantum transition matrix $W_t(m,m')$. This definition will allow us to derive important fluctuation relations in next sections.

Let us consider a new dynamics in the time interval $\tilde{\tau}=[0,t]$ obeying

$$\dot{\tilde{P}}_{\tilde{\tau}}(m) = \sum_{\tilde{m}'} \tilde{W}_{\tilde{\tau}}(\tilde{m}',\tilde{m})\tilde{P}_{\tilde{\tau}}(\tilde{m}'), \quad (63)$$

where the rates are related to the previous rates in the following way:

$$\tilde{W}_{\tilde{\tau}}(\tilde{m}',\tilde{m}) = W_{t-\tilde{\tau}}(\tilde{m}',\tilde{m}). \quad (64)$$

The backward dynamics may have an arbitrary initial condition $\tilde{P}_0(\tilde{n}_0)$. We now define the following trajectory for this backward dynamics:

$$\tilde{n}_{(\tilde{\tau})} = \tilde{n}_0 \to \tilde{n}_1 \to \tilde{n}_2 \to \cdots \to \tilde{n}_N = n_N \to n_{N-1} \to n_{N-2} \to \cdots \to n_0, \quad (65)$$

where the jumps between $\tilde{n}_{j-1}$ and $\tilde{n}_j$ occur at time $\tilde{\tau}_j=t-\tau_{N-j+1}$ and where $\tilde{n}_j=n_{N-j}$. Because this dynamics and the dynamics (40) both span the same configuration space, summing over all trajectories of the backward process is equivalent to summing over all the trajectories of the original process. The trajectory (65) is depicted in Fig. 1. The probability of this trajectory is evaluated in Appendix B and reads

$$\mu_B[\tilde{n}_{(\tilde{\tau})}] = \tilde{P}_0(n_N)\left[\prod_{j=1}^{N} \exp\left(-\int_{\tau_{j-1}}^{\tau_j} d\tau' \sum_m W_{\tau'}(n_{j-1},m)\right) \times W_{\tau_j}(n_j,n_{j-1})\right] \exp\left(-\int_{\tau_N}^{t} d\tau' \sum_m W_{\tau'}(n_N,m)\right). \tag{66}$$

We now calculate the ratio of the forward (62) and backward (66) probabilities. Noticing that the exponentials cancel, we find the fundamental result of the paper:

$$r(t) = \ln \frac{\mu_F[n_{(\tau)}]}{\mu_B[\tilde{n}_{(\tilde{\tau})}]} = \ln \frac{P_0(n_0)}{\tilde{P}_0(n_N)} - \Delta s_e(t), \tag{67}$$

where the trajectory entropy flow is

$$\Delta s_e(t) \equiv s_e(t) - s_e(0) = \int_0^t d\tau \dot{s}_e(\tau) = \sum_{j=1}^{N} \ln \frac{W_{\tau_j}(n_j,n_{j-1})}{W_{\tau_j}(n_{j-1},n_j)}. \tag{68}$$

In analogy with the classical results of Seifert [19], we can now derive the various fluctuation theorems by specific choices of initial conditions for the backward trajectories. By choosing $\tilde{P}_0(n_N) = P_t(n_N)$ and using the trajectory entropy

$$\Delta s(t) \equiv s(t) - s(0) = \ln \frac{P_0(n_0)}{P_t(n_N)}, \tag{69}$$

Eq. (67) becomes

$$r(t) = \ln \frac{\mu_F[n_{(\tau)}]}{\mu_B[\tilde{n}_{(\tilde{\tau})}]} = \Delta s(t) - \Delta s_e(t) = \Delta s_i(t), \tag{70}$$

where $\Delta s_i(t) = s_i(t) - s_i(0)$ is the trajectory entropy production.

Equation (67) was first derived by Crooks [17] for classical stochastic processes and later generalized by others [5,19]. We have shown that this relation may be extended to quantum systems. In the classical case the time dependence of the rates is exclusively due to the external driving. In the quantum case it is also due to the quantum evolution of the density matrix itself. However, in both cases, the backward dynamics is obtained by reversing the time dependence of the rate matrix.

VII. QUANTUM INTEGRAL FLUCTUATION THEOREM

Summing over all possible trajectories of the backward process is equivalent to summing over all possible trajectories of the original process $\Sigma_{\tilde{n}_{(\tilde{\tau})}} = \Sigma_{n_{(\tau)}}$. By averaging Eq. (67) over all possible trajectories, we find

$$1 = \sum_{\tilde{n}_{(\tilde{\tau})}} \mu_B[\tilde{n}_{(\tilde{\tau})}] = \sum_{n_{(\tau)}} \mu_B[\tilde{n}_{(\tilde{\tau})}] = \sum_{n_{(\tau)}} \mu_F[n_{(\tau)}] e^{-r(t)} = \langle e^{-r(t)} \rangle. \tag{71}$$

This integral fluctuation theorem [19] is valid for any choice of $P_0(n_0)$ and $\tilde{P}_0(n_N)$ in Eq. (67). Using the fact that $\langle e^x \rangle \geq e^{\langle x \rangle}$ this relation also means that on average the quantity $r(t)$ is always non-negative, $\langle r(t) \rangle \geq 0$. Choosing $\tilde{P}_0(n_N) = P_t(n_N)$ we have $r(t) = \Delta s_i(t)$ and we show again [see text below Eq. (47)] that the ensemble averaged trajectory entropy production is always non-negative, $\langle \Delta s_i(t) \rangle \geq 0$.

VIII. QUANTUM FLUCTUATION THEOREM FOR A STEADY STATE

We consider a subsystem subjected to nonequilibrium constraints, whose dynamics is described by a QME of the form (34). When the subsystem is in a steady state, its density matrix does not evolve in time and the rates in Eq. (40) are time independent. An example of such system could be a two-level atom driven by a coherent single-mode field on resonance (in the dipole approximation and in the rotating wave approximation) described by Bloch equations (see p. 154 of Ref. [34]). The basis set of the forward process is time independent and is the same as the basis set of the backward process. By definition, we have

$$p_F(R(t)) = \langle \delta(R(t) - r_F(t)) \rangle_F = \sum_{n_{(\tau)}} \mu_F[n_{(\tau)}] \delta(R(t) - r_F(t)). \tag{72}$$

Using Eq. (67), we can write

$$\begin{aligned} p_F(R(t)) &= \sum_{n_{(\tau)}} \mu_B[n_{(\tau)}] e^{r_F(t)} \delta(R(t) - r_F(t)) \\ &= \sum_{n_{(\tau)}} \mu_B[n_{(\tau)}] e^{R(t)} \delta(R(t) - r_F(t)) \\ &= \langle \delta(R(t) + r_B(t)) \rangle_B e^{R(t)} \\ &= p_B(-R(t)) e^{R(t)}, \end{aligned} \tag{73}$$

where to go from the second to the third line, we used $r_F(t) = -r_B(t)$ which comes from Eq. (67). When $\tilde{P}_0(n_N) = P_t(n_N)$ and therefore Eq. (70) holds, Eq. (73) becomes a fluctuation theorem for the entropy production:

$$p_F(\Delta S_i(t)) = p_B(-\Delta S_i(t)) e^{\Delta S_i(t)}. \tag{74}$$

This relation shows that at steady state, the ratio of the probability to observe a given entropy production during a froward process and the probability to observe the same entropy production with a minus sign during the backward process is given by the exponential of the entropy production. This is the most familiar form of the fluctuation theorem. In the infinite-time limit, if the subsystem has a finite number of levels (this condition is usually implicitly assumed in QME theory), $\Delta S(t)$ will be bounded and $\Delta S_i(t) \underset{t\to\infty}{=} \Delta S_e(t)$ so that Eq. (74) also becomes a fluctuation theorem for the entropy flow and therefore also for the heat. For completeness, we give in Appendix C a different derivation of a fluctuation theorem similar to Eq. (74) and which is not restricted to steady states.

IX. HEAT AND WORK FOR QUANTUM TRAJECTORIES

If we use the relation (48) together with the definition of the trajectory entropy flow (68), we find that the heat associated with a single trajectory is given by

$$\begin{aligned}\tilde{q}_S(t) &\equiv \beta^{-1}\Delta s_e(t)\\ &= \sum_{j=1}^{N}[\langle n_j|\hat{H}_S(\tau_j)|n_j\rangle - \langle n_{j-1}|\hat{H}_S(\tau_j)|n_{j-1}\rangle].\end{aligned} \tag{75}$$

The interpretation of this result is that the heat flowing to the environment results from transitions between the subsystem states n_j.

The energy associated with a trajectory is a state function and only depends on the initial and final states of the trajectory:

$$\begin{aligned}\Delta e_S(t) &= \langle n_N|\hat{H}_S(t)|n_N\rangle - \langle n_0|\hat{H}_S(0)|n_0\rangle\\ &= \sum_{j=1}^{N}[\langle n_j|\hat{H}_S(\tau_j)|n_j\rangle - \langle n_{j-1}|\hat{H}_S(\tau_j)|n_{j-1}\rangle]\\ &= \tilde{w}_S(t) + \tilde{q}_S(t).\end{aligned} \tag{76}$$

The work is therefore given by

$$\begin{aligned}\tilde{w}_S(t) &= \Delta e_S(t) - \tilde{q}_S(t)\\ &= \sum_{j=1}^{N}[\langle n_{j-1}|\hat{H}_S(\tau_j)|n_{j-1}\rangle - \langle n_{j-1}|\hat{H}_S(\tau_{j-1})|n_{j-1}\rangle].\end{aligned} \tag{77}$$

The work thus results from the time evolution of the Hamiltonian (due to the driving force) along the states n_j of the subsystem between the transitions. It is interesting to point out the parallel between our description of heat and work in the $\{|m_t\rangle\}$ basis set and the adiabatic basis description of Ref. [22]. In the latter the work comes from the evolution along the adiabatic states and the heat comes from the transitions between the adiabatic state. This can be understood by comparing Eqs. (A9) and (A10) with Eqs. (A12) and (A13).

X. QUANTUM JARZYNSKI RELATION

We assume that the subsystem is initially at equilibrium with respect to the Hamiltonian $\hat{H}_S(0)=\hat{H}_S(\lambda(0))$ and is therefore described by a canonical distribution. The system is then driven out of equilibrium by turning the driving force from $\lambda(0)$ to $\lambda(t')$ at time t'. After t' the driving force stop evolving. On long time scales after t'—say, t $(t\gg t')$—the system is again at equilibrium in a canonical distribution but now with respect to $\hat{H}_S(t)=\hat{H}_S(\lambda(t))$. We choose

$$P_0(n_0) = \frac{e^{-\beta\langle n_0|\hat{H}_S(0)|n_0\rangle}}{Z_0},$$

$$P_t(n_N) = \frac{e^{-\beta\langle n_N|\hat{H}_S(t)|n_N\rangle}}{Z_t}, \tag{78}$$

where $Z_0=\Sigma_n\exp[-\beta\langle n_0|\hat{H}_S(0)|n_0\rangle]$ and $Z_t=\Sigma_n \times\exp[-\beta\langle n_N|\hat{H}_S(t)|n_N\rangle]$. Notice that $\{|n_0\rangle\}$ $[\{|n_t\rangle\}]$ is now the eigenbasis of $\hat{H}_S(0)$ $[\hat{H}_S(t)]$.

The free energy difference between the initial and final states is given by

$$\Delta F(t) = F(t) - F(0) = -\beta^{-1}\ln\frac{Z_t}{Z_0}. \tag{79}$$

Using Eq. (75) which defines the heat of a single subsystem trajectory, we can write Eq. (70) as

$$\Delta s_i(t) = -\ln P_t(n_N) + \ln P_0(n_0) - \beta\tilde{q}_S(t). \tag{80}$$

Using now Eqs. (78), (79), (76), and (77), we can rewrite Eq. (80) as

$$\Delta s_i(t) = -\beta\Delta F(t) + \beta\tilde{w}_S(t). \tag{81}$$

Finally, by inserting Eq. (81) into the integral fluctuation theorem (71) where $r(t)=\Delta s_i(t)$, we find the quantum Jarzynski relation

$$e^{-\beta\Delta F(t)} = \langle e^{-\beta\tilde{w}_S(t)}\rangle. \tag{82}$$

XI. CONCLUSIONS

We have presented a unified derivation of a quantum integral fluctuation theorem, a quantum steady-state fluctuation theorem, and the quantum Jarzynski relation for a driven subsystem interacting with its environment and described by a QME. This generalizes earlier results obtained for quantum systems. By recasting the solution of the QME in a BDME form with time-dependent rates for the eigenvalues of the subsystem density matrix, we naturally define quantum trajectories and their associated entropy, heat, and work and study their fluctuation properties. The connection between the trajectory quantities which naturally enter our formulation and measurable quantum trajectory quantities is still an open issue. Deriving quantum fluctuation relations without having to assume QME's, which do not correctly account for strong subsystem-environment entanglement, is an exiting perspective.

ACKNOWLEDGMENTS

The support of the National Science Foundation (Grant No. CHE-0446555), NIRT (Grant No. EEC 0303389), and the National Institute of Health (Grant No. GM59230-05) is gratefully acknowledged. M.E. is also *collaborateur scientifique* at the FNRS Belgium.

APPENDIX A: BASIS DEPENDENCE OF HEAT AND WORK

We consider a system with a time-dependent Hamiltonian $\hat{H}(t)$ in the Schrödinger picture described by the density ma-

trix $\hat{\rho}(t)$. The evolution equation of $\hat{\rho}(t)$ is not necessarily unitary.

The energy of the system is given by

$$\langle \hat{H} \rangle \equiv \mathrm{Tr}\hat{H}(t)\hat{\rho}(t) = \sum_{aa'} \langle a_t|\hat{H}(t)|a_t'\rangle\langle a_t'|\hat{\rho}(t)|a_t\rangle, \quad \text{(A1)}$$

where $\{|a_t\rangle\}$ is an arbitrary time-dependent basis set. The energy change of the system can be written as

$$\Delta E(t) \equiv \int_0^t d\tau \frac{d\langle \hat{H}(\tau)\rangle}{d\tau} = W(t) + Q(t) = \tilde{W}(t) + \tilde{Q}(t), \quad \text{(A2)}$$

where the heat and work are given by

$$\dot{Q}(t) \equiv \mathrm{Tr}\hat{H}(t)\dot{\hat{\rho}}(t), \quad \text{(A3)}$$

$$\dot{W}(t) \equiv \mathrm{Tr}\dot{\hat{H}}(t)\hat{\rho}(t), \quad \text{(A4)}$$

in a time-independent basis, and by

$$\dot{\tilde{Q}}(t) \equiv \sum_{aa'} \langle a_t|\hat{H}(t)|a_t'\rangle \frac{d}{dt}[\langle a_t'|\hat{\rho}(t)|a_t\rangle], \quad \text{(A5)}$$

$$\dot{\tilde{W}}(t) \equiv \sum_{aa'} \frac{d}{dt}[\langle a_t|\hat{H}(t)|a_t'\rangle]\langle a_t'|\hat{\rho}(t)|a_t\rangle, \quad \text{(A6)}$$

in a time-dependent basis.

How does $Q(t)$ $[W(t)]$ relate to $\tilde{Q}(t)$ $[\tilde{W}(t)]$? We find

$$\dot{\tilde{Q}}(t) = \dot{Q}(t) - \dot{A}(t),$$

$$\dot{\tilde{W}}(t) = \dot{W}(t) + \dot{A}(t), \quad \text{(A7)}$$

where

$$\dot{A}(t) = -\sum_{aa'} \langle a_t|\hat{H}(t)|a_t'\rangle[\langle \dot{a}_t'|\hat{\rho}(t)|a_t\rangle + \langle a_t'|\hat{\rho}(t)|\dot{a}_t\rangle]$$
$$= \sum_{aa'} [\langle \dot{a}_t|\hat{H}(t)|a_t'\rangle + \langle a_t|\hat{H}(t)|\dot{a}_t'\rangle]\langle a_t'|\hat{\rho}(t)|a_t\rangle. \quad \text{(A8)}$$

We have used the fact that $\langle a_t|\dot{a}_t'\rangle = -\langle \dot{a}_t|a_t'\rangle$ which come from $\frac{d}{dt}(\langle a_t|a_t'\rangle) = 0$.

If we consider the time-dependent basis set which diagonalizes the instantaneous density density matrix $\{|a_t\rangle\} = \{|m_t\rangle\}$, where $\langle m_t|\hat{\rho}(t)|m_t'\rangle = P_t(m)\delta_{mm'}$, we have

$$\dot{\tilde{Q}}(t) = \sum_m \langle m_t|\hat{H}(t)|m_t\rangle \dot{P}_t(m) = \dot{Q}(t) - \dot{A}(t), \quad \text{(A9)}$$

$$\dot{\tilde{W}}(t) = \sum_m \frac{d}{dt}[\langle m_t|\hat{H}(t)|m_t\rangle]P_t(m) = \dot{W}(t) + \dot{A}(t), \quad \text{(A10)}$$

where

$$\dot{A}(t) = \sum_{mm'} \langle m_t|\hat{H}(t)|m_t'\rangle\langle m_t'|\dot{m}_t\rangle[P_t(m) - P_t(m')]$$
$$= \sum_m [\langle \dot{m}_t|\hat{H}(t)|m_t\rangle + \langle m_t|\hat{H}(t)|\dot{m}_t\rangle]P_t(m). \quad \text{(A11)}$$

If we consider the time-dependent basis diagonalizing the instantaneous Hamiltonian (adiabatic basis) $\{|a_t\rangle\} = \{|i_t\rangle\}$, where $\langle i_t|\hat{H}(t)|i_t'\rangle = \epsilon_i(t)\delta_{ii'}$, we have

$$\dot{\tilde{Q}}'(t) = \sum_i \epsilon_i(t)\frac{d}{dt}[\langle i_t|\hat{\rho}(t)|i_t\rangle] = \dot{Q}(t) - \dot{A}'(t), \quad \text{(A12)}$$

$$\dot{\tilde{W}}'(t) = \sum_i \dot{\epsilon}_i(t)\langle i_t|\hat{\rho}(t)|i_t\rangle = \dot{W}(t) + \dot{A}'(t), \quad \text{(A13)}$$

where

$$\dot{A}'(t) = -\sum_i \epsilon_i(t)[\langle \dot{i}_t|\hat{\rho}(t)|i_t\rangle + \langle i_t|\hat{\rho}(t)|\dot{i}_t\rangle]$$
$$= \sum_{ii'} [\epsilon_i(t) - \epsilon_{i'}(t)]\langle i_t|\dot{i}_t'\rangle\langle i_t'|\hat{\rho}(t)|i_t\rangle. \quad \text{(A14)}$$

It is interesting to notice the similarity between the two bases $\{|m_t\rangle\}$ and $\{|i_t\rangle\}$. In both cases, the heat results from changes in the population of the states (and therefore from transitions between states) [see Eqs. (A9) and (A12)] and the work from the evolution of the Hamiltonian along the states [see Eqs. (A10) and (A13)]. Using Eqs. (A11) with (A14) one gets

$$\dot{A}(t) - \dot{A}'(t) = \dot{\tilde{W}}(t) - \dot{\tilde{W}}'(t)$$
$$= \dot{\tilde{Q}}'(t) - \dot{\tilde{Q}}(t)$$
$$= \sum_{i,m} P_t(m)\epsilon_i(t)\frac{d}{dt}(|\langle i_t|m_t\rangle|^2). \quad \text{(A15)}$$

Let us assume now that the density matrix of the system obeys the von Neumann equation

$$\dot{\hat{\rho}}(t) = -i[\hat{H}(t), \hat{\rho}(t)] = \hat{\hat{\mathcal{L}}}(t)\hat{\rho}(t), \quad \text{(A16)}$$

whose solution reads

$$\hat{\rho}(t) = \hat{\hat{\mathcal{U}}}(t)\hat{\rho}(0) = \hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t), \quad \text{(A17)}$$

where

$$\hat{\hat{\mathcal{U}}}(t) = \exp_+\left\{\int_0^t d\tau \hat{\hat{\mathcal{L}}}(\tau)\right\},$$

$$\hat{U}(t) = \exp_+\left\{-i\int_0^t d\tau \hat{H}(\tau)\right\}. \quad \text{(A18)}$$

The evolution operator [superoperator] $\hat{U}_T(t)$ $[\hat{\hat{\mathcal{U}}}_T(t)]$ is unitary. In this case, the expressions in the basis $\{|m_t\rangle\}$ simplify to

$$\tilde{Q}(t) = Q(t) = 0, \quad \text{(A19)}$$

$$\widetilde{W}(t) = W(t) = \Delta E(t). \tag{A20}$$

This is due to the fact that

$$P_0(m) = \langle m_0|\hat{\rho}(0)|m_0\rangle = \langle m_0|\hat{U}^\dagger(t)\hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t)\hat{U}(t)|m_0\rangle = \langle m_t|\hat{\rho}(t)|m_t\rangle = P_t(m). \tag{A21}$$

This means that no heat is produced by the driving force for a unitary evolution. This is reasonable since there is no environment. The only way in which the energy of the system may increase is via the work done on the system. Notice that in the adiabatic basis both $\widetilde{W}'(t)$ and $\widetilde{Q}'(t)$ are finite for a unitary evolution.

APPENDIX B: PROBABILITY OF THE BACKWARD TRAJECTORY

The probability of a backward trajectory $\tilde{n}_{(\tilde{\tau})}$ reads

$$\mu_B[\tilde{n}_{(\tilde{\tau})}] = \widetilde{P}_0(\tilde{n}_0)\left[\prod_{j=1}^{N}\exp\left(-\int_{\tilde{\tau}_{j-1}}^{\tilde{\tau}_j} d\tilde{\tau}'\sum_{\tilde{m}}\widetilde{W}_{\tilde{\tau}'}(\tilde{n}_{j-1},\tilde{m})\right)\times\widetilde{W}_{\tilde{\tau}_j}(\tilde{n}_{j-1},\tilde{n}_j)\right]\exp\left(-\int_{\tilde{\tau}_N}^{t} d\tilde{\tau}'\sum_{\tilde{m}}\widetilde{W}_{\tilde{\tau}'}(\tilde{n}_N,\tilde{m})\right), \tag{B1}$$

where $\tilde{\tau}_0=0$ and $\tilde{\tau}_{N+1}=t$. Using Eqs. (64) and (65) and $\tilde{\tau}_j=t-\tau_{N-j+1}$, we can rewrite this probability as

$$\mu_B[\tilde{n}_{(\tilde{\tau})}] = \widetilde{P}_0(n_N)\left[\prod_{j=1}^{N}\exp\left(-\int_{t-\tau_{N-j+2}}^{t-\tau_{N-j+1}} d\tilde{\tau}'\sum_{m} W_{t-\tilde{\tau}'}(n_{N-j+1},m)\right)W_{\tau_{N-j+1}}(n_{N-j+1},n_{N-j})\right]\exp\left(-\int_{t-\tau_1}^{t} d\tilde{\tau}'\sum_{m} W_{t-\tilde{\tau}'}(n_0,m)\right). \tag{B2}$$

Using the change of variables $\tau=t-\tilde{\tau}$, we get

$$\mu_B[\tilde{n}_{(\tilde{\tau})}] = \widetilde{P}_0(n_N)\exp\left(-\int_0^{\tau_1} d\tau'\sum_{m} W_{\tau'}(n_0,m)\right)\left[\prod_{j=1}^{N}\exp\left(-\int_{\tau_{N-j+1}}^{\tau_{N-j+2}} d\tau'\sum_{m} W_{\tau'}(n_{N-j+1},m)\right)W_{\tau_{N-j+1}}(n_{N-j+1},n_{N-j})\right]. \tag{B3}$$

With the help of $j=N-j_{old}+2$, Eq. (B3) finally becomes Eq. (66).

APPENDIX C: QUANTUM FLUCTUATION THEOREM FOR UNCORRELATED SUBSYSTEM AND BATH

We derive a general quantum fluctuation theorem (not restricted to steady states) for a driven quantum subsystem in contact with its environment. The derivation is similar to that of Monnai in [27] and is given for completeness.

We assume weak coupling between the subsystem and environment and that the environment is infinitely large so that at all times the density matrix of the total system (subsystem plus environment) can be written as

$$\hat{\rho}(t) = \hat{\sigma}(t)\hat{\rho}_B^{eq}, \tag{C1}$$

where $\hat{\rho}_B^{eq}=e^{-\beta\hat{H}_B}/Z_B$ is the time-independent equilibrium reduced density matrix of the environment and $\hat{\sigma}(t)$ the time-dependent reduced density matrix of the subsystem. Assuming the form (C1) is not very different from assuming that the subsystem density matrix obeys a QME since the QME derivation implicitly assumes an invariant environment density matrix (e.g., the Born approximation [34]).

Let us define the basis $\{|m_t b\rangle\}$, where $\{|m_t\rangle\}$ diagonalize the subsystem density matrix at time t and where $\{|b\rangle\}$ diagonalize the time-independent environment Hamiltonian. The probability to go from $|m_0 b\rangle$ at time 0 to $|m_t b'\rangle$ at time t is given by

$$\mu_F[|m_0 b\rangle \to |m_t b'\rangle] = \langle m_0 b|\hat{\sigma}(0)\hat{\rho}_B^{eq}|m_0 b\rangle \times|\langle m_t b'|\hat{U}(t)|m_0 b\rangle|^2, \tag{C2}$$

where $\hat{U}(t)$ is the unitary evolution operator of the total system. The probability of the backward process to go from $|m_t b'\rangle$ at time t to $|m_0 b\rangle$ at time 0 by the time reversed evolution [27,30,40] is given by

$$\mu_B[|m_t b'\rangle \to |m_0 b\rangle] = \langle m_t b'|\hat{\sigma}(t)\hat{\rho}_B^{eq}|m_t b'\rangle \times|\langle m_t b'|\hat{U}(t)|m_0 b\rangle|^2. \tag{C3}$$

We therefore have that

$$\frac{\mu_F[|m_0 b\rangle \to |m_t b'\rangle]}{\mu_B[|m_t b'\rangle \to |m_0 b\rangle]} = \frac{P_0(m_0)}{P_t(m_t)}e^{-\beta Q_{bb'}}, \tag{C4}$$

where $P_0(m_0)=\langle m_0|\hat{\sigma}(0)|m_0\rangle$, $P_t(m_t)=\langle m_t|\hat{\sigma}(t)|m_t\rangle$, and $Q_{bb'}=E_b-E_{b'}$.

The entropy of a state m_t is defined as

$$s(t) = -\ln P_t(m_t). \tag{C5}$$

This definition makes sense because $\{|m_t\rangle\}$ diagonalizes $\hat{\sigma}(t)$, so that by averaging over the different states we recover the von Neumann entropy. The entropy difference between the initial and final states of the subsystem starting at time 0 in $|m_0\rangle$ and ending at time t in $|m_t\rangle$ is given by

$$\Delta s(s_0,s_t;t) = \ln \frac{P_0(m_0)}{P_t(m_t)}. \tag{C6}$$

The entropy production of this same process is given by

$$\Delta s_i(m_0,m_t,b,b';t) = \ln \frac{P_0(m_0)}{P_t(m_t)} - \frac{Q_{bb'}}{T}, \tag{C7}$$

because one assumes that the entropy flow difference is given by

$$\Delta s_e(b,b';t) = \frac{Q_{bb'}}{T}. \tag{C8}$$

Using Eqs. (C6)–(C8), Eq. (C4) becomes

$$\ln \frac{\mu_F[|m_0 b\rangle \to |m_t b'\rangle]}{\mu_B[|m_t b'\rangle \to |m_0 b\rangle]} = \Delta s_i(m_0,m_t,b,b';t) = \Delta s(s_0,s_t;t) - \Delta s_e(b,b';t). \tag{C9}$$

This result is the analog of our fundamental relation (70). By averaging the probabilities over all possible initial and final states, we get the general fluctuation theorem

$$\begin{aligned} p(\Delta S_i(t)) &= \sum_{m_0,m_t,b,b'} \mu_F[|m_0 b\rangle \to |m_t b'\rangle] \\ &\quad \times \delta(\Delta S_i(t) - \Delta s_i(m_0,m_t,b,b';t)) \\ &= \sum_{m_0,m_t,b,b'} \mu_B[|m_0 b\rangle \to |m_t b'\rangle] e^{\Delta s_i(m_0,m_t,b,b';t)} \\ &\quad \times \delta(\Delta S_i(t) - \Delta s_i(m_0,m_t,b,b';t)) \\ &= p(-\Delta S_i(t)) e^{\Delta S_i(t)}. \end{aligned} \tag{C10}$$

This result agrees with Eq. (74) and is not restricted to steady states. This approach is based on the time reversal invariance of the evolution of the total system and does not provide a trajectory picture. We note that in the total system space, the heat (or the entropy flow) going from the subsystem to the environment only depends on the end points and not on the path itself. If one derive a Jarzynski relation from this result as in [27], the work is found to be path independent. When considering the reduced dynamics of the system alone, as done in this paper, these quantities become path dependent and a trajectory picture is provided.

[1] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); Phys. Rev. E **56**, 5018 (1997).
[2] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
[3] C. Jarzynski, J. Stat. Mech.: Theory Exp. 2004, P09005.
[4] E. G. D. Cohen and D. Mauzerall, J. Stat. Mech.: Theory Exp. 2004, P07006.
[5] C. Maes, Séminaire Poincaré **2**, 29 (2003).
[6] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[7] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995); J. Stat. Phys. **80**, 931 (1995).
[8] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994); **52**, 5839 (1995); **53**, 5808 (1996).
[9] G. N. Bochkov and Yu. E. Kuzovlev, Zh. Eksp. Teor. Fiz. **72**, 238 (1977) [Sov. Phys. JETP **45**, 125 (1977)]; **76**, 1071 (1979) [**49**, 543 (1979)]; Physica A **106**, 443 (1981); **106**, 480 (1981).
[10] R. L. Stratonovich, *Nonlinear Nonequilibrium Thermodynamics II* (Springer, Berlin, 1994).
[11] C. Jarsynski, J. Stat. Phys. **98**, 77 (2000).
[12] J. Kurchan, J. Phys. A **31**, 3719 (1998).
[13] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[14] D. J. Searles and D. J. Evans, Phys. Rev. E **60**, 159 (1999).
[15] P. Gaspard, J. Chem. Phys. **120**, 19 (2004); D. Andrieux and P. Gaspard, *ibid.* **121**, 6167 (2004).
[16] P. Gaspard, J. Stat. Phys. **117**, 599 (2004); e-print cond-mat/0603382.
[17] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[18] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[19] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[20] S. Yukawa, J. Phys. Soc. Jpn. **69**, 2367 (2000).
[21] H. Tasaki, e-print cond-mat/0009244.
[22] S. Mukamel, Phys. Rev. Lett. **90**, 170604 (2003).
[23] W. De Roeck and C. Maes, Phys. Rev. E **69**, 026115 (2004).
[24] J. Kurchan, e-print cond-mat/0007360.
[25] T. Monnai and S. Tasaki, e-print cond-mat/0308337.
[26] W. De Roeck and C. Maes, e-print cond-mat/0406004.
[27] T. Monnai, Phys. Rev. E **72**, 027102 (2005).
[28] C. Jarzynski and D. K. Wojcik, Phys. Rev. Lett. **92**, 230602 (2004).
[29] A. E. Allahverdyan and Th. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).
[30] I. Callens, W. De Roeck, T. Jacobs, C. Maes, and K. Netocny, Physica D **187**, 383 (2004).
[31] F. Haake, *Statistical Treatment of Open Systems*, Springer Tracts in Modern Physics No. 66 (Springer, New York, 1973).
[32] H. Spohn, Rev. Mod. Phys. **53**, 569 (1980).
[33] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer, Berlin, 2000).
[34] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).
[35] The trace invariance used in Eq. (7) could be a delicate matter for systems with a continuous spectrum where the trace may diverge. This can be always formally addressed by using a dense quasicontinuum. Physically, Eq. (7) implies energy conservation.
[36] S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
[37] D. Kondepudi and I. Prigogine, *Modern Thermodynamics* (Wiley, Chichester, 1998).
[38] S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (Dover, New York, 1984).
[39] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 2nd ed. (North-Holland, Amsterdam, 1997).
[40] T. Yacobs and C. Maes, Phys. Mag. **27**, 119 (2005).