Model studies on the quantum Jarzynski relation

Jens Teifel and Günter Mahler

Institut für Theoretische Physik 1, Universität Stuttgart, 70550 Stuttgart, Germany
(Received 12 June 2007; published 30 November 2007)

We study the quantum Jarzynski relation for driven quantum models embedded in various environments. We do so by generalizing a proof presented by Mukamel [Phys. Rev. Lett. **90**, 170604 (2003)] for closed quantum systems. In this way, we are able to prove that the Jarzynski relation also holds for a bipartite system with microcanonical coupling. Furthermore, we show that, under the assumption that the interaction energy remains constant during the whole process, the relation is valid even for canonical coupling. The same follows for open quantum systems at high initial temperatures up to third order of the inverse temperature. Our analytical study is complemented by a numerical investigation of a special model system.

I. INTRODUCTION

If a system is driven far away from thermal equilibrium by an external force, its behavior can no longer be described by linear response theory or other near-equilibrium approximations. Fluctuations may dominate the evolution of the system. Recently, some astonishingly general theorems, which make exact statements, have been found [1–5]. For classically described systems there exist several proofs for various kinds of Hamilton dynamics [2,5].

One faces a different situation for systems the dynamics of which are described by quantum mechanics. Mostly for very restricted situations or only for very special model systems have those theorems been shown to hold [6–12] so far. However, there exists a general proof of the Jarzynski relation, a nonequilibrium work theorem, for closed quantum systems [13]. Generalizations to open quantum systems have been discussed only for restricted model systems or for various constraints on the microscopic dynamics [11,14].

One major problem is that the proofs of the so-called group of nonequilibrium fluctuation theorems as well as the work fluctuation theorems relate properties of trajectories from forward and backward processes with each other. Trying to transfer these proofs to quantum systems brings up the problem of trajectories in the quantum case.

Another, possibly even more severe problem is the question of what exactly has to be understood by fluctuations in a quantum system. In quantum mechanics the state of a system is fully described by its density operator. Without any measurement, it is ambiguous to speak about fluctuations, because the probability of a measurement outcome of a certain quantity might vary, but not the untouched quantity itself. Fluctuations of one-time quantities are not ambiguous as far as their measurement is concerned. If, however, the relevant observable of two-time quantities, like work, for example, does not commute with itself at different times, then even the measurement of fluctuation becomes ambiguous. Every measurement changes the state of the system, thereby influencing its evolution. In the case of work fluctuation theorems, the measurement has naturally to be included since one needs to measure the amount of work performed on the system. This measurement will then yield different values for each repetition and is thus said to fluctuate. So in this sense we have no problem speaking of fluctuations.

This is why we try the following ansatz. We concentrate on a work fluctuation theorem. In order to circumvent the problem of the definition of a quantum trajectory, we use the quite general proof for closed quantum systems [13], and try to generalize it to open quantum systems without having to define any quantum trajectory (in the sense of stochastic unraveling, i.e., continuous measurement). This will be done by splitting up the exponential function of the partition sum into its system and environmental part, respectively. There is no need to specify the underlying dynamics any further. It is not necessary to measure the system during the process since we rely on measurements at the beginning and the end of the process only. The idea of a two-time measurement approach was first developed by Kurchan [6].

First, we start by a brief review of the Jarzynski relation in Sec. II. Then, in Sec. III we will discuss conditions under which we will prove that the Jarzynski relation is valid. This analytical study will then be supplemented by a discussion of our numerical results in Sec. IV.

II. JARZYNSKI RELATION

If we perform a process on a system initially in a canonical state, the system will, in general, be driven out of equilibrium. Therefore, the average work $\bar{W}$ we have to bring up in order to perform the process will exceed the free energy difference ΔF between initial and final states:

$$\bar{W} \geq \Delta F. \tag{1}$$

In 1997, Jarzynski came up with a remarkable relation that connects the nonequilibrium variable W with an equilibrium property of the system, ΔF [1]:

$$\overline{e^{-\beta W}} = e^{-\beta \Delta F} = \frac{Z(t)}{Z(0)}. \tag{2}$$

This relation holds no matter how far the system is driven out of equilibrium. It allows one to access the free energy difference between two states via nonequilibrium measurements. This is very useful for experiments which cannot be carried out in quasiequilibrium [15]: Note that the right-hand side does not depend on how we get from the initial to the final state.
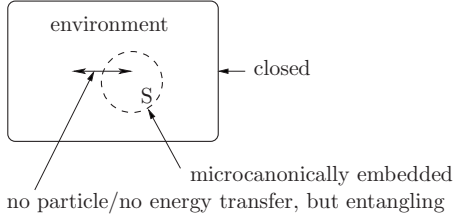
FIG. 1. Microcanonical coupling: between a small subsystem S and its environment. Total system is a closed compound system, initially in the canonical state $\hat{\varrho}(0) = e^{-\beta \hat{H}(0)}/Z(0)$.

The Jarzynski relation has been generalized to closed quantum systems by Mukamel [13] under the condition that the system is initially in a canonical state.

III. ANALYTICAL RESULTS

We consider a bipartite system, split into the so-called system and environment, respectively. An arbitrary process is performed on the system, leading to an explicitly time-dependent Hamilton operator. The compound system is considered closed and the complete Hamilton operator reads

$$\hat{H}(t) = \hat{H}^S(t) \otimes \hat{1}^C + \hat{1}^S \otimes \hat{H}^C + \sigma \hat{H}^{\text{int}}, \tag{3}$$

where $\hat{H}^S(t)$ denotes the Hamiltonian of the system, $\hat{H}^C$ the operator for the environment, and $\hat{H}^{\text{int}}$ the interaction Hamiltonian with the coupling constant σ, determining the strength of the coupling. In the following, we investigate different coupling scenarios.

A. Microcanonical coupling

We consider a microcanonical coupling [16] (see Fig. 1) first. This means that neither particle nor energy exchange between the system and the environment is allowed. Note that because of entanglement between the system and the environment the work performed on the system will, nevertheless, be influenced by the environment. This is due to the fact that the off-diagonal elements of the reduced density operator of the system will be damped away. Since these elements are crucial for transitions within the system, the work should, indeed, be influenced by the microcanonical coupling.

Since no energy transfer between system and environment is allowed, we require that at any time t the Hamilton operator of the environment and the system, respectively, have to be separate constants of motion:

$$[\hat{H}(t), \hat{H}^C] = 0, \quad [\hat{H}(t), \hat{H}^S(t)] = 0. \tag{4}$$

Using $[\hat{H}^S(t), \hat{H}^C] = 0$ we readily arrive at

$$[\hat{H}^C, \hat{H}^{\text{int}}] = 0, \quad [\hat{H}^S(t), \hat{H}^{\text{int}}] = 0. \tag{5}$$

The quantum Jarzynski relation for closed systems states that [cf. (2)]

$$\overline{e^{-\beta W}} = \frac{Z(t)}{Z(0)} \tag{6}$$

holds for arbitrary processes. Our aim is to prove that this relation also holds for this type of open quantum system, i.e., for microcanonical coupling. To this end, we investigate the relation of the partition sums:

$$Z(t) = \text{Tr}(e^{-\beta \hat{H}(t)}) = \text{Tr}(e^{-\beta[\hat{H}^S(t) \otimes \hat{1}^C + \hat{1}^S \otimes \hat{H}^C + \hat{H}^I]}). \tag{7}$$

Here we have introduced the abbreviation $\hat{H}^I \equiv \sigma \hat{H}^{\text{int}}$. Since the Hamilton operators commute with each other according to (5), we can split the exponential function, arriving at

$$Z(t) = \text{Tr}(e^{-\beta \hat{H}^S(t)} \otimes e^{-\beta \hat{H}^C} e^{-\beta \hat{H}^I}). \tag{8}$$

Using the definition of a canonical density operator $\hat{\varrho}_{\text{can}} \equiv e^{-\beta \hat{H}}/Z$ we rewrite

$$Z(t) = \text{Tr}[\hat{\varrho}^S_{\text{can}}(t) \otimes \hat{\varrho}^C_{\text{can}} e^{-\beta \hat{H}^I}] Z^S(t) Z^C. \tag{9}$$

For the time dependency of the trace we find, based on the Liouville-von Neumann equation,

$$\frac{d}{dt}\left(\frac{Z(t)}{Z^S(t)Z^C}\right) = -i \, \text{Tr}\{[\hat{H}(t'), \hat{\varrho}^S_{\text{can}}(t) \otimes \hat{\varrho}^C_{\text{can}}] e^{-\beta \hat{H}^I}\}. \tag{10}$$

Making use of the cyclic property of the trace and the commutator relations (5), we end up with

$$\frac{d}{dt} \text{Tr}[\hat{\varrho}^S_{\text{can}}(t) \otimes \hat{\varrho}^C_{\text{can}} e^{-\beta \hat{H}^I}] = 0 \; \forall \; t. \tag{11}$$

Thus, we have in particular

$$\text{Tr}[\hat{\varrho}^S_{\text{can}}(t) \otimes \hat{\varrho}^C_{\text{can}} e^{-\beta \hat{H}^I}] = \text{Tr}[\hat{\varrho}^S_{\text{can}}(0) \otimes \hat{\varrho}^C_{\text{can}} e^{-\beta \hat{H}^I}]. \tag{12}$$

From this we conclude, using Eq. (9) in Eq. (6),

$$\overline{e^{-\beta W}} = \overline{e^{-\beta W^S}} = \frac{Z^S(t)}{Z^S(0)}. \tag{13}$$

The work on the compound system equals that on the open one since no energy exchange between the system and the environment is allowed, and therefore the substitution in Eq. (13) on the left-hand side is justified. Thus the Jarzynski equation holds for microcanonical coupling.

B. Canonical coupling: Constant interaction energy

Next, we turn to systems which are allowed to exchange energy with their surroundings. The Hamilton operator is again of the form (3). We assume the interaction energy to be a constant of motion:

$$[\hat{H}(t), \hat{H}^I] = 0 \Rightarrow [\hat{H}^S(t) + \hat{H}^C, \hat{H}^I] = 0. \tag{14}$$

This assumption is motivated by the fact that, for a sufficiently clear separation between system and environment, the interaction energy should be very small compared to the energy of the system and the environment, respectively. Therefore, its change should also be considered to be negli-

gible. The proof then runs in complete analogy to that of the microcanonical coupling scenario. We start using the Jarzynski relation (2) for closed systems. The partition sum can be written as in Eq. (9) using the commutator relation (14):

$$Z(t) = \text{Tr}(e^{-\beta(\hat{A}+\hat{B})}) \overset{(14)}{=} \text{Tr}(e^{-\beta\hat{A}}e^{-\beta\hat{B}}), \quad (15)$$

with $\hat{A} \equiv \hat{H}^S(t) \otimes \hat{1}^C + \hat{1}^S \otimes \hat{H}^C$ and $\hat{B} \equiv \hat{H}^I$. Using $[\hat{A},\hat{B}]=0$ and resubstituting $\hat{A}$ and $\hat{B}$, we easily arrive at

$$Z(t) = \text{Tr}[\hat{\varrho}_{\text{can}}^S(t) \otimes \hat{\varrho}_{\text{can}}^C e^{-\beta\hat{H}^I}]Z^S(t)Z^C, \quad (16)$$

as in the case of microcanonical coupling. Again, we investigate the time dependency of the trace using the Liouville-von Neumann equation and the cyclic property of the trace,

$$\frac{d}{dt}\left(\frac{Z(t)}{Z^S(t)Z^C}\right) = -i \, \text{Tr}\{\hat{\varrho}_{\text{can}}^S(t) \otimes \hat{\varrho}_{\text{can}}^C[e^{-\beta\hat{H}^I},\hat{H}(t')]\}. \quad (17)$$

Using the commutator relation (14) we conclude that $[e^{-\beta\hat{H}^I},\hat{H}(t')]=0 \;\forall\; t'$. Thus, the relation (11) also holds for this scenario and therefore we conclude using Eq. (16) and relation (11) in Eq. (2) that we still have

$$\overline{e^{-\beta W}} = \frac{Z^S(t)}{Z^S(0)}. \quad (18)$$

This proves that the Jarzynski relation also holds for canonical coupling which conserves the interaction energy.

Note that the work ΔW performed on the compound closed system equals the work ΔW_{sub} done on the subsystem S, i.e., $\Delta W = \Delta W_{\text{sub}}$ [14]. This is, in general, different from the energy change of the subsystem during this process since the subsystem also exchanges energy with its environment,

$$\Delta E_{\text{sub}} = \Delta W_{\text{sub}} + \Delta Q_{\text{sub}}, \quad (19)$$

whereas the energy change of the total, closed system still equals the work performed on it:

$$\Delta E_{\text{tot}} = \Delta W. \quad (20)$$

So the relation (18) is applicable if, in an experiment, one measures $\Delta W = \Delta W_{\text{sub}}$, which is the work one has to supply in order to perform a chosen process.

So Eq. (18) can be rewritten as

$$\overline{e^{-\beta W^S}} = \frac{Z^S(t)}{Z^S(0)}. \quad (21)$$

C. Canonical coupling: High-temperature limit

Finally, we consider systems with a relatively high temperature T, i.e., we have a small inverse temperature β at the beginning. We can then expand the exponent of the partition sum according to the Baker-Campbell-Hausdorff formula and related identities. In the following, we will use the following abbreviations:

$$A \equiv \hat{H}^S(t) \otimes \hat{1}^C + \hat{1}^S \otimes \hat{H}^C \quad \text{and} \quad B \equiv \sigma\hat{H}^{\text{int}}. \quad (22)$$

We thus have

$$e^{-\beta A}e^{-\beta B} = e^{-\beta(A+B)+(\beta^2/2)[A,B]} + O(\beta^3), \quad (23)$$

which can be rewritten as

$$e^{-\beta A}e^{-\beta B} = e^{-\beta(A+B)}e^{(\beta^2/2)[A,B]}. \quad (24)$$

On the other hand,

$$e^{-\beta B}e^{-\beta A} = e^{-\beta(B+A)}e^{(\beta^2/2)[B,A]} + O(\beta^3). \quad (25)$$

We combine Eq. (24) with Eq. (25), and using

$$[e^{-\beta(B+A)}, e^{(\beta^2/2)[B,A]}] = 0 + O(\beta^3) \quad (26)$$

we arrive at

$$\underbrace{e^{-\beta A}e^{-\beta B}}_{(24)}e \underbrace{{}^{-\beta B}e^{-\beta A}}_{(25)} = e^{-\beta(A+B)}e^{(\beta^2/2)[A,B]}e^{(\beta^2/2)[B,A]}e^{-\beta(B+A)}$$

$$= e^{-2\beta(A+B)}. \quad (27)$$

Thus, we find

$$Z(t) = \text{Tr}(e^{-\beta(A+B)}) \overset{(27)}{=} \text{Tr}(e^{-(\beta/2)A}e^{-(\beta/2)B}e^{-(\beta/2)B}e^{-(\beta/2)A})$$

$$= \text{Tr}(e^{-\beta A}e^{-\beta B}), \quad (28)$$

or, resubstituting A and B,

$$Z(t) = \text{Tr}(e^{-\beta[\hat{H}^S(t)\otimes\hat{1}^C+\hat{1}^S\otimes\hat{H}^C]}e^{-\beta\hat{H}^I}). \quad (29)$$

We now can proceed as above:

$$Z(t) = \text{Tr}[\hat{\varrho}^S(t) \otimes \hat{\varrho}^C e^{-\beta\hat{H}^I}]Z^S(t)Z^C. \quad (30)$$

We again investigate the time dependency of the trace using the Liouville-von Neuman equation:

$$\frac{d}{dt}\left(\frac{Z(t)}{Z^S(t)Z^C}\right) = -i \, \text{Tr}\{[\hat{H}(t'),\hat{\varrho}^S(t) \otimes \hat{\varrho}^C]e^{-\beta\hat{H}^I}\}$$

$$= -i \, \text{Tr}\{[\hat{H}^S(t') \otimes \hat{1}^C + \hat{1}^S \otimes \hat{H}^C,\hat{\varrho}^S(t)$$

$$\otimes \hat{\varrho}^C]e^{-\beta\hat{H}^I}\} = -i \, \text{Tr}\{[\hat{H}^S(t') \otimes \hat{1}^C,\hat{\varrho}^S(t)$$

$$\otimes \hat{\varrho}^C]e^{-\beta\hat{H}^I}\} = -i \, \text{Tr}\{[\hat{H}^S(t'),\hat{\varrho}^S(t)]$$

$$\otimes \hat{\varrho}^C e^{-\beta\hat{H}^I}\}, \quad (31)$$

where we have used the cyclic property of the trace in the first equality in order to get rid of the interaction term of the Hamiltonian. If we have $[\hat{H}^S(t'),\hat{H}^S(t)]=0 \;\forall\; t'$, as for a time-dependent Zeeman splitting of a spin, for example, we immediately arrive at

$$\frac{d}{dt}\left(\frac{Z(t)}{Z^S(t)Z^C}\right) = 0 \quad (32)$$

and therefore

$$\frac{Z(t)}{Z(0)} = \frac{Z^S(t)}{Z^S(0)} = \overline{e^{-\beta W}} = \overline{e^{-\beta W^S}}. \quad (33)$$

So, for high initial temperatures of the open quantum system, the Jarzynski relation holds if the Hamiltonian of the system commutes at different times. This can be shown to hold also in third order of β (see the Appendix).
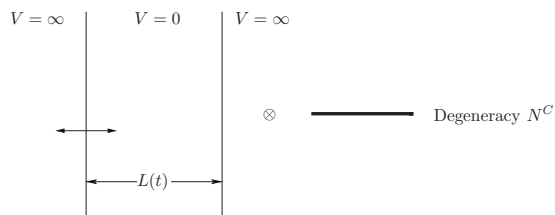
FIG. 2. Model system for our numerical investigations. On the left-hand side the quantum well with a movable wall is depicted. The well is coupled to an environment with a single energy level of high degeneracy.

From the derivation of this result one might expect that the Jarzynski relation would not hold in this form for arbitrary β, but that it could fail, and that the deviation of the average $\overline{e^{-\beta W}}$ from the relation of the partition sums should become larger the smaller is the inverse temperature β. This expectation will be tested numerically.

IV. NUMERICAL RESULTS

As a simple model system we choose the standard potential well with width L containing one particle (particle in a box). Its spectrum is truncated at level n (cutoff). The well is coupled *microcanonically* to an environment (see Fig. 2). To ensure that no energy is exchanged between system and environment, we model the surroundings as a single, highly degenerate, energy level [16].

One wall of this well (the left one, say) is now supposed to be movable, allowing for dilation and compression of the well [well width $L(t)$]. This time dependence will model the working process. The particle is prepared initially in a canonical state. Then the time evolution of the system is calculated numerically under pure Schrödinger dynamics. From this we can infer the work distribution for any control function $L(t)$. The discrete values of W result from the discrete spectral energy differences. The nonmonotonic behavior of $P(W)$ is a combined effect of the spectrum and the initial occupation probabilities. We use this $P(W)$ in order to compute the average $\overline{e^{-\beta W}}$. This result is then compared with the relation of the partition sums $Z^S(t)/Z^S(0)$, which can easily be computed independently for this model. We define $\overline{e^{-\beta W}} = \gamma_J Z^S(t)/Z^S(0)$ with γ_J being a test factor. If the Jarzynski relation is correct, the test factor should always equal 1 within numerical accuracy.

Because we want to consider the work done on the compound, closed system, we define the work as $W \equiv \Delta E$, since the work simply equals the energy change of the closed system. Since we will, in general, find the system in different states for each measurement, the work is said to fluctuate. However, the definition of work and the corresponding fluctuations is still a matter under dispute [17].

A. Closed systems

For our investigation of the closed system we choose the interaction strength with the environment to be zero.
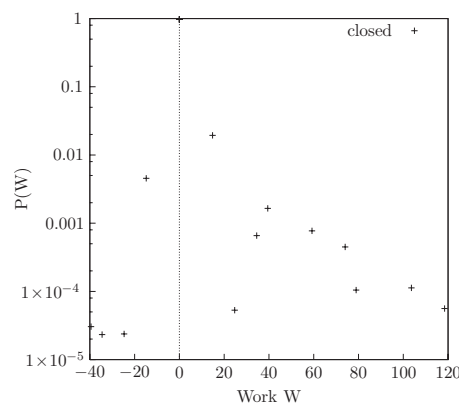

FIG. 3. Distribution of work $P(w)$ for a closed system. A circular process with $\hat{H}(t_{\mathrm{final}}) = \hat{H}(0)$ and a maximal width of $L_{\max} = 2L(0)$ was chosen. The wall has been moved with constant velocity $\pm v$. The velocity of the process as well as the unit energy is in arbitrary units. Probabilities smaller than 1×10^{-5} are suppressed; cutoff at $n = 5$ is justified by a low initial temperature. The average work $\bar{W}$ is larger than zero. There is a finite chance to gain work while performing this process, as can be seen from the points to the left of the dashed line $W = 0$.

Our numerics show that, within numerical errors, for any process realization and any process velocity chosen, we indeed have $\gamma_J = 1$, which means that the Jarzynski relation is found to be valid. This result holds for different wall displacement velocities, different initial temperatures, or different well widths. See Fig. 3 for a typical result for the work distribution obtained.

B. Microcanonical coupling

Next, we turn to systems with $\sigma \neq 0$. In order to check whether the Jarzynski relation holds, we need to consider a process for which the final Hamiltonian differs from the initial one; otherwise one trivially would have $Z(t)/Z(0) = Z^S(t)/Z^S(0) = 1$. For now, we focus on expansions of the quantum well.

We find that the work distribution is slightly affected by the microcanonical environment (see Fig. 4). For better comparison, the work distribution for the closed system is depicted, too. We get $\gamma_J = 1$ within our numerical accuracy, indicating that the Jarzynski relation is valid.

This also holds when changing the coupling strength, enlarging the environment by increasing the degeneracy, or changing the initial temperature. Note that the distribution of work for the closed system is, indeed, though only slightly, different from the one obtained for the same boundary conditions under microcanonical coupling.

Due to the interaction between system and environment the energy levels of the system are "smeared out." Despite the fact that the interaction is weak, the system's energy levels are each split into many energy levels with slightly different energy eigenvalues according to the number of energy levels in the environment and their respective degeneracies. Some of these energy levels are resolved within numerical accuracy, leading to slightly distinct values of work W.
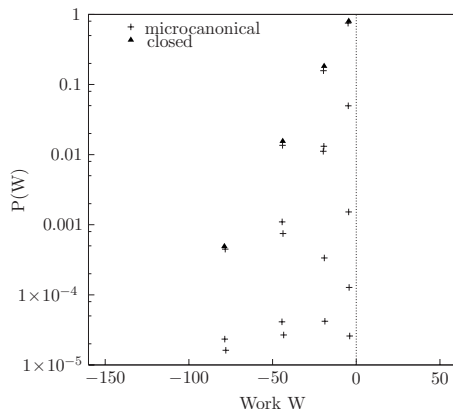
FIG. 4. Distribution of work $P(w)$ for a closed and a microcanonical system, respectively. An expansion with $L(t_{final})=2L(0)$ was chosen. The wall has been moved with constant velocity v. The degeneracy was $N^C=150$ and the coupling strength $\sigma=0.015$ for the microcanonical coupling. For the closed system, σ was set to zero. The velocity of the process as well as the unit energy are in arbitrary units. Probabilities smaller than 1×10^{-5} are suppressed; cutoff at $n=5$.

This might lead to the impression that we have multiple values of work probabilities $P(W)$ for a given work W in Figs. 4–7 for the open system. This is not the case since, for exactly—in the numerical sense—equal work values, the probabilities add up to a single probability $P(W)$.

C. Canonical coupling

We finally turn to a canonical bath, investigating whether the Jarzynski relation still holds in the high-temperature limit, as predicted. To this end, we couple a canonical bath to the quantum well. It is modeled in such a way that it allows every possible transition in our quantum well when in resonance [16]. Its degeneracy grows with increasing energy. For now, we focus on expansions of the quantum well, which is
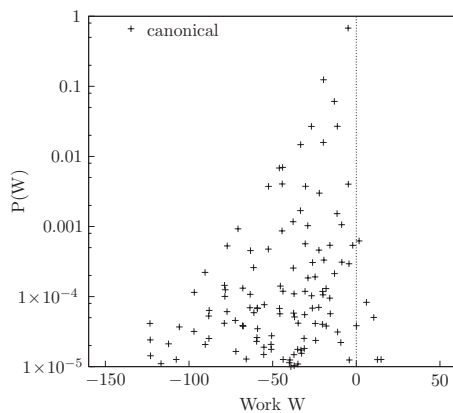

FIG. 5. Distribution of work $P(w)$ for a canonical system. An expansion with $L(t_{final})=2L(0)$ and $\sigma=0.015$ was chosen. The wall has been moved with constant velocity v. The velocity of the process as well as the unit energy are in arbitrary units. Probabilities smaller than 1×10^{-5} are suppressed; cutoff at $n=5$.
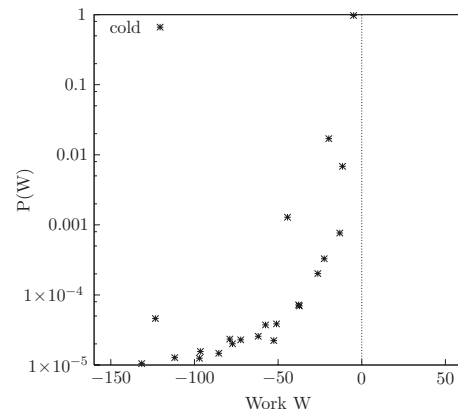

FIG. 6. Same as Fig. 5, but with low initial temperature.

effectively cooled down by these dilations. This will cause energy to flow from the bath into the system, trying to restore the initial temperature, and we expect that transitions to higher energy levels will be facilitated. First, we have chosen a moderate initial temperature, such that only the first two energy levels are significantly occupied, resulting typically in a work distribution as depicted in Fig. 5.

We find that Eq. (33) does not hold; the Jarzynski relation of this form is not strictly valid. We have for our test factor $\gamma_J=0.999$, which quantifies the still small deviation from the Jarzynski relation. From Sec. III C we suppose that the lower the temperature the bigger this deviation. Indeed, if we choose the initial temperature to be very low, such that almost only the ground state is occupied, we have $\gamma_J=0.98$. Moreover, we expect the work distribution to be very narrow since transitions from the ground state into higher ones are relatively unlikely (see Fig. 6).

Next, we choose a very high initial temperature, such that almost every energy level has the same probability to be occupied, in order to check whether the Jarzynski relation is valid as predicted. The work distribution is, as expected,
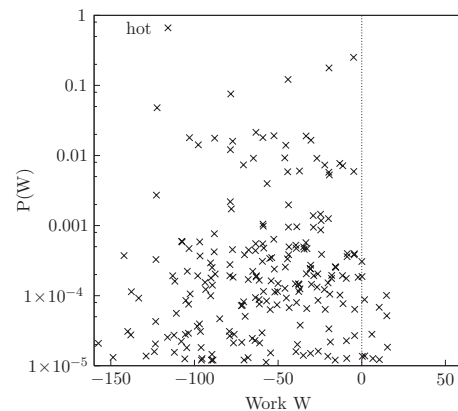

FIG. 7. Same as Fig. 5, but with high initial temperature.

rather broad for these temperatures (see Fig. 7). Nevertheless, we find that $\gamma_J = 1$ within numerical accuracy. This means that Eq. (33) is satisfied for high initial temperatures, confirming our high-temperature limit.

V. CONCLUSION

Starting from the quantum version of the Jarzynski relation for closed systems, we have shown that this relation can be generalized to open quantum systems of different coupling types. The key idea has been to split the exponent of the partition sum into parts for the system and environment, respectively. This is, in general, not possible, since the system and environmental Hamiltonians do not commute with the interaction part. This fact would result in additional terms containing the commutator of these Hamiltonians.

Nevertheless, we have been able to split up the partition function under the selected conditions, since then the correction terms cancel.

We have stressed that the problem of fluctuation is intimately related to some measurement scheme. Further investigations should be guided by the observation that in quantum mechanics it seems more natural to consider the fluctuation of a (subsystem) energy rather than of work proper [18].

APPENDIX: THIRD-ORDER EXPANSION IN TERMS OF $\beta \ll 1$

We want to extend the result of Sec. III C to the third power of β. We proceed analogously as above, trying to split up the exponential function. For $O(\beta^4) = 0$ we arrive at

$$e^{-\beta A} e^{-\beta B} = e^{-\beta(A+B)+(\beta^2/2)[A,B]} e^{-(\beta^3/12)([A,[A,B]]+[B,[A,B]])}$$

(A1)

and

$$e^{-\beta B} e^{-\beta A} = e^{(\beta^3/12)([A,[A,B]]+[B,[A,B]])} e^{-\beta(B+A)-(\beta^2/2)[A,B]}.$$

(A2)

Combining Eq. (A1) with Eq. (A2), we have

$$\underbrace{e^{-\beta A} e^{-\beta B}}_{(A1)} e \underbrace{e^{-\beta B} e^{-\beta A}}_{(A2)} = e^{-\beta(A+B)+(\beta^2/2)[A,B]} e^{-\beta(B+A)-(\beta^2/2)[A,B]}$$

$$\equiv e^{-\beta C_+} e^{-\beta C_-},$$

(A3)

with $C_+ \equiv -\beta(A+B)+(\beta^2/2)[A,B]$ and $C_- \equiv -\beta(A+B)-(\beta^2/2)[A,B]$. We investigate the resulting product, introducing the abbreviation $[C_+, C_-] \equiv D_\pm$:

$$e^{-\beta C_+} e^{-\beta C_-} = e^{-\beta(C_++C_-)+(\beta^2/2)D_\pm-(\beta^3/12)([C_+,D_\pm]+[C_-,D_\pm])} + O(\beta^4).$$

(A4)

Since $C_+ = C_- + \beta^2[A,B]$, we have

$$D_\pm = -\beta^3([A,[A,B]]+[B,[A,B]])$$

(A5)

$$\Rightarrow [C_\pm, D_\pm] \overset{O(\beta^4)}{=} 0.$$

(A6)

This leaves us with

$$e^{-\beta C_+} e^{-\beta C_-} = e^{-\beta(C_++C_-)+(\beta^2/2)D_\pm} = e^{-2\beta(A+B)} e^{(\beta^2/2)D_\pm} + O(\beta^4)$$

(A7)

and analogously

$$e^{-\beta C_-} e^{-\beta C_+} = e^{(\beta^2/2)D_\mp} e^{-2\beta(A+B)} + O(\beta^4),$$

(A8)

with $D_\mp \equiv [C_-, C_+] = -D_\pm$. Combining these two relations, we have

$$e^{-\beta C_+} e^{-\beta C_-} e^{-\beta C_-} e^{-\beta C_+} = e^{-4\beta(A+B)}.$$

(A9)

Therefore we have

$$\text{Tr}(e^{-\beta A} e^{-\beta B}) = \text{Tr}(e^{-(\beta/2)A} e^{-(\beta/2)B} e^{-(\beta/2)B} e^{-(\beta/2)A})$$

$$\overset{(A3)}{=} \text{Tr}(e^{-(\beta/2)C_+} e^{-(\beta/2)C_-})$$

$$= \text{Tr}(e^{-(\beta/4)C_+} e^{-(\beta/4)C_-} e^{-(\beta/4)C_-} e^{-(\beta/4)C_+})$$

$$\overset{(A9)}{=} \text{Tr}(e^{-\beta(A+B)}) = Z.$$

(A10)

The rest of the proof for the Jarzynski relation to hold runs analogously to Sec. III C.

[1] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[2] C. Jarzynski, J. Stat. Phys. **98**, 77 (2000).
[3] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[4] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[5] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[6] J. Kurchan, e-print arXiv:cond-mat/0007360.
[7] T. Monnai, Phys. Rev. E **72**, 027102 (2005).
[8] H. Tasaki, e-print arXiv:cond-mat/0009244.
[9] T. Monnai and S. Tasaki, e-print arXiv:cond-mat/0308337.
[10] A. E. Allahverdyan and Th. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).
[11] C. Jarzynski and D. K. Wójcik, Phys. Rev. Lett. **92**, 230602 (2004).
[12] W. De Roeck and C. Maes, Phys. Rev. E **69**, 026115 (2004).
[13] S. Mukamel, Phys. Rev. Lett. **90**, 170604 (2003).
[14] M. Esposito and S. Mukamel, Phys. Rev. E **73**, 046129 (2006).
[15] F. Ritort, Poincare Sem. **2**, 195 (2003).
[16] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics—Emergence of Thermodynamic Behavior Within Composite Quantum Systems* 1st ed. (Springer-Verlag, Berlin, 2004).
[17] A. Engel and R. Nolte, Europhys. Lett. **79**, 10003 (2007).
[18] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102(R) (2007).