PHYSICAL REVIEW E **89**, 042113 (2014)

Entropy increase in K-step Markovian and consistent dynamics of closed quantum systems

Jochen Gemmer[1,*] and Robin Steinigeweg[2]
[1]*Department of Physics, University of Osnabrück, D-49069 Osnabrück, Germany*
[2]*Institute for Theoretical Physics, Technical University Braunschweig, D-38106 Braunschweig, Germany*
(Received 21 January 2014; published 7 April 2014)

We consider sequences of measurements implemented by positive operator valued measures (POVMs). Starting from the assumption that these sequences may be described as consistent and Markovian, even and especially for closed quantum systems, we identify properties of the equilibrium state that coincide with the properties of typical pure quantum states. We define a physical entropy that converges against the standard entropies in the approach to equilibrium. Furthermore, strict limits to its possible decrease are derived on the basis of Renyi entropies. It is demonstrated that Landauer's principle follows directly from these limits. Since the above assumptions are rather strong, we exemplify the fact that they may nevertheless apply by checking them numerically for some transition paths in a concrete model.

DOI: 10.1103/PhysRevE.89.042113 PACS number(s): 05.30.−d, 05.70.Ln, 03.65.Yz

I. INTRODUCTION

Roughly 100 years after the beginning of its systematic microscopic interpretation, the origin of thermodynamics is still under dispute (see, e.g., Refs. [1–3], and references therein). It is, however, an empirical fact that macroscopic systems behave according to the laws of thermodynamics, and they are routinely viewed as large quantum systems. Accordingly, already in early works on quantum mechanics [4–8] the question about the relationship between quantum mechanics and thermodynamics arose. Meanwhile many concepts have been discussed such as "typicality" [2,5,9–12], "pure state quantum statistical mechanics" [13–15], "eigenstate thermalization hypothesis" [5,16–19], "thermal environment coupling" [20–22], and many more.

In most of the more recent approaches entropy as a concept does not play a crucial role. In the context of typicality it has been shown that pure states yielding high von Neumann entropies for a small subsystem are in some sense by far the most frequent ones (see, e.g., Refs. [11,23]) but such a statement in itself bares no rigorous consequences on entropy dynamics. In other typicality approaches that do not focus on subsystems entropy is not even mentioned [12].

In some of the "pure state quantum statistical mechanics" approaches it is shown that (under some conditions on the model and the initial pure state) the reduced state of some small system is, at most points in time during an interval of unspecified but possibly very large length, close to a maximum-entropy state (see, e.g., Ref. [14]). While this has some implications on entropy dynamics, it does not exclude large entropy oscillations during a long period of time as long as they eventually die out and disappear for the largest part of the above interval. In nature, however, it appears that entropy decreases are, if they appear visibly at all, seriously limited by fluctuation theorems (see, e.g., Ref. [24]). Or, in plain language, entropy always and time-locally increases predominantly. Such findings cannot be inferred from "pure state quantum statistical mechanics."

Another recent approach to the reconciliation of entropy increase and microscopic dynamics is the concept of "stochastic thermodynamics." The basis of this approach is usually Markovian master or Langevin equations along with the stochastic trajectories corresponding to them [25]. Entropy productions along those trajectories are defined involving the concrete transition rates. General statements on positivity of entropy production [25] and fluctuation theorems [26] result. However, this concept is centered very much on entropy production rather than entropy itself. Since temporal integration of the entropy production involves the details of the stochastic dynamics, it is apparently rather challenging to tell whether the actual entropy will converge against standard equilibrium entropies in standard scenarios in the long-time limit. Investigations along these lines of two macroscopic objects exchanging heat with each other exist [27] but resort to the assumption of local Gibbs states. Furthermore (other than in Ref. [27]) stochastic thermodynamics usually rely on a system-environment partition concept, which makes the application to closed quantum systems challenging. Another question is if and in which sense quantum dynamics may be mapped onto master equations. In standard cases of open quantum systems this is well known [28], but in general it is more complicated. Investigations in this direction, however, exist [29–31].

It is the purpose of the paper at hand to progress in the direction of a (quantum) entropy that is defined in equilibrium and nonequilibrium and can be shown to increase predominantly. To this end we adopt and develop an approach by Penrose [32], while combining it with the concept of consistent histories. Within this approach the entropy has the following general features: (1) The entropy is always defined as a function of the actual state $\hat{\rho}(t)$ of a system, a prechosen set of measurement outcomes or "properties" as termed by Joos [33] (represented by positive operators such as $\hat{A}_n^\dagger \hat{A}_n$) and possibly the Hamiltonian $\hat{H}$ but not on transition rates, etc. No notion of a subsystem or any kind of environment is invoked. (2) Rigorous limits on possible decreases of the entropy may be formulated. (3) If the entropy reaches a maximum in the long-time limit, this maximum coincides with the standard Boltzmann entropy in equilibrium. Thus, in those cases extensitivity of entropy etc. follows in the long-time limit.

*jgemmer@uos.de

The above claimed features of the entropy are rather strong. In order to demonstrate them, we resort to some rather strong assumptions as well: (1) We require probabilities of measurement sequences of the above properties to form "consistent histories," i.e., fulfill the third Kolmogorov axiom. (2) We assume that the probability to obtain a certain measurement outcome or property in the future only depends on a finite set of measurement outcomes from the finite past. We call this "K-step Markovianity." While it is close to impossible that those assumptions apply exactly to a closed, finite quantum system, they may apply in an approximate sense, and this is what we focus on.

The consistency assumption essentially establishes that the visible dynamics (measurement outcomes corresponding to the $\hat{A}_n^\dagger \hat{A}_n$) of a system do not depend on whether or not one actually watches it. While this seems natural in the context of everyday experience, it is clearly violated for certain quantum experiments such as the double slit etc. Thus, the consistency assumption implies a certain degree of "classicality." Also the Markovianity claim seems natural from everyday experience. It is, however, anything but obvious why it should apply to any generic class of quantum systems at all. Thus, as already mentioned, both assumptions are rather strong, and justifying them in any generality is beyond the scope of this paper. In order to demonstrate, however, that both may apply at all, we present concrete numerical examples based on a class of Hamiltonians with a finite number of (≈ 1000) eigenstates to which both assumptions already apply to good approximation. Furthermore, the consistency and Markovianity appear to increase with growing system sizes.

In the following we explain the organization of this paper, thereby highlighting in which sense it goes beyond existing literature. While Penrose requires quantum observables to commute with each other at any time, i.e., $[\hat{P}_n(t),\hat{P}_m(t')] = 0$, we do not assume this commutativity, not even approximately. Instead of this we resort to the concept of consistent histories. This is along the lines of Refs. [34] or [35]. The latter work makes some contact with relaxation issues since it applies the formalism to, e.g., the Caldeira-Legett model. While we will refer to the concept of consistent histories quite frequently, it is important to notice that for no part of the work at hand is the interpretation of quantum mechanics of any relevance. We use the consistent-history formalism simply as a "in some sense more elaborate version of the Copenhagen interpretation" [36]. Consistent histories are routinely introduced on the basis of projective measurements. However, since projective measurements of, e.g., positions are in quantum mechanics in conflict with finite energies (nevertheless projective measurements form the basis of the considerations in Ref. [32]), we present a generalization of consistent histories to POVMs in Sec. II that appears, according to, e.g., Ref. [37], not to be present in the literature. We consider this generalization conceptually relevant, even though it is not explicitly used in the remainder of this paper. While the considerations in Ref. [32] are restricted to situations where probabilities for future measurement outcomes may be determined from the last, most recent measurement only, we generalize to scenarios in which the last K measurements are relevant. Thus, Sec. III is essentially dedicated to the introduction of our notion of K-step Markovianity and the concrete translation of K-step Markovian, consistent quantum dynamics into stochastic processes. We identify fixed points as well as limits on transition rates in Sec. III. Based on Renyi divergences, irreversibility and a concrete notion of entropy are introduced in Sec. V. In Sec. VI we go beyond Ref. [32] by demonstrating that the fixed point found in Sec. III just from consistency and Markovianity is actually in accord with the recently much discussed concept of quantum typicality. While the nondecrease of entropy in Ref. [32] is essentially based on considerations of Kullback-Leibler divergences, we demonstrate that this does not suffice to obtain Landauer's principle in Sec. V. To this end the previously introduced Renyi divergences are necessary. Section VIII is somewhat detached from the previous sections. It consists only of a simple numerical example for the emergence of consistency and Markovianity within closed, finite quantum systems.

II. CONSISTENT HISTORIES AND POVM MEASUREMENTS

The concept of consistent histories is needed in the context of this approach in order to map the dynamics of measurement outcomes onto stochastic processes; cf. Sec. III. The concept of POVMs [38], which generalizes projective von Neumann measurements to nonprojective measurements, will not appear explicitly in any calculation after this section. It is, however, important for the applicability of the approach at hand in general. This is due to the same reason which motivated the introduction of POVMs to quantum mechanics in the first place: There are scenarios for which projective measurements are simply unphysical. For example, a projective position measurement of any particle would, strictly speaking, send its mean energy to infinity, regardless of the mass of the particle and the length of the measured interval. While in practice probably most macroscopic measurements are close to projective measurements, the impact of small deviations from "clean projectiveness" should be considered. To those ends, we aim at formally combining the concepts of POVM measurements with the concept of consistent histories. While this formal combination requires only a modification of the "decoherence functional" [39], it does not appear to be present in the literature [37]. (In Ref. [35] it is explicitly mentioned that it would be convenient to be able to go beyond projective measurements in consistent histories; however, the approach is not worked out in detail.)

In POVM measurements a measurement outcome or property n is routinely associated with a (positive) operator $\hat{A}_n^\dagger \hat{A}_n$. Since some outcome has to result, one requires (completeness)

$$\sum_{n=1}^{N} \hat{A}_n^\dagger \hat{A}_n = \hat{1}. \tag{1}$$

The probability of getting the outcome n is given by $P(n) = \mathrm{Tr}\{\hat{A}_n^\dagger \hat{A}_n \hat{\rho}\}$, where $\hat{\rho}$ is the density operator of the full system. For brevity, we introduce two "superoperations": the analog of measurement projection,

$$\mathcal{A}_{nm}\,\hat{\rho} := \hat{A}_n\,\hat{\rho}\,\hat{A}_m^\dagger \tag{2}$$

(note that $\mathcal{A}_{nn}$ always maps a positive operator onto a positive operator), and the time evolution,

$$\mathcal{U}\,\hat{\rho} := \hat{U}(\tau)\,\hat{\rho}\,\hat{U}^\dagger(\tau), \tag{3}$$

where $\hat{U}(\tau)$ is a unitary operation that propagates for time τ. Note that $\hat{U}(\tau)$ is not required to conserve energy. Given those definitions, the probability $P(ijk,\hat{\rho})$ to measure, e.g., first i, then j, then k with time steps τ in between, and starting from an initial state $\hat{\rho}$, may be denoted as

$$P(ijk,\hat{\rho}) := \mathrm{Tr}\{\mathcal{A}_{kk}\,\mathcal{U}\,\mathcal{A}_{jj}\,\mathcal{U}\mathcal{A}_{ii}\,\hat{\rho}\}. \tag{4}$$

(This is sometimes referred to as "Wigner's formula" [40].) It is an acceptable definition insofar as $0 \leqslant P(ijk,\hat{\rho}) \leqslant 1$ and $\sum_{ijk} P(ijk,\hat{\rho}) = 1$. For simplicity, we consider possible measurements only at times t that are multiple integers of τ. While this restriction is not crucial, it turns out to be convenient. In order to apply now the concept of consistent histories to the probabilities for sequences of measurement outcomes, we define another operator $\hat{A}_0$:

$$\hat{A}_0 := \hat{1} - \sum_{n=1}^{N} \hat{A}_n. \tag{5}$$

For purely projective measurements this definition is unnecessary since, if all $\hat{A}_n$ are projectors, clearly $\hat{A}_0 = 0$. As an obvious consequence of this definition, we get

$$\sum_{n=0}^{N} \hat{A}_n = \hat{1}. \tag{6}$$

Now the equivalence to the consistency condition in the consistent-history approach may be formulated based on the following equality. Since this will be the main subject below, we only consider the situation where there is no information at all prior to the first measurement, i.e., $\hat{\rho} = \hat{1}/d$, where d is the (relevant) dimension of the system,

$$P(i\#k,\hat{1}/d) = \sum_{j=1}^{N} P(ijk,\hat{1}/d) \tag{7}$$

$$+ \sum_{j,l=0;j\neq l}^{N} \mathrm{Tr}\{\mathcal{A}_{kk}\,\mathcal{U}\,\mathcal{A}_{jl}\,\mathcal{U}\,\mathcal{A}_{ii}\,\hat{1}/d\} \tag{8}$$

$$+ \mathrm{Tr}\{\mathcal{A}_{kk}\,\mathcal{U}\,\mathcal{A}_{00}\,\mathcal{U}\,\mathcal{A}_{ii}\,\hat{1}/d\}, \tag{9}$$

where the # character is meant to indicate that no measurement is performed at the corresponding point in time. If the sum in Eq. (8) and the term in Eq. (9) were zero, i.e., if the equation consisted only of Eq. (7), then the probabilities as defined in Eq. (4) would obey the third Kolmogorov axiom on the additivity of probabilities of disjoint events. Or, in plain language, if a property can be reached from another property via different paths, the probabilities of those paths simply add up. Below we are going to assume that this indeed either holds exactly or, more realistically, at least to a degree at which deviations are negligible. However, whether or not contributions from Eqs. (8) and (9) vanish or not ultimately depends on the measurement operators $\hat{A}_n$ and, through $\mathcal{U}$, on the Hamiltonian. Thus, in Sec. VIII we provide numerical evidence that for a concretely given Hamiltonian, represented by a finite-dimensional matrix, and concretely given $\hat{A}_n$ the sum in Eq. (8) appears to go to zero as the Hilbert-space dimension d goes to infinity. The sum in Eq. (8) is equivalent to what is called the "decoherence functional" in the context of consistent histories. The term in Eq. (9) is not present in standard discussions of consistent histories. Whether or not it is small in principal has to be checked for any given model and given properties $\hat{A}_n$. It is, however, plausible that this term is small as long as the $\hat{A}_n$ are reasonably close to projectors.

III. TRANSITION PROBABILITIES, MARKOVIANITY, AND STOCHASTIC PROCESSES

Given the "consistency condition" derived in the previous section, we now aim at mapping the quantum dynamics onto stochastic processes in this section. To those ends, we introduce the definition of another quantity Ω which will turn out to be closely connected to the below introduced entropy:

$$\Omega(ijk) := \mathrm{Tr}\{\mathcal{A}_{kk}\,\mathcal{U}\,\mathcal{A}_{jj}\,\mathcal{U}\,\mathcal{A}_{ii}\,\hat{1}\}. \tag{10}$$

(Note that those Ω are all non-negative.) This may be viewed as being proportional to the probability to measure (ijk) without any prior knowledge, i.e., $\Omega(ijk) = d\,P(ijk,\hat{1}/d)$. For later reference, we note here that "summing over an index shortens the history"; i.e., by virtue of Eq. (1) we get, e.g.,

$$\sum_{i=1}^{N} \Omega(ijk) = \Omega(jk). \tag{11}$$

The conditional probability to get some outcome l in an upcoming measurement, given that one has observed a certain sequence of outcomes in previous measurements of, e.g., three measurements (ijk) but without any knowledge whatsoever prior to the first measurement, is now obtained within the above framework of POVMs and consistent histories from Eqs. (4) and (10) as

$$w(l|ijk) = \frac{\Omega(ijkl)}{\Omega(ijk)}. \tag{12}$$

Obviously the conditional probabilities are always fractions of Ω in which the sequence in the argument of the Ω in the numerator is one element longer than the sequence in the argument of the Ω in the denominator. Evidence from physical experience suggests that the probability of measuring some outcome l may be independent of measurement outcomes obtained in a distant past. Thus, those conditional probabilities may become at some point independent of the lengths of the sequences. This allows for a definition of K-step Markovianity: If there is one-step Markovianity, which is often simply called "Markovianity" only, the last past measurement outcome should be relevant for the conditional probability:

$$\frac{\Omega(\ldots ijkl)}{\Omega(\ldots ijk)} = \frac{\Omega(kl)}{\Omega(k)} = w(l|k). \tag{13}$$

In the case of two-step Markovianity we require, respectively,

$$\frac{\Omega(\ldots ijkl)}{\Omega(\ldots ijk)} = \frac{\Omega(jkl)}{\Omega(jk)} = w(l|jk), \tag{14}$$

and so on. For the remainder of this paper we are going to assume that K-step Markovianity holds with K arbitrarily large but finite. This means that is unnecessary to keep an infinite record from the past in order to come up with the best possible prediction on the future. While such a statement may appear very natural, it is a challenging task to explain why this should hold for a great variety of systems and properties. In this paper we are not going to address this question in any generality. However, in Sec. VIII we provide numerical evidence that for a concretely given Hamiltonian, represented by a finite-dimensional matrix, and concretely given $\hat{A}_n$ the deviation from one-step Markovianity appears to go to zero as the Hilbert-space dimension goes to infinity. Thus, the set of finite-dimensional closed quantum systems that are in the above sense Markovian can, at least, be shown to be not empty. (Note that the term "Markovian" is not unambiguously defined in the literature. For another definition of Markovianity see, e.g., Ref. [41], and references therein.)

Thus, for the case of one-step Markovianity and decoherent dynamics as described below Eq. (7) one may set up a time-discrete stochastic process which describes the measurable dynamics as

$$p(j;t+\tau) = \sum_i w(i|j)\, p(j;t). \tag{15}$$

Given the definition (13), it is evident that $0 \leqslant w(i|j) \leqslant 1$, and, considering Eq. (1), one also gets $\sum_i w(i|j) = 1$. Hence, Eq. (15) has the properties of a standard stochastic process. Note that Eq. (15) does not necessarily feature detailed balance. This will only be the case if $\Omega(ji)/\Omega(ij) = f(i)/g(j)$, where f,g are some real, non-negative functions. However, whether or not this applies is irrelevant for the validity of the remainder of this paper. In the case of "more-than-one"-step Markovianity the mapping of the measurable dynamics onto a stochastic process is slightly more complex. For simplicity, we only address two-step Markovianity, but K-step Markovianity may be dealt with in the same way. We resort to the method of going from two-step Markovianity back to one-step Markovianity by using a new set of states. Let $w(lm|jk)$ denote the probability that, after jk have been measured, the next measurement will yield m and the last measurement l. Of course, this can only be nonzero if $k = l$, but, for reasons which will become clear below, it is convenient to formally define the two-step Markovian transition probabilities this way. Thus, for the case of two-step Markovianity and, again, decoherent dynamics as described below Eq. (7) one may set up a time-discrete stochastic process which describes the measurable dynamics as

$$p(lm;t+\tau) = \sum_{jk} w(lm|jk)\, p(jk;t). \tag{16}$$

From Eq. (14) follows the explicit form of $w(lm|jk)$:

$$w(lm|jk) = \delta_{kl}\, w(m|jk). \tag{17}$$

Again, from Eq. (14) it is evident that $0 \leqslant w(lm|jk) \leqslant 1$ and, considering Eq. (1), one also gets $\sum_{lm} w(lm|jk) = 1$. Hence, Eq. (16) also has the properties of a standard stochastic map.

To shorten notation, we label all possible ordered sequences of such K results by Greek letters, i.e., $(\dots ijk) := \alpha$. In the following such sequences will be sometimes called "properties." Many of the standard examples of nonequilibrium thermodynamics are one-step Markovian, e.g., the decay of the temperature difference between two macroscopic objects in thermal contact is a one-step Markovian (macroscopically deterministic) process. However, a damped (not overdamped) harmonic oscillator is described by a two-step Markovian process if one restricts oneself to position measurements only. It may be one-step Markovian if the properties $\hat{A}_n$ are constructed to encode positions and momenta. However, since measurements of velocity are usually done by measuring subsequent positions and the time elapsed in between, it is convenient to argue within a framework which includes experiments based on position measurements only.

IV. EQUILIBRIUM STATE AND UPPER BOUNDS ON TRANSITION PROBABILITIES

Although the above stochastic processes do not necessarily fulfill detailed balance, their equilibrium states may be determined. This is most conveniently demonstrated by "guessing" the equilibrium state and plugging it into the stochastic process. Our guess for the equilibrium state is

$$p_\alpha^{\text{eq}} \propto \Omega(\alpha). \tag{18}$$

We check the guess by inserting it into Eqs. (15) and (16) and demanding that it should be identically reproduced by the transition matrix. Thus, for one-step Markovianity we get

$$\begin{aligned} \Omega(j) &= \sum_i w(j|i)\,\Omega(i), \\ \Omega(j) &= \sum_i \frac{\Omega(ij)}{\Omega(i)}\,\Omega(i), \\ \Omega(j) &= \sum_i \Omega(ij), \\ \Omega(j) &= \Omega(j). \end{aligned} \tag{19}$$

The last line follows from Eq. (11). Obviously, the guess from Eq. (18) is correct. The same consideration for two-step Markovianity reads:

$$\begin{aligned} \Omega(lm) &= \sum_{jk} w(lm|jk)\,\Omega(jk), \\ \Omega(lm) &= \sum_{jk} \delta_{kl}\,\frac{\Omega(jkm)}{\Omega(jk)}\,\Omega(jk), \\ \Omega(lm) &= \sum_{jk} \delta_{kl}\,\Omega(jkm), \\ \Omega(lm) &= \sum_{j} \Omega(jlm), \\ \Omega(lm) &= \Omega(lm). \end{aligned} \tag{20}$$

Again, the guess from Eq. (18) is obviously correct. Note that all statements below do not depend on whether or not there may be more invariant equilibrium states in addition to the one given by Eq. (18).

In the case of consistent, one-step Markovian dynamics, the equilibrium state does not even depend on $\mathcal{U}$ or $\hat{H}$. It only depends on the "measured properties" since in this case one simply gets $p_n^{\text{eq}} = \text{Tr}\{\hat{A}_n^\dagger\,\hat{A}_n\}$. This framework accounts for the fact that in equilibrium, e.g., the probabilities for various distributions of energy onto two subsystems in thermal contact are entirely independent of the specific details or, within reasonable bounds, the strength of the contact. They only depend on the subsystems; the interaction would only enter through $\mathcal{U}$ and is thus irrelevant.

Next we establish bounds to the transition probabilities between different properties based on the properties themselves. In the case of one-step Markovian dynamics, this means that the bounds hold irrespective of the specific $\mathcal{U}$. Let α be the sequence of the last K measurement outcomes and β the sequence of the last K outcomes after one more measurement has been done. (This means that β is essentially α but only "backshifted" by one, with the last one erased, and with a new most recent measurement outcome.) Then the transition probability $w(\beta|\alpha)$ to go from α to β is limited by

$$w(\beta|\alpha) \leqslant \frac{\Omega(\beta)}{\Omega(\alpha)}. \tag{21}$$

Demonstrating the validity of Eq. (21) is rather simple. Using the property notation, the first lines of Eqs. (19) and (20) read:

$$\Omega(\beta) = \sum_{\alpha} w(\beta|\alpha)\,\Omega(\alpha) \tag{22}$$

or

$$1 = \sum_{\alpha} w(\beta|\alpha)\,\frac{\Omega(\alpha)}{\Omega(\beta)}. \tag{23}$$

Since the transition probabilities as well as the Ω are all non-negative, Eq. (21) directly follows. This result is certainly important for establishing irreversibility since, if, say, $\Omega(\beta) \gg \Omega(\alpha)$, it is well possible that $w(\beta|\alpha) = 1$. However, the backward transition probability must be small, i.e., bound by $w(\alpha|\beta) \leqslant \Omega(\alpha)/\Omega(\beta)$.

Based on Eq. (21) it may appear reasonable to define entropy simply as a mean Ω or a monotonous function of it, e.g.,

$$S := \sum_{\alpha} p_\alpha \ln \Omega(\alpha). \tag{24}$$

While such a function indeed plays an important role in the approach at hand, it is by itself not strictly nondecreasing, as will be demonstrated in the next section; cf. especially Eq. (29).

V. ENTROPY AS A NONDECREASING QUANTITY

The fact that the above scenario cannot happen, even though it is not in conflict with Eq. (21), may be seen from another consideration. This consideration also leads to the introduction of an entropy and expresses the sense in which this entropy is nondecreasing.

The following analysis involves the Renyi divergences $D_a(P||Q)$. The latter are defined as

$$D_a(P||Q) := \frac{1}{a-1} \ln \sum_{\alpha} \left(\frac{p_\alpha}{q_\alpha}\right)^a q_\alpha \tag{25}$$

for $a > 0$. Although the Renyi divergences are no metrics, they may to some extent be viewed as distances between two probability distributions $P := \{p_\alpha\}, Q := \{q_\alpha\}$. It has been shown that in stochastic processes of the type of Eqs. (15) and (16) the Renyi divergences between some actual state and an equilibrium state cannot increase, which means

$$D_a[P(t+i\tau)||Q] \leqslant D_a[P(t)||Q] \tag{26}$$

with i being some positive integer. For a derivation of Eq. (26) see, e.g., Ref. [32] or the Appendix. This statement holds regardless of whether or not detailed balance holds. From Eq. (25) it is plain to see that rescaling the $\{q_\alpha\}$ by some factor only results in an additive constant C to the Renyi divergences, which does not alter its "nonincrease" property (26). Since, according to Eq. (18), the equilibrium probabilities p_α^{eq} of the stochastic processes considered here are, up to a factor, given by $\Omega(\alpha)$, we may rewrite Eq. (25) as

$$D_a(P||Q) = \frac{1}{a-1} \ln \sum_{\alpha} \left[\frac{p_\alpha}{\Omega(\alpha)}\right]^a \Omega(\alpha) + C_a. \tag{27}$$

One important Renyi divergence is obtained for $a \to 1$. It is well known that in this limit the Renyi divergence converges against the Kullback-Leibler divergence [42]. Hence, we get

$$D_1(P||Q) = \sum_{\alpha} p_\alpha\,[\ln p_\alpha - \ln \Omega(\alpha)] + C_1. \tag{28}$$

From this follows directly that a function L_1 defined as

$$L_1 := \sum_{\alpha} -p_\alpha \ln p_\alpha + p_\alpha \ln \Omega(\alpha) \tag{29}$$

is strictly nondecreasing under stochastic maps as defined by Eqs. (15) and (16), i.e.,

$$L_1(t+i\tau) \geqslant L_1(t). \tag{30}$$

This motivates the interpretation of Eq. (29) as consisting of two additive parts: One is a "lack-of-information entropy" $S_{\mathrm{Sh}} := -\sum_\alpha p_\alpha \ln(p_\alpha)$, which is just the Shannon entropy measuring the uncertainty that comes with the distribution of probabilities onto the different possible measurement outcomes. The other part is a "mean-property entropy" $\overline{S}_{\mathrm{pr}} := \sum_\alpha p_\alpha \ln \Omega(\alpha)$ that measures the entropy associated with the specific property or measurement outcome. In order for the latter to apply, we eventually define the property entropy as

$$S_{\mathrm{pr}}(\alpha) := \ln \Omega(\alpha). \tag{31}$$

This entropy is the physical entropy associated with a "macrostate" or property of a system, regardless of the probability with which it may possibly occur. If, e.g., the property of the system is just expressed by the fact that the system contains an energy from a narrow energy window, then $\Omega(\alpha) = \mathrm{Tr}\{\hat{P}_E\}$, where $\hat{P}_E$ is a projector onto this energy window. Hence, up to factor k_B, in this case S_{pr} is just the standard microcanonical Boltzmann entropy. Thus, Eq. (30), which is one of the (in)equalities establishing irreversibility within this framework, implies the following statements: A process that reduces the lack of information, i.e., leads to a well-predictable property if there is a lot of uncertainty about the initial property is necessarily a process that increases the mean-property entropy $\overline{S}_{\mathrm{pr}}$. Also the reverse applies: The mean-property entropy may be reduced in a process; however, in this case the predictability of the final outcome is lost since the lack-of-information entropy must increase. The mean-property entropy $\overline{S}_{\mathrm{pr}}$ in itself is not strictly increasing. Thus, it may fluctuate along a sequence of measurements, e.g., in the sense of a fluctuation theorem.

It is worth mentioning here that Eq. (29) is only one condition on the dynamics that follows from Eq. (27), namely, for $a = 1$. Thus, the full and concrete meaning of "irreversibility" within this framework is not comprised in Eq. (29) alone but in all conditions that follow from Eq. (27) for all a. Some of them may be of importance in certain contexts, as will be discussed below. The idea of having not only one but many inequalities in order to establish irreversibility is very much in line with other recent approaches such as, e.g., Ref. [43].

VI. CONSISTENCY, MARKOVIANITY, AND TYPICALITY

Starting from the assumption that the quantum dynamics should be consistent and Markovian in the senses described in Secs. II and III, the specific equilibrium probabilities for the properties (18) have been derived. This should be compared to results from the field of equilibration in closed quantum systems, in which neither consistency nor Markovianity are taken into account. One of these results states that the overwhelming majority of states (w.r.t. the Haar measure) in some high-dimensional Hilbert space features very similar expectation values of operators with spectral widths of order unity [5]. Thus, even closed quantum systems may "apparently" equilibrate since, as long as the wave function ventures through regions in Hilbert space that are filled with states from the above majority, the expectation values of the above operators hardly show any dynamics. These findings are sometimes called typicality and have been demonstrated in various numerical examples, e.g., in Refs. [44–46]. If this applies, the above typical expectation values (to which systems in this case equilibrate) are given by $\mathrm{Tr}\{\hat{O}\}/d$ for some observable $\hat{O}$. Computing the equilibrium probabilities within the approach at hand from Eq. (18) yields

$$p_\alpha^{\mathrm{eq}} = \frac{\Omega(\alpha)}{\sum_\beta \Omega(\beta)}. \tag{32}$$

With Eqs. (10) and (1), this becomes

$$p_\alpha^{\mathrm{eq}} = \frac{\Omega(\alpha)}{d}. \tag{33}$$

The implication of this is best illustrated in the case of one-step Markovianity, where we simply have $\Omega(\alpha) = \mathrm{Tr}\{\hat{A}_n^\dagger \hat{A}_n\}$. Hence, in this case we get

$$p_n^{\mathrm{eq}} = \frac{\mathrm{Tr}\{\hat{A}_n^\dagger \hat{A}_n\}}{d}, \tag{34}$$

which is just the typical expectation value of the observable $\hat{O} = \hat{A}_n^\dagger \hat{A}_n$ corresponding to the property n. Due to Eq. (1), the spectral widths of $\hat{A}_n^\dagger \hat{A}_n$ are always of order unity since the eigenvalues are positive but bound by one. Thus, one may conclude the following: Whenever a closed quantum system is consistent and Markovian in the senses described in Secs. II and III, then, once equilibrium is reached, its wave function has to venture through the (large) part of the Hilbert space filled with typical quantum states. Thus, consistency and Markovianity imply the applicability of the typicality concept while the reverse it not necessarily true.

VII. DERIVATION OF LANDAUER'S PRINCIPLE

In this section we intend to demonstrate that Landauer's principle follows directly within the approach at hand without the necessity of invoking equilibrium environments, etc. Landauer's principle essentially states that it is impossible to erase or, more precisely, reset one bit of information without generating at least $\Delta S = \ln 2$ of physical entropy. Usually this is achieved by performing an amount $W \geqslant T \ln 2$ of work and then transferring the latter to a heat bath at temperature T [47]. Within the approach at hand, of course, the property entropy S_{pr} is the physical entropy. So the question here is: Is it possible to reset one bit without increasing $\overline{S_p}$? The information must be encoded in the properties. One bit requires the system to feature at least two measurable properties, say, $\alpha = 1, 2$. Before the reset the state of the bit is unknown, i.e., $p_1^{\mathrm{ini}},\ p_2^{\mathrm{ini}} = 1/2$. After the reset the state of the bit should be unambiguously fixed, e.g., $p_1^{\mathrm{fin}} = 1,\ p_2^{\mathrm{fin}} = 0$. In order to not generate any $\overline{S}_{\mathrm{pr}}$, we assume that all $\Omega(\alpha)$ are the same, i.e., $\Omega(\alpha) = \mathrm{const}$. Thus, in this reset process we would decrease the lack-of-information entropy by $\Delta S_{\mathrm{Sh}} = -\ln 2$. Since we do not increase the mean-property entropy $\overline{S}_{\mathrm{pr}}$, this is forbidden by Eq. (30) as Landauer's principle requires. Only if we had $\Omega(1) \geqslant 2\Omega(2)$, this was possible. Given only Eq. (30), however, another similar process, that has been first mentioned by R. Alicki and is now frequently addressed by C. Bennett, is possible. Assume that we do not require the bit to exhibit property 1 after the reset with absolute certainty but that we allow for a small failure rate ϵ, i.e., we require only $p_1^{\mathrm{fin}} = 1 - \epsilon$. The "excess probability" ϵ could then be uniformly distributed onto $X \geqslant 2^{1/\epsilon}$ different other properties. Then the process is not in conflict with Eq. (30), and it appears as if one could do the resetting up to an arbitrary confidence level $\epsilon \to 0$ without generating any physical entropy $\overline{S}_{\mathrm{pr}}$. This, however, is not the case. If one considers D_∞ rather than D_1 as done in Eq. (28), one finds

$$D_\infty(P||Q) = \ln\left[\max_\alpha \frac{p_\alpha}{\Omega(\alpha)}\right] + C_\infty. \tag{35}$$

Since all Renyi divergences are nonincreasing under stochastic maps, it is obvious that L_∞ defined as

$$L_\infty := \max_\alpha \frac{p_\alpha}{\Omega(\alpha)} \tag{36}$$

cannot increase. This, however, is in conflict with the "approximate resetting" as described above. In this approximate resetting we would have $L_\infty^{\mathrm{fin}} - L_\infty^{\mathrm{ini}} = 1/2 - \epsilon$, which is an increase, and thus it cannot occur. So, taking all the conditions that arise from Eq. (27) into account, Landauer's principle is demonstrated to follow.

VIII. NUMERICAL EXAMPLE FOR THE EMERGENCE OF CONSISTENCY AND MARKOVIANITY IN A FINITE CLOSED QUANTUM SYSTEM

So far, all results in the paper at hand have been derived by starting from the assumption of consistency and Markovianity as defined in Secs. II and III. Thus, an evident and important question is if and under what conditions on the model and the properties these assumptions are justified. Consistency is often attributed to the influence of some external systems like explicit measurement apparatuses, baths, etc. Markovianity is also questionable since finite, closed quantum systems always feature finite (quasi-) recurrence times.

While we are at present unable to state the conditions under which consistency and Markovianity emerge in general, we will give in this section a concrete numerical example for the occurrence of the latter for some specific transitions. These numerical findings will not establish consistency and Markovianity in general; moreover, they will not even demonstrate the latter for all possible transitions in the model defined below. However, the numerics will show how consistency and Markovianity approximately emerge in the limit of a high density of states (DOS) for some transitions. Thus, the below numerics are intended to demonstrate that it is at least not impossible for consistency and Markovianity to emerge approximately in a very simple and comparatively small closed quantum system. More far-reaching results in that direction are considered to be important but left for further research.

Our numerics are based on a model class which is designed to represent a very simple closed quantum system featuring exponential relaxation of some expectation values. (This model and its dynamics have already been analyzed, e.g., in Ref. [48], an almost identical version first appeared in Ref. [49].) All operators are given on the level of discrete finite matrices.

The eigenvalues of some "unperturbed Hamiltonian" $\hat{H}_0$ form two somehow distinguishable but spectrally identical "bands," both of width $\delta\epsilon$ and both with equidistant level spacing. Furthermore, there is a "perturbation" $\hat{V}$ consisting of transition operators representing transitions between the two bands. The full Hamiltonian $\hat{H} = \hat{H}_0 + \lambda\hat{V}$ reads

$$\hat{H}_0 = \sum_{i=0}^{n-1} \frac{i}{n-1}\,\delta\epsilon\,|i\rangle\langle i| + \sum_{j=0}^{n-1} \frac{j}{n-1}\,\delta\epsilon\,|j\rangle\langle j|\,,$$

$$\hat{V} = \sum_{i,j=0}^{n-1} v_{ij}\,|i\rangle\langle j| + \text{H.c.}, \tag{37}$$

$$\frac{1}{n^2}\sum_{i,j=0}^{n-1} |v_{ij}|^2 = 1,$$

where the complex v_{ij} are chosen as random Gaussian numbers, which are normalized according to the last equation of (37). The i label states in band 1, the j in band 2. Obviously, λ quantifies the overall perturbation strength. The observed property is just the system's occupation of either the first or the second band. Thus, we only have two property operators:

$$\hat{A}_1 := \sum_{i=0}^{n-1} |i\rangle\langle i|, \quad \hat{A}_2 := \sum_{j=0}^{n-1} |j\rangle\langle j|. \tag{38}$$

Obviously, both $\hat{A}$ are simply projectors. Thus, according to Eq. (5), we get $\hat{A}_0 = 0$. We are testing for one-step Markovianity here; hence, the probabilities for the properties are simply labeled $p_1 := \text{Tr}\{\hat{A}_1\,\hat{\rho}\}$, $p_2 := \text{Tr}\{\hat{A}_2\,\hat{\rho}\}$. This system has been shown before to exhibit dynamics that are in accord with the dynamics as generated by Eq. (15), even in the limit of very small τ. Thus, the dynamics may conveniently be compared to the dynamics of a master equation of the following form:

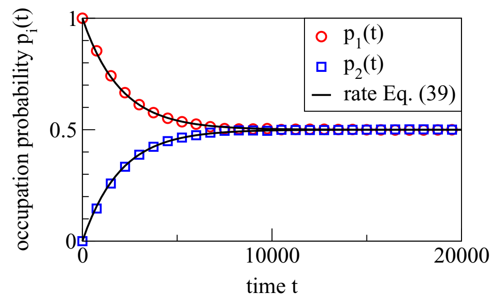

FIG. 1. (Color online) Comparison of the dynamics of the occupation probabilities $p_1(t)$ and $p_2(t)$ of the first and second band, respectively, as resulting from the Schrödinger equation with the same dynamics as resulting from the rate equation (39). Obviously there is a good agreement. For specific model parameters, see text.

$$\dot{p}_1 = -R\,(p_1 - p_2), \quad \dot{p}_2 = -R\,(p_2 - p_1). \tag{39}$$

Such a comparison is illustrated in Fig. 1. The initial state in this analysis is $\hat{\rho}(0) = \hat{A}_1/n$. Other parameters are $n = 800$, $\lambda = 5 \times 10^{-5}$, and $\delta\epsilon = 0.05$. The rate R is taken to be $R := 2\pi\lambda^2 n/\delta\epsilon = 2.5 \times 10^{-4}$, which may be inferred from Fermi's Golden Rule; cf. Ref. [48].

Obviously, there is a good agreement. Note, however, that this agreement does not proof consistency and Markovianity in itself, but it makes the model (37) a promising candidate to find the latter.

Next we check consistency directly. We choose to consider the transition $1 \to 2$ during two time steps τ. From Eq. (7) we find that, in the case of perfect consistency, one would get

$$1 - \frac{P(1\#2, \hat{1}/n)}{P(112, \hat{1}/n) + P(122, \hat{1}/n)} = 0, \tag{40}$$

which in terms of Ω reads

$$1 - \frac{\Omega(1\#2)}{\Omega(112) + \Omega(122)} = 0, \tag{41}$$

where the # character again indicates that no measurement is performed at the corresponding point in time. If consistency holds only approximately, the expression on the l.h.s. of Eq. (41) should still be small compared to unity. Hence, we call this expression "nonconsistency." Of course, this nonconsistency may depend on the specific random numbers that were used to construct the Hamiltonian. Therefore, for each set of parameters $(\lambda, n, \delta\epsilon)$ that is considered below, 10 different random Hamiltonians have been constructed and the nonconsistencies have been numerically calculated. The results are then given in terms of the mean and the standard deviation over these 10 realizations corresponding to the sets of parameters; see Fig. 2. (Means are indicated by symbols, standard deviations by the vertical bars.) These parameters have been chosen to systematically increase the DOS $(n/\delta\epsilon)$, but keeping the relaxation rate R as well as the band width $\delta\epsilon$ fixed. The fixed values are again $\delta\epsilon = 0.05$, $R = 2.5 \times 10^{-4}$. Obviously, nonconsistency is already rather low at a moderate DOS but decreases further with growing

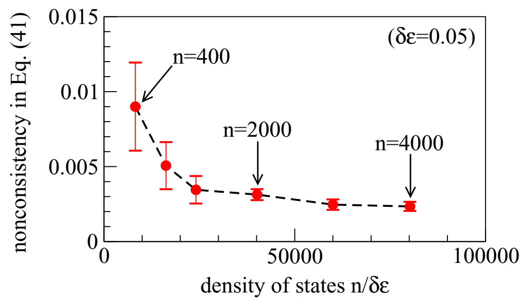

FIG. 2. (Color online) Numerical results on the measure of nonconsistency in the "two-macrostate" model (37). Central symbols indicate means, vertical bars spreadings for different random realizations of the model. Obviously, nonconsistency decreases with increasing the density of states. For specific model parameters, see text.

DOS. Thus, Fig. 2 suggests that histories may indeed become practically consistent in the limit of a large DOS.

Now we turn to a numerical testing of Markovianity. We proceed in a way similar to the one employed for the analysis of consistency above. We choose to consider the transition path ($1 \to 1 \to 2$). In the case of perfect one-step Markovianity, one would get from Eqs. (13) and (14)

$$\frac{\Omega(112)}{\Omega(11)} = \frac{\Omega(12)}{\Omega(1)}, \tag{42}$$

which may be rearranged as

$$1 - \frac{\Omega(12)\,\Omega(11)}{\Omega(112)\,\Omega(1)} = 0. \tag{43}$$

We call the l.h.s. of Eq. (43) "non-Markovianity" and calculate it still for 10 different realizations of the Hamiltonian for each given set of parameters $(\lambda, n, \delta\epsilon)$. As before, one-step Markovianity is approximately given if the non-Markovianity is small compared to one. The results are displayed in Fig. 3.

Obviously, non-Markovianity is already rather low at a moderate DOS but decreases further with growing DOS. Thus, Fig. 3 suggests that the transition path ($1 \to 1 \to 2$) may indeed become practically Markovian in the limit of a large DOS.

While all these numerical considerations can of course not establish consistency and Markovianity in general, they demonstrate exemplarily that both may emerge very well from the coherent quantum dynamics of a closed, moderately sized system. While the model addressed here is rather abstract, similar results have been found for Heisenberg spin models (comprising no random numbers) made of 16–32 spins [44,45].

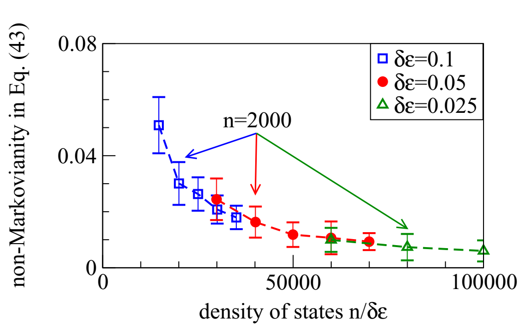

FIG. 3. (Color online) Numerical results on the measure of non-Markovianity in the "two-macrostate" model (37). Also non-Markovianity decreases with increasing the density of states. For specific model parameters, see text.

IX. SUMMARY, CONCLUSION, AND OUTLOOK

In the paper at hand we addressed the question of irreversibility in quantum mechanics. Our approach was based on the choice of a set of "properties" or (possibly but not necessarily projective) measurement operators such as POVMs and the dynamics of the system being unitary. Our point of departure was the assumption that a sequence of chosen measurement outcomes from the past is sufficient to predict the probabilities of measurement outcomes in the future. Furthermore, we assumed that the probability dynamics do not depend on whether or not the corresponding measurements are actually performed (Markovianity and consistency). Based on these assumptions, a set of equilibrium probabilities was identified, which is, in the case of simple Markovianity, independent of the specific unitary dynamics. This equilibrium set is compatible with the one suggested by quantum typicality, which establishes a connection between these different concepts. The actual set of probabilities can only approach but never depart from the equilibrium state w.r.t. the Renyi divergences. Based on this fact, a physical entropy and an entropy quantifying the lack of knowledge about the expected measurement outcome were formulated. The sum of both was rigorously shown to be nondecreasing. This implies, in plain language, that processes the results of which are well predictable from past measurements cannot occur unless these results represent states of equal or higher physical entropy. Furthermore, the validity of Landauer's principle directly follows. In order to exemplify that consistency and Markovianity may indeed emerge in closed quantum systems, corresponding numerical results were presented. We intend to do a similar but more detailed numerical investigation on a more realistic system, i.e., the Heisenberg spin model discussed in Ref. [45].

While the line of reasoning is to a large extend similar to the arguments given in previous works, e.g., Ref. [32], it is a specific feature of the work at hand that the measurement operators do not need to commute with each other, not even approximately. This may allow for a substantial weakening of the consistency condition: Since any density operator may entirely be represented by a sum of (noncommuting) positive operators, coherences may simply be included in the above set of properties. In this way, dynamics that appear as coherent or "inconsistent" w.r.t. a given set of properties may appear consistent w.r.t. an enlarged set of properties. We consider this an interesting line of research for the near future.

ACKNOWLEDGMENTS

J. Gemmer (J. G.) sincerely thanks R. Tumulka, S. Goldstein, J. Lebowitz, and E. Kotschunz for many extensive and fruitful discussions without which the work at hand could not have been done. Furthermore, J. G. is indebted to S. Wehner for bringing the Renyi entropies to his attention. We also thank J. Anders for valuable remarks.

APPENDIX

The appendix is dedicated to a derivation of Eq. (26). We give this derivation here to make the paper at hand sufficiently self-contained. It may be found in almost the same form in Ref. [32].

We start by considering the properties of functions H given as

$$H(t) := \sum_\alpha q_\alpha\, \phi(x_\alpha), \quad x_\alpha := \frac{p_\alpha(t)}{q_\alpha}. \tag{A1}$$

Here the $p_\alpha(t)$ are the actual probabilities at time t of some stochastic process, and the q_α are its invariant (equilibrium) probabilities. ϕ is any convex function. There are many definitions of convex functions; the one convenient for our purposes states that the graph of a convex function never lies below any of its tangents,

$$\phi(x) \geqslant \phi(x_0) + (x - x_0)\phi'(x_0), \tag{A2}$$

where ϕ' denotes the first derivative of ϕ. Let $\tilde{x}_\beta := p_\beta(t+\tau)/q_\beta$ be the "time-propagated" (w.r.t. the stochastic process) analog of x_α:

$$\tilde{x}_\beta = \sum_\alpha m(\beta|\alpha)\, x_\alpha, \quad m(\beta|\alpha) := w(\beta|\alpha)\,\frac{q_\alpha}{q_\beta}. \tag{A3}$$

Since $w(\beta|\alpha)$ is a stochastic map, i.e., $\sum_\beta w(\beta|\alpha) = 1$, we get from the above definition of $m(\beta|\alpha)$

$$q_\alpha = \sum_\beta q_\beta\, m(\beta|\alpha). \tag{A4}$$

Furthermore, since the map $w(\beta|\alpha)$ identically reproduces the equilibrium probabilities $\{q_\alpha\}$, we find

$$\sum_\alpha m(\beta|\alpha) = 1. \tag{A5}$$

Consider now the (negative) change of the function H during the time step τ:

$$H(t) - H(t+\tau) = \sum_\alpha q_\alpha\, \phi(x_\alpha) - \sum_\beta q_\beta\, \phi(\tilde{x}_\beta). \tag{A6}$$

To make use of the convexity of ϕ, we transform both sums on the r.h.s. into double sums. First, using Eq. (A4) and, second, using Eq. (A5), we get

$$H(t) - H(t+\tau) = \sum_{\alpha,\beta} q_\beta\, m(\beta|\alpha)\,[\phi(x_\alpha) - \phi(\tilde{x}_\beta)]. \tag{A7}$$

This may also be written as

$$\begin{aligned} &H(t) - H(t+\tau) \\ &\quad = \sum_{\alpha,\beta} q_\beta\, m(\beta|\alpha)\{\phi(x_\alpha) - [\phi(\tilde{x}_\beta) + (x_\alpha - \tilde{x}_\beta)]\phi'(\tilde{x}_\beta)\}. \end{aligned} \tag{A8}$$

The part that is added compared to Eq. (A7) vanishes by virtue of Eqs. (A4) and (A5). (For this to hold, the fact that specifically $\phi'(\tilde{x}_\beta)$ appears is irrelevant, any function of only β would do.) By comparing Eq. (A8) to Eq. (A2) and identifying x_α: x, $\tilde{x}_\beta$: x_0, it is plain to see that

$$H(t) - H(t+\tau) \geqslant 0 \tag{A9}$$

as long as ϕ is convex.

Let ϕ be a function of the type $\phi(x) = x^a$. Then for $a > 1$ and $x \geqslant 0$ the function ϕ is clearly convex. However, with this choice, the Renyi divergence (25) may be written as

$$D_a(P||Q) = \frac{1}{a-1}\ln H. \tag{A10}$$

Thus, with Eq. (A9), Eq. (26) follows. It may be convenient to consider the case of $a = 1$ separately. In this case, as mentioned in Sec. V, the Renyi divergence converges against the Kullback-Leibler divergence. Hence, in this case we may simply choose $\phi(x) = -\ln x$ (which is convex) to find

$$D_1(P||Q) = H, \tag{A11}$$

and again Eq. (26) follows.

[1] J. Uffink, in *Handbook for the Philosophy of Physics*, edited by J. Butterfield and J. Earman (Elsevier, Amsterdam, 2007), pp. 924–1074.
[2] S. Popescu, A. J. Short, and A. Winter, Nature Phys. **2**, 754 (2006).
[3] G. P. Beretta, J. Phys.: Conf. Ser. **237**, 012004 (2010).
[4] H. Pauli, *Festschrift zum 60. Geburtstage A. Sommerfelds*, edited by P. Debye (Hirzel, Leipzig, 1928).
[5] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Eur. Phys. J. H **35**, 173 (2010).
[6] E. Schrödinger, *Statistical Thermodynamics* (Cambridge University Press, Cambridge, 1948).
[7] L. Van Hove, Physica **21**, 517 (1954).
[8] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998).
[9] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[10] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Phys. Rev. Lett. **96**, 050403 (2006).
[11] J. Gemmer and G. Mahler, Eur. Phys. J. B **31**, 249 (2003).
[12] P. Reimann, Phys. Rev. Lett. **99**, 160404 (2007).
[13] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. Lett. **100**, 190403 (2008).
[14] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E **79**, 061103 (2009).
[15] A. Riera, C. Gogolin, and J. Eisert, Phys. Rev. Lett. **108**, 080402 (2012).
[16] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991).
[17] M. Srednicki, Phys. Rev. E **50**, 888 (1994).

[18] M. Rigol, V. Dunjiko, and M. Olshanii, Nature (London) **452**, 854 (2008).

[19] S. Dubey, L. Silvestri, J. Finn, S. Vinjanampathy, and K. Jacobs, Phys. Rev. E **85**, 011141 (2012).

[20] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics, Vol. 286 (Springer, Berlin, 1987).

[21] A. O. Caldeira and A. J. Leggett, Physica A **121**, 587 (1983).

[22] F. Jin, H. De Raedt, S. Yuan, M. I. Katsnelson, S. Miyashita, and K. Michielsen, J. Phys. Soc. Jpn. **79**, 124005 (2010).

[23] J. Gemmer, A. Otte, and G. Mahler, Phys. Rev. Lett **86**, 1927 (2001).

[24] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).

[25] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).

[26] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).

[27] C. Jarzynski and D. K. Wojcik, Phys. Rev. Lett. **92**, 230602 (2004).

[28] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[29] M. Esposito and S. Mukamel, Phys. Rev. E **73**, 046129 (2006).

[30] M. Esposito, K. Lindenberg, and C. Van den Broeck, New J. Phys. **12**, 013013 (2010).

[31] M. Esposito and P. Gaspard, Phys. Rev. E **76**, 041134 (2007).

[32] O. Penrose, *Foundations of Statistical Mechanics: A Deductive Treatment* (Pergamon Press, Oxford, 1970).

[33] D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H. D. Zeh, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin, 1996).

[34] R. Griffiths, in *Physical Origins of Time Asymmetry*, edited by J. J. Halliwell, J. P. Perez Mercader, and W. H. Zurek (Cambridge University Press, Cambridge, 1994).

[35] H. F. Dowker and J. J. Halliwell, Phys. Rev. D **46**, 1580 (1992).

[36] R. Omnès, *Understanding Quantum Mechanics* (Princeton University Press, Princeton, 1999).

[37] E. D. Chisolm, J. Phys.: Conf. Ser. **237**, 012004 (2009).

[38] K. Kraus, *States, Effects, and Operations*, Lecture Notes in Physics, Vol. 190 (Springer, Berlin, 1983).

[39] R. Griffiths, J. Stat. Phys. **36**, 219 (1984).

[40] E. P. Wigner, Am. J. Phys. **31**, 6 (1963).

[41] E.-M. Laine, J. Piilo, and H.-P. Breuer, Phys. Rev. A **81**, 062115 (2010).

[42] T. van Erven and P. Harremoës, arXiv:1206.2459.

[43] F. G. S. L. Brandao, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, arXiv:1305.5278.

[44] H. Niemeyer, D. Schmidtke, and J. Gemmer, Europhys. Lett. **101**, 10010 (2013).

[45] H. Niemeyer, K. Michielsen, H. De Raedt, and J. Gemmer, Phys. Rev. E **89**, 012131 (2014).

[46] P. Borowski, J. Gemmer, and G. Mahler, Eur. Phys. J. B **35**, 255 (2003).

[47] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).

[48] C. Bartsch, R. Steinigeweg, and J. Gemmer, Phys. Rev. E **77**, 011119 (2008).

[49] M. Esposito and P. Gaspard, Phys. Rev. E **68**, 066113 (2003).