

## Comparing classical and quantum equilibration

Artur S. L. Malabarba,<sup>1</sup> Terry Farrelly,<sup>2</sup> and Anthony J. Short<sup>1</sup>

<sup>1</sup>*H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*

<sup>2</sup>*Institut für Theoretische Physik, Leibniz Universität, Appelstraße 2, 30167 Hannover, Germany*

(Received 19 May 2016; published 15 September 2016)

By using a physically relevant and theory independent definition of measurement-based equilibration, we show quantitatively that equilibration is easier for quantum systems than for classical systems, in the situation where the initial state of the system is completely known (a pure state). This shows that quantum equilibration is a fundamental aspect of many quantum systems, while classical equilibration relies on experimental ignorance. When the state is not completely known (a mixed state), this framework also shows that quantum equilibration requires weaker conditions.

DOI: [10.1103/PhysRevE.94.032119](https://doi.org/10.1103/PhysRevE.94.032119)

### I. INTRODUCTION

For over a century, physicists have sought to understand the emergence of apparently irreversible equilibration from reversible microscopic dynamics. Results over the past few years have shown equilibration for general quantum systems under very weak assumptions [1–6]. In classical mechanics, results on equilibration rely on assumptions such as ergodicity, mixing dynamics, mixed initial states, and coarse graining of the measurements [7–10].

Recently, Reimann and Evstigneev [7] analyzed equilibration in classical and quantum theory by evaluating observable expectation values and comparing them to a parameter representing experimental precision. Then, considering mixed initial states, they were able to compare the conditions necessary for equilibration under each case, showing that they require a very different set of conditions.

Our work complements and extends the work in [7], using a theory independent definition of equilibration to compare the two scenarios, for both pure and mixed initial states. Following [3,4,6], our definition employs a generalized distinguishability that incorporates the full outcome distribution of a measurement, corresponding to its ability to help us distinguish a time-evolving state from a time-invariant equilibrium state.

On the classical side, we show that pure systems equilibrate only when the measurement coarse-graining partitions the state space in a very specific (and very uneven) way. On the other hand, quantum equilibration needs only that the measurement be coarse grained at all, with no restriction on how it partitions the state space. Then, using a parameter to characterize measurements on both theories, we are able to show that the range of values that permit classical equilibration is very close to the range that guarantees equilibration in any theory. Meanwhile, quantum equilibration is possible for a much wider range of this parameter. Furthermore, when the initial state of the system is taken to be mixed, our approach corroborates the previous results by Reimann and Evstigneev [7].

Note that the definitions in this paper are kept as general as possible. For example, our definition of classical mechanics (Sec. III A) covers both Newtonian and relativistic classical mechanics. However, our classical and quantum equilibration results concern only reversible dynamics. Hence those results do not apply to irreversible dynamics such as those generated by stochastic maps.

### II. DEFINITIONS

Although we will specifically consider quantum and classical dynamics, we start our discussion by talking about equilibration and distinguishability strictly in terms of measurements and outcome probabilities, which could be applied to almost any theory. As such, we need to restate some of our definitions in general terms, without relying on quantum observables or density matrices.

We consider the possible states of a system to be elements in a compact convex space  $\mathcal{S}$  (herein, a state space). This includes classical mechanics in a bounded phase space (for example, particles in a finite volume with a velocity cutoff) and finite-dimensional quantum theory. Convexity of the state space corresponds to our ability to prepare probabilistic mixtures of states and compactness rules out many situations in which equilibration does not occur (such as a free particle moving on an infinite line). When we refer to functions on  $\mathcal{S}$  being linear, this is with respect to convex mixtures in this space.

We compare different states using only the information provided by measurements, as defined below.

*Definition 1 (measurement).* Given a state space  $\mathcal{S}$  and  $N \in \mathbb{N}$ , a measurement with  $N$  outcomes is any set of  $N$  linear functions  $\mathcal{M} = \{p_j | j = 1, \dots, N\}$ , where

$$p_j : \mathcal{S} \rightarrow [0, 1],$$

$$\rho \mapsto p_j(\rho)$$

satisfy  $\sum_{j=1}^N p_j(\rho) = 1 \forall \rho \in \mathcal{S}$ . Each of these functions represents the probability of obtaining the  $j$ th outcome when measuring  $\mathcal{M}$  on  $\rho$ .

The maximum information that can be gathered about a state, using measurements, is the probability of each particular outcome.

*Definition 2 (distinguishability).* Given two possible states of a system,  $\rho$  and  $\sigma \in \mathcal{S}$ , and a measurement  $\mathcal{M}$  with  $N$  outcomes, the distinguishability between  $\rho$  and  $\sigma$  according to  $\mathcal{M}$  is

$$D_{\mathcal{M}}(\rho, \sigma) = \frac{1}{2} \sum_{j=1}^N |p_j(\rho) - p_j(\sigma)|, \quad (1)$$

where  $D_{\mathcal{M}} : \mathcal{S} \times \mathcal{S} \rightarrow [0, 1]$ .

The distinguishability is defined this way so that, after performing the measurement, the distinguishability quantifies the average probability of successfully guessing which state the system was in [3], according to

$$p_{\text{success}} = \frac{1}{2} + \frac{1}{2} D_{\mathcal{M}}(\sigma, \rho). \quad (2)$$

When  $D_{\mathcal{M}}(\sigma, \rho) = 0$  the measurement does not provide information that helps to distinguish  $\sigma$  from  $\rho$ . On the other hand, when  $D_{\mathcal{M}}(\sigma, \rho) = 1$  the states are perfectly discriminated by this measurement. This provides a physical and practical meaning to the distinguishability, i.e., if  $D_{\mathcal{M}}(\sigma, \rho) \approx 0$  then  $\sigma$  and  $\rho$  are experimentally indistinguishable.

In order to talk about equilibration, we also need a notion of time evolution. We describe it via a function on  $\mathcal{S} \times \mathbb{R}_{\geq 0}$ , which is linear on  $\mathcal{S}$  and whose value represents the state  $\rho$  evolved by some time  $t \geq 0$ ,

$$T : \mathcal{S} \times \mathbb{R}_{\geq 0} \rightarrow \mathcal{S}, \\ (\rho, t) \mapsto T_t(\rho),$$

where  $T_0(\rho) = \rho$ . For short, we will abbreviate  $T_t(\rho)$  as  $\rho_t$ . Note that, in general,  $T$  does not need to be time independent or reversible. However, we will be particularly interested in reversible evolutions, for which there exists a function  $T^{-1}$  (defined in the same way as  $T$ ) such that  $T_t^{-1}(T_t(\rho)) = T_t(T_t^{-1}(\rho)) = \rho$ . We will use the assumption of reversibility when we talk about classical mechanics, but general time evolution is enough to define equilibration and even prove our first theorem.

**Definition 3 (equilibration).** Given a state  $\rho$ , a time evolution  $T$ , and  $0 \leq \epsilon < 1$ , we say that  $\rho$  equilibrates up to  $\epsilon$  (or  $\epsilon$ -equilibrates) under this evolution with respect to a measurement  $\mathcal{M}$  when both averages  $\omega = \langle \rho_t \rangle$  and  $\langle D_{\mathcal{M}}(\rho_t, \omega) \rangle$  exist and satisfy

$$\langle D_{\mathcal{M}}(\rho_t, \omega) \rangle \leq \epsilon. \quad (3)$$

Here  $\langle \cdot \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt$  denotes the time average.

This definition applies regardless of the dynamics that govern the evolution of  $\rho_t$ , i.e., it is theory independent, so we must account for the possibility of the averages not existing. Fortunately, in the particular cases of interest to us the dynamics guarantee the convergence and existence of the averages. Since we only consider compact state spaces here, in quantum mechanics the time average equals a decoherence in a finite-dimensional energy basis and in classical mechanics the averages converge by Birkhoff's theorem [8]. Note that in what follows we will not discuss the time scale for equilibration, which may be very long [4–6].

When  $\omega = \langle \rho_t \rangle$  does exist, we call it the equilibrium state. Since the probabilities are linear functions, the probabilities on  $\omega$  can be written as  $p_j(\omega) = \langle p_j(\rho_t) \rangle$ .<sup>1</sup> In addition, we have that  $\langle \rho_t \rangle = \langle \rho_{t+\tau} \rangle$  for all  $\tau \in \mathbb{R}$  and so the equilibration of  $\rho$  also implies the equilibration of  $\rho_t$  for all  $t$ .

Furthermore, note how it is only required that  $\rho$  be close to  $\omega$  for most times and not for all large-enough

times. This condition is much weaker and more physically meaningful, first because it does not preclude the possibility of recurrence in the time evolution (which is possible in Hamiltonian mechanics) and second because it has been shown experimentally that systems do fluctuate away from equilibrium [11,12]. This approach is well established in the field of quantum equilibration [1,3–6,13–15] and is also being used to describe classical equilibration [7,16].

In particular, Werndl and Frigg [16] define a macroscopic state (a region of the state space  $\mathcal{S}$ ) as an  $\alpha$ - $\delta$ -equilibrium state if the fraction of time spent inside it is greater than or equal to  $\alpha$  for all states in  $\mathcal{S}$  except for a fraction  $\delta$  of them.<sup>2</sup> This definition is more oriented towards the macroscopic aspect, a top-down approach that defines an equilibrium macrostate in terms how much time the microstates spend inside it. Meanwhile we define equilibration in terms of the initial microstate and the measurement probabilities, a bottom-up approach. Still, the two are related in their definition of equilibration in terms of most times.

Finally, it is also useful to define the following.

**Definition 4 (pure and mixed states).** A state  $\psi \in \mathcal{S}$  is pure if and only if it cannot be written as a convex sum of other states in  $\mathcal{S}$  (i.e., the pure states are the extreme points of  $\mathcal{S}$ ). We denote by  $\mathcal{P}$  the set of all pure states in  $\mathcal{S}$ . A state is mixed if it is not pure.

We now show that reversible time evolutions must preserve purity, i.e.,  $\psi \in \mathcal{P}$  implies  $\psi_t \in \mathcal{P} \forall t$ . This is because otherwise one could find a  $\psi \in \mathcal{P}$  such that

$$T_t(\psi) = q\rho + (1-q)\rho', \quad (4)$$

where  $\rho, \rho' \in \mathcal{S}$ , with  $\rho \neq \rho'$ , and  $0 < q < 1$ . Applying  $T^{-1}$  to both sides of this equation, we find

$$\psi = qT_t^{-1}(\rho) + (1-q)T_t^{-1}(\rho'), \quad (5)$$

which means that  $T_t^{-1}(\rho) = \psi$  and  $T_t^{-1}(\rho') = \psi$ . However, this would imply that  $\rho = T_t(\psi) = \rho'$ , which contradicts the assumption that  $\rho \neq \rho'$ .

### III. GENERAL EQUILIBRATION

Here we show that a very uneven measurement coarse graining (with respect to the state space explored by the system) is a sufficient condition for equilibration in any theory where the average  $\langle \rho_t \rangle$  exists. As explained above, this includes quantum and classical mechanics.

**Theorem 1 (sufficiency).** Take any  $\rho \in \mathcal{S}$ , any time evolution  $\rho_t$  such that  $\omega = \langle \rho_t \rangle$  and  $\langle D_{\mathcal{M}}(\rho_t, \omega) \rangle$  exist, and any measurement  $\mathcal{M}$ . Here  $\rho_t$   $\epsilon$ -equilibrates under  $\mathcal{M}$  if

$$\max_j p_j(\omega) \geq 1 - \frac{\epsilon}{2}. \quad (6)$$

*Proof.* First, without loss of generality, we label as 1 the outcome that satisfies Eq. (6), hence  $p_1(\omega) \geq 1 - \frac{\epsilon}{2}$ . Then it follows that  $p_1(\rho_t) - p_1(\omega) \leq 1 - p_1(\omega) \leq \frac{\epsilon}{2}$ . Additionally, we define the positive part of a function  $f(t)$

<sup>1</sup>Note that, strictly speaking, we could replace the condition that  $\omega$  exists with the slightly weaker condition that  $\langle p_j(\rho_t) \rangle$  exist for all  $j$ , but we use  $\omega$  for convenience and to link with previous literature.

<sup>2</sup>They call it  $\alpha$ - $\epsilon$ -equilibrium, but we have replaced  $\epsilon$  with  $\delta$  here to avoid confusion with our own  $\epsilon$ , which actually corresponds to their  $\alpha$ .

as  $f(t)^+ = \frac{1}{2}[|f(t)| + f(t)]$ . Rearranging this equation, we obtain  $|f(t)| = 2f(t)^+ - f(t)$ . This leads to

$$\begin{aligned} & \langle |p_1(\rho_t) - p_1(\omega)| \rangle \\ &= \langle 2[p_1(\rho_t) - p_1(\omega)]^+ - [p_1(\rho_t) - p_1(\omega)] \rangle \\ &= 2\langle [p_1(\rho_t) - p_1(\omega)]^+ \rangle, \\ &\leq 2\langle [1 - p_1(\omega)]^+ \rangle \\ &= 2\langle [1 - p_1(\omega)] \rangle, \\ &\leq \epsilon, \end{aligned} \quad (7)$$

where in the second line we have used  $\langle p_1(\rho_t) \rangle = p_1(\omega)$ . In the third line we have used  $p_1(\rho_t) \leq 1$  and the fact that  $f(t) \leq g(t) \Rightarrow f(t)^+ \leq g(t)^+$ . We also have

$$\sum_{j=2}^N p_j(\omega) = 1 - p_1(\omega) \leq \frac{\epsilon}{2}, \quad (8)$$

which in turn leads to

$$\begin{aligned} \langle D_{\mathcal{M}}(\rho_t, \omega) \rangle &= \frac{1}{2} \langle |p_1(\rho_t) - p_1(\omega)| \rangle \\ &+ \frac{1}{2} \sum_{j=2}^N \langle |p_j(\rho_t) - p_j(\omega)| \rangle \\ &\leq \frac{\epsilon}{2} + \frac{1}{2} \sum_{j=2}^N (p_j(\rho_t) + p_j(\omega)) \\ &= \frac{\epsilon}{2} + \sum_{j=2}^N p_j(\omega) \leq \epsilon. \end{aligned} \quad (9)$$

This result says that one will always observe equilibration if the measurement being used is bad enough, i.e., if one of the outcomes is predominantly more likely than all the others most of the time. We will herein refer to these as highly uneven measurements.

It is important to understand that this applies to any state, pure or mixed, of any theory that matches the definitions above, be it quantum, classical, or otherwise. In other words, if a measurement satisfies this assumption for a given state and a given time evolution, it is so bad at distinguishing the time evolving state from the equilibrium state that one is guaranteed to have equilibration regardless of any other properties of the theory. Below we take a similar approach to study what conditions are *necessary* for equilibration under each theory.

### A. Classical equilibration

In both classical and quantum mechanics, mixed states represent a lack of knowledge regarding the parameters of the system. Thus, we start by studying the case where the initial state is pure, so any subsequent equilibration is strictly a consequence of the theory and not of previous ignorance.

Below, we show a necessary condition for classical equilibration of pure states that is very similar to the sufficient condition above, which means that classical pure states only equilibrate when the measurement is very bad in a very specific way. In contrast, for quantum mechanics, we provide

a counterexample showing that the same condition is not necessary.

In order to define classical dynamics for our purposes, we only need three of its properties. The first property is that time evolution is reversible. The second defining characteristic is that for pure states, at any point in time, the outcome of any measurement is completely determined. The third is that time averages exist.

*Definition 5 (classical mechanics).* A given combination of state space  $\mathcal{S}$ ,  $N$ -outcome measurement  $\mathcal{M}$ , and time evolution  $T$  obeys classical mechanics only if  $T$  is reversible, the averages  $\omega = \langle \rho_t \rangle$  and  $\langle D_{\mathcal{M}}(\rho_t, \omega) \rangle$  exist for any initial state, and  $\exists \xi : \mathcal{P} \rightarrow \{1, 2, \dots, N\}$  such that

$$p_j(\psi) = \delta_{j, \xi(\psi)} \forall \psi \in \mathcal{P}, \quad (10)$$

with  $j = 1, \dots, N$  and  $p_j \in \mathcal{M}$ .

Consequently, one has  $p_j(\langle \psi_t \rangle) = \langle \delta_{j, \xi(\psi_t)} \rangle$ . Of course, binary measurement probabilities are not all that defines classical mechanics; there are many properties (especially of the time evolution) that are not being taken into account here. However, since the theorem below is a necessity theorem showing how hard equilibration is, adding further constraints to our definitions cannot make equilibration any easier.

#### 1. Classical equilibration of pure states

The following theorem then says that a classical pure state will only equilibrate with respect to  $\mathcal{M}$  if the evolving state spends most of its time inside the support of a single outcome of  $\mathcal{M}$ .

*Theorem 2 (classical necessity).* A classical pure state  $\psi$  may  $\epsilon$ -equilibrate with respect to  $\mathcal{M}$  only if

$$\max_j \langle p_j(\psi_t) \rangle = \max_j p_j(\omega) \geq 1 - \epsilon, \quad (11)$$

where  $\omega = \langle \psi_t \rangle$ .

*Proof.* First, recall that  $\forall p_j \in \mathcal{M}$ ,

$$p_j(\psi_t) = \begin{cases} 1, & j = \xi(\psi_t) \\ 0, & j \neq \xi(\psi_t). \end{cases} \quad (12)$$

Then it is straightforward to see that  $|p_j(\psi_t) - p_j(\omega)|$  must be either  $p_j(\omega)$  or  $1 - p_j(\omega)$ , depending on the value of  $j$ . That can be written as

$$|p_j(\psi_t) - p_j(\omega)| = [1 - p_j(\omega)]\delta_{j, \xi(\psi_t)} + p_j(\omega)[1 - \delta_{j, \xi(\psi_t)}]. \quad (13)$$

Then, since  $\langle \delta_{j, \xi(\psi_t)} \rangle = \langle p_j(\psi_t) \rangle = p_j(\omega)$ , one has that

$$\langle |p_j(\psi_t) - p_j(\omega)| \rangle = 2p_j(\omega)[1 - p_j(\omega)]. \quad (14)$$

This implies that

$$\begin{aligned} \langle D_{\mathcal{M}}(\psi_t, \omega) \rangle &= \sum_{j=1}^N p_j(\omega)[1 - p_j(\omega)] \\ &= 1 - \sum_{j=1}^N p_j(\omega)^2. \end{aligned} \quad (15)$$

The  $\epsilon$ -equilibration condition is then written as

$$1 - \sum_{j=1}^N p_j(\omega)^2 < \epsilon$$

$$\Rightarrow 1 - \epsilon < \sum_{j=1}^N p_j(\omega)^2 < \max_j p_j(\omega). \quad (16)$$

■

Note how similar the inequality in Eq. (11) is to Eq. (6). The sufficiency theorem, Theorem 1, shows how uneven a measurement needs to be so that even the most stubborn of systems must equilibrate under it. The necessity theorem, Theorem 2, shows that any classical measurement that allows pure states to  $\epsilon$ -equilibrate is at most an  $\epsilon/2$  away from being one of these exceptionally uneven measurements. That is, pure states in classical mechanics are among the hardest of all systems to equilibrate.

## 2. Classical equilibration of mixed states

In the case of Hamiltonian dynamics, the theorem below is a statement that equilibration will be achieved when (i) there is a chaotic subspace of  $\mathcal{P}$  and (ii) the initial state can be represented as a mixture of states mostly within this chaotic subspace. This theorem is an extension of the results in [7], applied to the distinguishability as defined in Eq. (1).

*Theorem 3.* A classical mixed state  $\rho$  will  $\epsilon$ -equilibrate with respect to  $\mathcal{M}$  if there exists a  $\mathcal{P}_c \subseteq \mathcal{P}$  such that (a) almost all pairs of initial states in  $\mathcal{P}_c$  yield measurement probabilities that are uncorrelated over time, that is,

$$\langle p_j(\psi_t) p_j(\psi'_t) \rangle = p_j(\langle \psi_t \rangle) p_j(\langle \psi'_t \rangle) \quad (17)$$

for any  $p_j \in \mathcal{M}$  and almost all (i.e., except for a set of measure zero) pairs of initial states  $(\psi, \psi') \in \mathcal{P}_c \times \mathcal{P}_c$ , and (b) the state  $\rho$  can be described as a mixture of pure states mostly contained in  $\mathcal{P}_c$ , i.e.,

$$\rho = \int_{\mathcal{P}} \mu(\psi) \psi \, d\psi, \quad (18)$$

where  $\mu(\psi)$  is a non-negative function satisfying  $\int \mu(\psi) d\psi = 1$ ,<sup>3</sup> such that

$$\int_{\mathcal{P}_c} \mu(\psi) d\psi \geq 1 - \delta, \quad (19)$$

where  $\delta = 2\epsilon^2/N \leq 1$ .

Note that  $\psi$  and  $\psi'$  can be very close. The reason  $\mathcal{P}_c$  is called the chaotic subspace is that after a long enough time, even an infinitesimal difference between these two states must become large enough to make their probabilities uncorrelated.

The following proof is similar to calculations by Reimann and Evstigneev [7], with the difference that we use the distinguishability instead of measurement expectation values.

<sup>3</sup>Note that by taking  $\mu(\psi)$  to be a function, we exclude probability distributions involving  $\delta$  functions, which could yield different results because  $\mathcal{I}_j^{\mathcal{P}_c \times \mathcal{P}_c}$  could be nonzero in that case.

*Proof.* First, we note that, for any  $\rho$  (mixed or pure) and any time evolution,

$$\begin{aligned} \langle D_{\mathcal{M}}(\omega, \rho_t) \rangle &= \frac{1}{2} \sum_{j=1}^N \langle |p_j(\rho_t) - p_j(\omega)| \rangle \\ &= \frac{1}{2} \sum_{j=1}^N \langle \sqrt{[p_j(\rho_t) - p_j(\omega)]^2} \rangle \\ &\leq \frac{1}{2} \sum_{j=1}^N \sqrt{\langle [p_j(\rho_t) - p_j(\omega)]^2 \rangle} \\ &\leq \frac{1}{2} \sqrt{N \sum_{j=1}^N \langle [p_j(\rho_t) - p_j(\omega)]^2 \rangle} \\ &= \frac{1}{2} \sqrt{N \sum_{j=1}^N \langle p_j(\rho_t)^2 \rangle - p_j(\omega)^2}. \end{aligned} \quad (20)$$

Now note that from (18),

$$p_j(\rho_t) = \int_{\mathcal{P}} p_j(\psi_t) \mu(\psi) d\psi, \quad (21)$$

$$p_j(\omega) = \int_{\mathcal{P}} p_j(\langle \psi_t \rangle) \mu(\psi) d\psi. \quad (22)$$

Defining

$$\begin{aligned} \mathcal{I}_j^R &= \iint_R [(p_j(\psi_t) p_j(\psi'_t)) - p_j(\langle \psi_t \rangle) p_j(\langle \psi'_t \rangle)] \\ &\quad \times \mu(\psi) \mu(\psi') d\psi d\psi', \end{aligned} \quad (23)$$

we obtain

$$\begin{aligned} &\sum_{j=1}^N \langle p_j(\rho_t)^2 \rangle - p_j(\omega)^2 \\ &= \sum_{j=1}^N \mathcal{I}_j^{\mathcal{P} \times \mathcal{P}} = \sum_{j=1}^N (\mathcal{I}_j^{\mathcal{P}_c \times \mathcal{P}_c} + \mathcal{I}_j^{(\mathcal{P} \times \mathcal{P}) \setminus (\mathcal{P}_c \times \mathcal{P}_c)}), \end{aligned} \quad (24)$$

where we have used the fact that  $\int_{R_1} + \int_{R_2} = \int_{R_1 \cup R_2}$  for any two disjoint regions  $R_1$  and  $R_2$ .

From Eq. (17) it is clear that

$$\sum_{j=1}^N \mathcal{I}_j^{\mathcal{P}_c \times \mathcal{P}_c} = 0. \quad (25)$$

From Eq. (19), combined with the inequality

$$\begin{aligned} &\sum_{j=1}^N \langle [p_j(\psi_t) p_j(\psi'_t)] - p_j(\langle \psi_t \rangle) p_j(\langle \psi'_t \rangle) \rangle \\ &\leq \sum_{j=1}^N \langle p_j(\psi_t) p_j(\psi'_t) \rangle \leq \sum_{j=1}^N \langle p_j(\psi_t) \rangle \leq 1, \end{aligned} \quad (26)$$

we have

$$\begin{aligned}
 \sum_{j=1}^N \mathcal{I}_j^{(\mathcal{P} \times \mathcal{P}) \setminus (\mathcal{P}_c \times \mathcal{P}_c)} &\leq \iint_{(\mathcal{P} \times \mathcal{P}) \setminus (\mathcal{P}_c \times \mathcal{P}_c)} \mu(\psi) \mu(\psi') d\psi d\psi' \\
 &\leq 1 - \iint_{\mathcal{P}_c \times \mathcal{P}_c} \mu(\psi) \mu(\psi') d\psi d\psi' \\
 &\leq 1 - (1 - \delta)^2 \\
 &\leq 2\delta - \delta^2 \\
 &\leq 2\delta,
 \end{aligned} \tag{27}$$

where the third line follows from Eq. (19). Substituting (25) and (27) back into (24), we obtain

$$\sum_{j=1}^N \langle p_j(\rho_t)^2 \rangle - p_j(\omega)^2 \leq 2\delta. \tag{28}$$

Finally, it follows from (20) that

$$\langle D_{\mathcal{M}}(\omega, \rho_t^\mu) \rangle \leq \sqrt{\frac{N\delta}{2}} \leq \epsilon. \tag{29}$$

■

### B. Quantum equilibration

Quantum states are represented by density matrices acting on a Hilbert space. The measurement  $\mathcal{M}$  is defined in terms of a set of positive operators  $\mathcal{O}_j$  acting on the same space, each corresponding to an outcome of the observable and satisfying  $\sum_{j=1}^N \mathcal{O}_j = 1$ . The probabilities are then defined by  $p_j(\rho) = \text{Tr}[\mathcal{O}_j \rho]$ .

In the quantum case, the condition of Theorem 1 is analogous to saying that one of the operators  $\mathcal{O}_j$  is close to the identity on almost all of the accessible state space (i.e., over the energy states that occur in  $\rho$ ). However, while this condition is sufficient, it is not at all necessary and even relatively fine-grained measurements will lead to equilibration.

As such, the following theorem is not a necessity statement like Theorem 2 in the classical case. Instead, we provide an alternative sufficiency theorem that applies to measurements not encompassed by Theorem 1 [those for which all  $p_j(\omega)$  are very small].

Remarkably, as has been well investigated [1,3–6,13,14,17,18], we can obtain the same bounds on quantum equilibration for both pure and mixed states. Here we provide an improved version of a bound known from the literature.

*Theorem 4.* The average distinguishability between a general quantum state  $\rho \in \mathcal{S}$  evolving via a static Hamiltonian and its time average with respect to an  $N$ -outcome measurement  $\mathcal{M}$  satisfies

$$\langle D_{\mathcal{M}}(\rho_t, \omega) \rangle \leq \frac{1}{2} \sqrt{D_G \frac{N-1}{d_{\text{eff}}}}, \tag{30}$$

where  $D_G$  is the maximum degeneracy among gaps in the system's spectrum and  $d_{\text{eff}}^{-1} = \sum_n \text{Tr}[\rho \Pi_n]^2$  is the effective dimension, with  $\Pi_n$  being the projector onto the  $n$ th energy eigenspace.

The proof of this theorem is a fairly straightforward improvement on a previous proof from the literature [4], so

we present it in the Appendix. Then it is easy to apply this result to the current definition of  $\epsilon$ -equilibration.

*Corollary 1.* Under the conditions of Sec. III B, a general quantum state  $\rho$  will  $\epsilon$ -equilibrate with respect to an  $N$ -outcome measurement  $\mathcal{M}$  if

$$N \leq 4 \frac{d_{\text{eff}}}{D_G} \epsilon^2 + 1, \tag{31}$$

where  $D_G$  and  $d_{\text{eff}}$  are defined as in Theorem 4.

The effective dimension  $d_{\text{eff}}$  is a recurring parameter in the field of quantum equilibration. It roughly quantifies how many energy eigenstates a system occupies with significant probability and is usually assumed to be very large.

Considering the effective dimension typically scales exponentially in the number of particles, it easily outgrows the number of outcomes of any conceivable physical measurement. In this case the system will equilibrate no matter how the outcomes partition the Hilbert space.

Here we see that the quantum conditions of a system with a small number of degenerate energy gaps and a state with high effective dimension achieve a similar effect as the classical condition of chaoticity of the initial state. In fact, apart from some small constants, the fraction  $\delta$  of the classical distribution that lies outside the chaotic subspace is under the same restriction as  $D_G/d_{\text{eff}}$ , which suggests these two quantities are analogs of each other.

### C. Multiple measurements

Instead of defining equilibration in terms of the distinguishability with respect to a single measurement, one can also take a set of  $K$  measurements and consider the maximum distinguishability among all of them at any point in time. This approach [4] describes the scenario where one is capable of performing different measurements on the system and always knows which one would be best at distinguishing the state at a given time from the equilibrium state (or has multiple copies of the system and performs all measurements).

In the language of this paper, we consider a set of  $K$  measurements  $\mathbb{M} = \{\mathcal{M}_i | i = 1, \dots, K\}$  and define the distinguishability with respect to this set of measurements as

$$D_{\mathbb{M}}(\rho_t, \omega) = \max_{\mathcal{M} \in \mathbb{M}} D_{\mathcal{M}}(\rho_t, \omega). \tag{32}$$

Given this change in the definition, it is useful to consider how each of the above theorems would have to be adapted to account for multiple measurements.

Theorems 1 and 3 and Corollary 1 still hold if we replace  $\epsilon$  with  $\epsilon/K$ , where their respective assumptions must hold for all  $\mathcal{M} \in \mathbb{M}$ . Then we have that

$$\begin{aligned}
 \langle D_{\mathbb{M}}(\rho_t, \omega) \rangle &= \left\langle \max_{\mathcal{M} \in \mathbb{M}} D_{\mathcal{M}}(\rho_t, \omega) \right\rangle \\
 &\leq \sum_{\mathcal{M} \in \mathbb{M}} \langle D_{\mathcal{M}}(\rho_t, \omega) \rangle \leq K \frac{\epsilon}{K}.
 \end{aligned} \tag{33}$$

Theorem 2 still holds if all  $\mathcal{M} \in \mathbb{M}$  satisfy Eq. (11), but the equation itself need not be changed.

#### IV. CONCLUSION

Here we have used the distinguishability to define an operational, theory-independent, notion of equilibration with respect to a given measurement. Although we have applied this to quantum and classical theory, it could be applied in other cases such as general probabilistic theories [19].

We first show a sufficient condition for equilibration in any theory, which depends on the largest average probability among the measurement outcomes. We find that a value of  $1 - \frac{\epsilon}{2}$  for this parameter is a sufficient condition for  $\epsilon$ -equilibration for any theory that fits our definitions, which simply means one always observes equilibration if the measurement being used is uneven enough. In order to achieve equilibration of pure states under classical Hamiltonian dynamics (simply by virtue of how classical measurements are defined), it is necessary that this parameter be at least  $1 - \epsilon$ , showing that classical mechanics is very similar to the worst possible case.

In contrast, we have also shown an alternative sufficient condition for quantum equilibration, which shows that it can happen even when this parameter is very close to 0. This quantitatively shows that equilibration of pure states is easier under quantum dynamics, at least where it pertains to measurement ignorance. Indeed, the quantum case seems closer to the classical mixed state case, where our results corroborate other recent results [7].

While it is difficult, if not impossible, to experimentally prepare large systems in pure states, these pure-state results are very important because they show that equilibration of microscopic systems is a fundamental aspect of physics, not just a consequence of ignorance. In particular, the results imply that it is much harder to keep quantum systems out of equilibrium than one might think, even for something as small as a nanoscale system. After all, even if one takes every possible action to reliably prepare it in a pure state and reliably isolate it from the environment, the measurement used to examine it still needs to have a number of outcomes exponential in the number of particles in the system, otherwise equilibration will be inevitable (assuming there are not very many degenerate energy gaps).

#### ACKNOWLEDGMENTS

A.J.S. acknowledges support from the Royal Society and FQXi through SVCF. A.S.L.M. acknowledges support from the CNPq. T.F. is grateful for support from the ERC grants QFTCMPS and SIQS and by the cluster of excellence EXC201 Quantum Engineering and Space-Time Research.

#### APPENDIX: THEOREM 4

For clarity, we start by proving a lemma for pure states and then use that to prove Theorem 4 for general states. As mentioned in the main text, this result is an improvement over previous bounds in the literature by a factor of  $\sqrt{N}$ . The step responsible for this improvement is Eq. (A6).

*Lemma 1.* Given a finite-dimensional Hilbert space  $\mathcal{H}$ , a projector  $P$ , and a pure initial state  $\rho : \mathcal{H} \rightarrow \mathcal{H}$  evolving under a Hamiltonian with energy levels  $E_n$ , then there exists an

energy basis  $\{|n\rangle\}$  such that

$$\langle \text{Tr}[P(\rho_t - \omega)]^2 \rangle \leq \sum_{n \neq j} \sum_{k \neq l} v_{nj} v_{kl}^* \delta_{G_{nj}, G_{kl}}, \quad (\text{A1})$$

where  $G_{nj} = E_n - E_j$ ,  $\omega = \langle \rho_t \rangle$ , and  $v_{nj} = \rho_{nj} P_{jn}$ , in which  $\rho_{nj} = \langle n | \rho | j \rangle$  and  $P_{nj} = \langle n | P | j \rangle$ .

*Proof.* First note that, since  $\rho$  is a pure state, even if some energy levels are degenerate, there is always an energy basis we can choose such that  $\rho$  only has support on one energy state inside each energy eigenspace. In this basis,  $G_{nj} = 0$  implies that either  $n = j$  or  $\rho_{nj} = 0$ .

Then let us write  $\rho(t) = \sum_{nj} \rho_{nj} e^{-iG_{nj}t} |n\rangle\langle j|$  and note that

$$\omega = \sum_{n=j} \rho_{nj} |n\rangle\langle j|,$$

$$\rho(t) - \omega = \sum_{n \neq j} \rho_{nj} e^{-iG_{nj}t} |n\rangle\langle j|. \quad (\text{A2})$$

Thus, for any projector  $P$  we have

$$\begin{aligned} & \langle |\text{Tr}[P(\rho(t) - \omega)]|^2 \rangle \\ &= \left\langle \left| \sum_{n \neq j} P_{jn} \rho_{nj} e^{-iG_{nj}t} \right|^2 \right\rangle \\ &= \sum_{n \neq j} \sum_{k \neq l} \rho_{nj} P_{jn} (\rho_{kl} P_{lk})^* \langle e^{-i(G_{nj} - G_{kl})t} \rangle, \\ &= \sum_{n \neq j} \sum_{k \neq l} \rho_{nj} P_{jn} (\rho_{kl} P_{lk})^* \delta_{G_{nj}, G_{kl}}. \end{aligned} \quad (\text{A3})$$

■  
*Theorem 5.* Given a finite-dimensional Hilbert space  $\mathcal{H}$ , any initial state  $\rho : \mathcal{H} \rightarrow \mathcal{H}$  evolving under a Hamiltonian with energy levels  $E_n$ , and any  $N$ -outcome measurement  $\mathcal{M}$ , the following equation holds:

$$\langle D_{\mathcal{M}}(\rho_t, \omega) \rangle \leq \frac{1}{2} \sqrt{D_G \frac{N-1}{d_{\text{eff}}}}, \quad (\text{A4})$$

where  $D_G$  is the maximum degeneracy of any gap in the spectrum.

*Proof.* We start by assuming that  $\rho$  is a pure state and then generalize to mixed states. This means we can use Lemma 1.

Next, the following steps are easier to follow if we label each energy gap by the indices  $\alpha = (n, j)$  and  $\beta = (k, l)$ , denote summing over  $\alpha$  as shorthand for summing over  $n \neq j$ , and define the Hermitian matrix  $M_{\alpha\beta} = \delta_{G_{nj}, G_{kl}}$ . Using these definitions, it is easy to see that the sums in Lemma 1 form an inner product  $\sum_{\alpha\beta} v_{\alpha}^* M_{\alpha\beta} v_{\beta} = |v^\dagger M v|$ . Therefore, we can use the Cauchy-Schwarz inequality to bound this sum by

$$\begin{aligned} & \langle \text{Tr}[P(\rho_t - \omega)]^2 \rangle \\ & \leq \sum_{\alpha\beta} v_{\alpha}^* M_{\alpha\beta} v_{\beta} \leq \|M\| \sum_{\alpha} |v_{\alpha}|^2 = \|M\| \sum_{n \neq j} |\rho_{nj} P_{jn}|^2 \\ & \leq \|M\| \sum_{n,j} |\rho_{nj} P_{jn}|^2 \leq \|M\| \sum_{n,j} |\rho_{jj} P_{jn} \rho_{nn} P_{nj}| \\ & = \|M\| \text{Tr}[P\omega P\omega], \end{aligned} \quad (\text{A5})$$

where we also used that  $\rho_{jj}\rho_{nn} \geq \rho_{nj}\rho_{jn}$  by positivity of the density matrix.

Now note that the left-hand side does not change if you subtract from the projector anything proportional to the identity. Therefore, we can write

$$\begin{aligned}
 & \langle \text{Tr}[P(\rho_t - \omega)]^2 \rangle \frac{1}{\|M\|} \\
 & \leq \text{Tr} \left[ \left( P - \frac{\mathbb{1}}{N} \right) \omega \left( P - \frac{\mathbb{1}}{N} \right) \omega \right] \\
 & \leq \text{Tr}[P\omega P\omega] - 2 \text{Tr} \left[ \frac{\mathbb{1}}{N} \omega P \omega \right] + \text{Tr} \left[ \frac{\mathbb{1}}{N} \omega \frac{\mathbb{1}}{N} \omega \right] \\
 & = \text{Tr}[P\omega P\omega] - \frac{2}{N} \text{Tr}[P\omega^2] + \frac{1}{N^2} \text{Tr}[\omega^2] \\
 & \leq \text{Tr}[P\omega^2] \left( 1 - \frac{2}{N} \right) + \frac{1}{N^2} \text{Tr}[\omega^2], \tag{A6}
 \end{aligned}$$

which leads to

$$\begin{aligned}
 \sum_{P \in \mathcal{M}} \langle \text{Tr}[P(\rho_t - \omega)]^2 \rangle & \leq \|M\| \text{Tr}[\omega^2] \left( 1 - \frac{2}{N} + \frac{1}{N} \right) \\
 & = \frac{\|M\|}{d_{\text{eff}}} \frac{N-1}{N}, \tag{A7}
 \end{aligned}$$

where we used  $\text{Tr}[\omega^2] = d_{\text{eff}}^{-1}$ .

Since  $M$  is a block diagonal matrix, where each block is composed of only 1's and is at most of size  $D_G$ , then the largest eigenvalue of  $M$  is at most  $D_G$  and we have  $\|M\| \leq D_G$ . This finally leads to

$$\begin{aligned}
 \langle D_{\mathcal{M}}(\rho_t, \omega) \rangle & = \frac{1}{2} \sum_{P \in \mathcal{M}} \langle \text{Tr}[P(\rho_t - \omega)] \rangle \\
 & \leq \frac{1}{2} \sum_{P \in \mathcal{M}} \sqrt{\langle \text{Tr}[P(\rho_t - \omega)]^2 \rangle}
 \end{aligned}$$

$$\begin{aligned}
 & \leq \frac{1}{2} \sqrt{N \sum_{P \in \mathcal{M}} \langle \text{Tr}[P(\rho_t - \omega)]^2 \rangle} \\
 & \leq \frac{1}{2} \sqrt{(N-1) \frac{\|M\|}{d_{\text{eff}}}} \\
 & \leq \frac{1}{2} \sqrt{D_G \frac{N-1}{d_{\text{eff}}}}. \tag{A8}
 \end{aligned}$$

To see that the same will hold for mixed states, take an ancillary Hilbert space  $\mathcal{A}$  with the same dimension as  $\mathcal{H}$  and use it to purify  $\rho$ . That is, define a pure state  $\rho'$  on  $\mathcal{H} \otimes \mathcal{A}$  such that  $\text{Tr}_{\mathcal{A}}[\rho'] = \rho$ , and define  $\omega' = \langle \rho' \rangle$ . Then it is straightforward to see that

$$\begin{aligned}
 \langle D_{\mathcal{M}}(\rho_t, \omega) \rangle & = \frac{1}{2} \sum_{P \in \mathcal{M}} \langle |\text{Tr}_{\mathcal{H}}[P(\rho_t - \omega)]| \rangle \\
 & = \frac{1}{2} \sum_{P \in \mathcal{M}} \langle |\text{Tr}_{\mathcal{H}}[P \text{Tr}_{\mathcal{A}}[\rho'_t - \omega']]| \rangle \\
 & = \frac{1}{2} \sum_{P \in \mathcal{M}} \langle |\text{Tr}[P \otimes \mathbb{1}_{\mathcal{A}}(\rho'_t - \omega')]| \rangle \\
 & = \langle D_{\mathcal{M}}(\rho'_t, \omega') \rangle \\
 & \leq \frac{1}{2} \sqrt{D'_G \frac{N-1}{d_{\text{eff}}(\rho')}} \tag{A9}
 \end{aligned}$$

where  $D'_G$  is the maximum degeneracy among energy gaps in  $\mathcal{H} \otimes \mathcal{A}$ .

At last, to reproduce Eq. (A8), simply choose a null Hamiltonian for the  $\mathcal{A}$  subspace ( $H_{\mathcal{A}} = 0$  and  $H' = H \otimes \mathbb{1}_{\mathcal{A}}$ ). Of course this choice leads to a highly degenerate energy spectrum, but, since  $\rho'$  is pure, this does not affect any of the quantities by the same argument used at the start of this proof. This has the consequence that  $D'_G = D_G$  and  $d_{\text{eff}}(\rho') = d_{\text{eff}}(\rho)$ . ■

- 
- [1] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, *Phys. Rev. E* **79**, 061103 (2009).
- [2] P. Reimann, Canonical thermalization, *New J. Phys.* **12**, 055027 (2010).
- [3] A. J. Short, Equilibration of quantum systems and subsystems, *New J. Phys.* **13**, 053009 (2011).
- [4] A. J. Short and T. C. Farrelly, Quantum equilibration in finite time, *New J. Phys.* **14**, 013063 (2012).
- [5] S. Goldstein, T. Hara, and H. Tasaki, Time Scales in the Approach to Equilibrium of Macroscopic Quantum Systems, *Phys. Rev. Lett.* **111**, 140401 (2013).
- [6] A. S. L. Malabarba, L. P. García-Pintos, N. Linden, T. C. Farrelly, and A. J. Short, Quantum systems equilibrate rapidly for most observables, *Phys. Rev. E* **90**, 012121 (2014).
- [7] P. Reimann and M. Evstigneev, Quantum versus classical foundation of statistical mechanics under experimentally realistic conditions, *Phys. Rev. E* **88**, 052114 (2013).
- [8] I. E. Farquhar, *Ergodic Theory in Statistical Mechanics* (Interscience, New York, 1964).
- [9] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics: Emergence of Thermodynamic Behavior Within Composite Quantum Systems*, Lecture Notes in Physics Vol. 784 (Springer, Berlin, 2009).
- [10] L. Sklar, *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics* (Cambridge University Press, Cambridge, 1995).
- [11] D. K. McDonald, *Noise and Fluctuations: An Introduction* (Wiley, New York, 1962).
- [12] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales, *Phys. Rev. Lett.* **89**, 050601 (2002).
- [13] S. Goldstein, T. Hara, and H. Tasaki, Extremely quick thermalization in a macroscopic quantum system for a typical nonequilibrium subspace, *New J. Phys.* **17**, 045002 (2015).
- [14] F. G. S. L. Brandão, P. Ćwikliński, M. Horodecki, P. Horodecki, J. K. Korbicz, and M. Mozrzyk, Convergence to equilibrium under a random Hamiltonian, *Phys. Rev. E* **86**, 031101 (2012).

- [15] N. Linden, S. Popescu, A. J. Short, and A. Winter, On the speed of fluctuations around thermodynamic equilibrium, *New J. Phys.* **12**, 055021 (2010).
- [16] C. Werndl and R. Frigg, Reconceptualising equilibrium in Boltzmannian statistical mechanics and characterising its existence, *Stud. Hist. Philos. Mod. Phys.* **49**, 19 (2015).
- [17] L. Masanes, A. J. Roncaglia, and A. Acín, Complexity of energy eigenstates as a mechanism for equilibration, *Phys. Rev. E* **87**, 032137 (2013).
- [18] L. C. Venuti, S. Yeshwanth, and S. Haas, Equilibration times in clean and noisy systems, *Phys. Rev. A* **87**, 032108 (2013).
- [19] J. Barrett, Information processing in generalized probabilistic theories, *Phys. Rev. A* **75**, 032304 (2007).