PHYSICAL REVIEW E **90**, 012121 (2014)

Quantum systems equilibrate rapidly for most observables

Artur S. L. Malabarba,[1] Luis Pedro García-Pintos,[2] Noah Linden,[2] Terence C. Farrelly,[3] and Anthony J. Short[1]

[1] *H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, United Kingdom*

[2] *School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom*

[3] *DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom*

(Received 27 February 2014; revised manuscript received 26 June 2014; published 22 July 2014)

Considering any Hamiltonian, any initial state, and measurements with a small number of possible outcomes compared to the dimension, we show that most measurements are already equilibrated. To investigate nontrivial equilibration, we therefore consider a restricted set of measurements. When the initial state is spread over many energy levels, and we consider the set of observables for which this state is an eigenstate, most observables are initially out of equilibrium yet equilibrate rapidly. Moreover, all two-outcome measurements, where one of the projectors is of low rank, equilibrate rapidly.

DOI: 10.1103/PhysRevE.90.012121 PACS number(s): 05.30.−d, 03.65.Yz

The topic of equilibration time scales has been of much interest lately [1–9]. Given that it has been shown that quantum systems equilibrate under rather general conditions [10–12], it is important to understand the time scale for the process. However, attempts to derive an upper bound on equilibration time have resulted in very large time scales. Short and Farrelly [1], for instance, obtain a very general bound, of which we give an improved derivation in Appendix A, which scales with the dimension of the system, typically exponentially in the number of particles.

A quantum system is said to undergo equilibration when its quantum state spends most of its time almost indistinguishable from a fixed (time-invariant) steady state. This is not the same as thermalization, in which the steady state is a Gibbs state. Thus, thermalization is a special case of equilibration, and understanding equilibration times is a key step in understanding thermalization times.

When one discusses quantum equilibration, it is common to refer to either subsystem equilibration [10,13], in which a small system equilibrates due to contact with a bath, or observable equilibration, in which a fully closed system appears to equilibrate due to the limited information offered by outcomes of a particular set of observables. The latter was initially shown by Reimann [11,14], as a statement that the expectation values of quantum observables stay predominantly close to a static value, and was later built on by Short [15], who showed that these results apply even if one considers all the information that can be gathered from the observable, instead of just the expectation value.

In this paper, we consider any finite-dimensional system and any Hamiltonian, and show that most N-outcome observables are initially in equilibrium (for N small compared to the dimension). To investigate time scales we therefore turn to a natural class of observables which are initially typically out of equilibrium, those with a definite initial value (i.e., observables for which the initial state is an eigenstate). We show that, for pure initial states spread over many energy levels, most of these observables equilibrate in very short times, in fact, most equilibrate essentially as fast as possible. Moreover, in the case of two-outcome observables where one of the projectors is of low rank, we show that all observables equilibrate fast (for any initial state spread over many energy levels).

As will be clear in Theorems 2 and 3, when referring to "typical" or "most" observables, we mean that in the context of the Haar measure. While this does include all observables of physical significance, they constitute a small fraction of all possible observables. Still, these results give us new insight into the problem of equilibration time scales, which was not available through the use of strict upper bounds. This also raises the question of what is special about physical measurements that makes them much slower than most measurements.

To obtain these results, we address the issue of equilibration time scales with respect to measurements composed of N outcomes. As a figure of merit for equilibration we will use the *distinguishability* $D_{\mathcal{M}}(\sigma,\rho)$ between two states σ and ρ according to an observable $\mathcal{M} = \{P_1,\dots,P_N\}$, where the projectors P_j represent the different outcomes of the measurement. We define it so that after performing the measurement, given full information about the two states being compared, the distinguishability quantifies the probability of successfully "guessing" which state the system was in [15], according to

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

where

$$D_{\mathcal{M}}(\sigma,\rho) = \frac{1}{2}\sum_j |\mathrm{Tr}[\sigma P_j] - \mathrm{Tr}[\rho P_j]|. \tag{2}$$

When $D_{\mathcal{M}}(\sigma,\rho) = 0$, the measurement does not provide information that helps to distinguish σ from ρ (one might as well toss an unbiased coin to decide). On the other hand, when $D_{\mathcal{M}}(\sigma,\rho) = 1$, the states are perfectly discriminated by this measurement.

In the special case in which the measurement has two outcomes (i.e., $\mathcal{M} = \{P, \mathbb{1} - P\}$), the distinguishability is given by $D_{\mathcal{M}}(\sigma,\rho) = |\mathrm{Tr}[\sigma P] - \mathrm{Tr}[\rho P]|$, and we will denote it by $D_P(\sigma,\rho)$.

Given an initial state ρ evolving under a Hamiltonian H, we say equilibration has taken place at time T_{eq} when, for some small constant $\varepsilon > 0$,

$$\langle D_{\mathcal{M}}(\rho_t,\omega)\rangle_T \leqslant \varepsilon, \quad \forall\, T > T_{\text{eq}}, \tag{3}$$

where $\langle f(t)\rangle_T = \frac{1}{T}\int_0^T f(t)dt$, $\omega = \lim_{T\to\infty}\langle\rho_t\rangle_T$ is the time-averaged state, and $\rho_t = e^{-iHt}\rho e^{iHt}$ is the evolved state. When equilibration does take place, ω is also called the equilibrium state.[1]

In words, equilibration with respect to a measurement has taken place when the time-averaged distinguishability for this measurement falls below a small value. Since the distinguishability is a positive quantity, this means that, for any $T > T_{\mathrm{eq}}$, the instantaneous state ρ_t is essentially indistinguishable from ω for almost all times in the interval $[0,T]$.

It is worth stating that we are not claiming thermalization, in which the equilibrium state is a thermal state. This definition of equilibration, while necessary for thermalization [10], is a separate issue and guarantees only the approach to a steady state.

Two parameters of the initial state that will be of importance to us are the effective dimension d_{eff} and the energy standard deviation σ_E, defined by

$$d_{\mathrm{eff}} = \left[\sum_{n=1}^{\tilde d} p_n^2\right]^{-1}, \qquad \sigma_E^2 = \sum_{n=1}^{\tilde d} p_n (E_n - \bar E)^2. \tag{4}$$

Here, $\tilde d$ is the number of distinct energy levels E_n with probabilities $p_n = \mathrm{Tr}[\mathcal{Q}_n\rho]$, where $\mathcal{Q}_n$ is the projector onto the eigenspace with energy E_n, and $\bar E = \sum_{n=1}^{\tilde d} p_n E_n$. The effective dimension gives an estimate of how many energy levels ρ occupies with significant probability,[2] and $d_{\mathrm{eff}} \gg 1$ is the key requirement for equilibration to take place, as we will see. The energy standard deviation, in turn, will take a primary role in the time scale of equilibration. Note that we use $\hbar = 1$ throughout.

Given any d-dimensional Hamiltonian and any initial state with support on many energy levels (high effective dimension), we present the following results regarding the time scale of equilibration:

(1) We show that *all* two-outcome measurements, for which one of the projectors P has small rank, have very short time scales ($T_{\mathrm{eq}} \sim \mathrm{rank}\, P/\sigma_E$).

(2) For N-outcome measurements, we prove that *most* measurements are already equilibrated, and that *most* measurements with a definite initial value (i.e., for which the initial state is a pure eigenstate) equilibrate with extremely short time scales ($T_{\mathrm{eq}} \sim 1/\sigma_E$). These results hold as long as $N \ll d$, but regardless of the rank of the projectors.

The first statement shows how restricting the measurements can lead to a whole family of observables that equilibrate fast, while the second one refers to the time scales of typical measurements (under the Haar measure), which are indeed also fast.

To understand why $1/\sigma_E$ is a fast time scale, one can consider the uncertainty relation between H and an observable O, which states [16] $2\sigma_E\sigma_O \geqslant |\langle[H,O]\rangle| = |\langle\dot O\rangle|$. Thus, the minimum necessary time for the expectation value of *any* observable to vary significantly is $\sigma_O/|\langle\dot O\rangle| \geqslant 1/(2\sigma_E)$.

Previous related results on equilibration time scales have been obtained by Goldstein, Hara, and Tasaki [8]. In that paper, the authors considered particular constructions of two-outcome measurements that take a very long or very short time to equilibrate. Interestingly, they find an example of a projector with high rank which equilibrates fast, albeit with very particular properties. Our first result proves fast equilibration of *all* small rank projectors, for *any* system with high d_{eff}. Moreover, we find that most observables also equilibrate fast (when $N \ll d$ and initial state is pure with high d_{eff}). On the issue of slow time scales, we give an alternative example to that in [8] of a measurement which equilibrates slowly ($T_{\mathrm{eq}} \gtrsim d_{\mathrm{eff}}/\sigma_E$), given a pure initial state with high d_{eff}.

We end with a short discussion. Our results can also be stated almost identically in terms of expectation value of observables (distinguishability is used here as it provides a stronger statement of equilibration).

I. FAST EQUILIBRATION

We now show that all systems with high effective dimension equilibrate fast with respect to the two-outcome measurement $\mathcal{M} = \{P, \mathbb{1} - P\}$, whenever either of the projectors is of sufficiently low rank. Given $K = \min\{\mathrm{rank}\, P, \mathrm{rank}(\mathbb{1} - P)\}$, we will show that the average distinguishability $\langle D_P(\rho_t,\omega)\rangle_T$ becomes small in a time of the order of K/σ_E. We start by defining the function η_ϵ:

Definition 1. Given any Hamiltonian with spectrum $\{E_j \mid j = 1,\dots,\tilde d\}$, a d-dimensional Hilbert space $\mathcal{H}$ and any state $\rho : \mathcal{H} \to \mathcal{H}$ with probabilities p_j associated to each energy level. For any $\epsilon > 0$,

$$\eta_\epsilon = \max_{E\in\mathbb{R}} \sum_{\substack{j:\\ E_j\in[E,E+\epsilon]}} p_j \tag{5}$$

is the maximum probability that can be found inside any energy interval of size ϵ.

This function is useful because it captures both the state's energy distribution and the Hamiltonian's energy spectrum. The theorem is then simply an upper bound on the finite-time average of the distinguishability $D_P(\rho_t,\omega)$ for any projector P.

Theorem 1 (Fast equilibration). For any initial state $\rho : \mathcal{H} \to \mathcal{H}$, any Hamiltonian, and any projector P where $K = \min\{\mathrm{rank}\, P, \mathrm{rank}(\mathbb{1} - P)\}$,

$$\langle D_P(\rho_t,\omega)\rangle_T \leqslant c\sqrt{\eta_{\frac{1}{T}} K}, \tag{6}$$

where $c = \frac{5\pi}{4}\sqrt{\frac{2}{1-e^{-2}}} + 1 \approx 6.97$.

Proof. Defining the Lorentzian average $\langle f(t)\rangle_{L_T} = \int_{-\infty}^{\infty} \frac{f(t)T}{T^2+(t-T/2)^2}\frac{dt}{\pi}$, any positive function f satisfies $\langle f\rangle_T \leqslant \frac{5\pi}{4}\langle f\rangle_{L_T}$.[3] Then, by use of the Cauchy-Schwarz inequality

[1] Note that this limit is well defined, and is obtained by decohering ρ in the energy eigenbasis, i.e., $\omega = \sum_n \mathcal{Q}_n\rho\mathcal{Q}_n$ where $\mathcal{Q}_n$ is the projector onto the nth energy eigenspace.

[2] In particular, if the state is spread evenly over N energies then $d_{\mathrm{eff}} = N$.

[3] Define $\Theta_T(t) = \frac{1}{T}$ for $t \in [0,T]$, and 0 otherwise. Then, $\langle f\rangle_T = \int_{-\infty}^{\infty}\Theta_T(t)f(t)dt$, and $\Theta_T(t) \leqslant \frac{5}{4}\frac{T}{T^2+(t-T/2)^2}$.

and the fact that $\mathrm{Tr}[\omega^2] \leqslant 1/d_{\mathrm{eff}}$[4]

$$\begin{aligned}\langle D_P(\rho_t,\omega)\rangle_T &= \langle|\mathrm{Tr}[P(\rho_t-\omega)]|\rangle_T \\ &\leqslant \langle \mathrm{Tr}[P\rho_t]\rangle_T + \mathrm{Tr}[P\omega] \\ &\leqslant \frac{5\pi}{4}\mathrm{Tr}[P\omega_{L_T}] + \sqrt{\mathrm{Tr}[\omega^2]\,\mathrm{Tr}[P^2]} \\ &\leqslant \frac{5\pi}{4}\sqrt{\mathrm{Tr}\big[\omega_{L_T}^2\big]\,\mathrm{Tr}[P^2]} + \sqrt{\mathrm{Tr}[\omega^2]\,\mathrm{Tr}[P^2]} \\ &\leqslant \frac{5\pi}{4}\sqrt{K\,\mathrm{Tr}\big[\omega_{L_T}^2\big]} + \sqrt{\frac{K}{d_{\mathrm{eff}}}}, \end{aligned} \tag{7}$$

where $\omega_{L_T} = \langle\rho_t\rangle_{L_T}$ and $K = \mathrm{rank}\,P$. Appendix B shows that

$$\mathrm{Tr}\big[\omega_{L_T}^2\big] \leqslant \frac{2\eta_{\frac{1}{T}}}{1-e^{-2}}. \tag{8}$$

Using the fact that $d_{\mathrm{eff}}^{-1} \leqslant p_{\max} \leqslant \eta_\epsilon, \forall\,\epsilon > 0$, where $p_{\max}$ is the maximum occupation probability of any energy level, results in Eq. (6). The reason we may take $K = \min\{\mathrm{rank}\,P, \mathrm{rank}(\mathbb{1}-P)\}$ is that $D_P(\rho_t,\omega) = D_{\mathbb{1}-P}(\rho_t,\omega)\ \forall\, t \in \mathbb{R}$. ■

The requirement of "large d_{eff}" mentioned in the introductory section is a consequence of $d_{\mathrm{eff}}^{-1} \leqslant \eta_\epsilon$ since $\eta_{\frac{1}{T}}$ cannot converge to a small value if d_{eff} is small.

Note that the quantity $\mathrm{Tr}[\omega_{L_T}^2]$, corresponding to the purity of the time-averaged state, is at the core of the equilibration process, dictating, for any given system, an upper bound on the time scale of equilibration (the reciprocal of this quantity acts like a time-dependent effective dimension, growing from 1 to d_{eff} as T increases). The right-hand side of Eq. (8) displays a bound on the purity which is easier to calculate than the purity itself (in fact, it is trivial if one knows the spectrum and the state) and whose tightness is discussed in the following.

This theorem proves that the measurement of a rank-1 projector equilibrates as soon as the energy interval $1/T$ is too small to contain a significant portion of the probabilities, which happens roughly when it is small compared to σ_E (which is a very short time scale). Conversely, a rank-K projector requires that the probabilities be K times smaller. For instance, if $\eta_{\frac{1}{T}} \sim \frac{1}{\sigma_E T}$, as we argue in the following, the time scale of equilibration is at most $\sim\frac{K}{\sigma_E}$.

A. Estimating η

We focus now our attention on Eq. (8), in order to compare how well $\eta_{\frac{1}{T}}$ upper bounds the purity, and to illustrate how easy it is to estimate η. Given a dense enough energy spectrum, we can approximate the probability distribution of the initial state by a continuous function $p(E)$ for which the maximum value is roughly

$$\max_E p(E) \sim \frac{a}{\sigma_E}, \tag{9}$$

where a is some constant which depends on the shape of the distribution. Since η_ϵ can always be upper bounded by $\epsilon \max_E p(E)$, we have

$$\eta_{\frac{1}{T}} \leqslant \frac{a}{\sigma_E T}, \tag{10}$$

as long as T is not large enough that the $\frac{1}{T}$ window only contains a few energy levels. In Appendix B we show that the above estimation is correct for the case of a Gaussian distribution for the energy probabilities, with $a \approx 0.40$ in this case.

II. TYPICAL MEASUREMENTS

Here, we prove two statements regarding typical two-outcome measurements composed of a projector of *any* rank, applied to any initial state and any Hamiltonian.

Theorem 2 (Typical two-outcome measurements are already equilibrated for any initial state). Take the rank-K projector P_U defined as the unitary transformation from an energy basis projector

$$P_U = UPU^\dagger = \sum_{n=1}^{K} U|n\rangle\langle n|U^\dagger, \tag{11}$$

with $U : \mathcal{H} \to \mathcal{H}$ unitary and $|n\rangle$ being energy eigenstates. The distinguishability between ρ_t and ω according to P_U (and its complement) averaged over all unitaries is

$$\langle D_{P_U}(\rho_t,\omega)\rangle_U \leqslant \sqrt{\frac{K}{d^2}\frac{d-K}{d+1}} \leqslant \frac{1}{2\sqrt{d+1}}. \tag{12}$$

Proof. The only necessary inequality is the first step, Jensen's inequality [17]

$$\langle D_{P_U}(\rho_t,\omega)\rangle_U \leqslant \sqrt{\langle D_{P_U}(\rho_t,\omega)^2\rangle_U}. \tag{13}$$

In Appendix C 1 we show that the average of the squared distinguishability can be exactly calculated to be

$$\langle D_{P_U}(\rho_t,\omega)^2\rangle_U = \frac{K}{d}\frac{d-K}{d^2-1}\mathrm{Tr}\big[\rho_t^2-\omega^2\big]. \tag{14}$$

Then, the fact that $\mathrm{Tr}[\omega^2] \geqslant 1/d$[5] implies that $\mathrm{Tr}[\rho_t^2-\omega^2] \leqslant (d-1)/d$ and leads to the first inequality of Eq. (12). The second inequality is obtained by setting $K = d/2$, which maximizes the expression. ■

This average result is relevant because $D_{P_U}(\rho_t,\omega)$ is a positive definite quantity. Thus, stating that its average is small necessarily implies that $D_{P_U}(\rho_t,\omega)$ is small for most P_U (in other words, it is strongly concentrated close to zero).

This result, however, does not make any statements about time scales, or the dynamics of equilibration. It is more relevant to study measurements which start out of equilibrium, and ask how fast they approach it. For this reason, Theorem 3 visits again the average distinguishability, but constrains the projector to contain $\rho(0) = \rho_0 = |\Psi\rangle\langle\Psi|$ as one of its terms

[4]The equality is easy to check for nondegenerate Hamiltonians or pure initial states, while the inequality is necessary for degenerate Hamiltonians with a mixed initial state.

[5]It is easy to see the trace is minimized when $\omega = \mathbb{1}/d$.

(note that for this theorem, we restrict the initial state to be pure). For that, we divide the Hilbert space between the span of the initial state and everything else $\mathcal{H} = \mathcal{H}' \oplus \rho_0$, where $\dim \mathcal{H} = d$ and $\dim \mathcal{H}' = d - 1$.

Theorem 3 (*Typical two-outcome measurements with a definite initial value equilibrate fast for any pure initial state with high*d_{eff}). Consider the projector given by

$$\Pi_U = \rho_0 + P_U, \quad P_U = UPU^\dagger, \tag{15}$$

where ρ_0 is the initial (pure) state, U is a partial unitary with $UU^\dagger = U^\dagger U = \mathbb{1}_{\mathcal{H}'}$, P is any rank-$(K-1)$ projector with support on $\mathcal{H}'$. The distinguishability between ρ_t and ω according to Π_U (and its complement) averaged over all unitaries on $\mathcal{H}'$ is

$$\langle D_{\Pi_U}(\rho_t,\omega)\rangle_U \leqslant D_{\rho_0}(\rho_t,\omega) + \frac{1}{2\sqrt{d-1}}, \tag{16}$$

where we have (from Theorem 1) that

$$\langle\langle D_{\Pi_U}(\rho_t,\omega)\rangle_U\rangle_T \leqslant c\sqrt{\eta_{\frac{1}{T}}} + \frac{1}{2\sqrt{d-1}} \tag{17}$$

decays very fast, with $c \approx 6$.

The details of the proof can be found in Appendix C 2. This result is enough to state that most two-outcome measurements (of any rank) containing the initial state equilibrate essentially as fast as the measurement of the rank-1 projector consisting of *only* the initial state.

To show that this class of observables is typically out of equilibrium initially (and thus equilibrates in a nontrivial way), we show in Appendix C 3 that the average initial distinguishability is given by

$$\langle D_{\Pi_U}(\rho_0,\omega)\rangle_U \geqslant \left(1 - \frac{K-1}{d-1}\right)\left(1 - \frac{1}{d_{\text{eff}}}\right), \tag{18}$$

and is therefore significantly above zero so long as the projector does not cover almost the entire space.

We now extend these results to multioutcome measurements.

Corollary 1 (*N-outcome generalization of Theorem 2*). Given $N \ll d$, the typical N-outcome measurement is already equilibrated. Describing the measurement by the positive operator valued measures (POVM) $\mathcal{M}_U = \{U^\dagger P_i U\}_{i=1,N}$, and using the result from Theorem 2, it is easy to see that

$$\begin{aligned}\langle D_{\mathcal{M}_U}(\rho_t,\omega)\rangle_U &= \frac{1}{2}\sum_{j=1}^{N}\langle D_{U^\dagger P_j U}(\rho_t,\omega)\rangle_U \\ &\leqslant \frac{1}{2}\sum_{j=1}^{N}\sqrt{\frac{K_j}{d^2}\frac{d-K_j}{d+1}} \\ &\leqslant \frac{1}{2}\sqrt{\frac{N}{d+1}},\end{aligned} \tag{19}$$

where $K_j = \text{rank}\, P_j$, and the second line is maximal for $K_j = d/N$.

Corollary 2 (*N-outcome generalization of Theorem 3*). Given $N \ll d$, typical out-of-equilibrium N-outcome measurements equilibrate fast for any pure initial state with high d_{eff}. Define the POVM $\mathcal{M}_U^{\rho_0} = \{\rho_0 + U^\dagger P_1 U, U^\dagger P_2 U, \ldots, U^\dagger P_N U\}$, with $\rho_0 + \sum U^\dagger P_n U = \mathbb{1}$. In Appendix C 4, we show that

$$\left\langle D_{\mathcal{M}_U^{\rho_0}}(\rho_t,\omega)\right\rangle_U \leqslant D_{\rho_0}(\rho_t,\omega) + \frac{1}{2}\sqrt{\frac{N}{d-1}}. \tag{20}$$

This means most measurements are already equilibrated even for a large number of outcomes as long as $N \ll d$, a physically reasonable assumption for systems composed of many particles given that the dimension d grows exponentially with the number particles. Furthermore, for any $N \ll d$, most measurements with a definite initial value (which are typically out of equilibrium initially) still equilibrate essentially as fast as a rank-1 projector.

III. SLOW EQUILIBRATION

This result complements the previous sections by showing that fast equilibration is not always the case. We find that for any pure system with high effective dimension it is always possible to define a measurement for which the equilibration time is tremendously long.

We do that by considering the projector $P_{\mathcal{H}_K}$ onto the subspace $\mathcal{H}_K$ defined by

$$\mathcal{H}_K = \text{span}\{|\psi(j\tau)\rangle \mid j = 0,\ldots,K-1\}, \tag{21}$$

with $\tau = 2\epsilon/\sigma_E$. We prove in Appendix D that, for any ϵ,

$$D_{P_{\mathcal{H}_K}}(\rho_t,\omega) \geqslant 1 - \epsilon^2 - \sqrt{\frac{K}{d_{\text{eff}}}}, \quad \forall\, t \in \left[0, K\tau - \frac{\epsilon}{\sigma_E}\right] \tag{22}$$

and

$$\left\langle D_{P_{\mathcal{H}_K}}(\rho_t,\omega)\right\rangle_{T\to\infty} \leqslant 2\sqrt{\frac{K}{d_{\text{eff}}}} \ll 1, \tag{23}$$

showing that the distinguishability is above some constant for a time that can be very long, but still equilibrates eventually.

The construction simply takes the subspace comprised of K sequential "snapshots" of the wave function, and makes sure that the time step between these snapshots is small enough such that the wave function does not move out during the intermediate times. The "*time is long*" statement holds because the necessity to take small steps is nothing compared to the very large number of steps we are allowed to include ($K \ll d_{\text{eff}}$). In Appendix D, we provide an example where this time is $\sim \frac{d_{\text{eff}}}{1000\sigma_E}$, which can easily be longer than the age of the universe.

IV. DISCUSSION

In this work, we have proved several properties regarding observable equilibration, all of which apply to any system capable of equilibration. First, we find an upper bound on the time scale of equilibration of *any* two-outcome measurement based on the rank of the projector that defines it, which turns out to be very fast for small ranks. We also find that typical measurements of any rank and any reasonable number of outcomes

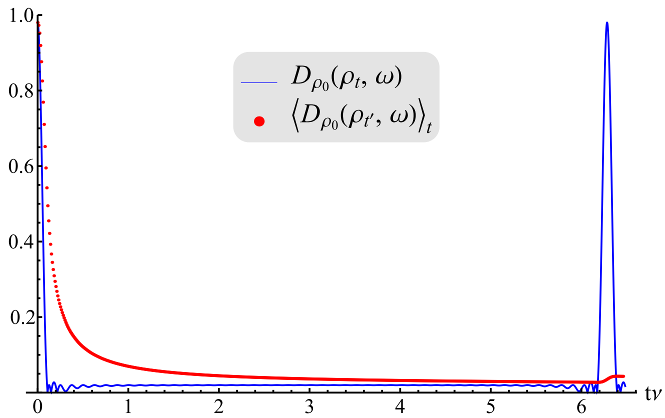

FIG. 1. (Color online) $D_{\rho_0}(\rho_t,\omega)$ and its finite-time average for a full period of the harmonic oscillator with level spacing ν and the initial condition is a pure state spread equally over the first 50 energy levels (irrespective of phases). Notice that, despite the revival (blue solid line), the projector still equilibrates (red dotted line).

are already equilibrated. To investigate time scales we then turn to a natural class of measurements which are typically initially out of equilibrium (those for which the initial state gives a definite value) and show that most of these measurements equilibrate fast, approximately as fast as a rank-1 projector. On the other hand, we construct a measurement which is extremely slow to equilibrate. This shows that, indeed, in order to obtain physically realistic time scales, one must restrict to further constraints on the measurements and/or on the system considered.

One characteristic that distinguishes this work from some previous results [1,10,11,13,15] is that there was no need to assume nondegenerate energy gaps in order to prove equilibration. To emphasize this, Fig. 1 plots an example of the distinguishability $D_{\rho_0}(\rho_t,\omega)$ of a harmonic oscillator (with highly degenerate gaps) against its time average. The function in the figure decays fast for large d, with $T_{\rm eq} \sim \frac{1}{d\nu}$, which implies that typical projectors equilibrate fast, as given by Eq. (16). Nevertheless, the function returns to its original value at multiple times of $T_{\rm rev} = \frac{2\pi}{\nu}$, times at which a full revival manifests. This does not conflict with equilibration because these revivals are so short that they cannot affect the average significantly.

The results described here aim to be general, by making statements as a function of the rank, and which apply to any system. However, there are specific cases which deserve special attention. When the measurement is restricted to a small subsystem of a complex many-body system, it is expected to equilibrate fast; however, Theorem 1 by itself does not lead to that conclusion since the outcomes of these measurements are of high rank. Moreover, typical measurements (in the Haar measure sense) need not necessarily represent physically relevant measurements. It would be interesting to study whether these results can be extended to typical measurements with certain constraints, for instance, measurements acting on a small subsystem.

Note added: Recently, we became aware of very recent independent work [18] which also addresses the issue of the rapid equilibration of quantum systems.

ACKNOWLEDGMENTS

We would like to thank A. Winter for helpful discussions. A.J.S. acknowledges support from the Royal Society. A.S.L.M. acknowledges support from the CNPq. Part of this work was done while the authors were at the Newton Institute programme on Mathematical Challenges in Quantum Information.

APPENDIX A: GENERAL BOUND ON TIME SCALES OF EQUILIBRATION

Following Short and Farrelly [1], we focus on the average distance between the expected value of some observable A and its infinite time average. As proved in a footnote in the main text,[6] the usual average can be bounded by the Lorentzian average, therefore,

$$\langle |{\rm Tr}[A(\rho_t-\omega)]|^2\rangle_T \leqslant \frac{5\pi}{4}\langle |{\rm Tr}[A(\rho_t-\omega)]|^2\rangle_{L_T}, \quad \text{(A1)}$$

where ρ_t is the instantaneous state and $\omega = \lim_{T\to\infty}\langle\rho_t\rangle_T$ is the infinite time-averaged state.

Denoting by E_j and $|j\rangle$ the eigenvalues and eigenvectors of the Hamiltonian, and assuming an initially pure state for simplicity, the evolved state is

$$\rho_t = \sum_{j,k=1}^{\tilde d} c_j c_k^* e^{-i(E_j-E_k)t}|j\rangle\langle k|. \quad \text{(A2)}$$

Defining the matrix elements $\langle j|A|k\rangle = A_{jk}$ one has

$$\begin{aligned}\langle |{\rm Tr}[A(\rho_t-\omega)]|^2\rangle_{L_T} &= \left\langle \left|\sum_{j\neq k}(c_k^* A_{kj} c_j)e^{-i(E_j-E_k)t}\right|^2\right\rangle_{L_T}\\ &= \sum_{j\neq k, n\neq l}(c_k^* A_{jk} c_j)(c_l^* A_{ln} c_n)^*\\ &\quad\times\langle e^{-i[(E_j-E_k)-(E_n-E_l)]t}\rangle_{L_T}. \quad \text{(A3)}\end{aligned}$$

Each energy gap can be labeled by $G_{(j,k)} = E_j - E_k$ with indexes $\alpha = (n,l)$ and $\beta = (j,k)$. In this way, we define a vector v and a Hermitian matrix M:

$$v_\alpha = v_{n,l} = c_l^* A_{ln} c_n, \quad M_{\alpha\beta} = \langle e^{i(G_\alpha - G_\beta)t}\rangle_{L_T}. \quad \text{(A4)}$$

With the above definitions, we can see

$$\begin{aligned}&\langle |{\rm Tr}[A(\rho_t-\omega)]|^2\rangle_{L_T}\\ &\quad= \sum_{\alpha\beta} v_\alpha^* M_{\alpha\beta} v_\beta \leqslant \|M\|\sum_\alpha |v_\alpha|^2\\ &\quad\leqslant \|M\|\sum_{i,j}|c_i|^2|c_j|^2|A_{ji}|^2 = \|M\|{\rm Tr}(A\omega A^\dagger\omega)\\ &\quad\leqslant \|M\|\sqrt{{\rm Tr}(A^\dagger A\omega^2){\rm Tr}(AA^\dagger\omega^2)}\\ &\quad\leqslant \|M\|\|A\|^2{\rm Tr}(\omega^2) = \|M\|\frac{\|A\|^2}{d_{\rm eff}}. \quad \text{(A5)}\end{aligned}$$

[6]Define $\Theta_T(t) = \frac{1}{T}$ for $t\in[0,T]$, and 0 otherwise. Then, $\langle f\rangle_T = \int_{-\infty}^{\infty}\Theta_T(t)f(t)dt$, and $\Theta_T(t) \leqslant \frac{5}{4}\frac{T}{T^2+(t-T/2)^2}$.

On the first step, from the definition of the spectral norm $\|M\|\|v\| \geqslant \|Mv\|$ so $\|v\|\|M\|\|v\| \geqslant \|v\|\|Mv\| \geqslant |v^\dagger M v|$. The other steps come from using the Frobenius inner product and Cauchy-Schwarz inequality, and from the fact that for pure states $\mathrm{Tr}[\omega^2] = 1/d_{\mathrm{eff}}$, with $d_{\mathrm{eff}} = (\sum_j |c_j|^2)^{-1}$.

By use of the identity $\langle e^{i\nu t}\rangle_{L_T} = e^{-|\nu|T} e^{i\nu T/2}$, we have $|M_{\alpha\beta}| = |\langle e^{i(G_\alpha - G_\beta)t}\rangle_{L_T}| = e^{-|G_\alpha - G_\beta|T}$ and, since M is a Hermitian matrix, standard results give

$$\|M\| \leqslant \max_\beta \sum_\alpha |M_{\alpha\beta}| = \max_\beta \sum_\alpha e^{-|G_\alpha - G_\beta|T}. \tag{A6}$$

We can now break the sum into intervals of width ϵ, centered around a given gap G_β. An interval ϵ can fit at most $\mathcal{N}(\epsilon)$ gaps which satisfy $(k+1/2)\epsilon > G_\alpha - G_\beta > (k-1/2)\epsilon$, which in turn implies $|G_\alpha - G_\beta| \geqslant (|k| - 1/2)\epsilon$. Therefore,

$$|M_{\alpha\beta}| \leqslant e^{-(|k|-1/2)\epsilon T}. \tag{A7}$$

For the case $k = 0$, we just use the fact that $|M_{\alpha\beta}| \leqslant 1$.

The sum is maximized by taking as many small values of $|k|$ as possible, and since there are $\tilde{d}(\tilde{d}-1)$ terms in total we have that

$$\begin{aligned}\max_\beta \sum_\alpha |M_{\alpha\beta}| &\leqslant \mathcal{N}(\epsilon)\left(1 + 2\sum_{k=1}^{\tilde{d}(\tilde{d}-1)} e^{-(k-1/2)\epsilon T}\right)\\ &= \mathcal{N}(\epsilon)\left(1 + 2e^{\epsilon T/2}\frac{e^{-\epsilon T}(e^{-\epsilon T \tilde{d}(\tilde{d}-1)} - 1)}{e^{-\epsilon T} - 1}\right)\\ &\leqslant \mathcal{N}(\epsilon)\left(1 + 2\frac{e^{-\epsilon T/2}}{1 - e^{-\epsilon T}}\right).\end{aligned} \tag{A8}$$

Finally, by using $\frac{1}{1-e^{-x}} \leqslant 1 + \frac{1}{x}$, we get

$$\begin{aligned}\langle |\mathrm{Tr}[A(\rho_t - \omega)]|^2\rangle_T &\leqslant \frac{5\pi}{2}\frac{\|A\|^2}{d_{\mathrm{eff}}}\mathcal{N}(\epsilon)\left(\frac{1}{2} + e^{-\epsilon T/2} + \frac{e^{-\epsilon T/2}}{\epsilon T}\right)\\ &\leqslant \frac{5\pi}{2}\frac{\|A\|^2}{d_{\mathrm{eff}}}\mathcal{N}(\epsilon)\left(\frac{3}{2} + \frac{1}{\epsilon T}\right).\end{aligned} \tag{A9}$$

Comparing this expression to the result in [1], there is an improvement in the bound of order $\sim \log_2(\tilde{d})$.

The result is taken in the original paper by Short and Farrelly as a stepping stone to obtain bounds on the time scale of equilibration with respect to the distinguishability. By the same procedure they take, we obtain that the average distinguishability for a set of measurements $\mathcal{M}$ satisfies

$$\langle D_{\mathcal{M}}(\rho_t, \omega)\rangle_T \leqslant \frac{\mathcal{S}(\mathcal{M})}{4}\sqrt{\frac{5\pi\mathcal{N}(\epsilon)}{2d_{\mathrm{eff}}}\left(\frac{3}{2} + \frac{1}{\epsilon T}\right)}, \tag{A10}$$

where $\mathcal{S}(\mathcal{M})$ is the total number of outcomes of all the possible measurements.

A simple estimate illustrates how long these bounds on the time scales still are. If one assumes the energy levels are more or less equally distributed, the minimum distance between energy gaps scales as $\epsilon_{\min} \leqslant \frac{\Delta U}{\tilde{d}^2}$, with ΔU being the total energy range. This gives an equilibration time that scales very roughly as $T_{\mathrm{eq}} \sim \frac{1}{d_{\mathrm{eff}}\epsilon} > \frac{\tilde{d}}{\Delta U}$, which is terribly long for systems composed of more than a few particles.

APPENDIX B: THEOREM 1

For a general initial state given by $\rho = \sum_{jk} \rho_{jk}|j\rangle\langle k|$, the purity of ω_{L_T} can be written as

$$\begin{aligned}\mathrm{Tr}\big[\omega_{L_T}^2\big] = \mathrm{Tr}[\omega_{L_T}\omega_{L_T}^\dagger] &= \mathrm{Tr}\Bigg[\sum_{n,m}\rho_{nm}\langle e^{-i(E_n - E_m)t}\rangle_{L_T}|n\rangle\langle m|\\ &\quad\times \sum_{j,k}\rho_{jk}^*\langle e^{i(E_j - E_k)t}\rangle_{L_T}|k\rangle\langle j|\Bigg]\\ &= \sum_{j,k}|\rho_{jk}|^2|\langle e^{-i(E_j - E_k)t}\rangle_{L_T}|^2\\ &\leqslant \sum_{j,k}\rho_{jj}\rho_{kk}|\langle e^{-i(E_j - E_k)t}\rangle_{L_T}|^2\\ &= \sum_{j,k}p_j p_k|\langle e^{-i(E_j - E_k)t}\rangle_{L_T}|^2,\end{aligned} \tag{B1}$$

where the previous to the last line is an equality for an initially pure state, and the inequality follows in general from positivity of the density operator.[7]

By use of the identity $\langle e^{i\nu t}\rangle_{L_T} = e^{-|\nu|T}e^{i\nu T/2}$ we can in turn see

$$\mathrm{Tr}\big[\omega_{L_T}^2\big] \leqslant \sum_{jk} p_j p_k e^{-2|E_j - E_k|T}, \tag{B2}$$

the above being an equality for pure states.

To see the connection to $\eta_{\frac{1}{T}}$, we define the function

$$g(x) = \begin{cases} 1, & \text{if } x \in [0,1) \\ 0, & \text{otherwise.}\end{cases} \tag{B3}$$

This definition is important because it allows us to upper bound the exponential as

$$e^{-|x|} \leqslant \sum_{n=0}^{\infty} e^{-n\delta} g\left(\frac{|x|}{\delta} - n\right), \quad \forall\, \delta > 0. \tag{B4}$$

So, we have

$$\begin{aligned}\mathrm{Tr}\big[\omega_{L_T}^2\big] &\leqslant \sum_{n=0}^{\infty} e^{-n\delta}\sum_j p_j \sum_k p_k\\ &\quad\times g\left(\frac{2|E_j - E_k|T}{\delta} - n\right)\\ &= \sum_{n=0}^{\infty} e^{-n\delta}\sum_j p_j \sum_{\substack{k:\\ (2|E_j - E_k|\frac{T}{\delta} - n)\in[0,1)}} p_k\\ &\leqslant \sum_{n=0}^{\infty} e^{-n\delta}\sum_j p_j\left[\sum_{\substack{k:\\ E_k\in I_-}} p_k + \sum_{\substack{k:\\ E_k\in I_+}} p_k\right]\end{aligned}$$

[7] $\langle v|\rho|v\rangle \geqslant 0$ for all $|v\rangle$, which applies in particular to $|v\rangle = a|j\rangle + b|k\rangle$, so $\left(\begin{smallmatrix}\rho_{jj} & \rho_{jk}\\ \rho_{kj} & \rho_{kk}\end{smallmatrix}\right)$ is positive too, and, since the determinant must be greater than or equal to zero, $|\rho_{jk}|^2 \leqslant \rho_{jj}\rho_{kk}$.

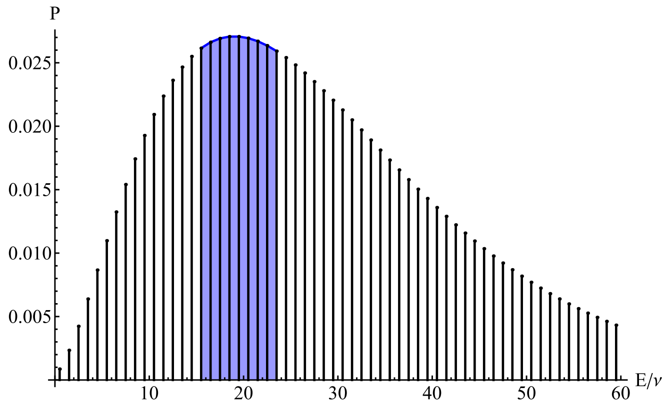

FIG. 2. (Color online) Graphic illustration of η_ϵ. The vertical lines display the probability distribution in energy space of a three-dimensional harmonic oscillator with energy levels $E_n = (n+1/2)\nu$, under the Boltzmann distribution (accounting for degeneracies) with temperature $= 10\nu$. The blue shaded region represents $\eta_{8\nu}$ (the maximum probability that can be found inside an energy interval of size 8ν).

$$\begin{aligned}&\leqslant \sum_{n=0}^{\infty} e^{-n\delta} \sum_j p_j(2\eta_{\frac{\delta}{2T}})\\&= \frac{2\eta_{\frac{\delta}{2T}}}{1-e^{-\delta}},\end{aligned} \tag{B5}$$

where $I_+ = [E_+, E_+ + \frac{\delta}{2T})$, $I_- = (E_- - \frac{\delta}{2T}, E_-]$, $E_\pm = E_j \pm \frac{n\delta}{2T}$, and the inequality in the penultimate line applies for any combination of n and j.

The important fact is that δ is ours to define, and determines the balance between the quotient and η. We can, for instance, set it to 2:

$$\mathrm{Tr}\big[\omega_{L_T}^2\big] \leqslant \frac{2\eta_{\frac{1}{T}}}{1-e^{-2}} < 2.32\eta_{\frac{1}{T}} \tag{B6}$$

and we know the system has equilibrated when the total probability inside any energy interval of size $1/T$ is small. We can also manually fix the energy interval with $\delta = 2T\Delta E$, so

$$\mathrm{Tr}\big[\omega_{L_T}^2\big] \leqslant \frac{2\eta_{\Delta E}}{1-e^{-2T\Delta E}}, \tag{B7}$$

which still leaves us with a free variable, but clearly singles out the time dependence.

In any case, we have that

$$\langle \mathrm{Tr}[\rho_t P]\rangle_T \leqslant \tfrac{5\pi}{4}\sqrt{K\,\tfrac{2\eta_{\frac{1}{T}}}{1-e^{-2}}} < 6\sqrt{\eta_{\frac{1}{T}}K}. \tag{B8}$$

For clarity, Fig. 2 displays a graphic illustration of η_ϵ.

1. Gaussian distribution example

Taking a pure initial state whose probability distribution can be approximated by the following continuous function of energy

$$p(E) = \frac{1}{\sqrt{2\pi}\sigma_E} e^{-\frac{E^2}{2\sigma_E^2}}, \tag{B9}$$

we show that the purity of ω_{L_T} can be approximated by

$$\mathrm{Tr}\big[\omega_{L_T}^2\big] \approx \frac{1}{2\sqrt{\pi}\sigma_E T}, \tag{B10}$$

for $\sigma_E T \gg 1$. Meanwhile, we also show that η satisfies

$$\frac{2\eta_{\frac{1}{T}}}{1-e^{-2}} < \frac{3.28}{2\sqrt{\pi}\sigma_E T}, \tag{B11}$$

and therefore

$$\eta_{\frac{1}{T}} < \frac{0.4}{\sigma_E T}. \tag{B12}$$

It is interesting to see that, despite the approximations taken, the function η gives good estimates for the purity of the average, with much simpler calculations.

a. Calculations

The Fourier transform of the energy distribution defined in Eq. (B9) is

$$\mu(t) = \int_{-\infty}^{\infty} p(E) e^{-iEt} dE = e^{-\frac{\sigma_E^2 t^2}{2}}. \tag{B13}$$

Thus, using the continuous version of Eq. (B2) and assuming a pure state,

$$\begin{aligned}\mathrm{Tr}\big[\omega_{L_T}^2\big] &= \iint_{-\infty}^{\infty} dE\, dE'\, p(E)p(E') e^{-2|E-E'|T}\\&= \iint dE\, dE'\, p(E)p(E') \int_{-\infty}^{\infty} \frac{dt}{\pi} \frac{e^{i(E-E')2t}T}{T^2+t^2}\\&= \int_{-\infty}^{\infty} \frac{dt}{\pi} \frac{T}{T^2+t^2} |\mu(2t)|^2\\&= \frac{T}{\pi}\int_{-\infty}^{\infty} \frac{e^{-4\sigma_E^2 t^2}}{T^2+t^2} dt\\&= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{-4\sigma_E^2 T^2 x^2}}{1+x^2} dx\\&= e^{4\sigma_E^2 T^2}[1-\mathrm{erf}(2\sigma_E T)],\end{aligned} \tag{B14}$$

where the second line uses the inverse Fourier transform of $e^{-2|E-E'|T}$.

We can use the expansion of the error function

$$\mathrm{erf}(x) = 1 - \frac{e^{-x^2}}{\sqrt{\pi}x} + e^{-x^2}\mathcal{O}(x^{-3}), \tag{B15}$$

which, for $\sigma_E T$ large enough, results in Eq. (B10).

b. Comparison to η

As we have established previously,

$$\mathrm{Tr}\big[\omega_{L_T}^2\big] < 2.32\eta_{\frac{1}{T}}. \tag{B16}$$

Since the density $p(E)$ is a Gaussian, η_ϵ is obviously at its center. Thus,

$$\begin{aligned}\eta_{\frac{1}{T}} &= \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{E^2}{2\sigma^2}} dE \leqslant \frac{1}{\sqrt{2\pi}\sigma T},\\ \mathrm{Tr}\big[\omega_{L_T}^2\big] &< \frac{3.28}{2\sqrt{\pi}\sigma T},\end{aligned} \tag{B17}$$

where the integral was trivially approximated by $\int_0^\epsilon p(E)dE < \epsilon p(0)$.

c. Infinite time limit

The upper bounds we calculated in the example for $\mathrm{Tr}[\omega_{L_T}^2]$ tend to zero as T tends to infinity, which seems to contradict the fact that for the infinite time average state ω, $\mathrm{Tr}[\omega^2]$ is not zero. The reason this occurs is that we need to be careful when averaging expressions such as $\sum p_j p_k \exp[-2i(E_j - E_k)t]$ over energy levels. The terms with $E_j = E_k$ in the integrals arising from averages over $p(E)$ do not contribute (they are of measure zero), whereas in the finite sum they did, giving $\sum_{j=k} p_j p_k = 1/d_{\mathrm{eff}}$. Another way of looking at this is that taking the continuous limit implies taking $d_{\mathrm{eff}} = \infty$.

APPENDIX C: TYPICAL PROJECTORS

Here, we prove Theorems 2 and 3.

1. Proof of Theorem 2

Here, we use the fact (well known from representation theory) that, for any operator M,

$$\langle U^{\otimes 2} M (U^{\otimes 2})^\dagger \rangle_U = \alpha \Pi_S + \beta \Pi_A, \tag{C1}$$

where $\Pi_S = (\mathbb{1}^{\otimes 2} + S)/2$ and $\Pi_A = (\mathbb{1}^{\otimes 2} - S)/2$ are the projectors onto the symmetric and antisymmetric subspaces of dimensions $\frac{d(d+1)}{2}$ and $\frac{d(d-1)}{2}$, respectively, and S is the swap operator on $\mathcal{H} \otimes \mathcal{H}$: $S|a\rangle|b\rangle = |b\rangle|a\rangle$. Given that $\mathrm{Tr}[SA \otimes B] = \mathrm{Tr}[AB]$, it is straightforward to see that, for $M = P^{\otimes 2}$,

$$\begin{aligned}\alpha &= \frac{\mathrm{Tr}[\Pi_S P^{\otimes 2}]}{\mathrm{Tr}[\Pi_S]} = \frac{1}{d(d+1)}[\mathrm{Tr}[P \otimes P] + \mathrm{Tr}[P^2]] \\ &= \frac{K(K+1)}{d(d+1)};\end{aligned} \tag{C2}$$

and similarly $\beta = \frac{\mathrm{Tr}[\Pi_A P_U^{\otimes 2}]}{\mathrm{Tr}[\Pi_A]} = \frac{K(K-1)}{d(d-1)}$.

Finally, going back to the distinguishability and using the fact that $\mathrm{Tr}[\rho_t \omega] = \mathrm{Tr}[\omega^2]$,

$$\begin{aligned}&\langle D_{P_U}(\rho_t,\omega)^2\rangle_U \\ &= \langle \{\mathrm{Tr}[P_U(\rho_t - \omega)]\}^2 \rangle_U \\ &= \mathrm{Tr}[\langle P_U \otimes P_U \rangle_U (\rho_t - \omega)^{\otimes 2}] \\ &= \mathrm{Tr}[(\alpha \Pi_S + \beta \Pi_A)(\rho_t - \omega) \otimes (\rho_t - \omega)] \\ &= \frac{\alpha+\beta}{2}(\mathrm{Tr}[\rho_t - \omega])^2 + \frac{\alpha - \beta}{2}\mathrm{Tr}[(\rho_t - \omega)^2] \\ &= \frac{K}{2d}\left(\frac{K+1}{d+1} - \frac{K-1}{d-1}\right)\left(\mathrm{Tr}[\rho_t^2] + \mathrm{Tr}[\omega^2] - 2\,\mathrm{Tr}[\rho_t\omega]\right) \\ &= \frac{K}{d}\frac{d-K}{d^2-1}\mathrm{Tr}[\rho_t^2 - \omega^2].\end{aligned} \tag{C3}$$

2. Proof of Theorem 3

To simplify notation in the proof, we will assume $d > 2$ and set $d' = \dim \mathcal{H}' = d - 1$ and $K' = \mathrm{rank}\, P_U = K - 1$. Again, the quantity we are interested in will be

$$\begin{aligned}\langle D_{\Pi_U}(\rho_t,\omega)\rangle_U &= \langle |\mathrm{Tr}[\Pi_U(\rho_t - \omega)]| \rangle_U \\ &\leqslant \sqrt{\langle (\mathrm{Tr}[\Pi_U(\rho_t - \omega)])^2 \rangle_U} \\ &= \sqrt{\mathrm{Tr}[\langle \Pi_U \otimes \Pi_U \rangle_U (\rho_t - \omega)^{\otimes 2}]}.\end{aligned} \tag{C4}$$

Using the results from Sec. C 1, it is easy to see that

$$\begin{aligned}&\langle \Pi_U \otimes \Pi_U \rangle_U \\ &= \langle \rho_0 \otimes \rho_0 + \rho_0 \otimes P_U + P_U \otimes \rho_0 + P_U \otimes P_U \rangle_U \\ &= \rho_0 \otimes \rho_0 + \rho_0 \otimes \langle P_U \rangle_U \\ &\quad + \langle P_U \rangle_U \otimes \rho_0 + \langle P_U \otimes P_U \rangle_U \\ &= \rho_0 \otimes \rho_0 + \frac{K'}{d'}\rho_0 \otimes \mathbb{1}' + \frac{K'}{d'}\mathbb{1}' \otimes \rho_0 \\ &\quad + \frac{K'(K'+1)}{d'(d'+1)}\Pi'_S + \frac{K'(K'-1)}{d'(d'-1)}\Pi'_A,\end{aligned} \tag{C5}$$

where $\mathbb{1}' = \mathbb{1} - \rho_0$, $\Pi'_S = \mathbb{1}'^{\otimes 2}\Pi_S \mathbb{1}'^{\otimes 2}$, and $\Pi'_A = \mathbb{1}'^{\otimes 2}\Pi_A \mathbb{1}'^{\otimes 2}$. So,

$$\begin{aligned}&\langle \mathrm{Tr}[\Pi_U(\rho_t - \omega)]^2 \rangle_U \\ &= \mathrm{Tr}[\langle \Pi_U \otimes \Pi_U \rangle_U (\rho_t - \omega)^{\otimes 2}] \\ &= f(t)^2 - 2f(t)^2\frac{K'}{d'} + \frac{K'(K'+1)}{d'(d'+1)}\mathrm{Tr}[\Pi_S A(t)] \\ &\quad + \frac{K'(K'-1)}{d'(d'-1)}\mathrm{Tr}[\Pi_A A(t)],\end{aligned} \tag{C6}$$

where $A(t) = [(\mathbb{1} - \rho_0)(\rho_t - \omega)(\mathbb{1} - \rho_0)]^{\otimes 2}$, and $f(t) = \mathrm{Tr}[\rho_0(\rho_t - \omega)]$. Since

$$\begin{aligned}\mathrm{Tr}[\Pi_S A(t)] &= \mathrm{Tr}\left[\frac{\mathbb{1}^{\otimes 2} + S}{2}[(\mathbb{1} - \rho_0)(\rho_t - \omega)(\mathbb{1} - \rho_0)]^{\otimes 2}\right] \\ &= \frac{1}{2}\mathrm{Tr}[(\mathbb{1} - \rho_0)(\rho_t - \omega)(\mathbb{1} - \rho_0)(\rho_t - \omega)] \\ &\quad + \frac{1}{2}(\mathrm{Tr}[\rho_t - \omega - \rho_0(\rho_t - \omega)])^2 \\ &\leqslant \frac{1}{2}\left(1 - \frac{1}{d_{\mathrm{eff}}}\right) + \frac{1}{2}[f(t)]^2,\end{aligned} \tag{C7}$$

where the last line is due to $\mathrm{Tr}[\Pi X \Pi X] \leqslant \mathrm{Tr}[X^2]$ for a projector Π (in this case, $\mathbb{1} - \rho_0$), which follows from the Cauchy-Schwarz inequality $\mathrm{Tr}[\Pi X \Pi X] \leqslant \sqrt{\mathrm{Tr}[X^2]\,\mathrm{Tr}[\Pi X \Pi X]}$. Thus, since

$$\begin{aligned}\mathrm{Tr}[\Pi_A A(t)] &= \mathrm{Tr}\left[\frac{\mathbb{1}^{\otimes 2} - S}{2}[(\mathbb{1} - \rho_0)(\rho_t - \omega)]^{\otimes 2}\right] \\ &\leqslant \frac{1}{2}\left(f(t)^2 - 1 + \frac{1}{d_{\mathrm{eff}}}\right),\end{aligned} \tag{C8}$$

we have

$$\begin{aligned}&\langle \mathrm{Tr}[\Pi_U(\rho_t - \omega)]^2 \rangle_U \\ &\leqslant f(t)^2\left(1 + \frac{1}{2}\frac{K'(K'+1)}{d'(d'+1)} + \frac{1}{2}\frac{K'(K'-1)}{d'(d'-1)} - 2\frac{K'}{d'}\right)\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\left(1-\frac{1}{d_{\text{eff}}}\right)\left[\frac{K'(K'+1)}{d'(d'+1)}-\frac{K'(K'-1)}{d'(d'-1)}\right]\\
&=f(t)^2\left(1-2\frac{K'}{d'}+\frac{K'}{d'}\frac{K'd'-1}{d'^2-1}\right)\\
&+\left(1-\frac{1}{d_{\text{eff}}}\right)\frac{K'}{d'}\frac{d'-K'}{d'^2-1}\\
&\leqslant f(t)^2+\frac{1}{4d'}. \qquad \text{(C9)}
\end{aligned}
$$

The last line can be derived from the fact that the first parentheses in the penultimate equation is maximized by $K'=0$, and the second term is maximized by $K'=d'/2$, along with $d_{\text{eff}}\leqslant d$ and $\frac{d'^2}{(d'-1)(d'+1)^2}\leqslant\frac{1}{d'}$.

3. Proof of typical initial distinguishability

To show that observables with a definite initial value are typically out of equilibrium (and thus undergo a nontrivial equilibration process) we consider the initial distinguishability between ρ_0 and ω for a measurement of Π_U, averaged over U. As before, we will set $d'=\dim\mathcal{H}'=d-1$ and $K'=\operatorname{rank}P_U=K-1$:

$$
\begin{aligned}
\langle D_{\Pi_U}(\rho_0,\omega)\rangle_U &= \langle|\text{Tr}[\Pi_U(\rho_0-\omega)]|\rangle_U\\
&=\langle(1-\text{Tr}[\Pi_U\omega])\rangle_U\\
&=1-\text{Tr}\left[\left[\rho_0+\frac{K'}{d'}(\mathbb{1}-\rho_0)\right]\omega\right]\\
&=\left(1-\frac{K'}{d'}\right)(1-\text{Tr}[\rho_0\omega])\\
&\geqslant\left(1-\frac{K-1}{d-1}\right)\left(1-\frac{1}{d_{\text{eff}}}\right), \qquad \text{(C10)}
\end{aligned}
$$

where the last line is an equality if ρ_0 is pure.

Note that because refining a measurement (by splitting one outcome into many) can only increase the distinguishability, it follows that

$$
\left\langle D_{\mathcal{M}_U^{\rho_0}}(\rho_t,\omega)\right\rangle_U\geqslant\left(1-\frac{K-1}{d-1}\right)\left(1-\frac{1}{d_{\text{eff}}}\right), \qquad \text{(C11)}
$$

where here K is the rank of the measurement projector containing ρ_0.

4. Proof of Corollary 2

Denoting by K_j the rank P_j, we have that

$$
\sum_j K_j=d' \qquad \text{(C12)}
$$

and

$$
\begin{aligned}
&\left\langle D_{\mathcal{M}_U^{\rho_0}}(\rho_t,\omega)\right\rangle_U\\
&=\frac{1}{2}\langle D_{\rho_0+P_{1U}}(\rho_t,\omega)\rangle_U+\frac{1}{2}\sum_{j=2}^{N}\langle D_{P_{jU}}(\rho_t,\omega)\rangle_U\\
&\leqslant\frac{1}{2}\sqrt{\langle D_{\rho_0+P_{1U}}(\rho_t,\omega)^2\rangle_U}+\frac{1}{2}\sum_{j=2}^{N}\sqrt{\langle D_{P_{jU}}(\rho_t,\omega)^2\rangle_U}.
\end{aligned} \qquad \text{(C13)}
$$

Following the proof in Appendix C 2 above, and using the fact that $1-\frac{1}{d_{\text{eff}}}\leqslant1-\frac{1}{d}=\frac{d'}{d'+1}$, leads to

$$
\left\langle D_{P_{jU}}(\rho_t,\omega)^2\right\rangle_U\leqslant f(t)^2\frac{K_j}{d'}\frac{K_jd'-1}{d'^2-1}+\frac{K_j}{d'+1}\frac{d'-K_j}{d'^2-1} \qquad \text{(C14)}
$$

and

$$
\begin{aligned}
&\langle D_{\rho_0+P_{1U}}(\rho_t,\omega)^2\rangle_U\\
&\leqslant f(t)^2\left(1-2\frac{K_1}{d'}+\frac{K_1}{d'}\frac{K_1d'-1}{d'^2-1}\right)+\frac{K_1}{d'+1}\frac{d'-K_1}{d'^2-1}\\
&\leqslant f(t)^2\left(1+\frac{K_1}{d'}\frac{K_1d'-1}{d'^2-1}\right)+\frac{K_1}{d'+1}\frac{d'-K_1}{d'^2-1}\\
&=f(t)^2+\left\langle D_{P_{1U}}(\rho_t,\omega)^2\right\rangle_U. \qquad \text{(C15)}
\end{aligned}
$$

By using the fact that $\sqrt{a+b}\leqslant\sqrt{a}+\sqrt{b}$ for $a,b\geqslant0$, this leads to

$$
\left\langle D_{\mathcal{M}_U^{\rho_0}}(\rho_t,\omega)\right\rangle_U\leqslant\frac{1}{2}|f(t)|+\frac{1}{2}\sum_{j=1}^{N}\sqrt{\left\langle D_{P_{jU}}(\rho_t,\omega)^2\right\rangle_U}. \qquad \text{(C16)}
$$

Through the method of Lagrange multiplier, it is easy to see that the sum in Eq. (C16), expressed in terms of the K_j's through Eq. (C14) and constrained by Eq. (C12), is maximized by taking all P_j to be of equal rank. This rank must then be $K_j=d'/N$. Substituting that into Eq. (C14), and using the inequalities $\frac{d'^2-N}{d'^2-1}<1$, $1-1/N<1$, and $d'^3\leqslant(d'+1)^2(d'-1)$,

$$
\begin{aligned}
&\left\langle D_{\mathcal{M}_U^{\rho_0}}(\rho_t,\omega)\right\rangle_U\\
&\leqslant\frac{1}{2}|f(t)|+\frac{1}{2}N\sqrt{\frac{f(t)^2}{N^2}\frac{d'^2-N}{d'^2-1}+\frac{d'/N}{d'+1}\frac{d'\left(1-\frac{1}{N}\right)}{d'^2-1}}\\
&\leqslant\frac{1}{2}|f(t)|+\frac{1}{2}\sqrt{f(t)^2+\frac{Nd'^2}{(d'+1)^2(d'-1)}}\\
&\leqslant|f(t)|+\frac{1}{2}\sqrt{\frac{N}{d'}}. \qquad \text{(C17)}
\end{aligned}
$$

APPENDIX D: SLOW EQUILIBRATION

The *slow equilibration* result can be rigorously stated as the following theorem.

Theorem 4 (Slow equilibration). Given any Hamiltonian, any pure state $|\psi(t)\rangle\in\mathcal{H}$ with effective dimension d_{eff}, any positive integer $K\ll d_{\text{eff}}$, and any $\epsilon>0$; take σ_E to be the standard deviation in energy of $|\psi\rangle$, and $P_{\mathcal{H}_K}$ to be the projector onto the subspace

$$
\mathcal{H}_K=\text{span}\{|\psi(j\tau)\rangle\mid j=0,\ldots,K-1\}, \qquad \text{(D1)}
$$

with $\tau = 2\epsilon/\sigma_E$; then the distinguishability satisfies the following two equations:[8]

$$D_{P_{\mathcal{H}_K}}(\rho_t,\omega) \geqslant 1 - \epsilon^2 - \sqrt{\frac{K}{d_{\text{eff}}}}, \quad \forall\, t \in \left[0, \frac{(2K-1)\epsilon}{\sigma_E}\right] \tag{D2}$$

and

$$\left\langle D_{P_{\mathcal{H}_K}}(\rho_t,\omega)\right\rangle_{T\to\infty} \leqslant 2\sqrt{\frac{K}{d_{\text{eff}}}} \ll 1 \tag{D3}$$

(it is above some constant for long times, but still equilibrates eventually).

Proof. Since τ is a very small time step, the overlap between $|\psi(0)\rangle$ and $|\psi(\tau)\rangle$ is nearly 1. To prove this, we write $|\psi(t)\rangle$ in the energy basis

$$|\psi(t)\rangle = \sum_n^{\tilde{d}} c_n e^{-iE_n t}|n\rangle \tag{D4}$$

and calculate its internal product with its initial state

$$\begin{aligned}
|\langle\psi(t)|\psi(0)\rangle|^2 &= \left|\sum_n^{\tilde{d}} |c_n|^2 e^{-iE_n t}\right|^2 \\
&= \sum_{nm}^{\tilde{d}} |c_n|^2|c_m|^2 \cos[(E_n - E_m)t] \\
&\geqslant 1 - \frac{t^2}{2}\sum_{nm}^{\tilde{d}} |c_n|^2|c_m|^2\left(E_n^2 + E_m^2 - 2E_nE_m\right) \\
&= 1 - (2\overline{E^2} - 2\overline{E}^2)\frac{t^2}{2} \\
&= 1 - \sigma_E^2 t^2,
\end{aligned} \tag{D5}$$

where σ_E is the standard deviation in energy. So, we have that

$$|\langle\psi(t)|\psi(0)\rangle|^2 \geqslant 1 - \epsilon^2 \tag{D6}$$

$\forall\, t$ such that $|t| \leqslant \tau/2 = \epsilon/\sigma_E$. This trivially implies

$$|\langle\psi(t)|\psi(t')\rangle|^2 \geqslant 1 - \epsilon^2 \tag{D7}$$

$\forall\, t,t'$ such that $|t - t'| \leqslant \tau/2$.

Meanwhile, $\mathcal{H}_K$ contains, by definition, all projectors $|\psi(j\tau)\rangle\langle\psi(j\tau)|$ for j up to $K-1$. Therefore, for any time t up to $(K - {}^1\!/_2)\tau$, the state $|\psi(t)\rangle$ is very close to one of these projectors.

In other words, there is always a value of $0 \leqslant j \leqslant K-1$ such that $|t - j\tau| \leqslant \tau/2$ and

$$\begin{aligned}
\mathrm{Tr}[\rho_t P_{\mathcal{H}_K}] &= \langle\psi(t)|[P_{j\tau} + P_{j\tau}^{\perp}]|\psi(t)\rangle \\
&\geqslant |\langle\psi(j\tau)|\psi(t)\rangle|^2 \geqslant 1 - \epsilon^2,
\end{aligned} \tag{D8}$$

where $P_t = |\psi(t)\rangle\langle\psi(t)|$ and $P_t^{\perp} = P_{\mathcal{H}_K} - P_t$. This directly leads to Eq. (D2):

$$\begin{aligned}
D_{P_{\mathcal{H}_K}}(\rho_t,\omega) &= \left|\mathrm{Tr}\left[P_{\mathcal{H}_K}(\rho_t - \omega)\right]\right| \\
&\geqslant \mathrm{Tr}\left[P_{\mathcal{H}_K}\rho_t\right] - \mathrm{Tr}\left[P_{\mathcal{H}_K}\omega\right] \\
&\geqslant 1 - \epsilon^2 - \sqrt{\frac{K}{d_{\text{eff}}}}.
\end{aligned} \tag{D9}$$

Equation (D3) is easily obtained from the Cauchy-Schwarz inequality

$$\begin{aligned}
\left\langle D_{P_{\mathcal{H}_K}}(\rho_t,\omega)\right\rangle_{T\to\infty} &= \langle|\,\mathrm{Tr}[P_{\mathcal{H}_K}(\rho_t - \omega)]|\rangle_{T\to\infty} \\
&\leqslant \left\langle \mathrm{Tr}\left[P_{\mathcal{H}_K}\rho_t\right] + \mathrm{Tr}\left[P_{\mathcal{H}_K}\omega\right]\right\rangle_{T\to\infty} \\
&= 2\,\mathrm{Tr}\left[P_{\mathcal{H}_K}\omega\right] \leqslant 2\sqrt{\mathrm{Tr}\left[{P_{\mathcal{H}_K}}^2\right]\mathrm{Tr}[\omega^2]} \\
&\leqslant 2\sqrt{\frac{K}{d_{\text{eff}}}} \ll 1.
\end{aligned} \tag{D10}$$

■

The role played by Eqs. (D2) and (D3) is simple. (i) The system obviously has not equilibrated, and is still distinguishable from its equilibrium state, as long as $D_{P_{\mathcal{H}_K}}(\rho_t,\omega)$ is significantly above zero. (ii) On the other hand, the rank of $P_{\mathcal{H}_K}$ is small enough that any system spread over many energy levels will equilibrate with respect to it.

For instance, if we take $K = d_{\text{eff}}/1000$ (which is extremely large) and $\epsilon = \frac{1}{2}$, we have

$$D_{P_{\mathcal{H}_K}}(\rho_t,\omega) \geqslant \frac{1}{2}, \quad \forall\, t \in \left[0, \frac{d_{\text{eff}}}{1000\sigma_E}\right]. \tag{D11}$$

For systems composed of many particles, we would typically expect $\sigma_E \sim \ln(d_{\text{eff}})$ leading to the system taking a time of order $\frac{d_{\text{eff}}}{\ln(d_{\text{eff}})}$ to equilibrate with respect to this measurement.

To illustrate how large this time scale can be, we describe now a simple example, the time scale depends only on σ_E and d_{eff}, and is largely independent on the details. Consider a system of L weakly interacting qubits with level spacing $\delta E = 10^{-18}$ J, the order of the excitation energy in atoms. Defining each qubit to have equal population on each level, simple calculations give $\sigma_E \approx \sqrt{L}\delta E$ and $d_{\text{eff}} \approx 2^L$, and we get $T_{\text{eq}}^{\text{slow}} > \frac{\hbar d_{\text{eff}}}{1000\sigma_E} \approx 2^L L^{-\frac{1}{2}} 10^{-19}$ s.[9] Then, taking as little as 125 qubits already gives $T_{\text{eq}}^{\text{slow}} \gtrsim 4.10^{17}$ s, nearly the age of the universe and increasing exponentially with L. In contrast, for the same number of particles, the average distinguishability of a typical measurement falls below 10^{-3} in a time scale of $T_{\text{eq}}^{\text{typ}} \lesssim \frac{6000^2\hbar}{\sigma_E} \approx 3\times 10^{-10}$ s. This typical time scale decreases with $L^{-\frac{1}{2}}$, becoming even smaller for macroscopic systems, and is obtained from Theorem 3 by assuming $\eta_{\frac{1}{T}} \lesssim 1/\sigma_E T$ as discussed in the main text.

Of course, the construction in Theorem 4 is not the only possibility and indeed an alternative construction is given in [8]. For instance, one can easily define measurements with

[8] The time range in Eq. (D2) can be increased to $2K\epsilon/\sigma_E$ by a slightly more complicated construction of $\mathcal{H}_K$.

[9] Throughout the paper, we choose units such that $\hbar = 1$. Only in this example we adopt S. I. units for illustration purposes.

a larger number of outcomes, which also obey Eq. (D2) for at least as long as $D_{P_{\mathcal{H}_K}}$ (see Appendix E). It is also worth mentioning that this theorem trivially extends to the existence of an observable and whose expectation value takes a long time to equilibrate since $P_{\mathcal{H}_K}$ is, of course, an observable. The distinguishability simply presents a stronger definition of equilibration.

APPENDIX E: EXTENSION TO N OUTCOMES

Theorem 1 can be generalized to N-outcome measurements $\mathcal{M} = \{P_1, \ldots, P_N\}$ with the bound $\langle D_{\mathcal{M}}(\rho_t,\omega)\rangle_T \leqslant \frac{c}{2}\sqrt{\eta_{\frac{1}{T}}}\sum_{i=1}^{N}\sqrt{k_i}$ where $k_i = \min\{\text{rank } P_i, d - \text{rank } P_i\}$.

There are several ways one could extend Theorem 4. One is to simply divide $\mathcal{H}_K$ into $N-1$ smaller subspaces. Then, one has $\mathcal{M} = \{P_{\mathcal{H}_{K1}}, \ldots, P_{\mathcal{H}_{KN-1}}, \mathbb{1} - P_{\mathcal{H}_K}\}$, and the resulting distinguishability

$$\begin{aligned} D_{\mathcal{M}}(\rho_t,\omega) &= \frac{1}{2}\sum_{n=1}^{N-1}\left|\mathrm{Tr}\left[P_{\mathcal{H}_{Kn}}(\rho_t-\omega)\right]\right| \\ &\quad + \frac{1}{2}\left|\mathrm{Tr}\left[(\mathbb{1}-P_{\mathcal{H}_K})(\rho_t-\omega)\right]\right| \\ &\geqslant \frac{1}{2}\left|\mathrm{Tr}\left[\sum_{n=1}^{N-1}P_{\mathcal{H}_{Kn}}(\rho_t-\omega)\right]\right| \\ &\quad + \frac{1}{2}\left|\mathrm{Tr}\left[(\mathbb{1}-P_{\mathcal{H}_K})(\rho_t-\omega)\right]\right| \\ &= D_{P_{\mathcal{H}_K}}(\rho_t,\omega) \end{aligned} \tag{E1}$$

takes at least as long to equilibrate as $D_{P_{\mathcal{H}_K}}$.

[1] A. J. Short and T. C. Farrelly, New J. Phys. **14**, 013063 (2012).
[2] L. Masanes, A. J. Roncaglia, and A. Acín, Phys. Rev. E **87**, 032137 (2013).
[3] F. G. S. L. Brandão, P. Ćwikliński, M. Horodecki, P. Horodecki, J. K. Korbicz, and M. Mozrzymas, Phys. Rev. E **86**, 031101 (2012).
[4] Vinayak and M. Žnidarič, J. Phys. A: Math. Theor. **45**, 125204 (2012).
[5] M. Cramer, New J. Phys. **14**, 053051 (2012).
[6] M. van den Worm, B. C. Sawyer, J. J. Bollinger, and M. Kastner, New J. Phys. **15**, 083007 (2013).
[7] A. Hutter and S. Wehner, Phys. Rev. A **87**, 012121 (2013).
[8] S. Goldstein, T. Hara, and H. Tasaki, Phys. Rev. Lett. **111**, 140401 (2013).
[9] E. J. Torres-Herrera and L. F. Santos, Phys. Rev. A **89**, 043620 (2014).
[10] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E **79**, 061103 (2009).
[11] P. Reimann, New J. Phys. **12**, 055027 (2010).
[12] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghi, Phys. Rev. E **81**, 011109 (2010).
[13] N. Linden, S. Popescu, A. J. Short, and A. Winter, New J. Phys. **12**, 055021 (2010).
[14] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008).
[15] A. J. Short, New J. Phys. **13**, 053009 (2011).
[16] E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1998).
[17] W. Rudin, *Real and Complex Analysis*, Mathematics Series (McGraw-Hill, New York, 1987).
[18] S. Goldstein, T. Hara, and H. Tasaki, arXiv:1402.0324.