


Modern concepts of quantum equilibration do not rule out strange relaxation dynamics

Lars Knipschild^{*} and Jochen Gemmer[†]*Department of Physics, University of Osnabrück, D-49069 Osnabrück, Germany*

 (Received 24 July 2019; revised manuscript received 13 September 2019; accepted 10 March 2020; published 15 June 2020)

Numerous pivotal concepts have been introduced to clarify the puzzle of relaxation and/or equilibration in closed quantum systems. All of these concepts rely in some way on specific conditions on Hamiltonians H , observables A , and initial states ρ or combinations thereof. We numerically demonstrate and analytically argue that there is a multitude of pairs H, A that meet said conditions for equilibration and generate some typical expectation-value dynamics, which means $\langle A(t) \rangle \propto f(t)$ approximately holds for the vast majority of all initial states. Remarkably we find that, while restrictions on the $f(t)$ exist, they do not at all exclude $f(t)$ that are rather adverse or strange regarding thermal relaxation.

DOI: [10.1103/PhysRevE.101.062205](https://doi.org/10.1103/PhysRevE.101.062205)

I. INTRODUCTION

Since the beginning of the previous century, numerous concepts have been suggested to account for the emergence of irreversible thermodynamics from the underlying reversible unitary quantum dynamics [1]. Despite its long history, this is an active field of research, even to date. With the paper at hand we intend to show that, while all these concepts are certainly cornerstones of our understanding of the origins of the second law, they are not yet sufficient to rule out certain types of dynamics of observables that may be perceived as being at odds with the arrow of time and which we address as strange dynamics. But before embarking on a detailed demonstration of their existence, we first name some valuable approaches to equilibration in closed quantum systems, which nonetheless fail to exclude said strange dynamics. This list is neither intended to be complete nor to serve as a full fledged introduction to the subject; readers well acquainted with its items may skip it.

Equilibration on average. Statement on temporal fluctuations of, e.g., expectation values. Let the Hamiltonian H with $H|\epsilon_j\rangle = \epsilon_j|\epsilon_j\rangle$ have a sufficiently low number of equal energy gaps $\epsilon_i - \epsilon_j$ (nonresonance condition). Let furthermore the effective dimension $d_{\text{eff}} := (\sum_j \langle \epsilon_j | \rho | \epsilon_j \rangle^2)^{-1}$ with ρ being the initial state, be large, i.e., $d_{\text{eff}} \gg 1$. Then excursions of an expectation value $\langle A(t) \rangle$ from some equilibrium value $A_{\text{eq}} := \int_0^T \langle A(t) \rangle dt / T$ are rare in the sense that $(\langle A(t) \rangle - A_{\text{eq}})^2 \ll \|A\|$ for most t from $[0, T]$ [2–4].

Eigenstate thermalization hypothesis (ETH). Assumption on H, A , more precisely on random numberlike properties of matrix elements $A_{ij} := \langle \epsilon_i | A | \epsilon_j \rangle$, see also Eq. (2). The applicability of the ETH ensures the above rareness of excursions from equilibrium and a fixed A_{eq} for all initial states from some energy shell (thermalization). The validity of the ETH

is closely related to quantum chaos, cf. corresponding item below [5,6].

Typicality. Statistical statement on properties of pure states. Let the dimension of some finite Hilbert space d be large, $d \gg 1$. Let A be an operator on this Hilbert space with a spectral variance of order unity. Furthermore $|\psi\rangle$ are pure states that drawn at random according to the unitary invariant Haar measure from this Hilbert space. Then, for the overwhelming majority of all $|\psi\rangle$, the expectation values $\langle \psi | A | \psi \rangle$ are very close to $\text{Tr}\{A\}/d$ [4,7–13]. In the absence of any further information on $|\psi\rangle$ it is thus rational to assume $\langle \psi | A(t) | \psi \rangle \approx \text{Tr}\{A\}/d$, independent of t .

Quantum chaos. An observable featuring a finite overlap with a conserved quantity cannot relax to an equilibrium value independently of its initial value [14]. Wigner-Dyson-type (rather than Poissonian) statistics of spacings of adjacent energy eigenvalues signal the absence of many nontrivial conserved quantities (integrability) [15].

Microreversibility. Principal property of the dynamics of a system that ensures that an operation like flipping of particle momenta indeed entails a kind of reversed dynamics. Fluctuation relations, which are often viewed as more detailed formulations of the second law, are routinely based on microreversibility [16]. While the abstract concept is more encompassing, a version of microreversibility is implemented if A, H are both real in a common basis, cf. Eq. (2).

Non-fine-tuned initial states. Some specific mathematical constructions of initial states ρ obviously allow for implementing dynamics that are at odds with regular thermal relaxation. Thus a nontrivial conflict with the arrow of time requires the occurrence of strange relaxation dynamics for non-fine-tuned initial states, i.e., states that do not require backwards evolving of nonequilibrium states, flipping of particle momenta, or complex conjugating of wave functions $|\psi(t)\rangle \rightarrow |\psi(t)\rangle^*$ (no Loschmidt operations), full control over each individual matrix element $\langle \epsilon_i | \rho | \epsilon_j \rangle$, etc. [16]

This paper is organized as follows. In Sec. II we elaborate on what is meant by strange dynamics and formulate our main claim. In Sec. III we present numerical examples of strange

*lknipschild@uos.de

†jgemmer@uos.de

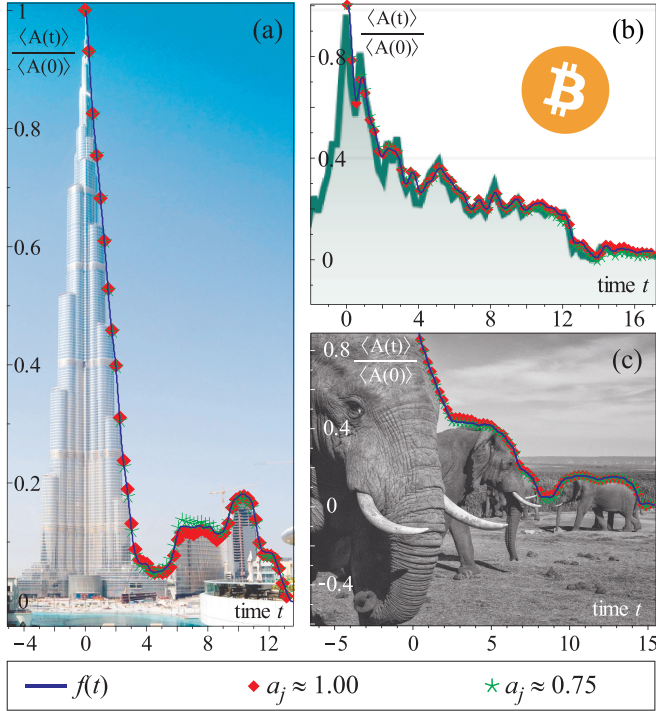


FIG. 1. H and A may be tailored such that $\langle A(t) \rangle / \langle A(0) \rangle$ follows almost any desired function $f(t)$ for a majority of initial states. Displayed data are for initial states that are eigenstates of A , i.e., $A|\psi(0)\rangle = a_j|\psi(0)\rangle$. Green and red symbols: $\langle A(t) \rangle / \langle A(0) \rangle$, blue line: desired function $f(t)$. (a) Burj Khalifa (original photo: *Burj Khalifa* by Joi is licensed under CC BY 2.0, <https://commons.wikimedia.org>). (b) Bitcoin price. (c) Elephants (original photo: *Family Of Elephants* by Javier Puig Ochoa is licensed under CC BY 3.0, <https://commons.wikimedia.org>).

dynamics that back up our main claim and explain how the numerical construction of these examples works. The physical relevance of the respective initial states underlying these dynamics is discussed in Sec. IV. In Sec. V we analytically argue for the validity of the main claim for (mixed) initial states that commute with the observable. These arguments suggest that the main claim is valid in the limit of large systems, which is supported by a numerical finite-size scaling in Sec. VI. In Sec. VII we generalize the result on certain classes of pure initial states that do not commute with the observable.

II. NOTION OF STRANGE DYNAMICS AND MAIN CLAIM

Prior to stating the main claim of the paper at hand we first establish the notion of strange dynamics: To comply with equilibration on average the considered expectation value $\text{Tr}\{A(t)\rho\} := \langle A(t) \rangle$ must take a fixed value A_{eq} for the vast majority of all instances in time. It may, however, nevertheless exhibit a behavior that is entirely unexpected in the context of relaxation dynamics. It could, e.g., follow the contour of some skyline before settling to A_{eq} , cf. Fig 1. Or it could, after having seemingly settled to A_{eq} , spike to a significantly off-equilibrium value at a time long compared to its initial relaxation time but much shorter than the Poincaré recurrence

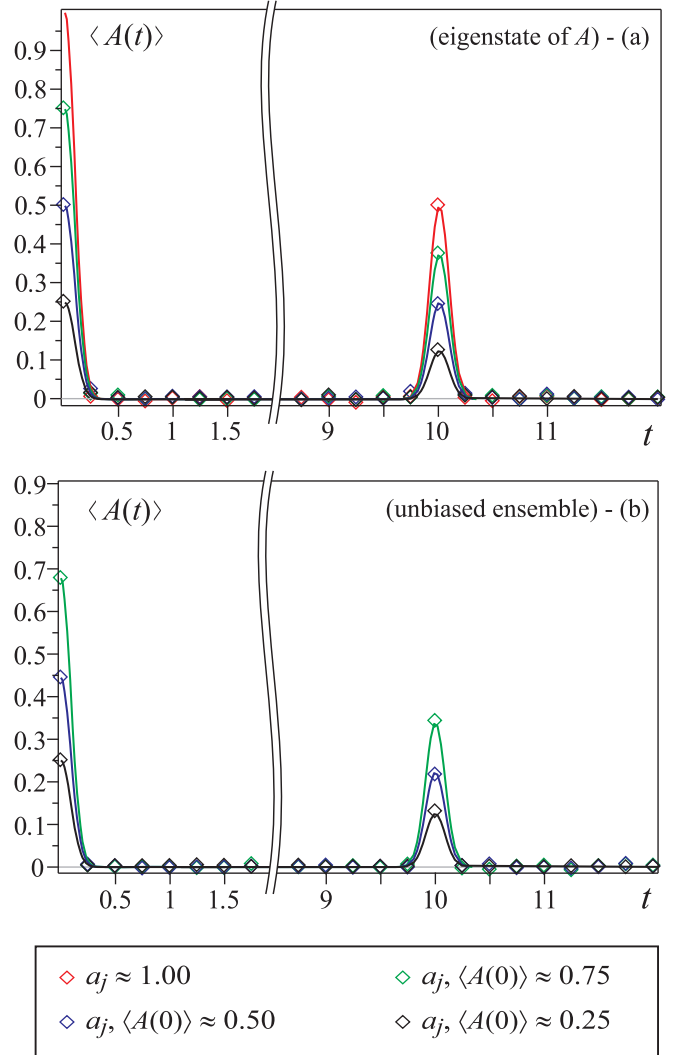


FIG. 2. $\langle A(t) \rangle$ exhibits a significant revival at an arbitrary pre-defined time τ_R (here $\tau_R = 10$) for two classes of initial states: (a) $|\psi(0)\rangle$ are eigenstates of A , (b) $|\psi(0)\rangle$ are samples of an unbiased pure state ensemble conditioned on a given $\langle A(0) \rangle$, cf. Eq. (3).

time, cf. Fig. 2. Any such unexpected dynamics we call strange dynamics without further rigorous definition.

Our main claim is as follows: It is possible to find a multitude of pairs H, A in accord with all above equilibration principles such that

$$\langle A(t) \rangle \approx \langle A(0) \rangle f(t) \quad (f(0) = 1), \tag{1}$$

where $f(t)$ must have a positive Fourier transform. Other than that $f(t)$ may essentially be freely chosen. Among the possible choices are plenty of strange dynamics, cf. Figs. 1 and 2. Most importantly the validity of Eq. (1) is claimed for a vast majority of all initial states ρ . This major set of initial states will be detailed below, see Secs. V, VII.

This claim is the more technical reformulation of the central statement, which forms the title of the present paper. (For some numerical evidence of the validity of Eq. (1) in concrete spin systems, albeit with nonstrange $f(t)$'s, see Ref. [17])

III. MAKING OF FIGS. 1 AND 2: HAMILTONIANS, OBSERVABLES, AND INITIAL STATES

Figures 1 and 2 show various expectation-value dynamics as resulting from the solution of the Schrödinger equation for fixed H, A (per panel in Fig. 1) but various ρ . The somewhat odd examples in Fig. 1 have simply been picked to substantiate the above claim that $f(t)$ may essentially be chosen at will, cf. Eq. (1). The setting in Fig. 2 is meant as a prime example of strange dynamics in the sense described above. Of course peculiar dynamics as presented in Figs. 1 and 2 require pairs H, A with specific properties. In the following we detail the construction of H, A , thereby unveiling their accordance with the cornerstone principles of thermal relaxation.

First a d -dimensional Hamiltonian H is defined by choosing d eigenvalues ϵ_j (examples in Figs. 1 and 2: $d = 20000$). To this end $d - 1$ energy gaps $l_j = \epsilon_{j+1} - \epsilon_j$ are drawn as i.i.d. random numbers from a pertinent Wigner-Dyson distribution. The spectrum is scaled to span an interval $[-E, E]$ (examples in Figs. 1 and 2: $E = 30$). Within this interval H has thus a constant density of states and exhibits Wigner-Dyson level statistics as expected for nonintegrable systems. Thus our modeling is in accord with the nonresonance condition and quantum chaos in the sense of the respective items in the introductory list in Sec. I. This Hamiltonian H may be viewed as the sector of, e.g., a many-body Hamiltonian that corresponds to a (narrow) energy window stretching from $-E$ to E . If the initial state lives (almost) entirely in this energy window, which is what we assume here, modeling of this sector suffices to compute the dynamics. Note that $\langle A(t) \rangle$ is fully determined by the spectra and the relative angles of the eigenvectors of the operators H, A, ρ . Thus, there is no need to specify the eigenvectors of H with respect to some computational basis for the purposes at hand.

Next we construct the observables A . To this end we first define $\tilde{f}(\omega)$ to be the real part of the Fourier transform of some desired, possibly strange, $f(t)$. Let furthermore ω_{\max} be some cutoff frequency ω_{\max} , such that $\tilde{f}(\omega)$ attains only negligible values at $|\omega| \geq \omega_{\max}$. Choose E, ω_{\max} such as to fulfill $E \gg \omega_{\max}$ (examples in Figs. 1 and 2: $\omega_{\max} \approx 3$). Furthermore $\tilde{f}(\omega)$ has to vary only negligibly on the scale of the level spacings l_j . While these conditions on $\tilde{f}(\omega)$ imply conditions on $f(t)$, these conditions become exceedingly mild, at sufficiently large d .

Now we specify the A 's in the energy eigenbasis $\{|\epsilon_i\rangle\}$ in full accord with the ETH [18] as

$$A_{jl} := \langle \epsilon_j | A | \epsilon_l \rangle = C_1 d^{-1/2} \sqrt{\tilde{f}(\epsilon_j - \epsilon_l)} R_{jl}, \quad (2)$$

where R_{jl} are normally i.i.d. random real numbers with zero mean and unit variance. C_1 is a constant, which we use to scale the extreme eigenvalues of A to $-a_{\min} \approx a_{\max} \approx 1$. To render A Hermitian, $\tilde{f}(\omega)$ must be nonnegative, which implies the condition on $f(t)$ mentioned below Eq. (1). To support our main claim it suffices that A as defined in Eq. (2) is not in conflict with the ETH. But what is more, numerous numerical studies found operators of local observables (currents, magnetizations, etc.) in the eigenbasis of the respective Hamiltonians (many-body lattice models) to essentially agree with the construction Eq. (2) [19–29], for more details see Appendix A. Choosing the A_{jl} real renders the setup microreversible in the

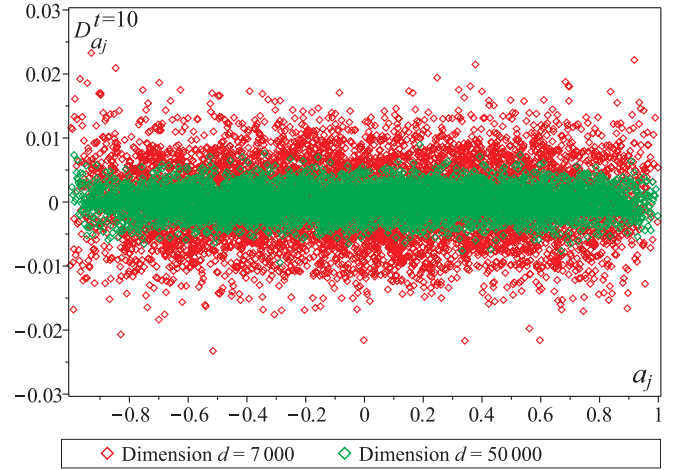


FIG. 3. The mean values of the errors and the variances are almost independent of the eigenstate, which served as the respective initial state.

sense defined in the respective item of the introductory list of equilibration principles.

Ultimately we aim at establishing that this construction of H, A renders Eq. (1) valid for practically all non-fine-tuned initial states ρ . However, for the numerical demonstration of this validity as displayed in Figs. 1 and 2, we chose initial states from the following two classes:

Observable eigenstates. Let $|a_j\rangle$ be an eigenstate of A , i.e., $A|a_j\rangle = a_j|a_j\rangle$. The dynamics in Figs. 1 and 2 (top panel) are computed for some sample initial eigenstates of A , i.e., $\rho = |a_j\rangle\langle a_j|$ for eigenvalues a_j , which may be inferred from the respective captions.

Unbiased ensemble states. The dynamics in Fig. 2 (bottom panel) correspond to states $\rho = |\psi\rangle\langle\psi|$ with

$$|\psi\rangle = \langle\phi|\phi\rangle^{-1/2}|\phi\rangle, \quad |\phi\rangle = (1 + \lambda A)^{-1/2}|\xi\rangle, \quad (3)$$

where $|\xi\rangle$ is a random vector unitarily drawn from the d -dimensional hypersphere in Hilbert space. λ has been chosen such as to obtain the respective $\langle A(0) \rangle$ as indicated in the caption of Fig. 2.

Before motivating the above choice of initial states we provide some more information on Figs. 1 and 2. For Fig. 1 the functions $f(t)$ have been extracted from the respective underlying pictures. (The pictures and contours have been chosen such as to make sure that the contours indeed feature positive Fourier transforms.) For Fig. 2 $f(t)$ is simply a pertinent mathematical function, the revival time τ_R of which is chosen as $\tau_R = 10$, which is much longer than the timescale of the initial decay but definitely much shorter than the Poincaré recurrence time, cf. Appendix B, making this example a prime instance of strange dynamics. The respective H, A have been constructed according to the above scheme. The respective Schrödinger equations have been numerically solved for the sample initial states as described above and in the captions. Obviously all computed data are in excellent agreement with our main claim, Eq. (1). (Even more numerical support for the latter comes from the data displayed in Fig. 3). All initial states feature effective dimensions of at least $d_{\text{eff}} \geq 4700$, i.e.,

are in accord with the principle of equilibration on average, for more detailed information see Appendix C.

IV. PHYSICAL SIGNIFICANCE OF THE INITIAL STATES UNDERLYING THE DISPLAYED DATA

Before explaining the inner workings of the above construction of H, A we comment on the physical significance of the initial states used in Figs. 1 and 2. Some physical relevance of the initial observable eigenstates $|a_j\rangle$ comes from their being identical to results of (preparatory) projective measurements of the monitored observable A at the beginning of the relaxation dynamics. More importantly, however, the validity of Eq. (1) for $\rho = |a_j\rangle\langle a_j|$ for all j necessarily entails the validity of Eq. (1) also for all initial states of the form $\rho = \sum_n c_n A^n$. The latter comprises, e.g., $\rho \propto \exp(\lambda A)$, which is the state of maximum von-Neumann entropy conditioned on some given $\langle A(0) \rangle$. This is a relevant class of initial states within the framework of Jayne's principle.

The states from the unbiased ensemble, cf. Eq. (3) represent an ensemble of pure states, which is entirely unbiased with respect to the unitary invariant Haar measure, under the condition of a given $\langle A(0) \rangle$ [30–32]. This ensemble thus is a very relevant class of initial states within the framework of pure state statistical mechanics.

V. VALIDITY OF THE MAIN CLAIM [EQ. (1)] FOR INITIAL STATES THAT COMMUTE WITH THE OBSERVABLE

In the following we explain why the above construction yields pairs H, A that render our main claim Eq. (1) valid for the above-mentioned class of initial states $\rho = \sum_n c_n A^n$ that also comprises the observable eigenstates $|a_j\rangle\langle a_j|$. Consider the relation

$$\text{Tr}\{A(t)A^N\} \approx \begin{cases} \propto \text{Tr}\{A(t)A\} \propto f(t), & \text{odd } N \\ 0, & \text{even } N \end{cases} \quad (4)$$

Obviously the validity of this relation entails the validity of Eq. (1) for all initial states of the above class. Accordingly Eq. (4) is the first pillar on which Eq. (1) rests. The validity of a subpart of Eq. (4), namely $\text{Tr}\{A(t)A\} \propto f(t)$ follows rather straightforward from Eq. (2) (for details see Appendix D). To fully validate Eq. (4) we employ a scheme suggested in Ref. [17]. As the argument is quite involved for large exponents N , we here restrict ourselves to $N = 2$ and $N = 3$. A comparable but more complex derivation for arbitrary but fixed N at large d can be found in Appendix E.

We start by writing out the correlation function for $N = 2$ explicitly,

$$\text{Tr}\{A(t)A^2\} = \sum_{a,b,c} A_{ab}A_{bc}A_{ca} e^{i(\epsilon_b - \epsilon_a)t}. \quad (5)$$

Given the matrix structure in Eq. (2), the biggest part of the addends in the sum are by construction (products of) independent random numbers with zero mean. Thus, to an accuracy set by the law of large numbers, summing the latter yields zero as well. There are, however, index combinations for which not all factors within the addends have vanishing

mean, namely, $c = a$. Focusing on these terms, we can write

$$\text{Tr}\{A(t)A^2\} \approx \sum_{a,b} |A_{ab}|^2 A_{aa} e^{i(\epsilon_b - \epsilon_a)t}. \quad (6)$$

While the numbers $|A_{ab}|^2$ do not have mean zero, the numbers A_{aa} do have zero mean. Furthermore for $a \neq b$ the $|A_{ab}|^2, A_{aa}, e^{i(\epsilon_b - \epsilon_a)t}$ are mutually independent stochastic variables, cf. Eq. (2). For $a = b$ these numbers are obviously not independent, however, in this case the sum in Eq. (6) is proportional to the third moment of the distribution of the A_{aa} , which vanishes according to Eq. (2). Exploiting these findings for both cases ($a \neq b$ as well as $a = b$) to evaluate Eq. (6) we obtain $\text{Tr}\{A(t)A^2\} \approx 0$, i.e., Eq. (4) for the even case $N = 2$.

Now we turn to $N = 3$, i.e.,

$$\text{Tr}\{A(t)A^3\} = \sum_{a,b,c,d} A_{ab}A_{bc}A_{cd}A_{da} e^{i(\epsilon_b - \epsilon_a)t}. \quad (7)$$

Again, the contributions of most addends approximately cancel each other upon summation. But also here there are exceptions, namely, the index combinations $c = a$ or $d = b$. Focusing on these terms, we find

$$\text{Tr}\{A(t)A^3\} \approx \sum_{a,b} (|A_{ab}|^2 \sum_c |A_{bc}|^2 + |A_{ac}|^2) e^{i(\epsilon_b - \epsilon_a)t}. \quad (8)$$

[Note that Eq. (8) erroneously counts the terms corresponding to $c = a$ and $d = b$ twice. However, as this overcounting error is of order d^{-1} , it becomes negligible at large d .] To proceed, consider the above sums over c first.

While these sums do not vanish, they are practically independent of a, b , due to the specific matrix structure of A , cf. Eq. (2). Thus, the respective sums may be replaced by a constant C_2 , i.e., $\sum_c |A_{bc}|^2 + |A_{ac}|^2 \approx C_2$. Inserting this into Eq. (8) yields $\text{Tr}\{A(t)A^3\} \approx C_2 \cdot \sum_{a,b} |A_{ab}|^2 e^{i(\epsilon_b - \epsilon_a)t}$. Comparing this to the exact relation $\text{Tr}\{A(t)A\} = \sum_{a,b} |A_{ab}|^2 e^{i(\epsilon_b - \epsilon_a)t}$ yields $\text{Tr}\{A(t)A^3\} \approx C_2 \text{Tr}\{A(t)A\}$, i.e., Eq. (4) for the odd case $N = 3$.

VI. ACCURACY OF THE MAIN CLAIM [EQ. (1)] FOR EIGENSTATES OF THE OBSERVABLE AS INITIAL STATES

Mainly due to the neglect of terms with random signs [cf. below Eq. (5)], Eq. (4) is not exact. To scrutinize the accuracy of Eq. (4) and hence Eq. (1) specifically with respect to growing dimensions d , the Schrödinger equation has been solved for the setup underlying Fig. 1(a), for all $|a_j\rangle$ ($d = 7000$), respectively, every tenth $|a_j\rangle$ ($d = 50000$) as initial states. Figure 3 shows the deviations from Eq. (1) at an exemplary point in time, here chosen as $t = 10$. These deviations are defined as

$$D_{a_j}^t := \langle a_j | A(t) | a_j \rangle - a_j f(t). \quad (9)$$

The errors have vanishing mean and a standard deviation that is almost independent of a_j . But, most importantly, the standard deviation decreases with the dimension d .

To numerically analyze this dependence, we consider the mean square of these errors (averaged over all eigenstates a_j

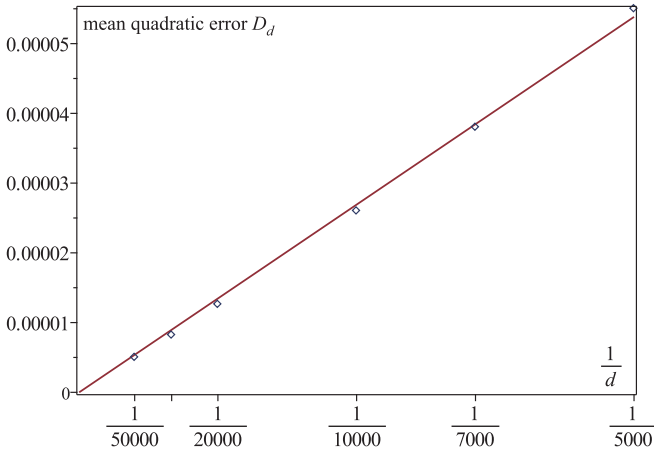


FIG. 4. The finite-size scaling of the mean square of the errors D_d indicates that these are proportional to $\frac{1}{d}$.

and over time):

$$D_d = \frac{1}{d} \sum_{j=1}^d \frac{1}{T} \int_0^T (D_{a_j}^t)^2 dt, \quad T = 13. \quad (10)$$

Plotting D_d over the reciprocal Hilbert-space dimension $1/d$ (see Fig. 4) reveals that the mean square of these errors vanishes as $\propto d^{-1}$. This is supported by an analytical reasoning (see Appendix E). These findings imply that Eq. (1) becomes exact in the limit of large d for all initial states of the form $\rho = \sum_n c_n A^n$, i.e., all initial states that commute with the observable A .

VII. VALIDITY OF THE MAIN CLAIM [EQ. (1)] FOR A CLASS OF INITIAL STATES THAT DO NOT COMMUTE WITH THE OBSERVABLE

Next we establish the validity of Eq. (1) for a set of (pure) initial states that are not of the form $\rho = \sum_n c_n A^n$, i.e., not functions of the observable A , and comprises the unbiased ensemble states [cf. Eq. (3)]. We call this set the r set and define its states $|\eta\rangle$ by

$$|\eta\rangle = \langle \chi | \chi \rangle^{-1/2} | \chi \rangle, \quad | \chi \rangle = \sqrt{r(A)} | \xi \rangle, \quad (11)$$

where $r(A)$ is a non-negative function that varies little on the scale of the eigenvalue spacings of A and $|\xi\rangle$ is a random vector drawn from the unitary invariant Haar measure. Typicality arguments may be used to show that for all $|\eta\rangle$, except for a fraction of at most $\propto d^{-1/2}$ the approximation

$$\langle \eta | A(t) | \eta \rangle \approx \text{Tr}\{r(A)A(t)\} \text{Tr}\{r(A)\}^{-1} \quad (12)$$

holds to very good accuracy. For a complete derivation see Ref. [32] and Appendix F. As $r(A)$ may be cast into the form $r(A) = \sum_n b_n A^n$, the combination of Eqs. (12) and (4) establishes the validity of Eq. (1) for all $|\eta\rangle$ except for the above fraction of size $\propto d^{-1/2}$. Thus Eq. (12) is the second pillar on which Eq. (1) rests. For the special case $r(A) = (1 + \lambda A)^{-1}$ the r set is identical to the unbiased ensemble, cf. Eq. (3), hence Eq. (12) explains the numerical findings displayed in Fig. 2 (bottom panel). Note that the r sets are very encompassing. The small fraction of order $d^{-1/2}$ of

initial states that does not comply with Eq. (12) for a suitable $r(A)$ thus corresponds to the set of fine-tuned initial states, which always exist but are excluded from the analysis at hand, c.f. last item of the introductory list of principles of equilibration. Note furthermore that the validity of Eq. (12) entails the validity of Eq. (1) for all initial states of the form $\rho = \int g(\eta) |\eta\rangle \langle \eta| d\eta$ with $g(\eta) \geq 0$ for the few (fine tuned) η to which Eq. (12) does not apply. Thus Eq. (1) also holds for a large set of mixed states.

VIII. OUTLOOK

While the body of this paper aims at pointing out an overlooked loophole in the current approach to equilibration in closed quantum systems, we eventually very briefly turn to possible closings of this loophole: While pairs H, A giving rise to strange dynamics definitely exist, these dynamics may not be stable under perturbations of the respective Hamiltonians [33–35]. Apart from that the ETH may miss some correlations that are in fact present in the matrices representing physical observables in physical systems [36]. These correlations may possibly rule out strange dynamics. Furthermore arguments are viable that are based on the condition of Hamiltonians being local [37]. Helpful insights may also come from clarifying the role of Markovianity in closed quantum systems [38].

ACKNOWLEDGMENTS

Stimulating discussions with M. Srednicki, K. Modi, M. Rigol, and W. Zurek during the QTHERMO18 program at the KITP are gratefully acknowledged. The authors also benefited from interacting with R. Steinigeweg, J. Richter, and R. Heveling. This work has been funded by the Deutsche Forschungsgemeinschaft (DFG), Grants No. 397107022 (GE 1657/3-1) and No. 355031190 within the DFG Research Unit FOR 2692. Furthermore this research was supported in part by the National Science Foundation under Grant No. NSF PHY-1748958.

APPENDIX A: EIGENSTATE THERMALIZATION HYPOTHESIS IN PHYSICAL MODELS

The similarity of local operators represented in the energy eigenbasis with random matrices, as implemented in Eq. (2), is at the heart of the ETH [18]. Such a similarity has numerically been observed frequently, see, e.g., Refs. [19–23, 25–29], it may, however, require a splitting of A into distinct symmetry sectors. The more specific form of A , namely the dependence of the envelope function $\tilde{f}(x)$ solely through the energy differences $\epsilon_j - \epsilon_l$ but not on individual energies, has approximately also been found for local observables in interacting lattice-particle models, see Refs. [17, 22, 39]. Currently discussed structural differences of local observables in physical chaotic systems from the ETH as formulated in [18] include non-Gaussian distributions of the matrix elements [40] and correlations between individual matrix elements [36]. While nonsystematic checking indicates that non-Gaussian distributions in Eq. (2) leave the validity of Eq. (4) unaltered, the impact of correlations is open and subject to further research. However, some evidence for the applicability of the

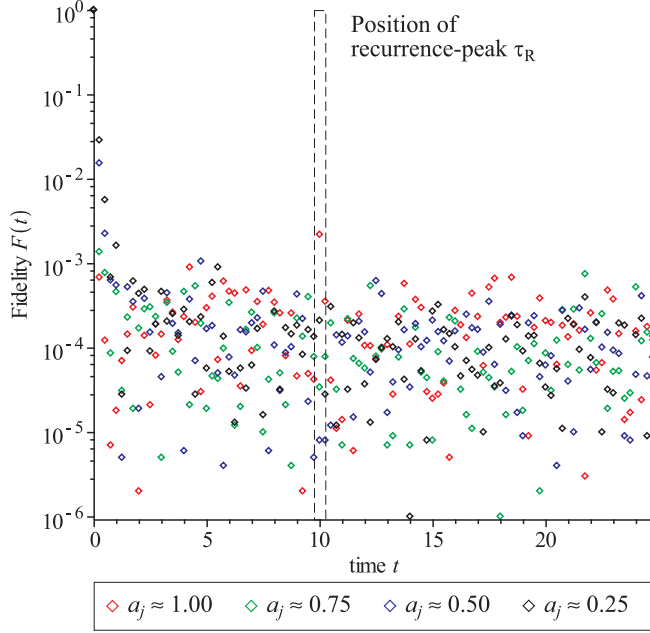


FIG. 5. For the dynamics of the eigenstates, corresponding to the eigenvalues a_j , the fidelity has been plotted over time. The fidelity rapidly decays and does not show any significant recurrences. Especially at the position of the recurrence peak τ_R (of the expectation-value dynamics) the fidelity is less than 0.01.

overall concept discussed in the paper at hand to currents in spin chains comes from Ref. [17].

APPENDIX B: DECAY OF FIDELITY

In this section we check the overlap of the evolved state with the initial state, to exclude that the revival peaks in Fig. 2 are due to Poincaré-recurrences. We therefore calculate the fidelity F of the time-evolved state $U(t)|a_j\rangle$ and the respective initial state $|a_j\rangle$:

$$F(t) = |\langle \psi_0 | U(t) | \psi_0 \rangle|^2. \quad (\text{B1})$$

$U(t)$ denotes the time-evolution operator for the time t . The fidelity rapidly decays during the initial relaxation of the expectation value (Fig. 5). There is no significant increase of $F(t)$ at any later time, including the revival time of the expectation value τ_R . Thus the second peak in the expectation-value dynamics (at time τ_R) is not due to any Poincaré recurrence.

APPENDIX C: EFFECTIVE DIMENSION OF VARIOUS INITIAL STATES

A sufficiently large effective dimension of the initial state is a necessary condition, when trying to prove the emergence of thermodynamic behavior in closed quantum systems on the basis of equilibration on average (see Sec. I). Table I lists the effective dimensions $d_{\text{eff}}(|\psi\rangle)$ of various initial states that are either eigenstates $|a_j\rangle$ of A or members of the unbiased ensemble $|\psi_a\rangle$, Eq. (3). For the latter the parameter λ has been chosen such that the expectation value of A for the ensemble average is equal to $\langle A \rangle = 0.25, 0.50, 0.70$. Obviously, $d_{\text{eff}}(|\psi\rangle)$ is rather large at all instances.

TABLE I. We calculated the effective dimension for various initial states. The data refers to the recurrence dynamics (2).

Initial State	Effective Dimension d_{eff}
$ a_j\rangle, a_j = 0.25$	6500
$ a_j\rangle, a_j = 0.50$	6500
$ a_j\rangle, a_j = 0.75$	6400
$ a_j\rangle, a_j = 1.00$	4700
$ \psi_a\rangle, \langle A(0) \rangle = 0.25$	9900
$ \psi_a\rangle, \langle A(0) \rangle = 0.50$	8900
$ \psi_a\rangle, \langle A(0) \rangle = 0.75$	7500

APPENDIX D: EXPLICIT CALCULATION OF THE AUTOCORRELATION FUNCTION

Within this section we calculate the autocorrelation function that follows from Eq. (2):

$$\text{Tr}\{A(t)A\} = C_1^2 \sum_{j,l} |a_{jl}|^2 \cos[(\epsilon_l - \epsilon_j)t] \quad (\text{D1})$$

$$\approx C_1^2 \sum_{j,l} \tilde{f}(\epsilon_l - \epsilon_j) \cos[(\epsilon_l - \epsilon_j)t] \quad (\text{D2})$$

$$\approx C_3 \int \tilde{f}(\omega) \cos(\omega t) d\omega \quad (\text{D3})$$

$$\propto f(t) \quad \text{for } t \geq 0. \quad (\text{D4})$$

Equation (D2) follows from the law of large numbers and for Eq. (D3) we exploit the uniform density of states of H and $E \gg \omega_{\text{max}}$. C_3 is a pertinent constant. In Eq. (D4) we used the definition of the positive Fourier transform. Hence this construction implemented by Eq. (2) essentially produces an autocorrelation function following predefined target dynamics $f(t)$ while being in accord with the ETH.

APPENDIX E: ANALYSIS OF THE VALIDITY OF EQ. (4)

According to Eq. (1) the expectation value dynamics of the observable A is proportional to its autocorrelation function for exceedingly many initial states. While this statement cannot be true for all initial states, we stress within this section that it holds for huge classes of physically relevant initial states.

In the main text the validity of (4) has been exemplarily proven for initial states $\rho \propto A^N$ with $N \in 2, 3$. In the first part of this section we extend this proof to arbitrary but fixed N at large d . This also ensures the validity of (4) for initial states that are analytical functions of A , e.g., $\exp(\lambda A)$, $r(A)$ for sufficiently large d . In the second part we numerically address the range of validity by checking the expectation-value dynamics of all eigenstates of A .

1. Evaluation of $\text{Tr}\{A(t)A^N\}$ for arbitrary but fixed N at large d

Before embarking on a concrete estimate of $\text{Tr}\{A(t)A^N\}$, the maximum N for which Eq. (4) needs to be established should be settled. This maximum N is $N = d$. To justify this,

consider the following reformulation of Eq. (4):

$$\sum_{n=1}^d \langle a_n | A(t) | a_n \rangle a_n^N \propto \text{Tr}\{A(t)A\}. \quad (\text{E1})$$

If this holds for $N \leq d$ it must hold for all $\langle a_n | A(t) | a_n \rangle$ individually, i.e.,

$$\langle a_n | A(t) | a_n \rangle \propto \text{Tr}\{A(t)A\} \quad (\text{E2})$$

since the a_n^N form a set of linearly independent functions of n . Having established this we are set to work towards Eq. (4)

To begin with, consider

$$\tilde{f}_{ab}^N := A_{ab} \cdot \sum_{i_1, \dots, i_{N-1}} (A_{bi_1} A_{i_1 i_2} \cdots A_{i_{N-2} i_{N-1}} A_{i_{N-1} a}), \quad (\text{E3})$$

where the addends are products of N matrix elements A_{ij} . Obviously \tilde{f}_{ab}^N is a Fourier component of $\text{Tr}\{A(t)A^N\}$ with time dependence proportional to $\exp(-i(\epsilon_b - \epsilon_a)t)$. Since A is essentially a random matrix [see Eq. (2)], most addends in Eq. (E4) are products of independent random numbers. As such they will be real random numbers themselves, with zero mean. Hence, to an accuracy set by the law of large numbers, these addends will sum up to zero. However, there are index combinations for which the respective addends are not just products of independent random numbers but necessarily real and positive. (These are also the only addends that would survive an averaging of Eq. (E4) over concrete implementations of A as may be inferred from Isserlis theorem [41].) First we focus exclusively on these addends to find the systematic part of $\text{Tr}\{A(t)A^N\}$. We come back to the random or fluctuating part below.

An index combination yields a sure positive, systematic contribution if and only if each individual matrix element A_{ij} in Eq. (E3) appears for an even number of times. Since for even N there is an odd number of matrix elements in Eq. (E3), the systematic contribution vanishes in this case. This already establishes Eq. (4) for even N . For N odd there are very many index combinations for which each individual matrix element appears for an even number of times. Consider first the two following types of such index combinations:

$$\tilde{\alpha}_{ab}^N := A_{ab} \cdot \sum_{i_1, \dots, i_{(N-1)/2}} (A_{bi_1} A_{i_1 i_2} \cdots A_{i_{(N-1)/2} b} A_{ba}), \quad (\text{E4})$$

$$\tilde{\beta}_{ab}^N := A_{ab} \cdot \sum_{i_1, \dots, i_{(N-1)/2}} (A_{ba} A_{ai_1} \cdots A_{i_{(N-1)/2} i_1} A_{i_1 a}). \quad (\text{E5})$$

These two contributions to \tilde{f}_{ab}^N feature the maximum number of free indices, i.e., indices that are summed over, under the condition that each matrix element has to appear at least twice. This number is $(N-1)/2$. There are much more sure positive index combinations, however, they all have at most $(N-3)/2$ free indices. The number of free indices is crucial since each free index gives rise to a multiplicity on the order of d to the respective contribution due to the corresponding summation. The number of index combinations that lead to sure positive index combinations with less than $(N-1)/2$ free indices depends on N . Their number grows (rapidly) with N . However, at any fixed N the contributions $\tilde{\alpha}_{ab}^N, \tilde{\beta}_{ab}^N$ will become more and more dominant with larger dimension d .

Above some d we may thus approximate

$$\tilde{f}_{ab}^N \approx \tilde{\alpha}_{ab}^N + \tilde{\beta}_{ab}^N. \quad (\text{E6})$$

While it is not obvious if this approximation is justified up to $N = d$, we focus on cases where Eq. (E6) is valid in the paper at hand. A more thorough analysis of the $N \approx d$ case is a subject for further research. Taking Eq. (E6) for granted we obtain

$$\tilde{f}_{ab}^N = A_{ab}^2 \cdot (P_b^N + P_a^N) \quad (\text{E7})$$

$$P_b^N := \sum_{i_1, \dots, i_{(N-1)/2}} (A_{bi_1} A_{i_1 i_2} \cdots A_{i_{(N-1)/2} b}).$$

P_b^N may be interpreted as a sum over certain paths on the set of the indices where each path features a corresponding weight: Each path has to start at b , it has to end at b and it must take each transition for an even number of times, i.e., at least twice. The total number of transitions is $N-1$. The number of different indices through which such a path ventures is the number of free indices, thus the paths with the largest number of different indices feature $(N-1)/2$ free indices, in accord with the above statement. The weight of each path is the product of all the squares A_{ij}^2 of the matrix elements corresponding to the transitions through which it went. Calculating (an estimate) of P_b^N is an ambitious endeavor, closely related to the derivation of Wigner's semicircle law for the spectra of random matrices. Fortunately there is no need to do this here. The following two observations suffice. (i) The statistical properties of the matrix A do not depend on individual indices, they only depend on the differences $\epsilon_i - \epsilon_j$. Hence, up to a (small) statistical error, P_b^N cannot depend on b (up to finite-size effects, cf. below). Much like the return probability of a particle in a disordered but homogenous medium does not depend on the starting point. (ii) For each path that ventures through the transition $a \leftrightarrow b$ two, four, etc. times, there are (at least) $\propto d$ paths that do not do so. Thus at large d , these paths have negligible weight. As a consequence \tilde{f}_{ab}^N has a dominant contribution proportional to A_{ab}^2 and only negligible contributions proportional to A_{ab}^4, A_{ab}^6 , etc. While the first observation is strictly correct in the limit of $\omega_{\max}/E \rightarrow 0$, it is not strictly correct outside this limit. If b is close to one of the edges, i.e., $\epsilon_b \approx \pm E$, the paths become affected by the vicinity to the edge. Much like the above return probability may be different if the starting point is sufficiently close to an edge of the disordered medium. Here we assume, however, that the resulting dependence of P_b^N is such that $P_b^N = P^N(\epsilon_b)$ does not change much on the scale of ω_{\max} . Equipped with these observations we now return to Eq. (E7). Recalling that \tilde{f}_{ab}^N are the Fourier components of the respective correlation functions yields:

$$\begin{aligned} & \text{Tr}\{A(t)A^N\}_{\text{sur. pos.}} \\ & \propto \sum_{a,b} A_{ab}^2 \exp(-i(\epsilon_a - \epsilon_b)t) (P_b^N + P_a^N). \end{aligned} \quad (\text{E8})$$

Employing the index transformation

$$\bar{\epsilon} := \frac{\epsilon_a + \epsilon_b}{2}, \quad \omega := \epsilon_a - \epsilon_b. \quad (\text{E9})$$

Equation (E8) may be rewritten as:

$$\begin{aligned} & \text{Tr}\{A(t)A^N\}_{\text{sur. pos.}} \\ & \propto \sum_{\bar{\epsilon}, \omega} A_{\bar{\epsilon}\omega}^2 \exp(-i\omega t) (P_{\bar{\epsilon}+\omega/2}^N + P_{\bar{\epsilon}-\omega/2}^N). \end{aligned} \quad (\text{E10})$$

We proceed by exploiting two facts: (i) the $A_{\bar{\epsilon}\omega}^2$ and the $P_{\bar{\epsilon}\pm\omega/2}^N$ are (approximately) uncorrelated; (ii) the $A_{\bar{\epsilon}\omega}^2$ depend on $\bar{\epsilon}$ only statistically, the systematic dependence is only on ω , cf. Eq. (2). Exploiting these facts allows to recast Eq. (E10) as

$$\begin{aligned} \text{Tr}\{A(t)A^N\}_{\text{sur. pos.}} & \propto \sum_{\omega} A_{\bar{\epsilon}\omega}^2 \exp(-i\omega t) \\ & \sum_{\bar{\epsilon}} (P_{\bar{\epsilon}+\omega/2}^N + P_{\bar{\epsilon}-\omega/2}^N). \end{aligned} \quad (\text{E11})$$

The second sum over $\bar{\epsilon}$ is approximately independent of ω ($\leq \omega_{\text{max}}$) unless a substantial fraction of the weight of the function $P_a^N(0)$ is concentrated within a range of width ω_{max} at the edges $-E, E$. Following the above observation (i), however, this is not to be expected. Hence Eq. (E10) may again be rewritten as

$$\text{Tr}\{A(t)A^N\}_{\text{sur. pos.}} \propto \sum_{\bar{\epsilon}, \omega} A_{\bar{\epsilon}\omega}^2 \exp(-i\omega t). \quad (\text{E12})$$

Realizing that the right-hand side is just the Fourier transform of $\text{Tr}\{A(t)A\}$ this yields

$$\text{Tr}\{A(t)A^N\}_{\text{sur. pos.}} \propto \text{Tr}\{A(t)A\} \quad (\text{E13})$$

and thus completes the justification of Eq. (4).

In the remainder we analyze the influence of the non-sure positive contribution, which mainly gives rise to deviations/fluctuations, i.e., the \approx relation in Eq. (4). We aim at estimating the scaling of these (squared) deviations with the dimension d . To this end we define $x_N(t)$:

$$x_N(t) := \frac{1}{d} (\text{Tr}\{A(t)A^N\} - \text{Tr}\{A(t)A^N\}_{\text{sur. pos.}}). \quad (\text{E14})$$

Note that the prefactor of d^{-1} renders $d^{-1}\text{Tr}\{A(t)A^N\}$ itself independent of d in the limit of large d . Again, we are eventually interested in the sure positive contributions to $|x_N(t)|^2$, denoted as $|x_N(t)|_{\text{sur. pos.}}^2$. The other contributions are expected to be of vanishing impact in the limit of large d . Writing $|x_N(t)|^2$ out explicitly yields

$$\begin{aligned} |x_N(t)|^2 & = \frac{1}{d^2} \sum_{\text{indices} \setminus \alpha, \beta} A_{ab} A_{b_1} \cdots A_{i_{N-2}a} \\ & \cdot A_{cd} A_{d_1} \cdots A_{j_{N-2}c} \\ & \cdot \exp(-i(\epsilon_a - \epsilon_b - \epsilon_c + \epsilon_d)t), \end{aligned} \quad (\text{E15})$$

where indices $\setminus \alpha, \beta$ stands for: all indices, i.e., index combinations, but without those that give rise to the $\tilde{\alpha}_{ab}^N, \tilde{\beta}_{ab}^N$ in Eq. (E4). This somewhat involved construction ensures the subtraction of the sure positive part as defined in Eq. (E14). Again, the sure positive contributions to $|x_N(t)|^2$ arise from index combinations for which all matrix elements appear to an even power, i.e., at least squared. Following the same scheme as in Eq. (E4) we find that the index combinations with the largest number of free indices are characterized by

$a = c, b = d, i_n = j_n$. This yields:

$$|x_N(t)|_{\text{sur. pos.}}^2 \approx \frac{1}{d^2} \sum_{\text{indices} \setminus \alpha, \beta} A_{ab}^2 A_{b_1}^2 \cdots A_{i_{N-2}a}^2. \quad (\text{E16})$$

The index combinations that would appear in Eq. (E16) but are excluded by the $\setminus \alpha, \beta$ are the ones for which each A_{ij}^2 appears at least twice (such as to form A_{ij}^4). Consequently the number of free indices corresponding to those excluded index combinations is of order $N/2$ while the number of the free indices of the nonexcluded index combinations in the summation of Eq. (E16) is of order N . Thus, creating only a negligible error, we may drop the exclusion of said index combinations, obtaining

$$|x_N(t)|_{\text{sur. pos.}}^2 \approx \frac{1}{d^2} \sum_{\text{indices}} A_{ab}^2 A_{b_1}^2 \cdots A_{i_{N-2}a}^2. \quad (\text{E17})$$

As, according to Eq. (2) the N matrix elements scale as $A_{ij}^2 \propto d^{-1}$ and there are N summations of d indices in Eq. (E17) we eventually find

$$|x_N(t)|_{\text{sur. pos.}}^2 \propto \frac{1}{d^2}. \quad (\text{E18})$$

While this result could in principle be compared to numerics directly, we resort here to a check of consistency of our much more detailed numerical findings with the result in Eq. (E18). Rather than addressing the $\text{Tr}\{A(t)A^N\}$ (for limited N) we numerically analyze dynamics of the form $\langle a_j | A(t) | a_j \rangle$. These data are more detailed in the sense that $\text{Tr}\{A(t)A^N\}$ may conveniently be computed from the set of all $\langle a_j | A(t) | a_j \rangle$. This way the consistency will eventually be demonstrated. We start, however, by postulating a specific form of the $\langle a_j | A(t) | a_j \rangle$, which is suggested by the numerical findings, cf. Fig. 3:

$$\langle a_j | A(t) | a_j \rangle \stackrel{!}{=} a_j f(t) + \frac{g_j(t)}{\sqrt{d}} Y_j, \quad (\text{E19})$$

where Y_j are independent random Gaussian numbers with zero mean and unit variance. $g_j(t)$ is a function that varies very mildly with j . Recalling Eq. (E1) and identifying $\sum_j a_j^{N+1} f(t) = \text{Tr}\{A(t)A^N\}_{\text{sur. pos.}}$ we re-express $x_N(t)$ based on Eq. (E19):

$$x_N(t) = \frac{1}{d} \sum_j a_j^N \frac{g_j(t)}{\sqrt{d}} Y_j. \quad (\text{E20})$$

Exploiting Eq. (E19) it is straightforward to compute

$$\langle |x_N(t)|^2 \rangle = \frac{1}{d^2} \sum_j a_j^{2N} \frac{g_j^2(t)}{d} \propto \frac{1}{d^2}. \quad (\text{E21})$$

Comparing this result to Eq. (E18) completes the demonstration of consistency and strongly supports the conjecture implemented by Eq. (E19).

2. Numerical checkup of (1) for eigenstates of A

We now turn to the numerical check of the expectation-value dynamics of the eigenstates of the observable. We generated the observable in such a way that its autocorrelation

function $\text{Tr}\{A(t)A\}$ is proportional to $g(t)$. For our numerical investigations we chose the reference function redrawing the contours of the Burj-Kalifa [$g(t)$]. Some dynamics are shown in Fig. 1. The expectation-value dynamics of each eigenstate $|a_j\rangle$ of A appear to be very similar to this autocorrelation function. To quantify the error we define the deviation at time t in the following way:

$$D'_{a_j} = \langle a_j|A(t)|a_j\rangle - a_j \frac{\text{Tr}\{AA(t)\}}{\text{Tr}\{A^2\}}. \quad (\text{E22})$$

Formally this quantity is very similar to $\frac{g_j(t)}{\sqrt{d}}Y_j$ defined in (E19), but while the properties of $\frac{g_j(t)}{\sqrt{d}}Y_j$ are simply conjectured, D'_{a_j} refers to numerical data. One aim in the following analysis is to show that the actual distribution of errors is compatible with the statistical properties of $\frac{g_j(t)}{\sqrt{d}}Y_j$.

We start by checking the dependence of the errors on the index of the eigenstate j . We therefore fix $t = 10$ and $d = 7000, 50000$ and plot the errors as a function of the eigenvalue a_j of the initial state $\rho_0 = |a_j\rangle\langle a_j|$ (Fig. 3).

Figure 3 indicates that there is no systematic dependence of the error on the position in the spectrum. Moreover the errors appear to be normal distributed. The numerics clearly show that the absolute errors decrease with dimension d . This dependence on d is studied in more detail in a finite-size scaling at the end of this section.

Up to now we focused on a single point in time. To drop this random choice, we define two new error measures: D_{a_j} , which quantifies the squared deviation of a dynamics averaged over time and D_d , which is the squared deviation averaged over time and over all eigenstates of A :

$$D_{a_j} = \frac{1}{M} \sum_{m=0}^{M-1} (D'_{a_j}{}^{m\Delta t})^2 \quad (\text{E23})$$

$$D_d = \frac{1}{d} \sum_{j=1}^d D_{a_j}. \quad (\text{E24})$$

$\Delta t = 0.25$ and $M = 200$ denote the (numerical) time step and the number of steps in time, respectively.

Figure 6 shows the averaged deviations of the expectation-value dynamics from the reference function for all eigenstates of A . The errors appear to be only marginally dependent on the position in the spectrum of A .

Figure 6 furthermore indicates that the deviations from the reference function decrease for larger dimension d . To address the dependence of the errors on the system size d we plot the averaged errors D_d as a function of the reciprocal dimension $1/d$ (Fig. 4).

This finite-size scaling indicates that the mean quadratic error D_d is proportional $\frac{1}{d}$. This suggests that the variances of the errors D'_{a_j} each are proportional to $\frac{1}{d}$. Thus the properties of the error distribution are in accord with the assumptions made in (E19), which in turn implies the $\frac{1}{d^2}$ -error scaling for $N \ll d$ found in (E21).

APPENDIX F: TYPICALITY

In this section a derivation of Eq. (12) is presented. It is similar to a comparable analysis in Ref. [32]. Consider a

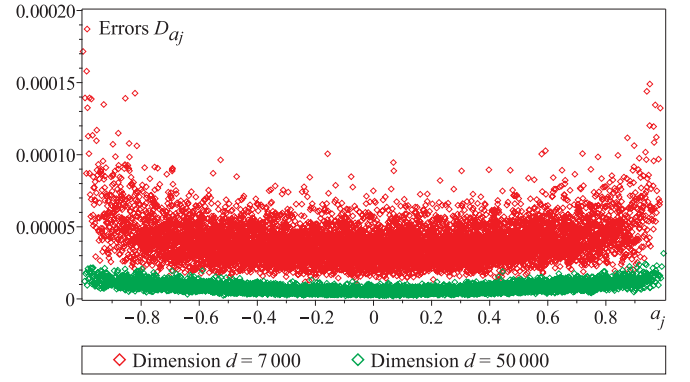


FIG. 6. The time-averaged quadratic error D_{a_j} only marginally depends on the eigenstate $\rho_0 = |a_j\rangle\langle a_j|$, which serves as the initial state. Furthermore the deviations from the reference function in the larger system $d = 50000$ are significantly smaller than the corresponding deviations for $d = 7000$.

ensemble of pure states given by

$$|\psi\rangle = \langle\phi|\phi\rangle^{-1/2}|\phi\rangle, \quad |\phi\rangle = \sqrt{r(A)}|\xi\rangle, \quad (\text{F1})$$

where $r(A)$ is a non-negative, smooth function, i.e., $|r(a_{j+1}) - r(a_j)|/|a_{j+1} - a_j| < d$, with respective expectation values $\langle\psi|A(t)|\psi\rangle$

$$\langle\psi|A(t)|\psi\rangle = \frac{\langle\xi|\sqrt{r(A)}A(t)\sqrt{r(A)}|\xi\rangle}{\langle\xi|r(A)|\xi\rangle}. \quad (\text{F2})$$

We analyze the statistical properties of the numerator first. Let the overbar $\overline{\dots}$ indicate the average over all $|\xi\rangle$ and $\sigma^2(\dots)$ the respective variance. Following [7,42] we obtain:

$$\overline{\langle\xi|\sqrt{r(A)}A(t)\sqrt{r(A)}|\xi\rangle} = \mathbb{E}[r(A)A(t)], \quad (\text{F3})$$

$$\sigma^2(\langle\xi|\sqrt{r(A)}A(t)\sqrt{r(A)}|\xi\rangle) = \quad (\text{F4})$$

$$\frac{\chi^2[\sqrt{r(A)}A(t)\sqrt{r(A)}]}{d+1} \leq \frac{\mathbb{E}[r(A)A^2]}{d+1}, \quad (\text{F5})$$

where $\mathbb{E}[\dots]$ denotes the mean and $\chi^2[\dots]$ the variance of the spectrum of the respective operator. As $\mathbb{E}[r(A)A^2]$ converges against a fixed value for large d and smooth $r(A)$, the variance $\sigma^2(\langle\xi|\sqrt{r(A)}A(t)\sqrt{r(A)}|\xi\rangle)$ has an upper bound that essentially scales as d^{-1} . Hence

$$\langle\xi|\sqrt{r(A)}A(t)\sqrt{r(A)}|\xi\rangle \approx \mathbb{E}[r(A)A(t)] = \frac{\text{Tr}\{r(A)A(t)\}}{d} \quad (\text{F6})$$

is a very good approximation for all $|\xi\rangle$ except for a fraction of size of at most $\propto d^{-1/2}$. We now perform an analogous analysis for the denominator of Eq. (F2):

$$\overline{\langle\xi|r(A)|\xi\rangle} = \mathbb{E}[r(A)], \quad \sigma^2(\langle\xi|r(A)|\xi\rangle) = \frac{\chi^2[r(A)]}{(d+1)}. \quad (\text{F7})$$

As $\chi^2[r(A)]$ converges against a fixed value for large d and smooth $r(A)$, the variance $\sigma^2(\langle\xi|r(A)|\xi\rangle)$ essentially scales

as d^{-1} . Hence

$$\langle \xi | r(A) | \xi \rangle \approx \mathbb{E}[r(A)] = \frac{\text{Tr}\{r(A)\}}{d} \quad (\text{F8})$$

is a very good approximation for all $|\xi\rangle$ except for a fraction of size of $\propto d^{-1/2}$. Inserting Eq. (F6) and Eq. (F6) into Eq. (F2)

yields

$$\langle \psi | A(t) | \psi \rangle \approx \frac{\text{Tr}\{r(A)A(t)\}}{\text{Tr}\{r(A)\}} \quad (\text{F9})$$

as a good approximation for all $|\psi\rangle$ except for a fraction of size of at most $\propto d^{-1/2}$. This establishes Eq. (12).

-
- [1] C. Gogolin and J. Eisert, Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems, *Rep. Prog. Phys.* **79**, 056001 (2016).
- [2] P. Reimann, Foundation of Statistical Mechanics Under Experimentally Realistic Conditions, *Phys. Rev. Lett.* **101**, 190403 (2008).
- [3] A. J. Short and T. C. Farrelly, Quantum equilibration in finite time, *New J. Phys.* **14**, 013063 (2012).
- [4] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, *Phys. Rev. E* **79**, 061103 (2009).
- [5] M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* **50**, 888 (1994).
- [6] J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* **43**, 2046 (1991).
- [7] S. Lloyd, Pure state quantum statistical mechanics and black holes, [arXiv:1307.0378](https://arxiv.org/abs/1307.0378).
- [8] P. Reimann, Typicality for Generalized Microcanonical Ensembles, *Phys. Rev. Lett.* **99**, 160404 (2007).
- [9] P. Reimann and J. Gemmer, Full expectation-value statistics for randomly sampled pure states in high-dimensional quantum systems, *Phys. Rev. E* **99**, 012126 (2019).
- [10] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, *Nature Phys.* **2**, 754 (2006).
- [11] H. Tasaki, Typicality of thermal equilibrium and thermalization in isolated macroscopic quantum systems, *J. Stat. Phys.* **163**, 937 (2016).
- [12] S. Sugiura and A. Shimizu, Canonical Thermal Pure Quantum State, *Phys. Rev. Lett.* **111**, 010401 (2013).
- [13] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, Canonical Typicality, *Phys. Rev. Lett.* **96**, 050403 (2006).
- [14] P. Mazur, Non-ergodicity of phase functions in certain systems, *Physica* **43**, 533 (1969).
- [15] F. Haake, Quantum signatures of chaos, in *Quantum Coherence in Mesoscopic Systems* (Springer, Berlin, 1991), pp. 583–595.
- [16] M. Campisi, P. Hänggi, and P. Talkner, Colloquium: Quantum fluctuation relations: Foundations and applications, *Rev. Mod. Phys.* **83**, 771 (2011).
- [17] J. Richter, J. Gemmer, and R. Steinigeweg, Impact of eigenstate thermalization on the route to equilibrium, *Phys. Rev. E* **99**, 050104(R) (2019).
- [18] M. Srednicki, The approach to thermal equilibrium in quantized chaotic systems, *J. Phys. A: Math. Gen.* **32**, 1163 (1999).
- [19] L. D’Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, *Adv. Phys.* **65**, 239 (2016).
- [20] W. Beugeling, R. Moessner, and M. Haque, Off-diagonal matrix elements of local operators in many-body quantum systems, *Phys. Rev. E* **91**, 012144 (2015).
- [21] R. Mondaini and M. Rigol, Eigenstate thermalization in the two-dimensional transverse field Ising model. ii. off-diagonal matrix elements of observables, *Phys. Rev. E* **96**, 012157 (2017).
- [22] S. Mukerjee, V. Oganesyan, and D. Huse, Statistical theory of transport by strongly interacting lattice fermions, *Phys. Rev. B* **73**, 035113 (2006).
- [23] A. R. Kolovsky and A. Buchleitner, Quantum chaos in the Bose-Hubbard model, *Europhys. Lett.* **68**, 632 (2004).
- [24] C. A. Parra-Murillo, J. Madroñero, and S. Wimberger, Quantum diffusion and thermalization at resonant tunneling, *Phys. Rev. A* **89**, 053610 (2014).
- [25] E. J. Torres-Herrera, J. Karp, M. Távora, and L. F. Santos, Realistic many-body quantum systems vs. full random matrices: Static and dynamical properties, *Entropy* **18**, 359 (2016).
- [26] C. Nation and D. Porras, Off-diagonal observable elements from random matrix theory: Distributions, fluctuations, and eigenstate thermalization, *New J. Phys.* **20**, 103003 (2018).
- [27] R. Mondaini, K. Mallayya, L. F. Santos, and M. Rigol, Comment on “Systematic Construction of Counterexamples to the Eigenstate Thermalization Hypothesis”, *Phys. Rev. Lett.* **121**, 038901 (2018).
- [28] R. Hamazaki and M. Ueda, Random-matrix behavior of quantum nonintegrable many-body systems with Dyson’s three symmetries, *Phys. Rev. E* **99**, 042116 (2019).
- [29] L. Foini and J. Kurchan, Eigenstate Thermalization and Rotational Invariance in Ergodic Quantum Systems, *Phys. Rev. Lett.* **123**, 260601 (2019).
- [30] B. V. Fine, Typical state of an isolated quantum system with fixed energy and unrestricted participation of eigenstates, *Phys. Rev. E* **80**, 051130 (2009).
- [31] M. P. Müller, D. Gross, and J. Eisert, Concentration of measure for quantum states with a fixed expectation value, *Commun. Math. Phys.* **303**, 785 (2011).
- [32] P. Reimann and J. Gemmer, Why are macroscopic experiments reproducible? Imitating the behavior of an ensemble by single pure states, *Physica A* **552**, 121840 (2020).
- [33] L. Dabelow and P. Reimann, Perturbed Relaxation of Quantum Many-Body Systems, *Phys. Rev. Lett.* **124**, 120602 (2020).
- [34] L. Knipschild and J. Gemmer, Stability of quantum dynamics under constant hamiltonian perturbations, *Phys. Rev. E* **98**, 062103 (2018).
- [35] J. Richter, F. Jin, L. Knipschild, H. De Raedt, K. Michielsen, J. Gemmer, and R. Steinigeweg, Exponential damping induced by random and realistic perturbations, [arXiv:1906.09268](https://arxiv.org/abs/1906.09268).
- [36] L. Foini and J. Kurchan, Eigenstate thermalization hypothesis and out of time order correlators, *Phys. Rev. E* **99**, 042139 (2019).

- [37] D. Nickelsen and M. Kastner, Classical Lieb-Robinson Bound for Estimating Equilibration Timescales of Isolated Quantum Systems, *Phys. Rev. Lett.* **122**, 180602 (2019).
- [38] P. Figueroa-Romero, K. Modi, and F. A. Pollock, Almost Markovian processes from closed dynamics, *Quantum* **3**, 136 (2019).
- [39] D. Jansen, J. Stolpp, L. Vidmar, and F. Heidrich-Meisner, Eigenstate thermalization and quantum chaos in the holstein polaron model, *Phys. Rev. B* **99**, 155130 (2019).
- [40] D. J. Luitz and Y. Bar Lev, Anomalous Thermalization in Ergodic Systems, *Phys. Rev. Lett.* **117**, 170404 (2016).
- [41] L. Isserlis, On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, *Biometrika* **12**, 134 (1918).
- [42] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics: Emergence of Thermodynamic Behavior Within Composite Quantum Systems*, Vol. 784 (Springer, Berlin, 2009).