View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Explore Bristol Research

Wilming, H., de Oliveira, T., Short, T., & Eisert, J. (2018). Equilibration time scales in closed many-body quantum systems. In F. Binder, L. A. Correa, C. Gogolin, J. Anders, & G. Adesso (Eds.), *Thermodynamics in the quantum regime - Recent Progress and Outlook* (Vol. 20). Springer Berlin Heidelberg. https://doi.org/10.1088/1367-2630/aab03b

License (if available):
CC BY

Link to published version (if available):
10.1088/1367-2630/aab03b

Link to publication record in Explore Bristol Research

PDF-document

This is the final published version of the article (version of record). It first appeared online via IOP at http://iopscience.iop.org/article/10.1088/1367-2630/aab03b/meta . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

PAPER • OPEN ACCESS

Equilibration time scales in closed many-body quantum systems

To cite this article: Thiago R de Oliveira *et al* 2018 *New J. Phys.* **20** 033032

View the article online for updates and enhancements.

Related content

- Thermalization and prethermalization in isolated quantum systems: a theoretical overview
Takashi Mori, Tatsuhiko N Ikeda, Eriko Kaminishi et al.

- Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems
Christian Gogolin and Jens Eisert

- The role of quantum information in thermodynamics—a topical review
John Goold, Marcus Huber, Arnau Riera et al.

IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

This content was downloaded from IP address 137.222.114.243 on 22/11/2018 at 16:57

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
18 May 2017

REVISED
7 February 2018

ACCEPTED FOR PUBLICATION
16 February 2018

PUBLISHED
29 March 2018

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Equilibration time scales in closed many-body quantum systems

Thiago R de Oliveira[1,5], **Christos Charalambous**[2], **Daniel Jonathan**[1], **Maciej Lewenstein**[2,3] **and Arnau Riera**[2,4]

[1] Instituto de Física, Universidade Federal Fluminense, Av. Gal. Milton Tavares de Souza s/n, Gragoatá, 24210-346, Niterói, RJ, Brazil
[2] ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Castelldefels (Barcelona), E-08860, Spain
[3] ICREA-Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, Barcelona, E-08010, Spain
[4] Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany
[5] Author to whom any correspondence should be addressed.

E-mail: tro@if.uff.br

Keywords: equilibration, unitary dynamics, quantum systems, foundations of statistical mechanics

Abstract

We show that the physical mechanism for the equilibration of closed quantum systems is dephasing, and identify the energy scales that determine the equilibration timescale of a given observable. For realistic physical systems (e.g those with local Hamiltonians), our arguments imply timescales that do not increase with the system size, in contrast to previously known upper bounds. In particular we show that, for such Hamiltonians, the matrix representation of local observables in the energy basis is banded, and that this property is crucial in order to derive equilibration times that are non-negligible in macroscopic systems. Finally, we give an intuitive interpretation to recent theorems on equilibration time-scales.

1. Introduction

There is currently a renewed interest in the derivation of statistical mechanics from the kinematics and dynamics of a closed quantum system [1]. In this approach, instead of assuming *a priori* that the system is in some mixed state, such as e.g. a micro-canonical ensemble, one describes it at all times using a pure state $|\psi(t)\rangle$. One then seeks to show that, under reasonable conditions, the system behaves *as if* it were described by a statistical ensemble. In this way the use of statistical mechanics can be justified without introducing additional external degrees of freedom, such as e.g. thermal 'baths'.

A central part of this programme has been to understand the process of *equilibration*, i.e., how a constantly-evolving closed quantum system can behave as if relaxing to a stable equilibrium. The main insight relies on the fact [2–4] that, if measurements are limited to small subsystems or restricted sets of observables, then 'typical' pure states of large quantum systems are essentially indistinguishable from thermal states. It can then be shown [5, 6] that under very general conditions on the Hamiltonian and nearly all initial states, the system will eventually equilibrate, in the sense that an (again, restricted) set of relevant physical quantities will remain most of the time very close to fixed, 'equilibrium' values. For example, given some observable A and a system of finite but arbitrarily large size, if its expectation value $\langle A(t)\rangle$ equilibrates, then it must do so around the infinite time average (see section 5.1 of [1])

$$\overline{A} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \langle A(t)\rangle \, \mathrm{d}t. \tag{1}$$

If the infinite-time average fluctuation of $\langle A(t)\rangle$ around $\bar{A}$ is small, then we say that the observable A equilibrates.

One major open question is to understand the time scale at which equilibration occurs in a given system, and in particular its scaling with respect to system parameters such as its size (number of degrees of freedom). Various authors have tackled this question, e.g. [7–19], producing upper bounds that imply finite-time equilibration in various contexts (see also the 'supplementary information' section in [17] for a brief survey of the literature).

© 2018 The Author(s). Published by IOP Publishing Ltd on behalf of Deutsche Physikalische Gesellschaft

Several of these results [8–13, 17, 18] are again obtained in a typicality framework: they estimate, in various different senses, the *average* equilibration time of evolutions. While in many cases these calculated averages can have an impressive correspondence to experimentally measured equilibration times [17, 18], this approach also has some inherent weaknesses. First of all, it is generally believed that many specific physical conditions that are realisable in Nature or in the lab can be very far from 'typical' (for example, the actual Hamiltonians in Nature tend to have a locality structure that may be absent from most members of a mathematically generated ensemble [18]). In addition, by averaging, one loses information about the physical properties that are relevant to the equilibration time scale of any one specific evolution.

Bounds on equilibration times without taking averages have also been obtained, but only for certain restricted classes of evolutions or observables, e.g. in Malabarba *et al* [11], Farrelly [16], Goldstein *et al* [20], Monnai [7], and Santos *et al* [19].

Finally, there are a few works by the Bristol group [14, 15] that derive general and rigorous bounds on the equilibration times of arbitrary observables, systems and initial states, without any ensemble averaging. However, the bound in [15] scales with the inverse of the minimum energy difference (gap) in the system's spectrum. For physically realistic systems it therefore increase exponentially with the system size, and cannot be a good estimate of the actual equilibration timescale—in particular since equilibration would then not occur in the thermodynamic limit (see section 7 for details). In contrast, a bound derived in [14] can be independent of the system's size, however only in a regime that requires it to be initially in a nearly completely mixed state, failing to give a physically reasonable estimate in the case of a closed system in a pure initial state.

In this work, we seek to identify the properties of a closed quantum system which are relevant for the equilibration timescale of a given (arbitrary) observable. Our approach is more heuristic than rigorous, but it allows us to estimate a timescale which, under reasonable circumstances, depends only weakly on the system size, and thus seems to capture the relevant physics. The main insight we rely on is that equilibration is due primarily to a process of *dephasing* between different Fourier components of the dynamical evolution—a point that was briefly made in a classic [21], but that has apparently not been fully appreciated by the current community.

Although our argument does not result in a rigorous bound such as those in [14, 15], nor in a definite average evolution such as in [17, 18], we are able to discuss how the equilibration timescale of a given observable A depends on the physical properties of the system. Specifically, we find that the coherences of the observable of interest in the energy basis, $\langle E_i|A|E_j\rangle$, play a fundamental role. More specifically, the equilibration time depends critically on the range of energy gaps $E_i - E_j$ for which these coherences have non-negligible values. In particular, if this range remains roughly constant as the system size increases, the same will be true for the equilibration time. As we discuss below, this indeed happens for many observables of interest in many-body systems. We illustrate these results with numerical simulations of a spin chain, finding a reasonable qualitative agreement. It should be noted that similar heuristic methods and conclusions have also recently been proposed in simultaneous, independent research by Wilming *et al* [22].

This point of view also gives a new understanding of some existing results. For instance, it allows us to identify the reason for the limitations of bounds and estimates such as those in [15, 20], which, while rigorous, can vastly overestimate the time scale at which equilibration occurs in realistic systems. In section 7, we argue that the reason for this behaviour is that these estimates ultimately rely on the wrong physical mechanism of equilibration, disregarding the crucial role played by dephasing.

Our main findings can be summarised as follows:

- In section 3 we discuss qualitatively why dephasing is the underlying mechanism of equilibration in closed quantum systems.
- In section 4 we develop a formalism based on the coarse-graining of functions in frequency space, which allows us to apply basic tools from Fourier transform theory, such as uncertainty relations, to equilibration related questions.
- In section 5 we determine the relevant energy scales that govern the equilibration time for a given observable, namely the energy fluctuations of the initial state and the bandwidth of the matrix of the observable in the energy basis. In particular, we give an independent proof of the fact [23] that local observables have banded matrices when written in the energy basis of a short-ranged spin Hamiltonian.
- In section 6 we illustrate our results with a numerical simulation of the XXZ model.
- In section 7 we discuss some implications of our results. In section 7.1, we reinterpret existing results from the point of view of our dephasing framework. In section 7.2 we discuss implications for the fields of quantum

chaos and integrability. In particular we present two models with identical eigenbases but different level statistics that have indistinguishable dynamics over realistic time-scales.

2. General setting and definition of the problem

Let us consider a closed system whose state is described by a vector in a Hilbert space of dimension d_T and whose Hamiltonian has a spectral representation

$$H = \sum_{k=1}^{d_E} E_k P_k, \tag{2}$$

where E_k are its energies and P_k the projectors onto its eigenspaces. Note that the sum runs over $d_E \leqslant d_T$ terms, since some eigenspaces can be degenerate.

We denote the initial state by $|\psi(0)\rangle$. If the Hamiltonian has degenerate energies, we choose an eigenbasis of H such that $|\psi(0)\rangle$ has non-zero overlap with only one eigenstate $|E_k\rangle$ for each distinct energy. Choosing units such that $\hbar = 1$, the state at time t is then given by

$$|\psi(t)\rangle = \sum_k c_k \mathrm{e}^{-\mathrm{i}E_k t}|E_k\rangle, \tag{3}$$

with $c_k \equiv \langle E_k|\psi(0)\rangle$. It is clear that $|\psi(t)\rangle$ evolves in the subspace spanned by $\{|E_k\rangle\}$ as if it were acted on by the non-degenerate Hamiltonian $H' = \sum_k E_k|E_k\rangle\langle E_k|$. In this case, if the system equilibrates, the equilibrium state must be

$$\omega = \sum_k |c_k|^2 |E_k\rangle\langle E_k|. \tag{4}$$

In this article, following a number of authors, [5, 6, 14, 15] we will study equilibration by focusing on observables. The idea is that a system can be considered in equilibrium if all experimentally relevant (typically, coarse-grained) observables A have equilibrated. In other words, we will focus on understanding how the expectation value $\langle A(t)\rangle$ approaches its equilibrium value $\mathrm{Tr}(A\omega)$. Note that, in order to even talk about equilibration time scales, we must assume that such a condition holds, i.e., that this observable sooner or later equilibrates.

Let us introduce the *time signal* of A, given the initial state $|\psi(0)\rangle$, as the distance of $\langle A(t)\rangle$ from equilibrium at time t

$$\begin{aligned} g(t) &:= \frac{1}{\Delta_A}(\langle\psi(t)|A|\psi(t)\rangle - \mathrm{Tr}(A\omega)) \\ &= \frac{1}{\Delta_A}\sum_{i\neq j}(c_j^* A_{ji} c_i)\mathrm{e}^{-\mathrm{i}(E_i - E_j)t}, \end{aligned} \tag{5}$$

where $A_{ij} := \langle E_i|A|E_j\rangle$ are the matrix elements of A in the energy eigenbasis, and $\Delta_A = a_{\max} - a_{\min}$ is the range of possible outcomes, being $a_{\max(\min)}$ the largest (smallest) eigenvalue of A. The denominator Δ_A is introduced to make the time signal dimensionless and satisfying $|g(t)| \leqslant 1$. Note that the time signals of two observables A and $A' = b(A - a_0)$ are identical for $a_0,\ b \in \mathbb{R}$.

We can conveniently rewrite the time signal as

$$g(t) = \sum_{\alpha\in\mathcal{G}} \nu_\alpha \mathrm{e}^{\mathrm{i}G_\alpha t}, \tag{6}$$

where $\alpha \in \mathcal{G} = \{(i, j)\colon i, j \in \{1, \ldots d_E\},\ i \neq j\}$ labels each energy gap $G_\alpha = (E_j - E_i)$ appearing in the system's spectrum, and where

$$\nu_\alpha = \nu_{(i,j)} = \frac{c_j^* A_{ji} c_i}{\Delta_A}. \tag{7}$$

We will refer to the complex number ν_α as the *amplitude* of the corresponding gap G_α, and to its normalised square modulus $q_\alpha := |\nu_\alpha|^2/\sum_\beta|\nu_\beta|^2$ as the *relevance* of G_α. Note that the set of relevances form a probability distribution over $\mathcal{G}$.

A physical interpretation of this normalisation factor can be given as follows [5, 6, 24]: note that, if the system has *non-degenerate gaps*, then the time-averaged fluctuations of the time signal

$$\langle|g|^2\rangle_T := \frac{1}{T}\int_0^T \mathrm{d}t|g(t)|^2 \tag{8}$$

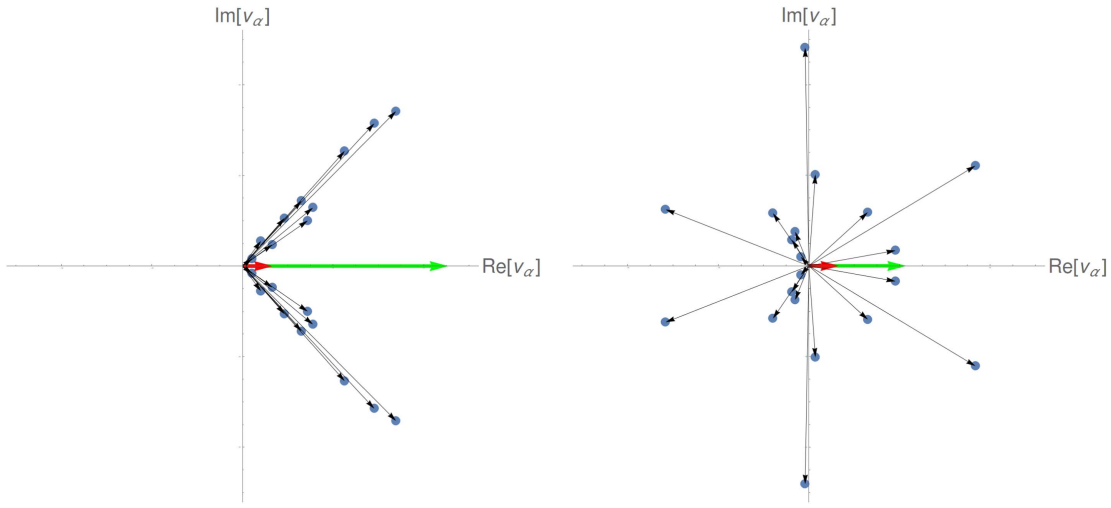

Figure 1. Illustration of the dephasing process of the complex terms $v_\alpha e^{iG_\alpha t}$ (blue dots/black arrows) of a time signal $g(t)$ (green arrow), see equation (6). On the left, the system is far from equilibrium, a short time $t = 0$, having been initialised with all v_α real, and $g(t)$ is substantial. Note that half of the complex terms have rotated clockwise, and the other half anti-clockwise, as expected from the symmetry of the set of gaps G_α. On the right, after a long time, the individual complex vectors have become spread out, and the system has equilibrated, with $g(t)$ becoming close to the typical fluctuation $\sqrt{\langle |g|^2 \rangle_\infty}$ (red arrow).

satisfy the limit

$$\langle |g|^2 \rangle_\infty := \lim_{T\to\infty} \langle |g|^2 \rangle_T = \sum_\alpha |v_\alpha|^2. \tag{9}$$

In other words, this quantity gives the infinite time average of (the square of) the deviation of $\langle A(t) \rangle$ from its equilibrium value. We consider that the observable A equilibrates if this quantity is small, in the sense that

$$\langle |g|^2 \rangle_\infty \lesssim \left(\frac{\delta A}{\Delta_A} \right)^2, \tag{10}$$

where δA is the experimentally available resolution. Furthermore, note that $\Delta_A/\delta A$ quantifies the the amount of different possible outcomes of a measurement and hence $\delta A/\Delta_A$ is expected to be much much smaller than 1. From now on, without loss of generality and for the sake of simplicity we will consider that $\Delta_A = 1$.

In [5, 6, 24], sufficient conditions for equilibration in the sense defined above are given by bounding the average distance from equilibrium equation (9). In particular, in [24] it is shown that for any Hamiltonian with non-degenerate gaps,

$$\sum_\alpha |v_\alpha|^2 \leqslant \frac{1}{d_{\text{eff}}}, \tag{11}$$

where the effective dimension is defined as $d_{\text{eff}} := 1/(\sum_k |c_k|^4) = 1/\operatorname{Tr}(\omega^2)$, which roughly speaking tells us how many eigenstates of the Hamiltonian participate in the superposition of the initial state. Thus, a large effective dimension, which is usually the case in many body systems (see appendix A and [25]), is sufficient to guarantee that condition (10) will be satisfied. Concerning the assumption of the Hamiltonian having non-degenerate gaps, it is shown in [15] that as long as there are not exponentially many degeneracies the argument stays the same.

3. Equilibration as dephasing

Let us consider a situation in which the initial state is out of equilibrium, i.e. the time signal of a given operator is initially significantly larger than the equilibrium value: $|g(0)| \gg \sqrt{\langle |g|^2 \rangle_\infty}$. For this to happen, the phases of the complex numbers v_α in the time signal equation (6) need to be highly synchronised. This case is presented pictorially in figure 1 (left) where the v_α's are depicted as points in the complex plane.

Result 1 (Equilibration is dephasing). Given a time signal $g(t) = \sum_\alpha v_\alpha e^{iG_\alpha t}$ with $v_\alpha = |v_\alpha| e^{i\theta_\alpha} \in \mathbb{C}$ being the initial amplitude of the gap $G_\alpha \in \mathbb{R}$, a necessary condition for the system to be initially out of equilibrium, i.e., $|g(0)|$ significantly larger than the typical equilibrium fluctuation $\sqrt{\langle |g|^2 \rangle_\infty}$, is that the initial phases θ_α are not isotropically distributed but significantly synchronised. More precisely, we quantify the distance from equilibrium as

$$|g(0)|^2 - \langle |g|^2 \rangle_\infty = 2 \sum_{\alpha<\beta} |v_\alpha||v_\beta| \cos(\theta_\alpha - \theta_\beta), \tag{12}$$

which becomes negligible when the phases θ_α are isotropically distributed.

Equation (12) follows from a straightforward calculation

$$|g(0)|^2 = \left(\sum_\alpha v_\alpha\right)\left(\sum_\beta v_\beta\right)^* = \langle |g|^2 \rangle_\infty + 2 \sum_{\alpha<\beta} |v_\alpha||v_\beta| \cos(\theta_\alpha - \theta_\beta). \tag{13}$$

To see that isotropic randomly distributed phases give $|g(0)|^2 \simeq \langle |g|^2 \rangle_\infty$, let $v_\alpha = |v_\alpha| e^{i\theta_\alpha}$ be a set of independent random complex variables with an isotropic probability distribution $p_\alpha(r, \theta) = p_\alpha(r) = \frac{1}{2\pi}\delta(r - r_\alpha)$, i.e. the random variable v_α has fixed modulus r_α and a random phase θ_α uniformly distributed around the circle. Then, the variance of the random variable $\sum_\alpha v_\alpha$ is

$$\mathrm{var}\left(\sum_\alpha v_\alpha\right) = \sum_\alpha \mathrm{var}(v_\alpha) = \sum_\alpha \langle |v_\alpha|^2 \rangle = \sum_\alpha |v_\alpha|^2, \tag{14}$$

where we have used the fact that the variance of a sum of independent random variables is the sum of variances and the first moments $\langle v_\alpha \rangle = 0$. That is, if the phases of v_α are uniformly distributed, then the typical initial value of the time signal is $|g(0)| = |\sum_\alpha v_\alpha| \simeq (\sum_\alpha |v_\alpha|^2)^{1/2}$.

In contrast, an out-of-equilibrium initial state, for a distribution of gaps with a non-zero dispersion and a large enough system, will evolve to an equilibrium state, represented by an isotropic cloud of points centred at the origin of the complex plane as shown in figure 1 (right). This mechanism for equilibration, or more generally the vanishing amplitude of a signal over time, is usually called *dephasing* and is a well-known feature in many different fields of physics [19, 26–32],

The time-scale necessary for the sum in equation (6) to dephase can be estimated with the following simple argument. Suppose for simplicity that all terms have the same phase at $t = 0$. One way to lower bound the time needed for these phases to spread around the whole range $(0, 2\pi)$ is to use the time it takes for the difference between the fastest and slowest phase to differ by 2π: $tG_{\max} - tG_{\min} = 2\pi$ which leads to $T_{\mathrm{eq}} \sim 2\pi/(G_{\max} - G_{\min})$. However, the slowest and fastest gaps are not necessarily very relevant to the sum, that is, they can have a relatively small amplitude v_α. For this reason, a better estimate is obtained by replacing the denominator $G_{\max} - G_{\min}$ with $2\sigma_G$, where we define the *gap dispersion* σ_G to be simply the standard deviation of the G_α when weighted by their respective relevances q_α, i.e.

$$\sigma_G^2 := \sum_\alpha q_\alpha (G_\alpha - \mu_G)^2. \tag{15}$$

Since for each gap $G_{(i,j)} = E_j - E_i$ there is also $G_{(j,i)} = -G_{(i,j)}$, then by symmetry the average gap is $\mu_G = 0$, and also $|v_{(i,j)}| = |v_{(j,i)}|$. By the argument made above, we can thus expect that, at least in cases where the distribution G_α is of a unimodal type, the equilibration time can be estimated by

$$T_{\mathrm{eq}} \sim \pi/\sigma_G. \tag{16}$$

Note that similar estimates are made in the various fields where dephasing is relevant, e.g. in [29, 31, 32]. In the next section we will give a more detailed justification for this estimate using standard tools from Fourier transform theory.

The physical properties that control the dispersion σ_G, and thus the equilibration time T_{eq}, can be identified once we note that the probabilities q_α are proportional to the $|v_\alpha|^2$. Since $v_\alpha \propto c_i^* c_j \langle E_i|A|E_j\rangle$ (equation (7)), then σ_G is determined by: (i) the probability distribution $|c_i|^2$ for the energies; (ii) the matrix-elements of the observable A in the energy basis; and (iii) the distribution of the values of the gaps $G_\alpha = E_j - E_i$ themselves.

In section 5 we identify the energy scales of the observable and the initial state which are relevant to determine the dispersion of gaps σ_G. Let us now anticipate the requirements on the initial state and the observable in order that σ_G does not diverge in the thermodynamic limit, i.e. when the system size $n \to \infty$ and the equilibration time T_{eq} does not vanish. We can identify two regimes where this will happen:

- First, there are observables A that equilibrate in finite time regardless of the details of the initial state. Note that, for large many-body systems, the variance of the density of states usually scales as $\sqrt{n}$, and for generic states the width of the distribution of $|c_i|^2$ is as wide as the energy spectrum, which increases with n. This implies that, unless restrictions are placed on A, the gap variance σ_G will have the same scaling and T_{eq} will vanish in the thermodynamic limit. In order to avoid this problem, the variance of the distribution of v_α should not increase with n. This requires that the distribution $\langle E_i|A|E_j\rangle$ should decrease (or become null) for large values of $E_j - E_i$—in other words, A needs to be *banded* in the energy basis.

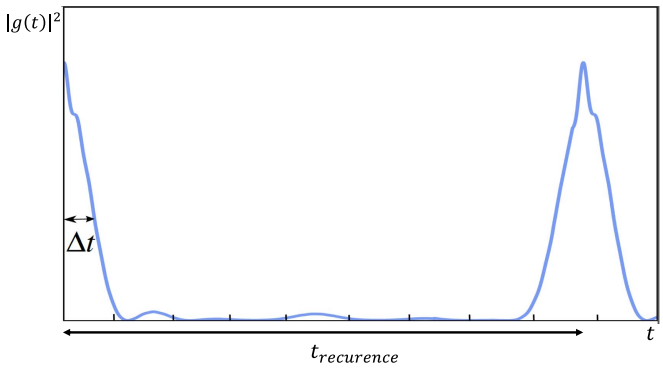

Figure 2. Example of a time signal which has a recurrence time. In the absence of recurrences, a good estimate the equilibration time scale T_{eq} would be a few multiples of Δt defined in equation (17).

- Conversely, suppose the initial state has support over an energy range that does not scale with n. This can happen, for example, in the case of a so-called 'local quench' [33], when a local subsystem of fixed dimension is excited regardless of the full size of the system. In this case, σ_G will also at worst be independent of n, and so even observables with long-range coherence between very different energies, which would otherwise equilibrate quickly, will now take a finite time T_{eq}.

Finally, it is also important to mention that the gap dispersion σ_G may actually decay to zero with n, leading to equilibration times that become very long. As an example of this situation consider two subsystems of increasing size n interacting through a spatially localised border of fixed size. The coherence in the energy basis of the interaction Hamiltonian is bounded by the operator norm of such an interaction, matching with the intuition that the stronger the interaction between systems, the faster the relaxation, and vice versa. By rescaling the global Hamiltonian, we can see that the interaction terms become relatively weaker as n grows and thereby the equilibration slower.

4. Fourier description of the dephasing framework

In this section we give further substance to the above heuristic argument to estimate the equilibration time-scale by means of Fourier transform techniques.

Let us first give a general idea of our approach. Suppose the time-signal $g(t)$ decayed more or less steadily to zero, and stayed there. If so, then a good estimate for the equilibration time scale would be given by a few multiples of the standard deviation Δt defined by

$$\Delta t^2 := \frac{1}{\|g\|_2} \int \mathrm{d}t |g(t)|^2 (t - \mu_t)^2 \tag{17}$$

with $\mu_t := \int \mathrm{d}t |g(t)|^2 t \big/ \|g\|_2$ (see figure 2).

In such a case, following the spirit of our previous heuristic example, it would also be tempting to estimate the order of magnitude of Δt by taking the inverse of the spectral variance $\Delta\omega$ of the signal. Indeed, this can be justified if we recall the standard uncertainty principle of Fourier analysis [34]

$$\Delta t \cdot \Delta\omega \geqslant 1/2. \tag{18}$$

Of course, this is a lower bound, that is saturated exactly only in the case of a Gaussian spectrum. However, it will be nearly saturated ($\Delta t \cdot \Delta\omega = c_1$, where c_1 is a constant of order 1), when the spectrum is unimodal and without long tails. In this case, we can also expect the time signal to decrease to a very small value after a time $c_2 \Delta t$, for some small multiple c_2 again of order 1. Taking both multiples together, we expect a good estimate for the equilibration time to be of order $\sim \frac{c_1 c_2}{2\Delta\omega}$.This is satisfied by equation (16), which we will therefore continue to take as our estimate.

Unfortunately, for finite systems, our initial assumption of steady decay does not apply. The time signal has recurrences, that is, a long time after the dephasing has occurred and the system has equilibrated, the phases get again aligned (synchronised) in the complex plane, and the signal regains strength (figure 2). In order to avoid this problem, in the next subsections we introduce a coarse-grained version of the signal spectrum, which dampens out the recurrences. This allows us to exploit the uncertainty principle to estimate the equilibration time-scales, as described above. Under some mild conditions, we show that the equilibration time-scale

estimated provided by this procedure coincides with the one previously given by the heuristic argument of points dephasing in the complex plane.

4.1. The frequency signal

We define the *frequency signal*, $\tilde{g}(\omega)$, as the Fourier transform of the time signal $g(t)$

$$\tilde{g}(\omega) \coloneqq \mathcal{F}[g](\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) \mathrm{e}^{-\mathrm{i}\omega t} \mathrm{d}t \tag{19}$$

which roughly speaking tells us the relevance with which every frequency contributes to the time signal. When both the time and frequency signals are square-integrable ($g,\ \tilde{g} \in L^2$), the standard uncertainty principle of equation (18) applies, where

$$\Delta\omega^2 \coloneqq \frac{1}{\|\tilde{g}\|_2} \int \mathrm{d}\omega |\tilde{g}(\omega)|^2 (\omega - \mu_\omega)^2 \tag{20}$$

with $\mu_\omega \coloneqq \int \mathrm{d}\omega |\tilde{g}(\omega)|^2 \omega / \|\tilde{g}\|_2$.

However, in the case of time signals such as in equation (6), $g(t) \notin L^2$, as can be seen from equation (9) (the integral diverges proportional to T). The same is true for the frequency signal, since

$$\tilde{g}(\omega) = \mathcal{F}[g](\omega) = \sum_\alpha \nu_\alpha \delta(\omega - G_\alpha), \tag{21}$$

where $\delta(x)$ is the Dirac delta distribution. Hence, the uncertainty principle in equation (18) cannot be directly applied.

It is worth noting here that, due to the finite range of energies present in our system, there is an asymmetry between the uncertainties in time and frequency of the signal $g(t)$. On the one hand, the uncertainty in frequency can still be well-defined. To see how, recall first that, for $g(t) \in L^2$, it is possible to write the moments of a frequency signal in terms of the corresponding time signal and its derivatives, e.g.

$$\langle \omega^2 \rangle = -\frac{\int_{-\infty}^{\infty} g^*(t) g''(t) \mathrm{d}t}{\int_{-\infty}^{\infty} |g(t)|^2 \mathrm{d}t}. \tag{22}$$

In our case, although each of these integrals diverges, their ratio does have a well-defined limit, in the sense that

$$\lim_{T\to\infty} -\frac{\int_{-T}^{T} g^*(t) g''(t) \mathrm{d}t}{\int_{-T}^{T} |g(t)|^2 \mathrm{d}t} = \frac{\sum_\alpha |\nu_\alpha|^2 G_\alpha^2}{\sum_\alpha |\nu_\alpha|^2}. \tag{23}$$

Taking then this limit as the appropriate definition of $\langle \omega^2 \rangle$ in this case, and noting that, in the same sense, $\langle \omega \rangle = 0$, we obtain that $\Delta\omega = \sqrt{\langle \omega^2 \rangle}$ is indeed precisely equal to the gap dispersion σ_G defined in equation (15).

On the other hand, though, the value of Δt diverges, even when taking limits in the same sense above. This can be understood physically due to the previously mentioned recurrences in the time signal. Indeed, $g(t)$ is a quasi-periodic function that experiences, over an infinitely large time interval, an infinite number of recurrences to a value arbitrary close to its initial one [35, 36].

4.2. The coarse-grained signal

We now define the notion of *coarse-graining*, in which we introduce a microscopic energy scale ϵ below which the fine-grained details of the spectrum are washed out. As we show below, this is done by replacing the discrete spectrum present in equation (21) by a suitable smooth version.

As previously mentioned, such coarse graining of the frequency signal dampens the time signal $g(t)$, removing the recurrences seen in figure 2 and making $g(t)$ and $\tilde{g}(\omega)$ belong to L^2. We shall see later that it will also allow us to exploit certain existing statements concerning the shapes of *energy densities* and *density of states* of realistic initial states and short-ranged local Hamiltonians [37], which will justify our assumption of quasi-saturating the uncertainty bound.

An important issue in the coarse-graining is obviously the choice of the energy scale ϵ. This will be discussed later in detail but we can already understand that in order to remove the recurrences, the discreteness of the frequency signal has to be removed, which implies an ϵ much larger than the separation between consecutive gaps.

Mathematically, the coarse-graining is accomplished by convolving the frequency signal with an appropriate *window function* $h_\epsilon(x)$, which is only nonzero over an interval of size $O(\epsilon)$. In our case, we find it convenient to choose $h_\epsilon(x) = C N_\epsilon(x)$, where

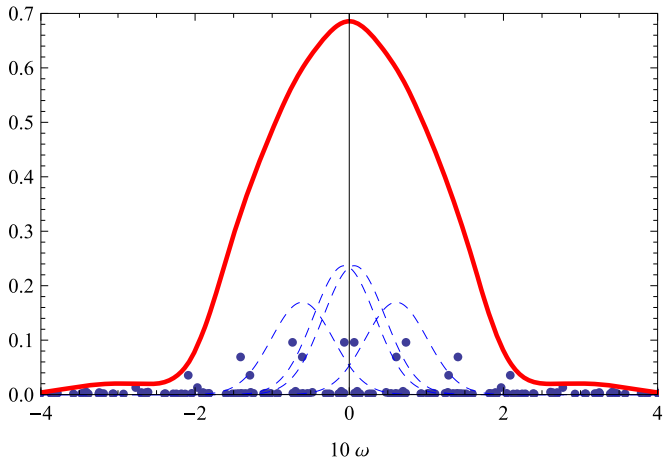

Figure 3. Illustration of the coarse-graining of a discrete gap spectrum with a Gaussian window function. The data here corresponds to the *XXZ* model that is studied in section 6, with $n = 12$ spins. The solid blue dots represent the amplitudes $\nu_\alpha(G_\alpha)$ of each gap (in this particular case they are all real and positive). The dashed blue lines illustrate a few of the corresponding weighted Gaussians ν_α $h_\epsilon(\omega - G_\alpha)$, where we have chosen $\epsilon = 0.4$ (in arbitrary frequency units). The solid red curve represents the full coarse-grained spectrum $\tilde{g}_\epsilon(\omega)$, obtained by summing these weighted Gaussians, according to equation (26). Note that, for this choice of ϵ, the width $\Delta\omega$ of the coarse-grained spectrum remains close to the dispersion σ_G of the original (discrete) gap spectrum. See figure 6 for a comparison of the corresponding coarse-grained time signal $g_\epsilon(t)$ with the exact one. Note finally that although, for simplicity, we use here a single numerical scale on the vertical axis, the ν_α are adimensional, whereas the continuous curves have physical dimension of time (with units that are the inverse of those used for ω and ϵ).

$$N_\sigma(x) := \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{24}$$

is the normalised Gaussian distribution centred at the origin and with standard deviation σ, and C a constant that is determined below.

With this spirit, the ϵ-coarse-grained version of the frequency signal is defined as

$$\tilde{g}_\epsilon(\omega) := (h_\epsilon * \tilde{g})(\omega) = \int_{-\infty}^{\infty} d\omega' h_\epsilon(\omega' - \omega)\tilde{g}(\omega'). \tag{25}$$

If $\tilde{g}(\omega)$ is given by equation (21), then the ϵ-coarse-grained frequency signal is

$$\tilde{g}_\epsilon(\omega) = \sum_\alpha \nu_\alpha\, h_\epsilon(\omega - G_\alpha). \tag{26}$$

In other words, coarse-graining corresponds to widening each Dirac-δ in the original spectral function into a Gaussian of $O(\epsilon)$ width (figure 3). Note that, in doing this, we remove fine details of the spectrum such as the *level statistics*. Furthermore, unlike $\tilde{g}(\omega)$, $\tilde{g}_\epsilon(\omega)$ is square integrable and lies in L^2:

$$\int d\omega |\tilde{g}_\epsilon(\omega)|^2 = \sum_{\alpha,\beta} \nu_\alpha \nu_\beta^*\, (h_\epsilon * h_\epsilon)(G_\alpha - G_\beta) = C^2 \sum_{\alpha,\beta} \nu_\alpha \nu_\beta^*\, N_{\sqrt{2}\epsilon}(G_\alpha - G_\beta) < \infty, \tag{27}$$

where we have used the fact that $(h_\epsilon * h_\epsilon)(x) = C^2 N_{\sqrt{2}\epsilon}(x)$.

A coarse-grained frequency signal defines a coarse-grained time signal given by

$$g_\epsilon(t) = \mathcal{F}^{-1}[\tilde{g}_\epsilon](t) = \mathcal{F}^{-1}[h_\epsilon](t) \cdot g(t) = \frac{C}{\sqrt{2\pi}} e^{-\frac{1}{2}\epsilon^2 t^2} g(t), \tag{28}$$

where we have used the convolution theorem for Fourier transforms. The constant C is fixed by imposing that the time signal is not affected by coarse-graining the frequency signal in time scales $t \ll \epsilon^{-1}$ i.e. $g_\epsilon(0) = g(0)$. This leads to $C = \sqrt{2\pi}$ and

$$h_\epsilon(\omega) = \sqrt{2\pi} N_\epsilon(\omega) = \frac{1}{\epsilon} e^{-\frac{\omega^2}{2\epsilon^2}}, \tag{29}$$

$$g_\epsilon(t) = e^{-\frac{1}{2}\epsilon^2 t^2} g(t). \tag{30}$$

Even if the time signal $g(t)$ equilibrates after some time, we know that it must eventually have recurrences. To determine the equilibration time-scale from $g_\epsilon(t)$, we need that ϵ^{-1} is much greater than the equilibration timescale, but much smaller than the recurrence timescale. In this way, we ensure that the coarse-grained time signal will be indistinguishable from the original one during the equilibration process, but unlike the latter will then decay to zero.

4.3. The coarse-grained density of relevant gaps

We now focus on the properties of the variance of gap values with respect to the coarse-grained frequency signal $\tilde{g}_\epsilon(\omega)$, i.e.

$$\Delta\omega_\epsilon^2 := \frac{1}{\|\tilde{g}_\epsilon\|_2}\int \mathrm{d}\omega |\tilde{g}_\epsilon(\omega)|^2(\omega - \tilde{\mu}_\omega)^2 \tag{31}$$

with $\tilde{\mu}_\omega := \int \mathrm{d}\omega |\tilde{g}_\epsilon(\omega)|^2\omega \big/ \|\tilde{g}_\epsilon\|_2$. We will refer to $\Delta\omega_\epsilon$ as the 'dispersion of relevant gaps', and to the weight

$$|\tilde{g}_\epsilon(\omega)|^2 = \sum_{\alpha,\beta} \nu_\alpha \nu_\beta^* h_\epsilon(\omega - G_\alpha) h_\epsilon(\omega - G_\beta) \tag{32}$$

as the *density of relevant gaps*.

Our goal is to show that, under a wide range of choices of ϵ and of physically relevant circumstances: (i) $\Delta\omega_\epsilon^2$ is very close to the gap dispersion σ_G^2 of the original signal, as defined in equation (15), and at the same time (ii) the inverse $\Delta\omega_\epsilon^{-1}$ is a good estimate for the equilibration time.

Note first that, by construction, for every gap $G_\alpha = E_j - E_i$ in equation (6) (and, by extension, equation (26)), its negative $G_{\bar{\alpha}} := E_i - E_j$ also appears, with $|\nu_{\bar{\alpha}}| = |\nu_\alpha|$. Hence the 'average gap' $\tilde{\mu}_\omega$ vanishes for any ϵ and the variance $\Delta\omega^2$ is equivalent to $\langle\omega^2\rangle$, i.e.

$$\Delta\omega_\epsilon^2 = \frac{\int \omega^2 |\tilde{g}_\epsilon(\omega)|^2 \mathrm{d}\omega}{\int |\tilde{g}_\epsilon(\omega)|^2 \mathrm{d}\omega}. \tag{33}$$

After some straightforward manipulation, using equations (29) and (32), we obtain

$$\Delta\omega_\epsilon^2 = \frac{\sum_\alpha |\nu_\alpha|^2 (G_\alpha^2 + \epsilon^2/2) + \frac{1}{4}\sum_{\alpha\neq\beta}\nu_\alpha\nu_\beta^*[(G_\alpha + G_\beta)^2 + 2\epsilon^2]\mathrm{e}^{-\frac{(G_\alpha - G_\beta)^2}{4\epsilon^2}}}{\sum_\alpha |\nu_\alpha|^2 + \sum_{\alpha\neq\beta}\nu_\alpha\nu_\beta^*\,\mathrm{e}^{-\frac{(G_\alpha - G_\beta)^2}{4\epsilon^2}}}. \tag{34}$$

It can be easily checked that, if $\epsilon \to 0$, this expression indeed reduces to equation (15). More specifically, $\Delta\omega_\epsilon^2$ will be very close to σ_G^2 for all $\epsilon \ll \min(G_\alpha - G_\beta)$. Indeed, in this limit the Gaussian window function $h_\epsilon(\omega)$ becomes negligibly thin with respect to the smallest separation between gaps, and the coarse-grained spectrum $\tilde{g}_\epsilon(\omega)$ resembles the original discrete spectrum $g(\omega)$. Precisely for this reason, however, this limit is of little use to our goals. Another way of putting this is that, for such small values of ϵ the coarse-grained time signal (equation (30)) does not have time to decay before the recurrence timescale of the original signal, which is of order $\min(G_\alpha - G_\beta)^{-1}$.

To make further progress at this point, it is necessary to assume some features about the amplitudes ν_α. Otherwise, a fine tuning between phases and modulus of ν_α can make $|\tilde{g}_\epsilon(\omega)|$ have an arbitrary behaviour, preventing any general statement concerning the equilibration of $g_\epsilon(t)$.

Inspired by [25], we will adopt a weak-typicality point of view: let us assume that the evolution we are considering is drawn from an ensemble for which the ν_α's are describable by some *smooth functions plus stochastic fluctuations*. Note that we do not assume a uniform ensemble over all states (or any one specific ensemble), as is the case in most typicality studies [8–13, 17, 18], merely one for which the resulting distribution over the ν_α's has some very general features which are described below. In the spirit of statistical physics, we basically *replace complexity by apparent randomness*. In most situations, the description of the gap relevances ν_α in terms of a smooth function plus stochastic fluctuations is a consequence of the energy level populations c_i and the matrix-elements of the observable A_{ij} having this same behaviour. In the next section we discuss under which conditions this is indeed the case.

In the following we show that the process of coarse-graining removes the fluctuations and makes the density of relevant gaps $|\tilde{g}_\epsilon(\omega)|^2$ have a smooth behaviour.

Result 2 (Coarse-grained frequency signal). Let us consider the amplitudes ν_α of the gaps G_α to be described by

$$\nu_\alpha = \nu(G_\alpha) + \delta\nu_\alpha, \tag{35}$$

where $\nu(\omega) = r(\omega)\mathrm{e}^{\mathrm{i}\theta(\omega)}$, with $r(\omega)$ and $\theta(\omega)$ two functions with a Lipshitz constant upper bounded by some $K \ll \epsilon^{-1}$, and where $\delta\nu_\alpha$ are independent random variables that average to zero $\langle\delta\nu_\alpha\rangle = 0$ and $\langle\delta\nu_\alpha\,\delta\nu_\beta\rangle = \gamma^2(G_\alpha)\delta_{\alpha\beta}$. The variance $\gamma^2(\omega)$ is a function that represents the strength of the fluctuations and also has a Lipshitz constant upper bounded by K. Then, with a high probability $\mathrm{erf}(m) \geqslant 1 - \exp(-m^2)$, the density of relevant gaps fulfils the following bound

$$|\tilde{g}_\epsilon(\omega) - \sqrt{2\pi}\,\nu(\omega)\rho_\epsilon(\omega)| \leqslant \rho_\epsilon(\omega)\left(c_1 K\epsilon + \pi^{1/4} m \frac{\gamma(\omega)}{\sqrt{\epsilon\rho_\epsilon(\omega)}}\right), \tag{36}$$

where $c_1 > 0$ is a constant, $\rho_\epsilon(\omega) = (\rho^* N_\epsilon)(\omega)$ is the *coarse-grained density of gaps*, with $\rho(\omega) = \sum_\alpha \delta(\omega - G_\alpha)$ being the density of gaps. The coarse-grained density of gaps $\rho_\epsilon(\omega)$ describes how many gaps are ϵ-close to the frequency ω. That is, the process of coarse-graining washes out the fluctuations $\delta\nu_\alpha$ turning the coarse-grained frequency signal into a smooth function

$$\tilde{g}_\epsilon(\omega) \simeq \sqrt{2\pi}\,\nu(\omega)\rho_\epsilon(\omega), \tag{37}$$

where the meaning of the approximation is made precise in equation (36).

The proof is tedious and non-illuminating so we give it in appendix B.

Note that the error in (36) has two components. The term $K\epsilon$ is due to the variation of the 'smooth' functions with Lipshitz constant smaller than K within an interval of width $O(\epsilon)$. The second component $\gamma(\omega)/\sqrt{\epsilon\rho_\epsilon(\omega)}$ is the fluctuation that shrinks according to the central limit theorem with $\sqrt{\epsilon\rho_\epsilon(\omega)}$, where $\epsilon\,\rho_\epsilon(\omega)$ is the number of gaps G_α within an interval $O(\epsilon)$.

In order for equation (36) to be meaningful and the error bound small, an optimisation over ϵ is needed. It is easy to see that for the error to be small, the parameter ϵ should be much larger than the spacing between consecutive gaps and much smaller than the inverse of the Lipshitz constants of the continuous functions $\nu(\omega)$ and $\gamma(\omega)$, i.e.

$$G_{\alpha+1} - G_\alpha \ll \epsilon \ll K^{-1}. \tag{38}$$

Here 'consecutive' refers to gaps ordered by size, and we use the informal notation '$G_{\alpha+1}$' as a shorthand for 'the gap immediately larger than G_α'. Note that, in practice, the lower bound above should be applied only for consecutive *relevant* gaps, i.e., gaps that whose amplitudes ν_α make a non-negligible contribution to the sum in equation (6).

For many-body systems, the number of gaps increases exponentially in the system size n, and energy differences between consecutive gaps shrink exponentially to zero in n. If for example these gaps all have roughly equal (exponentially small) relevances q_α, then ϵ can also be taken exponentially to zero, as long as this is done at a slower rate than the difference in gaps. This makes the time signal and its coarse-grained version indistinguishable in any realistic time-scale.

In sum, if the gap relevances ν_α can be described by a continuous part plus a fluctuating part, as in equation (35), then the coarse-grained density of relevant gaps is a smooth function given by

$$|\tilde{g}_\epsilon(\omega)|^2 \simeq 2\pi|\nu(\omega)|^2\rho_\epsilon(\omega)^2. \tag{39}$$

This will be particularly useful in the following section, where we apply these ideas in the case of many body systems described by short-ranged Hamiltonians.

It is worth noting that the factorisation in equation (37) is also automatically obtained if one assumes, as is often done, that in the thermodynamic limit $n \to \infty$ the discrete gap spectrum may be replaced by a smooth continuous gap density. i.e. taking

$$g(t) = \sum_\alpha \nu_\alpha e^{iG_\alpha t} \to \int \nu(\omega) e^{i\omega t}\rho_G(\omega)\,d\omega. \tag{40}$$

Comparing with the definition in equation (19), we see that in this case the frequency signal is again $\tilde{g}(\omega) = \sqrt{2\pi}\;\nu(\omega)\rho_G(\omega)$. This indicates that the assumptions we have made concerning the smoothness of ν_α are not severe. However, the point of attaining this relation via coarse-graining, while maintaining a finite dimension n, is that it allows us to control how the equilibration timescale scales with the system size, and in particular to understand how this scaling depends on the energy scaling of the relevant observable. We turn to this question in the next section.

5. Relevant energy scales and equilibration time scales for local Hamiltonians

In this section, we focus on the particular but relevant case of short-range Hamiltonians and initial states that have a finite correlation length. For such systems, we express the density of relevant gaps in terms of the energy density of the initial state and the function that describes the matrix-elements of the observable. By doing so, we identify the energy scales that determine the dispersion of gaps. We find that there are mainly two relevant energy scales: the energy fluctuations of the initial state and the bandwidth of the matrix of the observable A in the Hamiltonian eigenbasis. In the case of systems globally out of equilibrium, only those observables that are banded in the Hamiltonian basis can be observed out of equilibrium for a non-negligible time.

Local Hamiltonian. Let us define a *short-ranged* or *local* Hamiltonian of a spin lattice system, i.e., acting on a Hilbert space $\mathcal{H} = \bigotimes_{x\in V}\mathcal{H}_x$ with $\dim(\mathcal{H}_x) = d$, as

$$H = \sum_{u \in \mathcal{E}} h_u, \tag{41}$$

where the locality structure is given by a graph $(V,\ \mathcal{E})$ with a vertex set V and edge set $\mathcal{E}$. The number of terms of the Hamiltonian is denoted by $n = |\mathcal{E}|$. We consider systems for which it is possible to define a sequence of Hamiltonians H_n of different sizes. This becomes trivial in the case of translational invariant systems and regular lattices, but also includes systems with disorder, and defects. The reason for introducing such a sequence of Hamiltonians H_n is that it allows us to define the thermodynamic limit. For simplicity, the subindex n is not explicitly written from now on.

Energy density of the initial state. The *energy density* $f(E)$ of an initial state $|\psi(0)\rangle = \sum_i c_i |E_i\rangle$ is defined by

$$f(E) \coloneqq \sum_{i=1}^{d_E} |c_i|^2 \delta(E - E_i), \tag{42}$$

and, what is more relevant for us, its coarse-grained version reads $f_\epsilon(E) \coloneqq (N_\epsilon * f)(E)$.

For the case of local Hamiltonians, the coarse-grained energy density of states with a finite correlation length has been proven to approach a Gaussian as the size n of the system increases (lemma 8 in [37]). More specifically, for all E,

$$\int_{-\infty}^{E} \mathrm{d}E' f_\epsilon(E') \to \int_{-\infty}^{E} \mathrm{d}E' \frac{1}{\sqrt{2\pi}\,\sigma_E} \mathrm{e}^{-\frac{1}{2}\frac{(E'-\mu_E)^2}{2\sigma_E^2}}, \tag{43}$$

where the Gaussian has mean $\mu_E = \langle\psi(0)|H|\psi(0)\rangle$ and standard deviation $\sigma_E^2 = \langle\psi(0)|(H - \mu_E)^2|\psi(0)\rangle$, and where the difference in these expressions falls, at worst, essentially as $1/\sqrt{n}$. In what follows, we only make use of the energy density inside integrals, so in practice equation (43) allows us to replace the energy density $f_\epsilon(E')$ by the corresponding Gaussian, with vanishing error.

Initial states that are globally out of equilibrium, e.g. globally quenched, have energy densities with mean and standard deviation that scale in the system size as $\sigma_E \propto \sqrt{n}$ and $\mu_E \propto n$.

If the system is at criticality and the correlations decay in a power law, this Gaussian shape cannot be guaranteed anymore. In any case, the energy fluctuations can still scale as $\sigma_E \propto \sqrt{n}$ as long as the power m of the decay is sufficiently fast, i.e. $m > D + 1$ where D is the spatial dimension of the lattice (see appendix A for details).

Matrix-elements of an observable in the energy basis. Taking again a weakly stochastic approach, in the same spirit of the argument used in result 2, we constrain ourselves to observables with off-diagonal matrix-elements in the Hamiltonian basis that can be described by a 'continuous' function $S(E, \omega)$ plus some fluctuations δA_{ij}

$$|A_{ij}|^2 = S\left(\frac{E_i + E_j}{2}, \frac{E_i - E_j}{2}\right) + \delta A_{ij}, \tag{44}$$

where this choice of writing the arguments of the function $S(E, \omega)$ will be shown to be convenient in the following section. As in the previous section, the Lipshitz constant of $S(E, \omega)$ is assumed to be bounded by K.

Note that the so called eigenstate thermalisation hypothesis (ETH) [21] can be seen as a particular case of this assumption (44). One popular version of the ETH [38] is an ansatz on the matrix-elements of an observable A in the Hamiltonian eigenbasis,

$$A_{ij} = \langle E_i|A|E_j\rangle = \mathcal{A}(E)\delta_{ij} + \rho(E)^{-1} f_{\mathrm{ETH}}(E, \omega) R_{ij}, \tag{45}$$

where $E = (E_i + E_j)/2$ and $\omega = (E_i - E_j)/2$. The functions $\mathcal{A}$ and $f_{\mathrm{ETH}}(E, \omega)$ are smooth functions of their arguments, and R_{ij} are complex numbers randomly distributed, each with zero mean and unit variance. The essential idea is that both in (44) and (45) the off-diagonal matrix-elements of the observable can be described by a smooth function plus fluctuations that vanish when coarse-graining.

Note that in our case, in contrast to ETH, we do not assume anything about the diagonal elements of the observable in the energy basis. This is due to the fact that we are not concerned about what is the equilibrium state of the system (whether is thermal or not), but only about how long the relaxation process takes.

Now that we have introduced the energy density $f_\epsilon(E)$ and the function $S(E, \omega)$ that describes the observable, we are ready to express the density of relevant gaps in terms of these functions:

Result 3 (Density of relevant gaps). Given a Hamiltonian and an initial state with populations $|c_i|^2$, let A be an observable whose matrix-elements in the Hamiltonian eigenbasis, $|A_{ij}|^2$, can be described by means of the smooth function $S(E, \omega)$ plus some fluctuations as in equation (44). Then, up to errors $O(\epsilon K)$, the density of relevant gaps can be written as

$$|\tilde{g}_\epsilon(\omega)|^2 = 2\pi\rho_\epsilon(\omega) \int_{-\infty}^{\infty} \mathrm{d}E f_\epsilon(E - \omega/2) f_\epsilon(E + \omega/2) S(E, \omega). \tag{46}$$

Furthermore, if the Hamiltonian is local and the initial state has a finite correlation length, such that the coarse-grained energy density $f_\epsilon(E)$ is the Gaussian (43), then

$$|\tilde{g}_\epsilon(\omega)|^2 = 2\pi\, N_{\sqrt{2}\sigma_E}(\omega) S(\omega) \rho_\epsilon(\omega), \tag{47}$$

where the function $S(\omega)$ is defined as

$$S(\omega) := \int \mathrm{d}E\, N_{\frac{\sigma_E}{\sqrt{2}}}(E - \mu_E) S(E, \omega), \tag{48}$$

with μ_E and σ_E the mean and standard deviation of the energy density.

Proof. Let us introduce the density

$$R_\epsilon(\omega) := \int_{-\infty}^{\infty} \mathrm{d}E\, f_\epsilon(E - \omega/2) f_\epsilon(E + \omega/2) S(E, \omega) \tag{49}$$

and plug in it the energy density $f_\epsilon(E) := \sum_{i=1}^{d_E} |c_i|^2 N_\epsilon(E - E_i)$. A straightforward calculation leads to

$$R_\epsilon(\omega) = \sum_{i,j} N_{\sqrt{2}\epsilon}(\omega - (E_i - E_j)) |c_i|^2 |c_j|^2 \int \mathrm{d}E\, S(E, \omega) N_{\frac{\epsilon}{\sqrt{2}}}\left(E - \frac{E_i + E_j}{2}\right), \tag{50}$$

where we have used that $N_\epsilon(E + \omega - E_i) N_\epsilon(E - \omega - E_j) = N_{\frac{\epsilon}{\sqrt{2}}}\left(E - \frac{E_i + E_j}{2}\right) N_{\sqrt{2}\epsilon}(\omega - (E_i - E_j))$. By considering that $S(E, \omega)$ is the smooth description of the matrix elements $|A_{ij}|^2$, we get, up to errors of order ϵK,

$$R_\epsilon(\omega) = \sum_\alpha |v_\alpha|^2 N_{\sqrt{2}\epsilon}(\omega - G_\alpha) = |v(\omega)|^2 \rho_{\sqrt{2}\epsilon}(\omega). \tag{51}$$

Putting equations (49) and (51) together with result 3 implies (46). Now for initial states with a Gaussian energy density as in equation (43), simple algebra leads to the identity

$$f_\epsilon(E - \omega/2) f_\epsilon(E + \omega/2) = N_{\frac{\sigma_E}{\sqrt{2}}}(E - \mu_E) N_{\sqrt{2}\sigma_E}(\omega). \tag{52}$$

Plugging (52) in (46) completes the proof.

The function $S(\omega)$ describes the average magnitude of the off diagonal matrix-elements $|A_{ij}|^2$ at a distance $\omega = E_i - E_j$ from the diagonal.

Result 3 and in particular equation (47) show that the density of relevant gaps $|\tilde{g}_\epsilon(\omega)|^2$ for local Hamiltonians and initial states with decaying correlations decomposes in the product of two densities: $S(\omega)\rho_\epsilon(\omega)$ and $N_{\sqrt{2}\sigma_E}(\omega)$. The density $N_{\sqrt{2}\sigma_E}(\omega)$ is a Gaussian with standard deviation controlled by the energy fluctuations of the initial state σ_E, and $S(\omega)\rho_\epsilon(\omega)$ is the density of the off diagonal elements of the observable A.

The dispersion of relevant gaps σ_G, which we expect to estimate the equilibration time, is then controlled by the smallest of the standard deviations of these two densities. In the case that the system is globally out of equilibrium, e.g., a global quench, the variance of the energy density of the initial state is extensive with the system size, and $\sigma_E \propto \sqrt{n}$. This implies the following statement:

Result 4 (Out of equilibrium observables in global quenches). Given a local Hamiltonian, and an initial state with clustering of correlations let A be an observable that can be described by a smooth function $S(E, \omega)$. Then, the only one way to avoid that the dispersion of relevant gaps σ_G associated to an observable A diverges in the macroscopic limit is that the matrix-representation of the observable A in the energy basis is banded. More specifically, the density $S(\omega)\rho_\epsilon(\omega)$ has a standard deviation σ_A that is independent of the system size

$$\lim_{n\to\infty} \sigma_A = \lim_{n\to\infty} \frac{\int_\epsilon^\infty \mathrm{d}\omega (\omega - \mu_A)^2 S(\omega) \rho_\epsilon(\omega)}{\int_\epsilon^\infty \mathrm{d}\omega\, S(\omega) \rho_\epsilon(\omega)} < \infty, \tag{53}$$

where μ_A is the first moment.

Note that a divergent dispersion of relevant gaps σ_G is expected to imply an equilibration time that tends to zero in the thermodynamic limit. In other words, observables which are not banded in their energy representation are expected to be always equilibrated, since the amount of time that they can be out of equilibrium is negligible.

It is worth mentioning that generic observables have a flat $S(\omega)$ and fulfil $\sigma_A \propto \sqrt{n}$ due to the domination of the density of gaps $\rho_\epsilon(\omega)$. Thus, they will have microscopically short equilibration times [20]. Observables with the property (53) turn out to be both rare and physically relevant.

In [39] it is shown for several concrete examples that indeed the matrix elements $S(\omega)$ decrease exponentially or super-exponentially with ω from a certain threshold independent of the system size (see also section 4.3.1.2 of

[38]). In the following result, we show that this is a property of any local observable, that is, operators that only act non-trivially on a finite region of the system.

Result 5 (Local operators are banded in the energy basis). Let us consider a local Hamiltonian

$$H = \sum_{(x,x')\in\mathcal{E}} h_{(x,x')} \tag{54}$$

acting on a Hilbert space $\mathcal{H} = \bigotimes_{x\in V}\mathcal{H}_x$ with $\dim(\mathcal{H}_x) = d$, with a locality structure given by a graph with a vertex set V and edge set $\mathcal{E}$. Then, the matrix elements in the energy eigenbasis of a local operator $\langle E_i|A_x|E_j\rangle$ acting on a site x fulfil the condition

$$|\langle E_i|A_x|E_j\rangle| \leqslant \|A_x\|\ \mathrm{e}^{\log\left(\frac{\mathrm{e}(E_i-E_j)}{J(1+\alpha)}\right)-c(E_i-E_j)/J}, \tag{55}$$

where $c = \log(1+\alpha^{-1})$ is the decay rate, $J = \max_{(x,x')\in\mathcal{E}}\|h\|_\infty$ is the strength of the local interactions and α is the *lattice animal constant* [40] of the graph $(V, \mathcal{E})$.

The proof of result 5 is presented in appendix C. In particular, note that the lattice animal constant mentioned above is a parameter that captures the connectivity of the underlying graph of the Hamiltonian. For D-dimensional cubic lattices, it can be bounded as $\alpha \leqslant 2D\mathrm{e}$ (lemma 2 in [40]).

In the period of finishing this manuscript we have been alerted to the existence of a result very similar to our result 5, due to Arad *et al* (theorem 2.1 in [23]). Both proofs are similar in spirit and give similar decay rates. For a D-dimensional cubic lattice with interactions in the edges, they obtain a decay rate of $1/(8D)$ while we get $\log(1+(2\mathrm{e}D)^{-1}) \simeq 1/(2\mathrm{e}D)$. The main difference is that while we bound the number of terms by counting lattice animals, they use a combinatoric argument.

Of course, this behaviour of being banded in the energy basis extends to global operators that can be decomposed into a sum of local terms, as well as for operators that are not local in real space but in momentum space when the Hamiltonian is also local in the momentum representation. Note that, indeed, most observables considered in the literature are of this type.

Let us now consider the relevant scenario of a *local quench* [33], in which the system is brought out of equilibrium in only a local region of the system. In such a case, the width of the energy density of the initial state is independent of the system size and related to the operator norm of the perturbation applied on the system. Unlike the global quench scenario, now even the observables that are not banded (and are initially out of equilibrium) will take a finite non-negligible time to relax. The equilibration timescale is then governed by whichever energy scale is smallest: the energy fluctuations of the state, or the dispersion σ_A of the observable. Note that our results also allow for having equilibration times that increase with the system size, as long as either σ_A and σ_E shrink with it.

6. Numerical example: the *XXZ* model

We illustrate our results using the *XXZ* model in a transverse field and with next-nearest-neighbour coupling. We choose to use this particular model since it is not integrable, and hence does not have an exponential number of degenerate gaps. The hamiltonian is $H = J\sum_i S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z + J_2 S_i^z S_{i+2}^z + h_z S_i^z$. As our equilibrating observable we choose the magnetisation density in the x direction, $M^x = \sum_i S_i^x/N$ and, as our initial state, the fully magnetised state in the x direction. In figure 4, we show the evolution of the time signal $|g_M(t)|^2$ of this observable, in the sense of equation (5). The calculations were done using full exact diagonalization of H with $\Delta = 0.5, J_2 = 1.0$ and $h_z = 0.2$, for various system sizes.

We expect $g_M(t)$ to go to zero when the system equilibrates. Indeed, this is what happens initially, for all system sizes, and we can notice that the equilibration time (the time when $|g_M(t)|^2$ becomes negligible, just before Jt=20) does not depend much on n. Furthermore, we can compare this value with our heuristic estimate for the equilibration time, $T_{\mathrm{eq}} \sim \pi/\sigma_G$, where we use equation (15) to calculate σ_G from the numerically obtained eigenvalues. The results are shown in table 1. We can see that the estimated equilibration times also depend only weakly on n, and are in good agreement with the timescale indicated by figure 4.

Of course, due to the small size of the simulated systems, the time signals $g_M(t)$ also exhibit strong fluctuations. However, one can already see that, as n increases, the size of these fluctuations tends to decrease, and their onset happens later. Our numerical results therefore seem to corroborate our expectation that the observable M^x does indeed equilibrate in the limit of large n, and that this equilibration does happen at a timescale roughly given by $T_{\mathrm{eq}} \sim \pi/\sigma_G$.

To better illustrate the dephasing mechanism behind the equilibration process, in figure 5 we plot the amplitudes $v_\alpha \mathrm{e}^{\mathrm{i}G_\alpha t}$ for this same situation, in the case $n = 10$. Starting from an initial condition where all the amplitudes are in phase (in this case, all real and negative), one can see them rotating at different speeds and

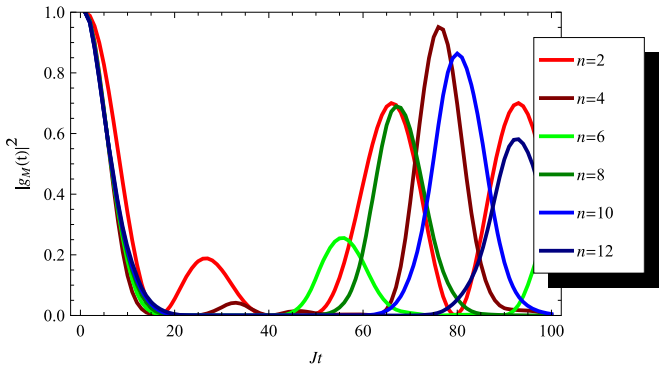

Figure 4. Fluctuation in the total magnetisation in the x direction, $|g_M(t)|^2$, for the XXZ model with next-nearest-neighbour coupling and external magnetic field ($\Delta = 0.5, J_2 = 1.0$ and $h_z = 0.2$). The different curves are for chains with different numbers n of spins.

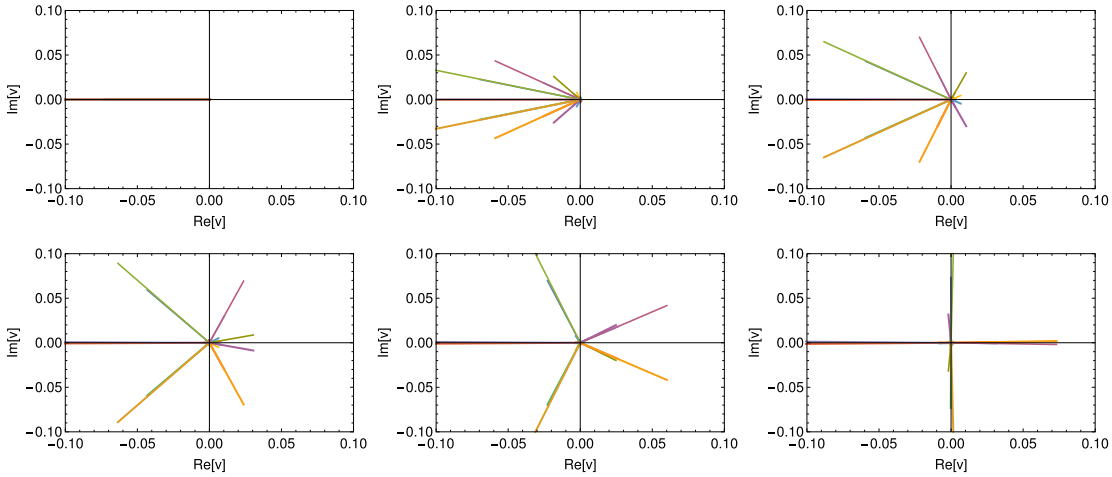

Figure 5. Evolution of each complex number $v_\alpha e^{iG_\alpha t}$ for our simulation of the XXZ model with $n = 10$. From upper left to the bottom right, (adimensional) time Jt goes from 0 to 20 in equal intervals.

Table 1. Estimated equilibration times T_{eq} for the XXZ model with next-nearest-neighbour coupling and an external magnetic field ($\Delta = 0.5, J_2 = 1.0$ and $h_z = 0.2$). The gap dispersion σ_G was obtained by explicitly calculating equation (15) from the numerically obtained energy spectrum.

n	$T_{eq} \sim \pi/\sigma_G$
2	21
4	19
6	20
8	22
10	23
12	24

becoming more spread out in the complex plane as time goes by, resulting in the decay seen in figure 4. Note that the approximate four-way symmetry exhibited at $Jt = 20$ (implying $g(t) \simeq 0$ at this time) is already a symptom of a future recurrence: clearly, after four times this interval, all of the amplitudes will have rotated approximately back to their initial position. Indeed, one can see in figure 4 that a recurrence occurs at around $Jt = 80$.

It is also instructive to compare the exact time evolution of the magnetisation with coarse-grained versions derived according to the procedures described in section 4. In figure 6 we plot again the exact time signal $g_M(t)$ for the chain with $n = 12$ spins (the dark blue line in figure 4), represented here by the black dotted curve. We

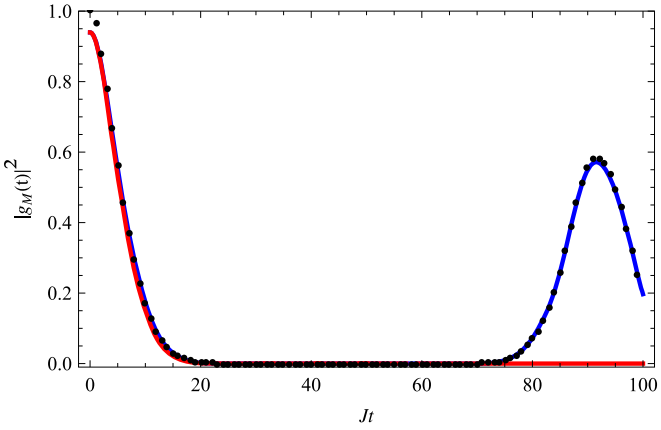

Figure 6. Comparison of exact and coarse-grained time signals for the magnetisation. The exact time signal $g_M(t)$ (black dots) is the same curve plotted in figure 4, for $n = 12$ spins. The full curves are coarse-grained time signals $g_M^\epsilon(t)$, obtained by Fourier transforming the coarse-grained frequency signals calculated according to equation (26), with $\epsilon = 0.4$ (red) and $\epsilon = 0.02$ (blue). By choosing a value of ϵ that is sufficiently large (but not too large), the coarse-grained signal reproduces the exact one during the equilibration phase, but suppresses later recurrences.

also plot two coarse-grained time signals $g_M^\epsilon(t)$, with $\epsilon = 0.4$ (red line) and $\epsilon = 0.02$ (blue line). These curves were obtained by Fourier transforming coarse-grained frequency spectra (such as the one in figure 3) that were numerically calculated according to equation (26) (i.e., we did not simply dampen the exact time signal using equation (30)). It can be seen that all three signals are essentially indistinguishable up to the equilibration time (the slight deviation close to $t = 0$ is an artifact of our having discarded terms with $\nu_\alpha < 10^{-4}$ when calculating the sum in equation (26)). We can see that, for small ϵ ($= 0.02$), the signals remain indistinguishable up to and including the recurrence time. However, by choosing a value of ϵ that is sufficiently large (0.4), the coarse-grained signal faithfully reproduces the exact one during the equilibration phase, but suppresses later recurrences. Note that ϵ must still be chosen sufficiently small ($\epsilon \ll \sigma_G$) in order to avoid suppressing the signal even before equilibration has occurred. As discussed in previous sections, under these circumstances we may use the width of the (square-integrable) coarse-grained signal as a measure of the equilibration time of the original signal.

Finally, let us remark that we also obtained similar results for other values of Δ and h_z, and also for the XY model. However, in the latter case, the fluctuations do not decrease exponentially with n, but only polynomially, since there are an exponential number of degenerate gaps.

7. Discussion

7.1. Reinterpretation of previous results

It is useful now to reinterpret some previous results on equilibration times from our dephasing point of view. For example, Short and Farrelly [15] obtain a rigorous upper bound for the equilibration time of any observable by studying the time-averaged fluctuations $\langle |g|^2 \rangle_T$ in equation (8) above. They are able to determine a value $T_0 \gtrsim d_E/\Delta E$, where ΔE is the range of energies in the system, and d_E the number of different energy levels, such that averages taken over intervals longer than T_0 are negligible. This implies that the equilibration time must be upper bounded by T_0. Unfortunately, for a typical many-body system with n degrees of freedom, ΔE scales only polynomially with n, while d_E scales exponentially with n—and thus so does T_0. In other words, although this upper bound is mathematically sound, it vastly overestimates the actual equilibration time of most observables. This suggests that its derivation must be incomplete, in the sense of missing or disregarding an essential physical ingredient [20].

We now argue that this ingredient is, in a word, dephasing. Roughly speaking, in the course of their derivation, the authors bound $\langle |g|^2 \rangle_T$ by separately bounding the absolute value of every term in its Fourier expansion, disregarding the interference between different terms due to dephasing. Each of these terms, which rotate according to $\exp[\mathrm{i}(G_\alpha - G_\beta)T]$ does gets individually dephased, due to the time averaging, but only on a time scale $t_{\alpha\beta} \sim 1/(G_\alpha - G_\beta)$. The bound in [15] hence corresponds to the amount of time needed for the

slowest term in the Fourier sum to average out over time. For many-body systems, the gaps G_α, and also the differences between different gaps, can be exponentially small in n, hence the exponentially long upper bound.

Although this bound is therefore much too large to be a reasonable estimate the equilibration time of most observables, it must be stressed that one can always construct a specific observable A which saturates it. This is not entirely unexpected, as the bound itself is observable-independent. In fact, in [20], Goldstein *et al* construct such an observable, by considering a direct sum of banded matrices in the energy basis, each of which has a bandwidth that is exponentially small in the system size. From the dephasing mindset, it is straightforward to understand what is going on in this example. Since each component of A is banded with an (exponentially) narrow bandwidth, with no coherences between different bands, the dispersion of relevant gaps, σ_G, is exponentially small. In this case, our estimate π/σ_G for the equilibration time becomes exponentially large, since the small difference in angular speeds means that it must take a long time until all points are dephased and isotropically distributed in the complex plane. Analogously, in [20] it is also shown that one can always construct an observable which equilibrates extremely fast, by defining an A which is far from banded, having coherences between vastly different energies. Again, the dephasing picture intuitively explains the reason for such quick equilibration.

We can also understand some of the results obtained in another approach to the problem of relaxation of many-body systems, namely the study of the survival probability given by the quantum fidelity $\mathcal{F}(t) := |\langle\psi(0)|\psi(t)\rangle|^2$, often in situations where the initial state $|\psi(0)\rangle$ is generated after a sudden displacement ('quantum quench') that brings the system out of equilibrium (see [19] and references therein for a review). Note that the quantum fidelity is, up to an additive constant, equivalent to the time evolution of the observable $A = |\psi(0)\rangle\langle\psi(0)|$. In this case, then, both the energy fluctuations of the initial state and the bandwidth of the observable written in the energy basis are σ_E. Hence, our considerations within the dephasing picture predict, for local interacting lattice systems, an equilibration time determined by σ_E, in agreement with the results of [19].

Finally, our outlook and conclusions are also compatible, and in some senses complementary, to recent remarkable results by Reimann *et al* [17, 18]. In these works, the time signal $g(t)$ in equations (5) and (6) above is rewritten as $g(t) = c\,F(t) + \xi(t)$, where (using our notations)

$$F(t) = \frac{1}{d_T(d_T - 1)} \sum_\alpha \mathrm{e}^{\mathrm{i}G_\alpha t}, \tag{56}$$

and c is a constant (=$g(0)$). Note that $F(t)$ depends only on the gap spectrum, and on the dimension of the system, but not on the initial condition, observable or eigenbasis of the Hamiltonian. It is then proven that, if one averages the time signal over certain ensembles of Hamiltonians with fixed spectra (i.e. varying only their eigenvectors) then both the average value and the standard deviation of $\xi(t)$ become extremely small in the thermodynamic limit. In other words, for *any* fixed initial condition and observable of a quantum system, and also any given fixed energy spectrum, the 'typical' time signal will always be of the form $g(t) = cF(t)$, up to negligible error. Comparing with equation (6) above, we can see that the effect of the ensemble average is to uniformize the different amplitudes v_α, both in amplitude and in phase. As a result, the 'typical' equilibration dynamics described by equation (56) becomes a pure dephasing process, i.e, a sum of uniform-length vectors in the complex plane, all initially pointing along the positive real axis, which then rotate at different speeds.

As we have argued above, dephasing/equilibration should then occur on a timescale of order π/σ_G. Indeed, this can be seen in the examples worked out in these references. For instance, one situation considered in [17] is a system described by a continuous 'microcanonical' density of energy eigenstates, of the form $\rho(E - y) = c\mathrm{e}^{-y/k_B T}$, where $y > 0$ and c is a normalisation constant. It is also assumed that the system's initial state has support restricted to a narrow energy band $[E - \Delta E, E]$. In this case $F(t)$ can be calculated exactly; in particular it is shown that, for $\Delta E \gg k_B T$, it has the form of a Lorentzian function,

$$F(t) = \frac{1}{1 + (\gamma t)^2}, \tag{57}$$

where $\gamma = k_B\,T/\hbar$.

Let us now analyse this result from the point of view of our approach. First of all, as we have noted in section 3, restricting the initial state to a finite energy band, and therefore a finite gap spectrum, ensures that, for all observables, dephasing/equilibration must occur in a finite timescale.

We can estimate this timescale from the (continuous) gap density for this model:

$$\rho_G(\omega) = \int \rho(E')\rho(E'+\omega)\mathrm{d}E' = \frac{\gamma c^2 \mathrm{e}^{-\Delta E/\gamma}}{2}\sinh\left[\frac{\Delta E - |\omega|}{\gamma}\right] \sim \frac{\gamma c^2}{2}\mathrm{e}^{-|\omega|/\gamma}, \tag{58}$$

where the last equality is valid for $\Delta E \gg k_B T$. Given that here the amplitudes v_α of each gap in the time signal are uniform, this expression is also proportional to the gap spectrum (i.e., the Fourier transform of $F(t)$), and we can calculate directly the gap dispersion

$$\Delta\omega = \frac{\int \omega^2 \rho_G^2(\omega)\mathrm{d}\omega}{\int \rho_G^2(\omega)\mathrm{d}\omega} = \frac{\gamma}{\sqrt{2}}.$$

Hence, we estimate an equilibration time $T_{\mathrm{eq}} \sim \frac{\pi\sqrt{2}}{\gamma} = \frac{\pi\sqrt{2}\hbar}{k_B T}$. It can be seen from equation (57), that, after $t = T_{\mathrm{eq}}$, the signal has decreased by about ~95% from its initial value, showing that this is indeed a fair estimate of the equilibration timescale. Note that for a time signal governed by equation (57), the standard deviation Δt defined in equation (17) equals $1/\gamma$, which is also the 'half-width at half maximum' of the Lorentzian[6]. Hence in this case the uncertainty relation, equation (18) is not far from being saturated, which is the condition we are relying on for our estimate.

As a final remark, it is also interesting to note that in this case the roles of the time signal and spectrum are exactly reversed from the well-known example of spontaneous decay of an excited atomic or nuclear state: there it is the spectrum that is a Lorentzian, and the time signal that is an exponential decay.

7.2. Equilibration time-scales and level statistics

A consequence of the coarse-graining machinery introduced above is that the fine-grained details of the spectrum do not affect the dynamics of the system up to very long time scales. More formally, the following result states that if two coarse-grained frequency signals are close in the trace norm, then their corresponding time signals must also be point-wise close for any $t < O(\epsilon^{-1})$:

Result 6 (Spectrum and dynamics). Let $g^{(1)}(t)$ and $g^{(2)}(t)$ be two time signals whose ϵ-coarse-grained Fourier transforms satisfy

$$\|\tilde{g}_\epsilon^{(1)} - \tilde{g}_\epsilon^{(2)}\|_1 \leqslant \delta_1 \tag{59}$$

for some $\delta_1 > 0$, and let some $0 < \delta_2 < 2$ set a distinguishably threshold. Then $g^{(1)}(t)$ and $g^{(2)}(t)$ are indistinguishable up to times $t \leqslant \sqrt{\delta_2}/\epsilon$, in the sense that

$$|g^{(1)}(t) - g^{(2)}(t)| < \delta_1 + \delta_2. \tag{60}$$

Proof. By using twice the triangular inequality, we have

$$|g^{(1)}(t) - g^{(2)}(t)| \leqslant |g^{(1)}(t) - g_\epsilon^{(1)}(t)| + |g_\epsilon^{(1)}(t) - g_\epsilon^{(2)}(t)| + |g_\epsilon^{(2)}(t) - g^{(2)}(t)|. \tag{61}$$

The first and third terms are analogous and are bounded by equation (30) and the fact that $|g(t)| \leqslant 1$. Concerning the second term, we note that, since $\tilde{g}_\epsilon(\omega) \in L^2$, then the uniform continuity statement of functional analysis allows us to write

$$\|g_\epsilon^{(1)} - g_\epsilon^{(2)}\|_\infty = \sup_t |g_\epsilon^{(1)}(t) - g_\epsilon^{(2)}(t)| \leqslant \|\tilde{g}_\epsilon^{(1)} - \tilde{g}_\epsilon^{(2)}\|_1 \leqslant \delta_1, \tag{62}$$

and putting everything together, we get

$$|g^{(1)}(t) - g^{(2)}(t)| \leqslant 2|1 - \mathrm{e}^{-\frac{1}{2}\epsilon^2 t^2}| + \delta_1. \tag{63}$$

Finally, the bound (60) is respected as long as $1/2\epsilon^2 t^2 \leqslant -\log(1 - \delta_2/2)$. The inequality $x \leqslant -\ln(1-x)$ for $0 < x < 2$ implies that $1/2\epsilon^2 t^2 \leqslant \delta_2/2 \leqslant -\log(1 - \delta_2/2)$. Thus, equation (60) is guaranteed to hold for times $t \leqslant \sqrt{\delta_2}/\epsilon$.

Result 6 shows that two frequency signals that become very similar once they are ϵ-coarse-grained have indistinguishable dynamics up to times ϵ^{-1}. This is particularly relevant in many-body systems where the separation between consecutive energy levels shrinks exponentially in the system size[7]. One can then consider the possibility of two many-body Hamiltonians with qualitatively different *level statistics*, for example one where the distribution of gaps $E_{i+1} - E_i$ between consecutive energy levels follows a Poisson distribution, and

[6] Note that here Δt refers to the standard deviation of the *square* of Lorentzian, not that of the Lorentzian itself, which is divergent.

[7] This is a consequence of the fact that in such systems the energy scales extensively (or even polynomially for some long range interactions) with the number of particles, while the dimension of the Hilbert space does so exponentially.

another with Wigner–Dyson statistics, giving rise to time signals that are nevertheless indistinguishable in practice for time up to and beyond the equilibration time.

In appendix D we present an example of this kind. We have estimated the one-norm distance between two coarse-grained frequency signals evolving according to two Hamiltonians with identical eigenbases but different level statistics (one Poissonian and the other Wigner–Dyson). For this case we find that

$$\|g_{\epsilon}^{(1)} - g_{\epsilon}^{(2)}\|_1 \leqslant \frac{C}{\epsilon}\sqrt{\frac{n^3}{d_{\text{eff}}}}, \tag{64}$$

where C is a numerical constant, n is the system size, ϵ the coarse-graining parameter and d_{eff} the effective dimension.

Note now that, assuming that d_{eff} increases exponentially in n ($d_{\text{eff}} \sim \exp(cn)$, for some constant c), then choosing $\epsilon \sim \exp(-cn/4)$ in equation (64) and using result 6 implies that the dynamics would not be affected up to times $t \sim \exp(cn/4)$. (Here we are also assuming, as in the discussion following equation (38), $\epsilon \gg E_{i+1} - E_i$, which also decrease exponentially. This requires choosing d_{eff} to increase sufficiently slowly (c sufficiently small)). If these conditions are met, we obtain time signals originating from two Hamiltonians with different level-statistics but that for all practical purposes display the same dynamical behaviour.

Recall now that it is a well-known conjecture that a Poissonian nearest-neighbour gap distribution is a manifestation of integrability, and Wigner–Dyson statistics are a signature of quantum chaos. Our example seems therefore to show that it is possible for both kinds of Hamiltonian to lead to identical time signals, with identical equilibration times, at least some specific cases.

It is less clear what will happen in a more realistic example in which the two Hamiltonians H_1 and H_2 with different level statistics also have different eigenbases (as is in practice always the case). In these situations the two Hamiltonians are related by a perturbation V which does not commute with the integrable Hamiltonian H_1 and where both H_1 and V have a locality structure of the type in (41). In such a scenario, the coarse-grained energy density of the initial state is not affected by the perturbation since it keeps the Hamiltonian local. Thus, any change in the one-norm distance between frequency signals, and thereby in the dynamical behaviour of the system, must come from a drastic change in the matrix-elements of the observable. A question that arises beyond the scope of this paper is then how integrability, non-integrability, and chaos can be identified in the behaviour of the matrix-elements of the observable in the energy basis.

7.3. The dephasing mindset for quadratic Hamiltonians

A relevant point to discuss is to which extent the results presented in this paper are valid for integrable models. In this respect, let us focus on quadratic (bosonic or fermionic) Hamiltonians that can be brought to a diagonal form $H = \sum_k \varepsilon(k) a_k^\dagger a_k$, where $a_k^{(\dagger)}$ are the annihilation (creation) operators that fulfil fermionic or bosonic commutation relations and $\varepsilon(k)$ is the dispersion relation. For such systems and quadratic observables, the time signal can also be written in the form of equation (6), where the sum now does not run over the gaps of the Hamiltonian, but over the gaps of the dispersion relation $\varepsilon(k)$. In [41] this is done in detail for the Caldeira–Leggett model (a quadratic system of harmonic oscillators). In sum, we see that the above formalism can also be applied to quadratic Hamiltonians, where, roughly speaking, the Hilbert space has been substituted by the space of modes.

8. Conclusions

In this work we have argued that equilibration in closed quantum systems should be understood as a process of dephasing of complex numbers in the complex plane. From this mechanism, we have heuristically estimated the equilibration time-scale as roughly the inverse of the dispersion of the relevant gaps. We have seen that, under physically relevant circumstances, the equilibration time-scale estimated in this way depends at most weakly on the system size, in agreement with realistic situations. Although our argument does not result in a rigorous bound, we claim that it captures the correct way in which the time-scale depends on the physical properties of the system. In particular, we have seen that the coherences of the observables of interest in the energy basis, $\langle E_i|A|E_j\rangle$, play a fundamental role: in order to attain a finite equilibration time for generic initial states, these coherences must become small as $E_i - E_j$ increases.

We have also observed that the size of the system only plays a role in the typical size of the fluctuations, but not in the time of equilibration, and thus small systems fail to equilibrate not because their equilibration time is large, but because their fluctuations are big. We illustrate these results with numerical simulations of spins chains.

Finally, we have applied the dephasing mindset to give an intuitive interpretation to earlier works on equilibration times. Our results satisfactorily reproduce many particular cases of determining equilibration time

scales found in the literature. Of course, further work is still required to put our claims on a more rigorous foundation. For example, we conjecture that in a sufficiently wide range of locally interacting n-body systems, the coarse-grained frequency signal $\tilde{g}_\epsilon(\omega)$ may itself approach a Gaussian in the limit of large n. In this case, the Heisenberg-like uncertainty principle we have been using to heuristically estimate the equilibration timescale would become close to saturated, and would therefore indeed be a bona-fide measure for it.

Acknowledgments

We would like to thank J Eisert and his co-authors in [22] for the coordination in the publication of this manuscript and their work, and A Molnar and H Wilming for letting us know about the existence of [23] which contains a very similar theorem to our result 5. TRO and DJ are supported by the Brazilian National Institute for Science and Technology of Quantum Information (INCT-IQ). TRO also thanks the National Counsel of Technological and Scientific Development (CNPq). CC, ML and AR thank financial support from Spanish MINECO (QIBEQI FIS2016-80773-P, FISICATEAMO FIS2016-79508-P and Severo Ochoa Grant No. SEV-2015-0522), Fundació Privada Cellex, Generalitat de Catalunya (Grant No. SGR 874, 875, and CERCA Programme) and the European Commission [EQuaM (FP7-ICT-2013-C No.323714), OSYRIS (ERC-2013-AdG No. 339106), QUIC (H2020-FETPROACT-2014 No. 641122), SIQS (FP7-ICT-2011-9 No. 600645)]. AR is also supported by the Beatriu de Pinós fellowship (BP-DGR 2013) and the CELLEX-ICFO-MPQ research fellowship.

Appendix A. Energy density and effective dimension of the initial state for local Hamiltonians

A.1. Energy uncertainty of the initial state

A standard way to bring quantum systems out of equilibrium is to initialise the system as the ground state of a local Hamiltonian, and change the Hamiltonian of the system sufficiently fast such that the state is kept unchanged. This procedure is called a quantum *quench* [33].

It is well known that the ground state of local Hamiltonians exhibit a clustering of correlations [42]. In fact, correlations between observables supported at different lattice points decay algebraically with the distance for critical systems (gapless Hamiltonians) and exponentially for systems off criticality (gapped Hamiltonians).

Consider a local Hamiltonian of the type (41) and a quantum state $|\psi\rangle \in \mathcal{H}$ with either exponentially or algebraically decaying correlations such that

$$|\langle h_u h_v\rangle - \langle h_u\rangle\langle h_v\rangle| \leqslant \frac{c}{d(u,v)^{D+1}}, \tag{A.1}$$

where $c > 0$, $d(u, v)$ is the graph distance between the edges u and v, and D is the spatial dimension in which the graph can be embedded. Then, the variance of the energy distribution is upper bounded by $\mathrm{var}_\psi(H_n) \leqslant c'n$ with $c' > 0$.

By using (A.1), it is easy to see that

$$\sum_{v\in E}(\langle h_u h_v\rangle - \langle h_u\rangle\langle h_v\rangle) \leqslant c' \;\forall\; u. \tag{A.2}$$

Hence, writing the variance in terms of the local terms of the Hamiltonian, and using the previous bound, one gets

$$\mathrm{var}_\psi(H_n) = \langle H^2\rangle_\psi - \langle H\rangle_\psi^2 = \sum_{u,v\in E}(\langle h_u h_v\rangle_\psi - \langle h_u\rangle_\psi\langle h_v\rangle_\psi) \leqslant cn. \tag{A.3}$$

A.2. Effective dimension

The effective dimension tells us how many eigenstates of the Hamiltonian contribute in the superposition of the initial state. In the previous section we have argued that the energy uncertainty of the initial state scales with $\sqrt{n}$ in the case where the system is brought out of equilibrium with a global quench, and is independent of the system size in the case where the system suffers a local quench.

Local Hamiltonians have an energy range that scales linearly in the system size while the dimension of the Hilbert space does so exponentially. This implies that the density of states scales exponentially, and so does the effective dimension.

Appendix B. Proof of result 2: coarse-grained frequency signal

In this section we provide the proof of result 2. The coarse grained Fourier transform of the time signal naturally decomposes

$$\tilde{g}_\epsilon(\omega) = \sum_\alpha \nu(G_\alpha) h_\epsilon(\omega - G_\alpha) + \sum_\alpha \delta\nu_\alpha h_\epsilon(\omega - G_\alpha). \tag{B.1}$$

The second term in the right-hand side is a random variable with variance given by

$$\delta\nu_\epsilon(\omega)^2 := \mathrm{var}\left(\sum_\alpha \delta\nu_\alpha h_\epsilon(\omega - G_\alpha)\right) = \sum_\alpha h_\epsilon^2(\omega - G_\alpha)\mathrm{var}(\delta\nu_\alpha) = \sum_\alpha h_\epsilon^2(\omega - G_\alpha)\gamma^2(G_\alpha).$$

By means of the density of gaps, the variance of the fluctuation term in equation (B.1) becomes

$$\delta\nu_\epsilon(\omega)^2 = \frac{\sqrt{\pi}}{\epsilon}\int \mathrm{d}\omega' N_{\epsilon/\sqrt{2}}(\omega - \omega')\rho(\omega')\gamma(\omega'), \tag{B.2}$$

where we have taken into account that $h_\epsilon(\omega)^2 = \sqrt{\pi}\,\epsilon^{-1}N_{\epsilon/\sqrt{2}}(\omega)$. Now we use the fact that $\gamma(\omega)$ is almost constant in an interval ϵ, i.e. the Lipshitz constant K of γ is such that $K\epsilon \ll 1$. The idea is to approximate $\delta\nu_\epsilon(\omega)^2$ by

$$\delta\nu_\epsilon(\omega)^2 \simeq \frac{\sqrt{\pi}}{\epsilon}\gamma^2(\omega)\int \mathrm{d}\omega' N_{\epsilon/\sqrt{2}}(\omega - \omega')\rho(\omega') = \frac{\sqrt{\pi}}{\epsilon}\gamma^2(\omega)\rho_{\epsilon/\sqrt{2}}(\omega). \tag{B.3}$$

More precisely, such approximation has 2 steps. The first step is to restrict the domain of integration around ω, where the Gaussian $N_{\epsilon/\sqrt{2}}(\omega - \omega')$ is centred. The error of such step can be upper bounded by

$$\left|\delta\nu_\epsilon(\omega)^2 - \frac{\sqrt{\pi}}{\epsilon}\int_{\omega-\Delta}^{\omega+\Delta}\mathrm{d}\omega' N_{\epsilon/\sqrt{2}}(\omega - \omega')\rho(\omega')\gamma(\omega')\right| \leqslant \frac{\sqrt{\pi}}{\epsilon}\rho_{\epsilon/\sqrt{2}}(\omega)\alpha\mathrm{e}^{-2\Delta^2/\epsilon^2}, \tag{B.4}$$

where we have used the Chernoff bound of the Gaussian distribution $\int_x^\infty \mathrm{d}y N_\sigma(y) \leqslant 2\mathrm{e}^{-x^2/\sigma^2}$ and $\alpha > 0$ is a constant that depends on the density of gaps.

The second step is considering the worst case in the variation of $\gamma(\omega)$ within the interval of integration $[\omega - \Delta, \omega + \Delta]$. To do so, let us consider two real positive functions $a(x)$ and $b(x)$, where $b(x)$ has Lipschitz constant K. Then, by using trivial calculus and the definition of Lipschitz constant, we get

$$\int_{\omega-\Delta}^{\omega+\Delta}\mathrm{d}x a(x)b(x) \leqslant \left(\max_{x\in[\omega-\Delta,\omega+\Delta]} b(x)\right)\int_{\omega-\Delta}^{\omega+\Delta}\mathrm{d}x a(x) \leqslant (b(\omega) + K\Delta)\int_{\omega-\Delta}^{\omega+\Delta}\mathrm{d}x a(x).$$

A straight forward application of the former statement leads to

$$\int_{\omega-\Delta}^{\omega+\Delta}\mathrm{d}\omega'\gamma^2(\omega')N_{\epsilon/\sqrt{2}}(\omega - \omega')\rho(\omega') - \gamma^2(\omega)\int_{\omega-\Delta}^{\omega+\Delta}\mathrm{d}\omega' N_{\epsilon/\sqrt{2}}(\omega - \omega')\rho(\omega') \leqslant \rho_{\epsilon/\sqrt{2}}(\omega)K\Delta.$$

By putting everything together, we get

$$\left|\delta\nu_\epsilon(\omega)^2 - \frac{\sqrt{\pi}}{\epsilon}\gamma^2(\omega)\rho_{\epsilon/\sqrt{2}}(\omega)\right| \leqslant \frac{\sqrt{\pi}}{\epsilon}\rho_{\epsilon/\sqrt{2}}(\omega)(2\alpha\mathrm{e}^{-2\Delta^2/\epsilon^2} + K\Delta). \tag{B.5}$$

Finally, an optimisation over Δ gives

$$\left|\delta\nu_\epsilon(\omega)^2 - \frac{\sqrt{\pi}}{\epsilon}\gamma^2(\omega)\rho_{\epsilon/\sqrt{2}}(\omega)\right| \leqslant \alpha'\epsilon^{-1}\rho_{\epsilon/\sqrt{2}}(\omega)K\epsilon, \tag{B.6}$$

where $\alpha' > 0$ contains all the numerical factors. Equation (B.6) implies a relative error in the of the order $K\epsilon$.

The standard deviation $\delta\nu_\epsilon(\omega)$ sets the order of magnitude for the fluctuations of the random variable. We will see that it is negligible in comparison with the first term in equation (B.1). Such a term can also be written in terms of the density of gaps $\rho(\omega)$

$$R_\epsilon(\omega) := \sum_\alpha \nu(G_\alpha)h_\epsilon(\omega - G_\alpha) = \int \mathrm{d}\omega'\rho(\omega')\nu(\omega')h_\epsilon(\omega - \omega'). \tag{B.7}$$

The function $R_\epsilon(\omega)$ can be shown to be close to $\sqrt{2\pi}\ \nu(\omega)\rho_\epsilon(\omega)$. More precisely, by following an analogous argument of the one above, it can be proven that

$$|R_\epsilon(\omega) - \sqrt{2\pi}\ \nu(\omega)\rho_\epsilon(\omega)| \leqslant \beta\rho_\epsilon(\omega)K\epsilon, \tag{B.8}$$

where $\beta > 0$ is an order one factor.

We are now ready to bound the difference $\tilde{g}_\epsilon(\omega) - \sqrt{2\pi}\nu(\omega)\rho_\epsilon(\omega)$. By using equation (B.1), definition (B.7) and the triangular inequality, we get

$$|\tilde{g}_\epsilon(\omega) - \sqrt{2\pi}\nu(\omega)\rho_\epsilon(\omega)| \leqslant |R_\epsilon(\omega) - \sqrt{2\pi}\nu(\omega)\rho_\epsilon(\omega)| + \left|\sum_\alpha \delta\nu_\alpha h_\epsilon(\omega - G_\alpha)\right|. \tag{B.9}$$

Finally, from equations (B.3) and (B.8), the error is bounded by

$$|\tilde{g}_\epsilon(\omega) - \sqrt{2\pi}\nu(\omega)\rho_\epsilon(\omega)| \leqslant \rho_\epsilon(\omega)\left(\beta K\epsilon + \frac{\pi^{1/4}\gamma(\omega)}{\sqrt{\epsilon\rho_\epsilon(\omega)}}\right). \tag{B.10}$$

Here, we have used that $\rho_{\epsilon/\sqrt{2}}(\omega) \leqslant \rho_{\epsilon/\sqrt{2}}(\omega)(1 + K\epsilon)$ where K also upper-bounds the Lipshitz constant of $\rho_\epsilon(\omega)$.

Appendix C. Proof of result 5: local operators are banded in the energy basis

In this section we prove result 5. We first introduce the imaginary time evolution of the observable

$$A_x(\beta) \coloneqq \mathrm{e}^{\beta H} A_x \mathrm{e}^{-\beta H}. \tag{C.1}$$

By Taylor expanding of the exponentials, it can be written as

$$A_x(\beta) = A_x + \sum_{m=1}^{\infty} \frac{\beta^m}{m!}[H, \ldots [H, A_x] \ldots], \tag{C.2}$$

where we have used that $[H, \ldots [H, A_x] \ldots] = \sum_{k=0}^{m}\binom{m}{k} H^k A_x (-H)^{m-k}$.

We use now the locality of the Hamiltonian given by its underlying graph $(V, \mathcal{E})$. In order to do so and for the sake of clarity, let us introduce some notation. We refer to edges $\lambda \in \mathcal{E}$ as *letters*, to the edge set $\mathcal{E}$ as an *alphabet*, and call the sequences of edges as *words*. For any *sub-alphabet* $F \subset \mathcal{E}$, we denote by F^l the set of words with letters in F and length l, where the length $|w|$ of a word w is the total number of letters it contains. In turn, the set of words with letters in F and arbitrary length l is defined as $F^* \coloneqq \bigcup_{l=0}^{\infty} F^l$. The size of a sub-alphabet F is denoted by $|F|$ and corresponds to the number of letters it contains. We call a sub-alphabet F an animal if it forms a set of *connected* edges, and a word *connected* or *cluster* $c \in \mathcal{E}^*$ if the set of letters in c is an animal, i.e., connected. That is, clusters are connected sequences of edges where the edges can also occur multiple times, while animals are connected sub-graphs (sets of connected edges without order or repetition). For $w \in E^*$ and any sub-alphabet $G \subset \mathcal{E}$, we write $G \subset w$ if every letter in G also occurs in w. When a sub-alphabet F contains at least one edge adjacent to a vertex x, we denote it by $F \ni x$.

With the above notation, the order n commutator can be written as

$$[H, \ldots [H, A_x] \ldots] = \sum_{w \in \mathcal{E}^m} [h_{w_1}, \ldots [h_{w_m}, A_x] \ldots], \tag{C.3}$$

where w_k is the kth letter of the word w and h_{w_k} the corresponding Hamiltonian term. Note that the only way for the w-commutator $[h_{w_1}, \ldots [h_{w_n}, A_x] \ldots]$ to be nonzero is if w forms a single cluster with at least one of its letters adjacent to the vertex x. More formally,

$$[H, \ldots [H, A_x] \ldots] = \sum_{\substack{F \subset \mathcal{E}: \\ |F| \leqslant n \\ F \ni x}} \sum_{\substack{w \in F^n: \\ F \subset w}} [h_{w_1}, \ldots [h_{w_n}, A_x] \ldots], \tag{C.4}$$

where the first sum runs over all the animals F that contain at least an edge adjacent to the vertex x and are smaller or equal than n. Then, the imaginary time evolution of the observable can be written as

$$A_x(\beta) = A_x + \sum_{\substack{F \subset \mathcal{E}: \\ F \ni x}} \sum_{\substack{w \in F^*: \\ F \subset w}} \frac{\beta^{|w|}}{|w|!}[h_{w_1}, \ldots [h_{w_{|w|}}, A_x] \ldots]. \tag{C.5}$$

We can upper bound the operator norm of $_{Ax}(\beta)$ as

$$\|A_x(\beta)\| \leqslant \|A_x\| \left(1 + \sum_{\substack{F \subset \mathcal{E}: \\ F \ni x}} \sum_{\substack{w \in F^*: \\ F \subset w}} \frac{(J\beta)^{|w|}}{|w|!}\right), \tag{C.6}$$

where we have used that for any operator O, $\|[h_{w_k}, O]\| \leqslant \|h_{w_k}\|\,\|O\|$, whose iteration implies $\|[h_{w_1}, \ldots [h_{w_{|w|}}, A_x] \ldots]\| \leqslant \|A_x\|\, J^{|w|}$. We recall now lemma 5 from [43] which states that

$$\sum_{w \in F^*: F \subset w} \frac{|\beta J|^{|w|}}{|w|!} = (\mathrm{e}^{|\beta J|} - 1)^{|F|} \tag{C.7}$$

and implies

$$\|A_x(\beta)\| \leqslant \|A_x\| \left(1 + \sum_{\substack{F\subset\mathcal{E}:\\ F\ni x}} (\mathrm{e}^{|\beta J|} - 1)^{|F|}\right). \tag{C.8}$$

By decomposing the sum over all animals which contain x according to their size, we get

$$\|A_x(\beta)\| \leqslant \|A_x\| \left(1 + \sum_{l=1}^{\infty}\sum_{\substack{F\subset\mathcal{E}:\\ |F|=l\\ F\ni x}} (\mathrm{e}^{\beta J} - 1)^{|F|}\right). \tag{C.9}$$

The number of lattice animals of size l is upper-bounded by means of the so called *lattice animal constant* α. Given a graph $(V, \mathcal{E})$ and denoting by a_l the number of lattice animals of size l containing the fixed vertex x, then the *animal constant* α is the smallest constant satisfying

$$a_l := \sum_{\substack{F\subset\mathcal{E}:\\ |F|=l\\ F\ni x}} 1 \leqslant \alpha^l. \tag{C.10}$$

Regular lattices have finite animal constants[44]. For example, the animal constant of a D-dimensional cubic lattice can be bounded as $\alpha \leqslant 2D\mathrm{e}$ (lemma 2 in [40]), where e is Euler's number.

Equation (C.10) allows us to bound (C.9) as

$$\|A_x(\beta)\| \leqslant \|A_x\| \sum_{l=0}^{\infty} (\alpha(\mathrm{e}^{\beta J} - 1))^l, \tag{C.11}$$

which is a geometric series with common ratio $\alpha(\mathrm{e}^{2\beta J} - 1)$. In order for the series to converge we require that β is such that $\alpha(\mathrm{e}^{\beta J} - 1) < 1$, for which

$$\|A_x(\beta)\| \leqslant \|A_x\| \frac{1}{1 - \alpha(\mathrm{e}^{\beta J} - 1)}. \tag{C.12}$$

We compute the absolute value of the matrix-element $\langle E_i|\cdot|E_j\rangle$ in equation (C.1)

$$\mathrm{e}^{\beta(E_i - E_j)} |\langle E_i|A_x|E_j\rangle| = |\langle E_i|A_x(\beta)|E_j\rangle| \leqslant \|A_x(\beta)\| \tag{C.13}$$

and, together with equation (C.12), it implies that

$$|\langle E_i|A_x|E_j\rangle| \leqslant \|A_x\| \frac{\mathrm{e}^{-\beta J(E_i - E_j)/J}}{1 - \alpha(\mathrm{e}^{\beta J} - 1)} \tag{C.14}$$

for any β such that $\alpha(\mathrm{e}^{\beta J} - 1) < 1$.

In order to get an explicit bound independent of β, we optimise over β. Given some energy difference $E_i - E_j$, we look for the β that minimises the upper bound of equation (C.14). To do so, it is useful to rewrite the bound in terms of a new parameter $z = \mathrm{e}^{\beta J}$ such that $|\langle E_i|A_x|E_j\rangle| \leqslant \|A_x\| f(z)$ with

$$f(z) := \frac{1}{\alpha(1 + \alpha^{-1} - z)z^{\omega}} \tag{C.15}$$

and $\omega = (E_i - E_j)/J$. By means of an optimisation over z, we find the minimum at $z = \frac{\omega}{\omega+1}\frac{\alpha+1}{\alpha}$ which fulfils the convergence condition $1 < z < 1 + \alpha^{-1}$ as long as $\omega > \alpha$. The minimum value of $f(z_*)$ becomes

$$f(z_*) = \frac{\omega+1}{\alpha+1}\left(\frac{\omega}{\omega+1}\right)^{-\omega}(1 + \alpha^{-1})^{-\omega} \leqslant \frac{\omega+1}{\alpha+1}\mathrm{e}(1 + \alpha^{-1})^{-\omega}, \tag{C.16}$$

where the inequality comes from the fact that $(\omega/(\omega+1))^{-\omega} \leqslant \mathrm{e}$. The matrix element $|\langle E_i|A_x|E_j\rangle|$ is then bounded by

$$|\langle E_i|A_x|E_j\rangle| \leqslant \|A_x\| \, \mathrm{e}^{\log\left(\frac{\mathrm{e}(E_i - E_j)}{J(1+\alpha)}\right) - c(E_i - E_j)/J}, \tag{C.17}$$

where $c = \log(1 + \alpha^{-1})$ is the decay rate.

Appendix D. One-norm distance between coarse-grained frequency signals of two systems which only differ in the level statistics

Two Hamiltonians $H^{(1)}$ and $H^{(2)}$ with identical eigenstates but eigenvalues with different level statistics give rise to time signals for equal initial states and observables given by

$$g^{(1)}(t) = \sum_{\alpha} \nu_{\alpha} e^{iG_{\alpha}^{(1)}t}, \tag{D.1}$$

and analogously for $g^{(2)}(t)$. Here, $G_{\alpha}^{(1)}$ are the gaps associated the spectrum of H_1, and the relevances ν_{α} are the same for both systems. These time-signals have Fourier transforms

$$\tilde{g}^{(1)}(\omega) = \sum_{\alpha} \nu_{\alpha} \delta(\omega - G_{\alpha}^{(1)}) \tag{D.2}$$

that once coarse grained read

$$\tilde{g}_{\epsilon}^{(1)}(\omega) = \sum_{\alpha} \nu_{\alpha} h_{\epsilon}(\omega - G_{\alpha}^{(1)}). \tag{D.3}$$

The one-norm distance between the frequency signals $\tilde{g}_{\epsilon}^{(1)}$ and $\tilde{g}_{\epsilon}^{(2)}$ is defined by

$$\|\tilde{g}_{\epsilon}^{(1)} - \tilde{g}_{\epsilon}^{(2)}\|_1 := \int d\omega |g^{(1)}(\omega) - g^{(2)}(\omega)|. \tag{D.4}$$

Let us now introduce the function

$$u_{\delta}(\omega) = 2\epsilon h_{\epsilon}(\omega) h_{\epsilon}(\delta) \sinh\left(\frac{\delta}{\epsilon}\frac{\omega}{\epsilon}\right). \tag{D.5}$$

By plugging (D.3), we can write the difference in frequency signals as

$$g^{(1)}(\omega) - g^{(2)}(\omega) = \sum_{\alpha} \nu_{\alpha}\ u_{\delta G_{\alpha}}(\omega - \bar{G}_{\alpha}), \tag{D.6}$$

where we have used that $u_{\delta G_{\alpha}}(\omega - \bar{G}_{\alpha}) = h_{\epsilon}(\omega - G_{\alpha}^{(1)}) - h_{\epsilon}(\omega - G_{\alpha}^{(2)})$ with $\delta G_{\alpha} = G_{\alpha}^{(1)} - G_{\alpha}^{(2)}$ and $\bar{G}_{\alpha} = (G_{\alpha}^{(1)} + G_{\alpha}^{(2)})/2$.

In order to understand how the variables $\bar{G}_{\alpha}$ and δG_{α} behave, let us introduce the spectrum of H_1 (with Poisson level statistics) by $\{E_k^{(1)}\}_{k=1}^{d_E}$, and the one of H_2 (with Wigner–Dyson level statistics) by $\{E_k^{(2)}\}_{k=1}^{d_E}$. The energy levels can be iteratively constructed as

$$E_{k+1} = E_k + s_k, \tag{D.7}$$

where $s_k\ \forall\ k$ are i.i.d. random variables sampled from a Poisson distribution in the integrable case and from a Wigner–Dyson distribution in the chaotic case. Both distributions are assumed to have the same mean μ_s and standard deviation $\sqrt{\mu_s}$. The gaps read then

$$G_{\alpha} = E_j - E_i = \sum_{k=i}^{j-1} s_k, \tag{D.8}$$

where we have assumed $i < j$. For $i \ll j$ we can apply the central limit theorem and obtain that the gaps are random variables

$$G_{\alpha} \sim N(\mu_s(j-i), \sqrt{\mu_s(j-i)}), \tag{D.9}$$

where $N(\mu, \sigma)$ is the Gaussian distribution with mean μ and standard deviation σ. The difference of gaps generated by different probability distributions

$$\delta G_{\alpha} := G_{\alpha}^{(1)} - G_{\alpha}^{(2)}, \sim N(0, \sqrt{2(j-i)\mu_s}). \tag{D.10}$$

Note the standard deviation increasing with the separation of the energy levels. This is indeed the behaviour of δG_{α} if both $s_k^{(1)}$ and $s_k^{(2)}$ are independent random variables. However, in our example the spectrum $E_k^{(2)}$ is built from $E_k^{(1)}$ with the single goal of changing its level-statistics. This can be made shifting the energy levels by an amount independent of the separation between them. One way to achieve so is to group the energy levels $\{E_k^{(1)}\}_{k=1}^{d}$ in different sets $\{E_k^{(1)}\}_{k=jL}^{(j+1)L-1}$ of consecutive L energy levels, where j labels the different sets. Then, the first and the last energy levels of every set are kept fixed and the other energy levels are shifted according to the other level-statistics. In such a case, the random variable δG_{α} is bounded by

$$\delta G_{\alpha} \leqslant K\sqrt{\mu_S} \tag{D.11}$$

with $K = \sqrt{2L}$ a constant independent of the system size.

One strategy to bound the one point distance between the coarse-grained frequency signals would be to use the triangular inequality as follows

$$|g_{\epsilon}^{(1)}(\omega) - g_{\epsilon}^{(2)}(\omega)| \leqslant \sum_{\alpha} |\nu_{\alpha}| |u_{\delta G_{\alpha}}(\omega - \bar{G}_{\alpha})| \tag{D.12}$$

and the one norm distance

$$\|g_{\epsilon}^{(1)} - g_{\epsilon}^{(2)}\|_1 \leqslant \sum_{\alpha} |\nu_{\alpha}| \int d\omega |u_{\delta G_{\alpha}}(\omega - \bar{G}_{\alpha})|. \tag{D.13}$$

In order to estimate the scaling of the $|v_\alpha|$ in the system size, let us consider that there are $d_{\rm eff}$ energy levels with a non-negligible occupation in the initial state. Thus, let us note that

$$\sum_\alpha |v_\alpha|^2 \sim d_{\rm eff}^2 |v_\alpha|^2 \leqslant \frac{1}{d_{\rm eff}}, \tag{D.14}$$

and therefore $|v_\alpha| \sim d_{\rm eff}^{-3/2}$.

By splitting the integrals and a convenient change of variables,

$$\begin{aligned}\int {\rm d}\omega |u_\sigma(\omega)| &= 2\left|\int_{-\infty}^{\frac{\sigma}{2}} {\rm d}\omega h_\epsilon(\omega) - \int_{-\infty}^{-\frac{\sigma}{2}} {\rm d}\omega h_\epsilon(\omega)\right| \\ &= 2\left|\int_{-\frac{\sigma}{2}}^{\frac{\sigma}{2}} {\rm d}\omega h_\epsilon(\omega)\right| \leqslant \frac{2\sigma}{\epsilon}.\end{aligned} \tag{D.15}$$

Note that for $\delta G_\alpha \leqslant K\sqrt{\mu_s}$ with $\mu_s \sim n/d_{\rm eff}$, the individual terms of the sum behave as $K\sqrt{n}/(\epsilon d_{\rm eff}^2)$. However, the sum over α contains $d_{\rm eff}^2$ many terms, and therefore the bound becomes useless. The reason for that is that, as it happened for the time signal, the quantity $|g^{(1)}(\omega) - g^{(2)}(\omega)|$ is small due to the cancellations between its different contributions $u_{\delta G_\alpha}(\omega)$.

Instead, let us exploit interference by applying the triangular inequality as follows

$$|g^{(1)}(\omega) - g^{(2)}(\omega)| \leqslant \sum_k \left| \sum_{\bar G_\alpha \in I_k} v_\alpha u_{\delta G_\alpha}(\omega - \bar G_\alpha) \right|, \tag{D.16}$$

where $I_k := [k\delta\omega, (k+1)\delta\omega)$.

For given k, the contribution to the one-norm behaves as

$$\int {\rm d}\omega \left| \sum_{\bar G_\alpha \in I_k} v_\alpha u_{\delta G_\alpha}(\omega - \bar G_\alpha) \right| \sim \sqrt{d_k}\,\langle v_\alpha\rangle_k \int {\rm d}\omega |u_{c\sqrt{\mu_s}}(\omega)| \leqslant \frac{K}{\epsilon}\sqrt{\frac{d_k \mu_s}{d_{\rm eff}^3}}, \tag{D.17}$$

where $d_k := |\{\bar G_\alpha \in I_k\}|$ is the number of gaps in the interval I_k and we have used $|v_\alpha| \sim d_{\rm eff}^{-3/2}$ and equation (D.15). Note that in (D.17) and due to interference between different $u_{\delta G_\alpha}(\omega - \bar G_\alpha)$ the scaling with the number of gaps within an interval I_k is $\sqrt{d_k}$ and not d_k. This behaviour has been inferred from numerical simulations and can be understood by means of the central limit theorem.

Finally, let us note that the amount of gaps in an interval I_k scales as $d_k \sim C_k d_{\rm eff}^2$ where $d_{\rm eff}^2$ is the total number of gaps, n the the system size, and C_k the fraction of gaps in the interval I_k. Thus, adding all the contributions of every interval

$$\|g_\epsilon^{(1)} - g_\epsilon^{(2)}\|_1 \leqslant \frac{C}{\epsilon}\sqrt{\frac{n^3}{d_{\rm eff}}}, \tag{D.18}$$

where we have used that the average level spacing μ_s scales as $\mu_s \sim K_2 n/d_{\rm eff}$ with C and K_2 positive constants independent of the system size.

References

[1] Gogolin C and Eisert J 2016 Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems *Rep. Prog. Phys.* **79** 056001
[2] Reimann P 2007 Typicality for generalized microcanonical ensembles *Phys. Rev. Lett.* **99** 160404
[3] Goldstein S, Lebowitz J L, Tumulka R and Zanghì N 2006 Canonical typicality *Phys. Rev. Lett.* **96** 050403
[4] Popescu S, Short A J and Winter A 2006 Entanglement and the foundations of statistical mechanics *Nat. Phys.* **2** 754–8
[5] Reimann P 2008 Foundations of statistical mechanics under experimentally realistic conditions *Phys. Rev. Lett.* **101** 3
[6] Linden N, Popescu S, Short A J and Winter A 2009 Quantum mechanical evolution towards thermal equilibrium *Phys. Rev.* E **79** 061103
[7] Monnai T 2013 Generic evaluation of relaxation time for quantum many-body systems: analysis of the system size dependence *J. Phys. Soc. Japan* **82** 044006
[8] Brandão F G S L, Ćwikliński P, Horodecki M, Horodecki P, Korbicz J K and Mozrzymas M 2012 Convergence to equilibrium under a random hamiltonian *Phys. Rev.* E **86** 031101
[9] Cramer M 2012 Thermalization under randomized local hamiltonians *New J. Phys.* **14** 053051
[10] Masanes L, Roncaglia A J and Acín A 2013 Complexity of energy eigenstates as a mechanism for equilibration *Phys. Rev.* E **87** 032137
[11] Malabarba A S L, García-Pintos L P, Linden N, Farrelly T C and Short A J 2014 Quantum systems equilibrate rapidly for most observables *Phys. Rev.* E **90** 012121
[12] Goldstein S, Hara T and Tasaki H 2014 The approach to equilibrium in a macroscopic quantum system for a typical nonequilibrium subspace arXiv:1402.3380
[13] Goldstein S, Hara T and Tasaki H 2015 Extremely quick thermalization in a macroscopic quantum system for a typical nonequilibrium subspace *New J. Phys.* **17** 045002
[14] García-Pintos L P, Linden N, Malabarba A S L, Short A J and Winter A 2017 Equilibration time scales of physically relevant observables *Phys. Rev.* X **7** 031027
[15] Short A J and Farrelly T C 2012 Quantum equilibration in finite time *New J. Phys.* **14** 013063

[16] Farrelly T 2016 Equilibration of quantum gases *New J. Phys.* **18** 073014
[17] Reimann P 2016 Typical fast thermalization processes in closed many-body systems *Nat. Commun.* **7** 10821
[18] Balz B N and Reimann P 2017 Typical relaxation of isolated many-body systems which do not thermalize *Phys. Rev. Lett.* **118** 190601
[19] Santos L F and Torres-Herrera E J 2017 Nonequilibrium quantum dynamics of many-body systems arXiv:1706.02031
[20] Goldstein S, Hara T and Tasaki H 2013 On the time scales in the approach to equilibrium of macroscopic quantum systems *Phys. Rev. Lett.* **111** 140401
[21] Srednicki M 1994 Chaos and quantum thermalization *Phys. Rev.* E **50** 901
[22] Wilming H, Goihl M, Krumnow C and Eisert J 2017 Towards local equilibration in closed interacting quantum many-body systems arXiv:1704.06291
[23] Arad I, Kuwahara T and Landau Z 2016 Connecting global and local energy distributions in quantum spin models on a lattice *J. Stat. Mech.* 2016 033301
[24] Short A J 2011 Equilibration of quantum systems and subsystems *New J. Phys.* **13** 053009
[25] Reimann P 2010 Canonical thermalization *New J. Phys.* **12** 055027
[26] Fonda L, Ghirardi G C and Rimini A 1978 Decay theory of unstable quantum systems *Rep. Prog. Phys.* **41** 587
[27] Robinett R W 2004 Quantum wave packet revivals *Phys. Rep.* **392** 1–119
[28] Eberly J H, Narozhny N B and Sanchez-Mondragon J J 1980 Periodic spontaneous collapse and revival in a simple quantum model *Phys. Rev. Lett.* **44** 1323
[29] Eberly J H, Narozhny N B and Sanchez-Mondragon J J 1981 Coherence versus incoherence: collapse and revival in a simple quantum model *Phys. Rev.* A **23** 236
[30] Allen L and Eberly J H 1987 *Optical Resonance and Two-Level Atoms* (New York: Dover)
[31] Lifshitz E M and Pitaevskii L P 1958 *Statistical Physics: Theory of the Condensed State* (*Course of Theoretical Physics* vol 9) (Oxford: Pergamon)
[32] Baym G and Pethick C 2004 *Landau Fermi-Liquid Theory: Concepts and Applications* (New York: Wiley-VCH)
[33] Polkovnikov A, Sengupta K, Silva A and Vengalattore M 2011 Colloquium: nonequilibrium dynamics of closed interacting quantum systems *Rev. Mod. Phys.* **83** 863
[34] Stein E M and Shakarchi R 2015 *Fourier Analysis: an Introduction* (Princeton, NJ: Princeton University Press)
[35] Eberlein W F 1949 Abstract ergodic theorems and weak almost periodic functions *Trans. Amer. Math. Soc.* **67** 217–40
[36] Corduneanu C 2009 *Almost Periodic Oscillation and Waves* (Berlin: Springer)
[37] Brandão F G S L and Cramer M 2015 Equivalence of statistical mechanical ensembles for non-critical quantum systems arXiv:1502.03263
[38] D'Alessio L, Kafri Y, Polkovnikov A and Rigol M 2016 From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics *Adv. Phys.* **65** 239–362
[39] Haque M, Beugeling W and Moessner R 2015 Off-diagonal matrix elements of local operators in many-body quantum systems *Phys. Rev.* E **91** 012144
[40] Miranda Y M and Slade G 2011 The growth constants of lattice trees and lattice animals in high dimensions *Electron. Commun. Probab.* **16** 129
[41] Perarnau-Llobet M, Wilming H, Riera A, Gallego R and Eisert J 2017 Fundamental corrections to work and power in the strong coupling regime arXiv:1704.05864
[42] Hastings M B and Koma T 2006 Spectral gap and exponential decay of correlations *Commun. Math. Phys.* **265** 781–804
[43] Kliesch M, Gogolin C, Kastoryano M J, Riera A and Eisert J 2014 Locality of temperature *Phys. Rev.* X **4** 031019
[44] Penrose M 1994 Self-avoiding walks and trees in spread-out lattices *J. Stat. Phys.* **77** 3