Typical thermalization of low-entanglement states

Christian Bertoni,[1] Clara Wassner,[1] Giacomo Guarnieri,[2,1] and Jens Eisert[1,3]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität, 14195 Berlin, Germany*
[2]*Department of Physics, University of Pavia, Via Bassi 6, 27100, Pavia, Italy*
[3]*Helmholtz Center Berlin, 14109 Berlin, Germany*

(Dated: April 24, 2024)

Proving thermalization from the unitary evolution of a closed quantum system is one of the oldest questions that is still nowadays only partially resolved. Several efforts have led to various formulations of what is called the eigenstate thermalization hypothesis, which leads to thermalization under certain conditions on the initial states. These conditions, however, are sensitive to the precise formulation of the hypothesis. In this work, we focus on the important case of low entanglement initial states, which are operationally accessible in many natural physical settings, including experimental schemes for testing thermalization and for quantum simulation. We prove thermalization of these states under precise conditions that have operational significance. More specifically, motivated by arguments of unavoidable finite resolution, we define a random energy smoothing on local Hamiltonians that leads to local thermalization when the initial state has low entanglement. Finally we show that such a transformation affects neither the Gibbs state locally nor, under generic smoothness conditions on the spectrum, the short-time dynamics.

Not long after their formulation, it became clear that the postulates of quantum mechanics were to a degree at odds with the principles of statistical mechanics [1]. On the one hand, closed quantum systems evolve unitarily, thus preserving information about their initial conditions. On the other hand, statistical mechanics is centered around concepts such as irreversibility and thermalization, which allow to describe many-body systems in terms of just a few macroscopic quantities and parameters that depend neither on the initial conditions nor on the microscopic details [2–6]. Both theories have nowadays provided correct predictions of uncountable experimental observations. Thermalization of closed many-body quantum systems is in particular observed to occur in practice with overwhelming evidence, also in quantum simulators which allow to probe their dynamics with high levels of precision [7–12]. This apparent contradiction is eased once one only requires the expectation value of a subset of "physical" observables, for example local observables, to agree with the predictions of statistical mechanics. As a matter of fact, in a lattice system, the state of the system reduced to local patches can relax to a thermal state while the global state evolves unitarily and remains pure. Nevertheless, a derivation of thermalization in this sense from the microscopic description of the dynamics has remained largely elusive. Efforts in this directions have led to various formulations of the *eigenstate thermalization hypothesis* (ETH) [13, 14], which posits that each eigenstate of a thermalizing Hamiltonian provides the same expectation values of physical observables as the ones provided by local thermal states. The ETH, however, is difficult to verify starting from a microscopic model. In addition, the precise conditions on the initial state and on the observables leading to thermalization are highly sensitive to the formulation of the hypothesis. On the other hand, *absence* of thermalization, as observed in integrable or localized systems [15–17], seems to be the occurrence requiring particular, carefully set-up conditions. This has led to the idea of *typicality*, i.e., that "most" physical systems obey statistical mechanics. This usually means that if the system (for instance, its Hamiltonian) is drawn at random from a reasonable set, with overwhelming probability thermalization occurs as expected, and while any given system under examination is usually not randomly drawn, it should behave like a typical one unless explicitly engineered not to. Works in this direction [18–22] have contributed to formalize the idea that thermalization is the rule, rather than the exception, even in isolated systems.

A drawback of many of these approaches, in the case of both the ETH and typicality, is that they often require the assumption that the system is prepared in an initial state which is concentrated around a given energy E. While all states with sub-extensive energy variance are concentrated around their average energy, it is often not clear how strong the concentration needs to be in order for thermalization to be guaranteed given either a formulation of ETH or a typicality transformation. One often assumes that the support of the state vanishes outside of a narrow micro-canonical energy window; depending on how narrow this window is such a state generally cannot be efficiently prepared. Crucially, this excludes important classes of physical and operationally meaningful initial states, such as product states and states with low entanglement. These states, in addition to being fundamental to the study of lattice systems, are commonly the only accessible initial states to thermalization experiments in quantum simulators or numerical examination of thermalization, for example in the context of probing many-body localization [23–27], and are often considered in works on relaxation and equilibration [28–30].

In this work, we provide this important missing piece of the puzzle by proving thermalization of low-entanglement states. We address this by studying states in an analog of the micro-canonical ensemble which is allowed to have support on the whole energy spectrum. Building upon and extending the work of Ref. [31], we show that, under precise and physically meaningful conditions, these states are locally equivalent to Gibbs states. We later show, generalizing the seminal work of Ref. [21], that such states crucially arise from the long-time dynamics of low-entanglement initial states under typical Hamiltonians. Our typicality transformation consists of a unitary operation that randomly mixes eigenstates with nearby energies. Related transformations have been introduced in Ref. [32]

to compute higher order correlation functions under a generalized version of the ETH stemming from local (in energy) unitary invariance, and this has been connected to the theory of free probability in Ref. [33]. In our work, we explicitly demonstrate how such a transformation leads to equilibrium states being locally indistinguishable from thermal states. Moreover, by bounding the energy difference between eigenstates that get mixed, we elucidate how the original Hamiltonian is related to the transformed one. More specifically, given a local Hamiltonian H and an initial low-entanglement state ρ, we define an ensemble of random Hamiltonians such that, if H' is a Hamiltonian drawn from this ensemble: 1. the Gibbs states of H and H' are locally indistinguishable, and 2. with overwhelming probability the equilibrium state resulting from the unitary evolution $\mathrm{e}^{-\mathrm{i}H't}\rho\mathrm{e}^{\mathrm{i}H't}$ is locally indistinguishable from said Gibbs state. In addition, we give an assumption under which the dynamics of H and H' are indistinguishable in the short-time regime. This new kind of Hamiltonian-typicality approach is rooted in physical grounds: it is unrealistic to have a system specified to infinite precision. This approach relaxes and overcomes the common idea that thermalization has to be proven for all Hamiltonians, but rather, it turns out to be typical for an overwhelming portion of meaningful settings.

Physical setting, equilibration, and thermalization. Throughout this work we consider a cubic D-dimensional lattice Λ with N sites. With each site i we associate a Hilbert space $\mathcal{H}_i$ of dimension d. H is a k-local Hamiltonian on Λ for a constant k, meaning that it is of the form $H = \sum_{i\in\Lambda} h_i$ where h_i has operator norm 1 and is only supported on sites j s.t. $\mathrm{d}(i,j) \leq k$. Here, $\mathrm{d}(\cdot,\cdot)$ is the standard Manhattan distance on the lattice. We will denote the spectral decomposition of H as $H = \sum_\nu E_\nu P_\nu$. For ease of presentation, in the main text we will assume the eigenstates to be non-degenerate, i.e., $\mathrm{tr}\,(P_\nu) = 1$ for all E_ν. In the Supplemental Material, we prove our results relaxing this assumptions. We will use standard O and Ω notation, and we will use $\tilde{O}, \tilde{\Omega}$ to denote asymptotic upper bounds where logarithmic factors are ignored, i.e. $f(h) = \tilde{O}(g(h))$ iff $f(h) = O(\log^r(h)g(h))$ for some r. Before discussing thermalization, we need to discuss a natural prerequisite: equilibration [6, 34–36]. Intuitively, a state equilibrates if after a finite amount of time the system reaches an invariant state. This means that if one looks at the whole evolution as time tends to infinity, the system will spend most of its time close to equilibrium, and hence the state at equilibrium should agree with the infinite-time averaged state

$$\rho_\infty^H = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t\, \rho^H(t), \tag{1}$$

with $\rho^H(t) = \mathrm{e}^{iHt}\rho\mathrm{e}^{-iHt}$. The system is then said to equilibrate if the fluctuations around the average for an observable A are small in the sense of

$$\Delta A_\infty := \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t\, \mathrm{tr}\left(A(\rho^H(t) - \rho_\infty^H)\right)^2 \xrightarrow{N\to\infty} 0. \tag{2}$$

Equilibration in this sense has been rigorously proven in a variety of settings [36–38]. Thermalization essentially consists of a stronger requirement of equilibration, where the equilibrium state must coincide with the thermal (Gibbs) state

$$g_\beta(H) = \frac{\mathrm{e}^{-\beta H}}{Z}, \quad Z = \mathrm{tr}\left(\mathrm{e}^{-\beta H}\right). \tag{3}$$

In particular, we will be interested in understanding how distinguishable a given state is from a Gibbs state, given only access to local observables. We introduce the quantity

$$D_l(\rho,\sigma) := \frac{1}{|\mathcal{C}_l|}\sum_{C\in\mathcal{C}_l} \|\rho_C - \sigma_C\|_1, \tag{4}$$

where $\mathcal{C}_l$ denotes the set of all hypercubes in the lattice of side length l and ρ_C is the state ρ reduced to C. If ρ, σ are translationally invariant, this is simply equal to the trace norm $\|\rho_C - \sigma_C\|_1$ for any $C \in \mathcal{C}_l$. Otherwise it measures the distinguishability of ρ and σ given access to observables of the form $\frac{1}{|\mathcal{C}_l|}\sum_{C\in\mathcal{C}_l} O_C$, such as typical average local quantities in statistical physics. We will then say that a state is *locally thermal* if $D_l(\rho, g_\beta(H))$ converges to 0 with N for some β. In the remainder of this work, we will make some assumptions on the Gibbs state of the system under examination, $g_\beta(H)$. First, we will require it to have exponential decay of correlations. A state ρ is said to have *exponential decay of correlations* if for some $\xi > 0$, for any two regions $X, Y \subset \Lambda$ and A, B supported on X and Y respectively,

$$|\langle AB\rangle_\rho - \langle A\rangle_\rho\langle B\rangle_\rho| \leq \|A\|\|B\|\mathrm{e}^{-\mathrm{d}(X,Y)/\xi}. \tag{5}$$

The Gibbs state $g_\beta(H)$ always has an exponential decay of correlations in one dimension [39], as well as in any higher dimension above a certain critical temperature [40] that depends only on a few parameters of the system. We will denote by σ the standard deviation of the energy, i.e., $\sigma^2 = \mathrm{tr}\left(g_\beta(H)H^2\right) - \mathrm{tr}\left(g_\beta(H)H\right)^2$, and assume $\sigma^2 \geq \Omega(N)$, which implies that the specific heat capacity is non-zero in the thermodynamic limit.

Conditions for thermality. The *micro-canonical ensemble* is commonly defined as the maximally mixed state supported in a narrow window around a fixed energy. This physically models maximal uncertainty with an energy constraint. However, arbitrary and even physically motivated states such as low-entanglement states, can have non-trivial support over the whole spectrum during the whole time evolution; these states will therefore never be micro-canonical in the sense above, particularly not at equilibrium. With the goal of overcoming this issue, we leverage techniques used to prove equivalence between micro-canonical and canonical ensemble. It has been proven in Ref. [31] that equivalence with a canonical ensemble is already achieved by states confined in a micro-canonical window that are not maximally mixed, but have sufficiently high entropy. We adapt this result to the situation in which the state is not confined in a window, but instead has a support over many such windows plus decaying tails. First of all, we call this a *generalized micro-canonical ensemble*, meant to capture the thermal behavior of states that are supported on regions of the spectrum larger than what the usual micro-canonical ensemble allows.

Definition 1 (Generalized micro-canonical ensemble (GmE))**.** *Let $[E-\Delta, E+\Delta]$ ($\Delta > 0$) denote an energy window centered around a value E and divided into K bins of various size δ_k, with $k = 1, \ldots, K$. Let $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_K)$; we define a* generalized micro-canonical ensemble *(GmE) to be the state of the form*

$$\omega := \omega(E, \Delta, \boldsymbol{\delta}, \mathbf{q}) = \sum_{k=1}^{K} q_k \omega_{\delta_k} \tag{6}$$

where ω_{δ_k} is the micro-canonical ensemble supported inside the window k, and where $\mathbf{q} = (q_1, \ldots, q_K)$ such that $\sum_{k=1}^{K} q_k = 1$.

This state therefore physically represents a statistical combination of micro-canonical ensembles; see Fig. 1. For the sake of simplicity, we will choose $\delta_k = \delta$ from now on. However, in the Supplemental Material we show that all our results still hold true if this assumption is relaxed. Before stating our first main result, we need to introduce the notion of the *Berry-Esseen* (BE) error, which quantifies the difference between a state written in the energy eigenbasis and a Gaussian distribution. More specifically, if Π_x is the projector onto all energy eigenstates with energy smaller than x, then the BE error of ρ with respect to H is defined as $\zeta_N = \sup_x |\mathrm{tr}\,(\rho\Pi_x) - G(x)|$, where $G(x)$ is the Gaussian distribution with the same mean and variance as ρ. It was proven that if ρ has exponential decay of correlations, then $\zeta_N \leq \tilde{O}(N^{-1/2})$ [31]. Simple examples saturating this bound are known, and this bound is expected to be saturated by certain non-thermalizing models. Nonetheless, under some more generic constraints, such as highly entangled eigenstates, a more favorable scaling is expected, even up to $\zeta_N \leq \mathrm{e}^{-\Omega(N)}$ [41]. From now on we denote by ζ_N the BE error with $\rho = g_\beta(H)$. In what follows, we will assume $\zeta_N \leq \tilde{O}(N^{-1/2-\kappa})$ for some $\kappa \geq 0$. This includes the worst case $\kappa = 0$.

Theorem 1 (Ensemble equivalence)**.** *Let H be a local Hamiltonian and β an inverse temperature for which the Gibbs state $g_\beta(H)$ has exponential decay of correlations, standard deviation $\sigma \geq \Omega(\sqrt{N})$ and Berry-Esseen error $\zeta_N \leq \tilde{O}(N^{-1/2-\kappa})$ for $\kappa \geq 0$. Let ω denote a GmE with Δ, δ satisfying*

$$\mathrm{e}^{\Delta^2/\sigma^2} \leq \tilde{O}\left(N^{\frac{1-\alpha}{D+1}}\right), \quad \Omega\left(N^{\frac{1-\alpha}{D+1}-\kappa}\right) \leq \delta \leq \sigma, \tag{7}$$

with $\alpha \in [0, 1)$ and such that $|E - E_\beta| \leq \sigma$. Then for any side length l such that $l^D \geq C_1\, N^{\frac{1}{D+1}-\gamma_1\alpha}$, the following holds

$$D_l(\omega, g_\beta(H)) \leq C_2 N^{-\gamma_2\alpha}, \tag{8}$$

with C_1, C_2 being system-size independent constants, and γ_1, γ_2 only depend on the dimension of the lattice D.

This first main result shows that, for appropriate choices Δ and δ, GmE states are locally indistinguishable from Gibbs states. A GmE state can be seen as a mixture of micro-canonical states spanning a range of temperatures and Theorem 1 shows that as long as its range is small enough, the state still looks thermal with a well-defined temperature. Notice that if $\kappa > 0$, i.e., if the BE error is better than the worst case scenario, δ can be chosen to decay with the system size.

Keeping in mind our initial goal of capturing equilibrium states resulting from natural and physically motivated initial states, it may seem artificial to consider only block-like states with sharp jumps between energy intervals. Therefore, postponing the discussion about their physicality to the next Section below, we first of all prove that the same ensemble equivalence holds if the state's structure gets more relaxed, i.e., if it is only approximately GmE in the sense precisely elucidated below.

Definition 2 (Approximate GmE)**.** *Let $E, \Delta, \boldsymbol{\delta}, \mathbf{q}$ be as in Definition 1. We define ω_η an approximate GmE if it is of the form*

$$\omega_\eta = p_\Delta \left(\sum_{k=1}^{K} q_k \tilde{\omega}_{\delta_k}\right) + (1 - p_\Delta)\rho_{\mathrm{tail}} \tag{9}$$

and its von Neumann entropy satisfies

$$\sum_{k=1}^{K} q_k (S(\omega_{\delta_k}) - S(\tilde{\omega}_{\delta_k})) \leq \eta, \tag{10}$$

with $\tilde{\omega}_{\delta_k}$ being defined on the Hilbert space spanned by the eigenstates in the k-th energy bin, and ρ_{tail} on the Hilbert space spanned by the eigenstates outside $[E - \Delta, E + \Delta]$.

This state represents a more physical version of a GmE state inside the energy window Δ, with decaying tails outside, that has an entropy η-close to the maximum one. Importantly, in the Supplemental Material we demonstrate that Theorem 1 holds true also for the approximate GmE and takes the form

$$D_l(\omega_\eta, g_\beta(H)) \leq C_2 N^{-\gamma_2\alpha} + 2(1 - p_\Delta), \tag{11}$$

with $\eta \leq N^{\frac{1-\alpha}{D+1}}$.

This shows that states which are concentrated around an energy regime and are sufficiently "smooth" are locally equivalent to Gibbs states. The results above are a generalization of the equivalence of ensembles result of Ref. [31], and their proof is presented in the Supplemental Material, Section I.

Thermalization via energy smoothing. The next question is whether (approximate) GmE states can actually be obtained from Hamiltonian evolution of isolated systems under natural, or typical, conditions. There are two main aspects we consider when talking about "natural" conditions: (i) the Hamiltonian responsible for the time evolution; (ii) the initial state of the system. Regarding (i), we consider *typical* Hamiltonians in a sense that we will make precise below, in order to exclude edge cases or fine-tuned Hamiltonians for which one does not expect thermalization (for instance, integrable models). Concerning (ii), previous typicality approaches have assumed the initial state to be confined in a well-defined energy interval, and have shown properties of the relaxation towards a micro-canonical ensemble in said interval. Here, instead, we start from the assumption of exponential decay of correlations, i.e., low entanglement between spatially separated regions, which we take as natural starting states for lattice systems. These states have

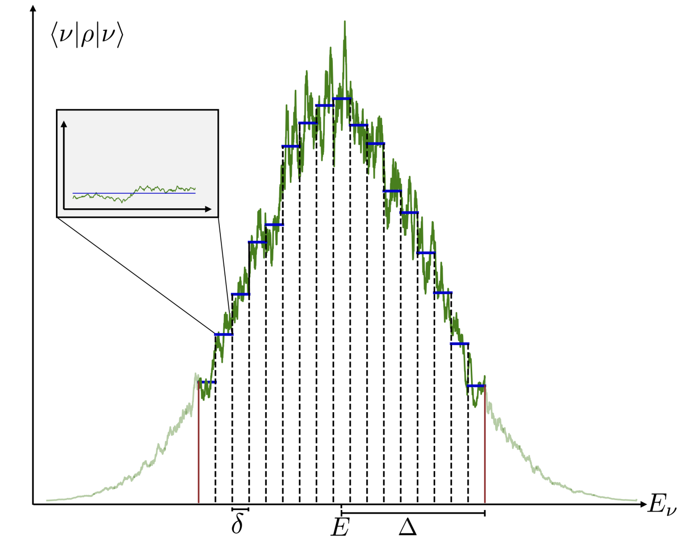

FIG. 1. Cartoon illustration of the setting of Definition 1 and Definition 2. In this example, the state is diagonal in the energy eigenbasis and it is represented as a probability distribution. The blue line is a GmE state, and the green line an approximate GmE state. The insert shows both states restricted to one of the windows.

been shown to have fast decaying tails in energy [42] which makes them ideal candidates to flow to approximately GmE states. Here, for simplicity of presentation, we focus on the case of product states, and leave the more general case of states with exponential decay of correlations and the proofs to the Supplemental Material, Section II.

Let us expand on the ensemble of typical Hamiltonians that we consider. Starting from any local Hamiltonian on the lattice, we divide its energy spectrum into energy intervals of equal width δ which we call I_k, for $k = 1, \cdots, K$. The eigenstates contained within each interval span a vector space which we call $\mathcal{W}_k$. We then consider unitaries of the form $U = \bigoplus_k U_k$, where U_k is drawn from the Haar measure of the unitary group acting on $\mathcal{W}_k$. This defines an ensemble of random unitaries which we denote as $\mathcal{E}(\delta)$. A typical Hamiltonian is then $UHU^\dagger$ for such a random unitary U. All these Hamiltonians have the same spectrum as the original local Hamiltonian H, and the randomization given by U is designed to preserve the expected energy of any state.

The following is a consequence of measure concentration and the results of Ref. [42] about energy tails of product states.

Lemma 1 (Approximate GmE at equilibrium)**.** *Let ρ be a product state and H be a local Hamiltonian. Let $\rho_\infty^{UHU^\dagger}$ be defined as in Eq. (1), where U is drawn from $\mathcal{E}(\delta)$. Consider the interval $I = [E-\Delta, E+\Delta]$ around $E = \mathrm{tr}\left(\rho UHU^\dagger\right)$ with $\Delta \geq \omega(\sqrt{N})$ an integer multiple of δ, then*

$$\rho_\infty^{UHU^\dagger} = p_\Delta \left(\sum_{k: I_k \subset I} q_k \tilde{\omega}_{\delta_k} \right) + (1-p_\Delta)\rho_{\mathrm{tail}} \tag{12}$$

with $p_\Delta \geq 1 - \mathrm{e}^{-c_1 \frac{\Delta^2}{N}}$, and for $r > 0$, with probability at least $1 - 2^{-r+1}$, we have

$$\sum_{k: I_k \subset I} q_k (S(\omega_{\delta_k}) - S(\tilde{\omega}_{\delta_k})) \leq r, \tag{13}$$

where c_1 is a system-size independent constant.

We have then the following consequence on typical thermalization.

Theorem 2 (Typical thermalization)**.** *Let H be a k-local Hamiltonian and ρ be a product state. Let $g_\beta(H)$ be the Gibbs state of H at inverse temperature β such that $|\mathrm{tr}\,(g_\beta(H)H) - \mathrm{tr}\,(\rho H)| \leq \sigma$. Assume $g_\beta(H)$ has exponential decay of correlations, $\sigma \geq \Omega(\sqrt{N})$, and $\zeta_N \leq \tilde{O}(N^{-1/2-\kappa})$. For any constant $\alpha \in [0,1)$, choosing $\delta = \Omega\left(N^{\frac{1-\alpha}{D+1}-\kappa}\right)$, with probability at least $1 - \exp(-c_2 N^{\frac{1-\alpha}{D+1}})$ drawing U at random from $\mathcal{E}(\delta)$, we have*

$$D_l(\rho_\infty^{UHU^\dagger}, g_\beta(H)) \leq C_2\, N^{-\gamma_2 \alpha} + \tilde{O}\left(N^{-\gamma_3(1-\alpha)}\right), \tag{14}$$

where $c_2, C_2, \gamma_2, \gamma_3$ are system-size independent constants.

In the Supplemental Material, Section II, we state and prove these results more generally for any state concentrated around its average, which includes states with exponentially decaying correlations. The consequences of this relaxation of the assumption is that the decay in the system size is slower. Theorem 2 shows that the equilibrium state is locally thermal; in the Supplemental Material, Section II, we show under mild spectral assumptions that the randomized Hamiltonian equilibrates with high probability to this state. Although it may seem strange at first glance that under the dynamics of $UHU^\dagger$ the state thermalizes to the Gibbs state of H and of $UHU^\dagger$, we prove that the Gibbs states of these two Hamiltonians are locally indistinguishable. More specifically, under the same assumptions as Theorem 2, for any U drawn from $\mathcal{E}(\delta)$ we have

$$D_l(g_\beta(H), g_\beta(UHU^\dagger)) \leq O(N^{-\gamma_4\alpha - \gamma_4\kappa}) \tag{15}$$

for system-size independent constants γ_4, γ_5. The proof may be found in the Supplemental Material, Section II, and easily generalizes to other choices of δ. As anticipated, the unitary ensemble $\mathcal{E}(\delta)$ is chosen in order to approximately preserve the energy of any state; this implies that $U \sim \mathcal{E}(\delta)$ approximately commutes with the Hamiltonian, and we show $\|H - UHU^\dagger\|_\infty \leq \delta$. For the choice of δ as in Theorem 2 we derive the following result

$$\|\mathrm{e}^{-iHt}\rho\mathrm{e}^{iHt} - \mathrm{e}^{-iH't}\rho\mathrm{e}^{iH't}\|_1 \leq 2t\, O\left(N^{\frac{1-\alpha}{D+1}-\kappa}\right), \tag{16}$$

with $H' = UHU^\dagger$. This means that the dynamics under U and $UHU^\dagger$ are indistinguishable up to a time $t^* \sim N^{\kappa - \frac{1-\alpha}{D+1}}$. If $\kappa > 0$, that is the BE error decays faster than the worst-case scenario, α can be chosen such that t^* increases with the system size. In other words H and $UHU^\dagger$ generate nearly the same dynamics for a time $t^* \sim \mathrm{poly}(N)$. Finally, we would like to emphasize that our rigorous approach allows to put on precise and solid ground some of the results obtained on equilibration in Ref. [43].

It is now worth noting that if both ρ and σ are translation-invariant, the averaging over regions in the definition of $D_l(\rho, \sigma)$ can be dropped, making the indistinguishability statement valid for any observable supported on an individual small region. We show that if the original Hamiltonian is translation-invariant, we recover this property to some extent in the equilibrium state of the perturbed Hamiltonian. More specifically, consider an observable A supported in $C \in \mathcal{C}_l$ and H translation-invariant. For U drawn from $\mathcal{E}(\delta)$ we show that, with probability at least $1 - \sum_k \exp(-cd_k\epsilon/\|A\|_\infty^2)$ for a constant c, it holds, with $d_k = \dim(\mathcal{W}_k)$,

$$\left|\mathrm{tr}\left(\left(g_\beta(H) - \rho_\infty^{UHU^\dagger}\right)A\right)\right| \leq \epsilon + D_l\left(\rho_\infty^{UHU^\dagger}, g_\beta(H)\right). \tag{17}$$

Details and proofs are available in the Supplemental Material, Section IV. Notice that the probability is large assuming that the number of eigenstates in each windows grows with the system size. By applying this to $\rho = g_\beta(UHU^\dagger)$ we also get translation invariance in the same sense for the randomized Gibbs state, that is, $\rho_\infty^{UHU^\dagger}$ can be replaced by $g_\beta(UHU^\dagger)$ in Eq. (17).

Dynamical thermalization. Turning to notions of *dynamical thermalization*, we have investigated the typical time-evolution of the expectation value of a generic observable A, $\langle A\rangle_\rho := \mathrm{tr}\,(A\rho)$, under the evolution generated by $UHU^\dagger$. In the Supplemental Material, Section III, we show that, with probability at least $1 - \sum_k \exp\left(-cd_k\epsilon/\|A\|_\infty^2\right)$ for a constant c, the time-evolution is bounded by

$$|\langle A\rangle_{\rho^{UHU^\dagger}(t)} - \langle A\rangle_{\rho_\infty^{UHU^\dagger}}| \leq \epsilon + R(t), \tag{18}$$

where $R(t)$ is a function of t depending on details of the spectrum of H, on A, and on ρ. Performing a similar analysis to the one of Ref. [21] to our ensemble, and assuming that the spectrum in each window can be well approximated by a suitably flat continuous spectrum, we show that

$$R(t) \sim \|A\|_\infty \frac{N^2}{\delta^2 t^2}. \tag{19}$$

Hence, under this physical assumption, thermalization up to some ϵ is reached after a time $\sim N^{O(1)}/\epsilon$.

Discussion and conclusion. In this work, we have made substantial progress in the long-standing research quest of proving thermalization from first principles of quantum mechanics, by showing thermalization of low-entanglement states under typical Hamiltonians. In particular, we have defined ensembles which naturally emerge from the time-evolution of low-entanglement states under typical Hamiltonian evolution and show that they are locally indistinguishable from the Gibbs ensemble. The typicality in the Hamiltonian is given by randomizing the eigenbasis locally in energy, and we show that this randomization does not affect the Gibbs state locally. Furthermore, we show that if the Gibbs state has a relatively generic smoothness property, the randomization does not affect the short-time dynamics in any discernible way.

On a higher level, this is one of the main physical messages of this work: it is unrealistic to assume that a Hamiltonian can be specified up to arbitrary precision. The ubiquity of thermalization may then be explained by considering that its absence requires instead a high precision in the specification of the system. In this light, proving thermalization for all, or a large class of, local Hamiltonians might not be required for explaining this observation.

The technical results established here are also expected to be helpful in the design of *quantum algorithms* for preparing Gibbs states [44–46]. Relaxing the assumption of requiring the entire state to be globally indistinguishable from a Gibbs state (as suggested in Ref. [46]) may well be helpful in this endeavour.

Several natural open questions remain. Importantly, the problem of identifying precise properties of the Hamiltonian achieving a lower than worst case BE error. For these Hamiltonians, our result predicts that a vast ensemble of other Hamiltonians exists that all have indistinguishable short time dynamics and which eventually thermalize. Furthermore, while our randomization achieves thermalization, more work may be needed about exploring the precise physical mechanisms that can be held responsible for them. On the one hand, being able to implement such transformations in a controlled way might have implications in the form of algorithms for Gibbs state preparation. On the other hand, understanding under what condition an uncertainty in either state or Hamiltonian preparation can translate into a random energy-preserving perturbation such as the one we defined might shed light on the stability of many-body localization under realistic conditions. It is the hope that this work contributes to understanding that under a "trembling hand", most systems follow the laws of quantum statistical mechanics.

Acknowledgements. We thank Henrik Wilming and Ingo Roth on discussions during early stages of this project. This work has been supported by the DFG (FOR 2724 and CRC 183), the BMBF (MUNIQCAtoms), the Quantum Flagship (PasQuans2), the FQXi, and the ERC (DebuQC).

[1] J. v. Neumann, Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik, Z. Phys. **57**, 30 (1929).
[2] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).
[3] E. T. Jaynes, Information theory and statistical mechanics. ii, Phys. Rev. **108**, 171 (1957).
[4] J. Eisert, M. Friesdorf, and C. Gogolin, Quantum many-body systems out of equilibrium, Nature Phys. **11**, 124 (2015).
[5] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Nonequilibrium dynamics of closed interacting quantum systems, Rev. Mod. Phys. **83**, 863 (2011).

[6] C. Gogolin and J. Eisert, Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems, Rep. Prog. Phys. **79**, 56001 (2016).

[7] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas, Nature Phys. **8**, 325 (2012).

[8] C. Neill, P. Roushan, M. Fang, Y. Chen, M. Kolodrubetz, Z. Chen, A. Megrant, R. Barends, B. Campbell, B. Chiaro, A. Dunsworth, E. Jeffrey, J. Kelly, J. Mutus, P. J. J. O'Malley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. Polkovnikov, and J. M. Martinis, Ergodic dynamics and thermalization in an isolated quantum system, Nature Phys. **12**, 1037 (2016).

[9] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Quantum thermalization through entanglement in an isolated many-body system, Science **353**, 794 (2016).

[10] G. Clos, D. Porras, U. Warring, and T. Schaetz, Time-resolved observation of thermalization in an isolated quantum system, Phys. Rev. Lett. **117**, 170401 (2016).

[11] F. H. B. Somhorst *et al.*, Quantum simulation of thermodynamics in an integrated quantum photonic processor, Nature Comm. **14**, 3895 (2023).

[12] T. Schweigler, M. Gluza, M. Tajik, S. Sotiriadis, F. Cataldini, S.-C. Ji, F. S. Möller, J. Sabino, B. Rauer, J. Eisert, and J. Schmiedmayer, Decay and recurrence of non-Gaussian correlations in a quantum many-body system, Nature Phys. **17**, 559 (2021).

[13] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).

[14] J. M. Deutsch, Eigenstate thermalization hypothesis, Rep. Prog. Phys. **81**, 082001 (2018).

[15] P. W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. **109**, 1492 (1958).

[16] J. Billy, V. Josse, Z. Zuo, A. Bernard, B. Hambrecht, P. Lugan, D. Clément, L. Sanchez-Palencia, P. Bouyer, and A. Aspect, Direct observation of anderson localization of matter waves in a controlled disorder, Nature **453**, 891 (2008).

[17] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Many-body localization, thermalization, and entanglement, Rev. Mod. Phys. **91**, 021001 (2019).

[18] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Canonical typicality, Phys. Rev. Lett. **96**, 050403 (2006).

[19] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, Nature Phys. **2**, 754 (2006).

[20] M. Cramer, Thermalization under randomized local Hamiltonians, **14**, 053051 (2012).

[21] P. Reimann, Typical fast thermalization processes in closed many-body systems, Nature Comm. **7**, 10821 (2016).

[22] L. Dabelow, P. Vorndamme, and P. Reimann, Thermalization of locally perturbed many-body quantum systems, Phys. Rev. B **105**, 024310 (2022).

[23] J. H. Bardarson, F. Pollmann, and J. E. Moore, Unbounded growth of entanglement in models of many-body localization, Phys. Rev. Lett. **109**, 017202 (2012).

[24] I. H. Kim, A. Chandran, and D. A. Abanin, Local integrals of motion and the logarithmic lightcone in many-body localized systems (2014), arXiv:1412.3073.

[25] R. Singh, J. H. Bardarson, and F. Pollmann, Signatures of the many-body localization transition in the dynamics of entanglement and bipartite fluctuations, **18**, 023046 (2016).

[26] B. Lu, C. Bertoni, S. J. Thomson, and J. Eisert, Measuring out quasi-local integrals of motion from entanglement, Commun. Phys. **7**, 17 (2024).

[27] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Many-body localization in a quantum simulator with programmable random disorder, Nature Phys. **12**, 907–911 (2016).

[28] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, Exact relaxation in a class of nonequilibrium quantum lattice systems, Phys. Rev. Lett. **100**, 030602 (2008).

[29] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Phys. Rev. E **79**, 061103 (2009).

[30] M. Gluza, J. Eisert, and T. Farrelly, Equilibration towards generalized Gibbs ensembles in non-interacting theories, SciPost Physics **7**, 038 (2019).

[31] F. G. S. L. Brandao and M. Cramer, Equivalence of statistical mechanical ensembles for non-critical quantum systems (2015), arXiv:1502.03263.

[32] L. Foini and J. Kurchan, Eigenstate thermalization hypothesis and out of time order correlators, Phys. Rev. E **99**, 042139 (2019).

[33] S. Pappalardi, L. Foini, and J. Kurchan, Eigenstate thermalization hypothesis and free probability, Phys. Rev. Lett. **129**, 170603 (2022).

[34] A. J. Short and T. C. Farrelly, Quantum equilibration in finite time, **14**, 013063 (2012).

[35] P. Reimann and M. Kastner, Equilibration of isolated macroscopic quantum systems, **14**, 043020 (2012).

[36] H. Wilming, M. Goihl, I. Roth, and J. Eisert, Entanglement-ergodic quantum systems equilibrate exponentially well, Phys. Rev. Lett. **123**, 200604 (2019).

[37] T. Farrelly, F. G. Brandão, and M. Cramer, Thermalization and return to equilibrium on finite quantum lattice systems, Phys. Rev. Lett. **118**, 140601 (2017).

[38] J. Haferkamp, C. Bertoni, I. Roth, and J. Eisert, Emergent statistical mechanics from properties of disordered random matrix product states, PRX Quantum **2**, 040308 (2021).

[39] H. Araki, Gibbs states of a one dimensional quantum lattice, Commun. Math. Phys. **14**, 120 (1969).

[40] M. Kliesch, C. Gogolin, M. Kastoryano, A. Riera, and J. Eisert, Locality of temperature, Phys. Rev. X **4**, 031019 (2014).

[41] K. S. Rai, J. I. Cirac, and A. M. Alhambra, Matrix product state approximations to quantum states of low energy variance (2023), arXiv:2307.05200.

[42] A. Anshu, Concentration bounds for quantum states with finite correlation length on quantum spin lattice systems, New J. Phys. **18**, 083011.

[43] P. Reimann, Transportless equilibration in isolated many-body quantum systems, New J. Phys. **21**, 053014 (2019).

[44] C. Rouzé, D. S. França, and Á. M. Alhambra, Efficient thermalization and universal quantum computing with quantum Gibbs samplers (2024), arXiv:2403.12691.

[45] C.-F. Chen, M. J. Kastoryano, F. G. Brandão, and A. Gilyén, Quantum thermal state preparation (2023), arXiv:2303.18224.

[46] A. Riera, C. Gogolin, and J. Eisert, Thermalization in nature and on a quantum computer, Phys. Rev. Lett. **108**, 080402 (2012).

Supplemental Material

I. EQUIVALENCE OF ENSEMBLES

The main purpose of this section is to prove Theorem 1 of the main text in a more general version. For what follows, we will need two auxiliary lemmas, both stated and proven in Ref. [1], the final result will follow from a similar argument to the one used to prove equivalence of statistical mechanics ensembles in the same work. The first one states that a state close in relative entropy to some other state with exponential decay of correlations will also be locally indistinguishable from that state.

Lemma 1 ([1], Proposition 2)**.** *Let ρ be a state with correlation length ξ and let τ be a state. Let $\mathcal{C}_l$ be the set of hypercubes in Λ with edge length l, if for $\epsilon > 0$*

$$S(\tau||\rho) + 3 + \epsilon^{\frac{D+1}{D+2}} \frac{2\xi \ln(d) l^D + l + 2}{\xi \ln(2)} + \epsilon^{\frac{D+1}{D+2}} \log(N) \leq \epsilon \left(\frac{N}{\ln(4)^D \xi^D} \right)^{\frac{1}{D+1}} \tag{1}$$

then

$$D_l(\tau, \rho) \leq 7\sqrt{\epsilon^{\frac{D+1}{D+2}}}. \tag{2}$$

This theorem has important consequences. In less formal terms, it implies that bounding the relative entropy between the two states bounds their local distinguishability, assuming some bounds on ϵ and the local regions:

Corollary 1 (Local indistinguishability from relative entropy closeness)**.** *Let ρ be a state with correlation length ξ and let τ be a state. Suppose that for some ϵ, we have*

$$S(\tau||\rho) \leq \epsilon N^{\frac{1}{D+1}}. \tag{3}$$

Let

$$\tilde{\epsilon} := \frac{4}{(\xi \ln(4))^{\frac{D}{D+1}}} \epsilon. \tag{4}$$

If

$$\tilde{\epsilon} \geq \left(\frac{\ln(4)^D \xi^D}{N} \right)^{\frac{D+2}{D+1}} (4 \log(N))^{D+2} \tag{5}$$

then for

$$l^D \leq \frac{1}{4} \frac{\xi \ln(2)}{2\xi \ln(d) + 3} \tilde{\epsilon}^{\frac{1}{D+2}} \left(\frac{N}{\xi^D \ln(4)^D} \right)^{\frac{1}{D+1}} \tag{6}$$

we have

$$D_l(\tau, \rho) \leq 7\sqrt{\tilde{\epsilon}^{\frac{D+1}{D+2}}}. \tag{7}$$

The second lemma is a variant of the Berry-Esseen theorem for quantum lattice systems. We denote with $\{|\nu\rangle\}_\nu$ and $\{E_\nu\}_\nu$ the eigenstates and corresponding eigenvalues of a local Hamiltonian H. If a state ρ is uncorrelated, the energy probability distribution $\langle \nu|\rho|\nu\rangle$ behaves like a sum of independent random variables, and converges to a normal distribution with the system size. The Berry-Esseen theorem for quantum lattice systems gives system-size dependent bounds on the speed of convergence of the distribution.

Definition 1 (Berry-Esseen error)**.** *Let ρ be a state and H a local Hamiltonian, define $\mu := \mathrm{tr}(\rho H)$ and $\sigma^2 = \mathrm{tr}\left(\rho(H-\mu)^2\right)$ and let the functions F and G be defined as*

$$F(x) = \sum_{\nu: E_\nu \leq x} \langle \nu|\rho|\nu\rangle, \qquad G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(y-x)^2}{2\sigma^2}}, \tag{8}$$

then the Berry-Esseen error is the supremum of the difference of these two cumulative distribution functions

$$\zeta_N(\rho, H) = \sup_{x \in \mathbb{R}} |F(x) - G(x)|. \tag{9}$$

Lemma 2 ([1], Lemma 8)**.** *Suppose* $g_\beta(H) = \frac{\mathrm{e}^{-\beta H}}{Z(\beta)}$ *has exponential decay of correlations, let* $\sigma^2 = \mathrm{tr}\left(g_\beta(H)H^2\right) - \mathrm{tr}\left(g_\beta(H)H\right)^2$ *and assume* $\sigma^2 \geq \Omega(N)$*. Then*

$$\zeta_N(g_\beta(H), H) \leq R\,\frac{\ln(N)^{2D}}{\sigma} \tag{10}$$

where R *is a constant depending on the correlation length, the locality of the Hamiltonian, and the dimension of the lattice.*

A tighter bound, without the logarithmic factor, holds for Gibbs states at sufficiently high temperature [2]. Throughout the document, we define $\zeta_N := \zeta_N(g_\beta(H), H)$. We are now ready to prove the following statement. This is a more general version of the main theorem of [1] for the case of the GmE and the proof follows from appropriately adapting the argument in that work.

Theorem 1 (Thermality condition)**.** *Let* H *be a local Hamiltonian and let*

$$\begin{aligned} &\frac{1}{6\sqrt{2\pi}}\sigma N^{\frac{-1}{D+1}} \geq \epsilon \geq \max\left\{\log\left(\frac{1}{\tilde{\zeta}_N}\right), \mathrm{e}^{1+\sqrt{5}}\ln(N)^{2D}\right\} N^{\frac{-1}{D+1}}, \\ &3\sqrt{2\pi}\,\mathrm{e}^2\sigma\tilde{\zeta}_N \leq \Delta \leq \sigma\sqrt{\ln\left(\frac{\epsilon N^{\frac{1}{D+1}}}{R\ln(N)^{2D}}\right)}, \end{aligned} \tag{11}$$

for some $\tilde{\zeta}_N \geq \zeta_N$ *such that* $\sigma\tilde{\zeta}_N \leq R\ln(N)^{2D}$*. Let* τ *be a state supported only on eigenstates with energy in* $I = [E+\Delta, E-\Delta]$ *for some energy* E*. Divide the interval* I *into subintervals* $[e_k, e_{k+1})$ *and suppose* $\delta_k = |e_{k+1} - e_k|$ *satisfies*

$$\delta_k^* := 3\sqrt{2\pi}\,\mathrm{e}^{\frac{1}{2}\left(\frac{\Delta(k)}{\sigma}+2\right)^2}\sigma\tilde{\zeta}_N \leq \delta_k \leq \sigma \tag{12}$$

where we have defined

$$\Delta(k) := \begin{cases} |e_k - E| & \text{if } E < e_k, \\ |e_{k+1} - E| & \text{if } E > e_{k+1}, \\ 0 & \text{if } E \in [e_k, e_{k+1}]. \end{cases} \tag{13}$$

Call the projector onto the eigenstates contained in the k*-th window* Π_k*. If for each* k

$$S(\tau) \geq \sum_k \mathrm{tr}\left(\Pi_k\tau\right)\log\left(\mathrm{tr}\left(\Pi_k\right)\right) - \epsilon N^{\frac{1}{D+1}}, \tag{14}$$

then for any l *such that*

$$l^D \leq \frac{1}{4}\left(\frac{4}{(\xi\ln(4))^{\frac{D}{D+1}}}\right)^{\frac{1}{D+2}} \frac{\xi\ln(2)}{2\xi\ln(d)+3}\epsilon^{\frac{1}{D+2}}\left(\frac{N}{\xi^D\ln(4)^D}\right)^{\frac{1}{D+1}}, \tag{15}$$

we have

$$D_l(\rho, g_\beta(H)) \leq 7\left(\frac{4\left(\beta\log(\mathrm{e})3\sqrt{2\pi}+1\right)}{(\xi\ln(4))^{\frac{D}{D+1}}}\right)^{\frac{1}{2}\frac{D+1}{D+2}}\epsilon^{\frac{1}{2}\frac{D+1}{D+2}} \tag{16}$$

for any β *such that* $|E - \mathrm{tr}\left(Hg_\beta(H)\right)| \leq \sigma$ *and* $g_\beta(H)$ *has exponential decay of correlations.*

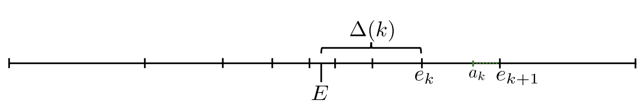

FIG. 1. A sketch of the interval Δ and the sub-intervals δ_k appearing in Theorem 1.

Proof. We would like to apply Lemma 1. We then need to bound the relative entropy $S(\tau||g_\beta(H))$. We begin by writing, by definition of relative entropy (here $\log$ denotes the logarithm in base 2),

$$S(\tau||g_\beta(H)) = -S(\tau) + \beta\log(\mathrm{e})\mathrm{tr}\,(H\tau) + \log(Z(\beta)). \tag{17}$$

In a next step, we have

$$\mathrm{tr}\,(H\tau) = \sum_k \mathrm{tr}\,(\Pi_k H \Pi_k \tau) \leq \sum_k \mathrm{tr}\,(\tau_k)\, e_{k+1}, \tag{18}$$

where we have defined $\tau_k := \Pi_k \tau \Pi_k$ and have used that $\Pi_k H \Pi_k \leq e_{k+1}$. We can then conclude that

$$S(\tau||g_\beta(H)) \leq -S(\tau) + \sum_k \mathrm{tr}\,(\tau_k)\,(\log(\mathrm{e})\beta e_{k+1} + \log(Z(\beta))). \tag{19}$$

We define

$$Z_k(\beta) := \sum_{E_\nu \in [a_k, e_{k+1}]} \mathrm{e}^{-\beta E_\nu} \tag{20}$$

for some $a_k \geq e_k$ to be defined later. Then we write for each k,

$$Z(\beta) = \frac{Z(\beta)}{Z_k(\beta)} Z_k(\beta). \tag{21}$$

At this point, notice that

$$Z_k(\beta) = \sum_{E_\nu \in [a_k, e_{k+1}]} \mathrm{e}^{-\beta E_\nu} \leq d_k \mathrm{e}^{-\beta a_k}, \tag{22}$$

where we have defined $d_k := \mathrm{tr}\,(\Pi_k)$. We then have

$$S(\tau||g_\beta(H)) \leq -S(\tau) + \sum_k \mathrm{tr}\,(\tau_k)\log(d_k) + \sum_k \mathrm{tr}\,(\tau_k)\log\left(\frac{Z(\beta)}{Z_k(\beta)}\right) + \beta\log(\mathrm{e})\sum_k \mathrm{tr}\,(\tau_k)\,\tilde{\delta}_k. \tag{23}$$

with $\tilde{\delta}_k := e_{k+1} - a_k$. We now need to show that $Z(\beta)/Z_k(\beta)$ is not too large for an appropriate choice of intervals. To this end, we write

$$\frac{Z_k(\beta)}{Z(\beta)} = \sum_{E_\nu \in [a_k, e_{k+1}]} \langle \nu | g_\beta(H) | \nu \rangle. \tag{24}$$

We proceed by using that the energy distribution of a thermal state is close to a Gaussian (Lemma 2). Define the functions F and G as

$$F(x) := \sum_{E_\nu \leq x} \langle \nu | g_\beta(H) | \nu \rangle, \qquad G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \mathrm{d}x\, \mathrm{e}^{-\frac{(x-E_\beta)^2}{2\sigma^2}} \tag{25}$$

with $E_\beta := \mathrm{tr}\,(H g_\beta(H))$. Then

$$\begin{aligned} \frac{Z_k(\beta)}{Z(\beta)} &= F(e_{k+1}) - F(a_k) \geq G(e_{k+1}) - G(a_k) - 2\mathrm{sup}|F(x) - G(x)| \\ &\geq G(e_{k+1}) - G(a_k) - 2\zeta_N. \end{aligned} \tag{26}$$

Recall that we wish to lower bound $Z_k(\beta)/Z(\beta)$. If $\tilde{\delta}_k$ is too small, $G(e_{k+1}) - G(a_k)$ will also be small and the lower bound in the above equation will be negative, hence we need to pick $\tilde{\delta}_k$ large enough such that $G(e_{k+1}) - G(a_k) \geq 3\tilde{\zeta}_N$ for some $\tilde{\zeta}_N \geq \zeta_N$ such that $\sigma\tilde{\zeta}_N \leq R\ln(N)^{2D}$, which will yield $Z_k(\beta)/Z(\beta) \geq \tilde{\zeta}_N$. The reason why we introduce $\tilde{\zeta}_N$ is that if ζ_N itself is too small, the lower bound $Z_k(\beta)/Z(\beta) \geq \zeta_N$ might be too loose later. We have

$$G(e_{k+1}) - G(a_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{a_k}^{e_{k+1}} \mathrm{d}x\, \mathrm{e}^{-\frac{(x-E_\beta)^2}{2\sigma^2}} \geq \tilde{\delta}_k \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\max_{x\in[a_k, e_{k+1}]} \frac{(x-E_\beta)^2}{2\sigma^2}\right). \tag{27}$$

Hence, in order for $G(e_{k+1}) - G(a_k) \geq 3\tilde{\zeta}_N$ we need to pick a_k such that

$$\tilde{\delta}_k \geq 3\sqrt{2\pi}\sigma\tilde{\zeta}_N \exp\left(\max_{x\in[a_k,e_{k+1}]} \frac{(x-E_\beta)^2}{2\sigma^2}\right). \tag{28}$$

We have

$$\frac{|x-E_\beta|}{\sqrt{2}\sigma} \leq \frac{|x-E|+|E-E_\beta|}{\sqrt{2}\sigma} \leq \frac{1}{\sqrt{2}}\left(2+\frac{\Delta(k)}{\sigma}\right) \tag{29}$$

where we have used that by assumption $|E - E_\beta| \leq \sigma$ and that since $x \in [a_k, e_{k+1}]$ we have

$$|x-E| \leq \Delta(k) + \delta_k \leq \Delta(k) + \sigma, \tag{30}$$

where we have imposed $\delta_k \leq \sigma$. We set

$$\tilde{\delta}_k = \delta_k^* := 3\sqrt{2\pi}\sigma\tilde{\zeta}_N \exp\left(\frac{1}{2}\left(\frac{\Delta(k)}{\sigma}+2\right)^2\right). \tag{31}$$

We have $G(e_{k+1}) - G(a_k) \geq 3\tilde{\zeta}_N$, and naturally, since $\tilde{\delta}_k \leq \delta_k$, in order to be able to choose $\tilde{\delta}_k = \delta_k^*$, we need $\delta_k \geq \delta_k^*$. Hence if

$$\delta_k^* \leq \delta_k \leq \sigma \tag{32}$$

we have

$$\frac{Z(\beta)}{Z_k(\beta)} \leq \frac{1}{\tilde{\zeta}_N}. \tag{33}$$

We need to ensure that the ranges for δ_k exist. Notice that since by definition $\Delta(k) \leq \Delta$

$$\frac{\Delta(k)}{\sigma} + 2 \leq \sqrt{\ln\left(\frac{\epsilon N^{\frac{1}{D+1}}}{R\ln(N)^{2D}}\right)} + 2 \leq \sqrt{2\ln\left(\frac{\epsilon N^{\frac{1}{D+1}}}{R\ln(N)^{2D}}\right)}, \tag{34}$$

where we for the last inequality we used

$$\epsilon \geq \mathrm{e}^{1+\sqrt{5}}\ln(N)^{2D}N^{\frac{-1}{D+1}}. \tag{35}$$

Hence,

$$\delta_k^* \leq 3\sqrt{2\pi}\epsilon N^{\frac{1}{D+1}} \frac{\sigma\tilde{\zeta}_N}{R\ln(N)^{2D}} \leq \frac{1}{2}\sigma \tag{36}$$

where we used the assumed upper bound on ϵ. Hence if for some k we have $\Delta(k) + \delta_k \geq \Delta$, that is, the lower bound on the next interval requires to leave I, we can simply redefine δ_{k-1} to extend all the way to the edge of I. We can now go back to equation Eq. (23) and we have

$$S(\tau||g_\beta(H)) \leq -S(\tau) + \sum_k \mathrm{tr}\left(\tau_k\right)\log(d_k) + \beta\log(\mathrm{e})3\sqrt{2\pi}\sigma\tilde{\zeta}_N\epsilon N^{\frac{1}{D+1}} + \log\left(\frac{1}{\tilde{\zeta}_N}\right). \tag{37}$$

Furthermore, we use

$$\epsilon \geq \log\left(\frac{1}{\tilde{\zeta}_N}\right)N^{\frac{-1}{D+1}} \tag{38}$$

which ensures that $\log\left(1/\tilde{\zeta}_N\right) \leq \epsilon N^{\frac{1}{D+1}}$, hence overall,

$$S(\tau||g_\beta(H)) \leq -S(\tau) + \sum_k \mathrm{tr}\left(\tau_k\right)\log(d_k) + \left(\beta\log(\mathrm{e})3\sqrt{2\pi}\sigma\tilde{\zeta}_N + 1\right)\epsilon N^{\frac{1}{D+1}}. \tag{39}$$

If we impose

$$\sum_k \mathrm{tr}\,(\tau_k)\log(d_k) - S(\tau) \leq \epsilon N^{\frac{1}{D+1}}, \tag{40}$$

we have that

$$S(\tau||g_\beta(H)) \leq (\beta\log(\mathrm{e})3\sqrt{2\pi}\sigma\tilde{\zeta}_N + 2)\epsilon N^{\frac{1}{D+1}} \leq (\beta\log(\mathrm{e})3\sqrt{2\pi} + 2)\epsilon N^{\frac{1}{D+1}}, \tag{41}$$

where we have used $\sigma\tilde{\zeta}_N \leq R\ln(N)^{2D}$. The result then follows from Corollary 1. □

Theorem 1 and its approximate version (Eq. (11) in the main text) follow from the following more general statement.

Theorem 2 (Ensemble equivalence)**.** *Let H be a local Hamiltonian and β be an inverse temperature for which the Gibbs state $g_\beta(H)$ has exponential decay of correlations and standard deviation $\sigma \geq \Omega(\sqrt{N})$. Suppose $\sigma\zeta_N \leq R\ln(N)^{2D}N^{-\kappa}$ for $B,\kappa \geq 0$ constant. Let ω denote an approximate GmE state with Δ,δ,η satisfying*

$$\begin{aligned}
&\mathrm{e}^{\Delta^2/\sigma^2} \leq \frac{N^{\frac{1-\alpha}{D+1}}}{R\ln(N)^{2D}},\\
&3\sqrt{2\pi}N^{\frac{1-\alpha}{D+1}-\kappa} \leq \delta \leq \sigma,\\
&\eta \leq N^{\frac{1-\alpha}{D+1}}
\end{aligned} \tag{42}$$

with $\alpha \in [0,1)$, such that $|E - E_\beta| \leq \sigma$. Then for any side length l such that $l^D \geq C_1 N^{\frac{1}{D+1}-\gamma_1\alpha}$, the following holds

$$D_l(\omega, g_\beta(H)) \leq C_2 N^{-\gamma_2\alpha} + 2(1-p_\Delta), \tag{43}$$

with C_1, C_2 being system-size independent constants, and γ_1,γ_2 only depend on the dimension of the lattice D.

Proof. Choose

$$\epsilon := N^{-\frac{\alpha}{D+1}} \tag{44}$$

with $\alpha < 1$ constant and set $\tilde{\zeta}_N := R\ln(N)^{2D}N^{-\kappa}/\sigma \geq \zeta_N$. Then we have

$$\Delta \leq \sigma\sqrt{\ln\left(\frac{N^{\frac{1-\alpha}{D+1}}}{R\ln(N)^{2D}}\right)} = \sigma\sqrt{\ln\left(\frac{\epsilon N^{\frac{1}{D+1}}}{R\ln(N)^{2D}}\right)}. \tag{45}$$

We would like to apply Theorem 1. For this, we set $\delta_k := \delta$, hence we need $\delta \geq \delta_k^*$ for all k. As shown in the proof of Theorem 1, for N sufficiently large we have by using $\Delta(k) \leq \Delta$

$$\delta_k^* \leq 3\sqrt{2\pi}\epsilon N^{\frac{1}{D+1}}\frac{\sigma\tilde{\zeta}_N}{R\ln(N)^{2D}} = 3\sqrt{2\pi}N^{\frac{1-\alpha}{D+1}-\kappa} \tag{46}$$

Then, for all k, choose

$$\delta = 3\sqrt{2\pi}N^{\frac{1-\alpha}{D+1}-\kappa}. \tag{47}$$

The subdivision then satisfies the assumptions of Theorem 1 for the chosen ϵ. Let $\tau = \sum_{k=1}^K q_k\tilde{\omega}_{\delta_k}$ such that $\omega = p_\Delta\tau + (1 - p_\Delta)\rho_{\mathrm{tail}}$. We have

$$\|\omega - \tau\|_1 = \|\omega - p_\Delta\tau + p_\Delta\tau - \tau\|_1 \leq 2(1-p_\Delta). \tag{48}$$

We then have

$$D_l(\omega, g_\beta(H)) \leq D_l(\tau, g_\beta(H)) + 2(1-p_\Delta) \tag{49}$$

Furthermore, by the concavity of the Von Neumann entropy

$$S(\tau) \geq \sum_k q_k S(\tilde{\omega}_{\delta_k}) \geq \sum_k q_k S(\omega_{\delta_k}) - N^{\frac{1-\alpha}{D+1}}. \tag{50}$$

Notice that $\mathrm{tr}\,(\Pi_k\tau) = q_k$ and $S(\omega_{\delta_k}) = \log(d_k)$. By Theorem 1,

$$D_l(\tau, g_\beta(H)) \leq C_2 N^{-\gamma_2\alpha}, \tag{51}$$

for C_2, γ_2 constants. □

Theorem 1 in the main text is the special case where $p_\Delta = 1$ and $\tilde{\omega}_{\delta_k} = \omega_{\delta_k}$.

II. THERMALIZATION UNDER TYPICAL HAMILTONIANS

We now apply our thermality condition to prove thermalization with high probability of appropriate states under a suitable slightly perturbed Hamiltonian. Before proving the theorems in the main text, we discuss tails decay for states with exponential decay of correlations. The following has been proven in Ref. [3].

Theorem 3 (Spectral tail bound)**.** *Let ρ be a state and H be a local Hamiltonian. Let Π_Δ be the projector onto all eigenstates of H with energy E such that $|E - \operatorname{tr}(\rho H)| \leq \Delta$. Then,*

1. *if ρ is a product state, and $\Delta \geq g_0\sqrt{N}$*

$$\operatorname{tr}((1-\Pi_\Delta)\rho) \leq \mathrm{e}^{-s_0\frac{\Delta^2}{N}}, \tag{52}$$

where g_0, s_0 are system size independent constants.

2. *if ρ has exponential decay of correlations, and $\Delta \geq g_\xi\sqrt{N}$*

$$\operatorname{tr}((1-\Pi_\Delta)\rho) \leq \mathrm{e}^{-s_\xi\left(\frac{\Delta^2}{N}\right)^{\frac{1}{D+1}}} \tag{53}$$

where g_ξ, s_ξ are system size independent constants depending on the correlation length ξ.

We repeat here the definition of the ensemble $\mathcal{E}(\delta)$: we divide the energy spectrum into intervals $\{I_k\}_{k=1}^K$ of equal energy width δ, the eigenstates in the k-th window span a vector space which we call $\mathcal{W}_k$. $\mathcal{E}(\delta)$ is then the ensemble of random unitaries of the form

$$U = \bigoplus_{k=1}^K U_k, \tag{54}$$

where U_k is drawn from the Haar measure on the unitary group acting on $\mathcal{W}_k$ independently for every k. Generally, we will denote by Π_k the projector onto $\mathcal{W}_k$, by W_k the set of indices ν such that $E_\nu \in I_k$, and by $d_k := |W_k|$ the number of eigenstates contained in each window. We will need to compute simple moments of Haar random unitaries. For an introduction we refer the reader to Ref. [4]. For convenience, let us prove a fact that will be useful later.

Lemma 3 (Perturbing Hamiltonians)**.** *Let U be drawn from $\mathcal{E}(\delta)$. Then*

$$\|H - UHU^\dagger\|_\infty \leq \delta. \tag{55}$$

Proof. Since both H and U commute with Π_k, we have

$$(UHU^\dagger - H)^2 = \sum_{k=1}^K \Pi_k (UHU^\dagger - H)^2 \Pi_k = \sum_{k=1}^K (UH_kU^\dagger - H_k)^2 \tag{56}$$

with $H_k := \Pi_k H \Pi_k$. Let ρ be a state and define $\rho_k := \Pi_k \rho \Pi_k$. We then have

$$\operatorname{tr}\left(\rho(UHU^\dagger - H)^2\right) = \sum_{k=1}^K \operatorname{tr}\left(\rho_k (UH_kU^\dagger - H_k)^2\right) \leq \sum_{k=1}^K \operatorname{tr}(\rho_k)\, \|UH_kU^\dagger - H_k\|_\infty^2, \tag{57}$$

where we have used Hölder's inequality and the fact that since $\rho_k > 0$, $\|\rho_k\|_1 = \operatorname{tr}(\rho_k)$. Define

$$m_k := e_k + \frac{\delta}{2} = \frac{e_k + e_{k+1}}{2}, \tag{58}$$

then

$$\|UH_kU^\dagger - H_k\|_\infty = \|UH_kU^\dagger - m_k\Pi_k + m_k\Pi_k - H_k\|_\infty \leq 2\|H_k - m_k\Pi_k\|_\infty \leq \delta^2. \tag{59}$$

Hence,

$$\operatorname{tr}\left(\rho(UHU^\dagger - H)^2\right) \leq \delta^2, \tag{60}$$

which shows $\|UHU^\dagger - H\|_\infty \leq \delta$ by the definition of operator norm. □

First, we prove Lemma 1 in the main text in the more general case where eigenvalues can be degenerate.

Lemma 4 (Approximate GmE at equilibrium)**.** *Let ρ be a product state and H be a local Hamiltonian. Let*

$$\rho_\infty^{UHU^\dagger} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{e}^{-iUHU^\dagger t}\rho \mathrm{e}^{iUHU^\dagger t} \tag{61}$$

where U is drawn from $\mathcal{E}(\delta)$. Consider the interval $I = [E-\Delta, E+\Delta]$ around $E = \mathrm{tr}\,(\rho H)$ with $\Delta \geq \omega(\sqrt{N})$ an integer multiple of δ, then

$$\rho_\infty^{UHU^\dagger} = p_\Delta \left(\sum_{k:I_k\subset I} q_k \tilde{\omega}_{\delta_k}\right) + (1-p_\Delta)\rho_{\mathrm{tail}} \tag{62}$$

with $p_\Delta \geq 1 - \mathrm{e}^{-c_1\frac{\Delta^2}{N}}$ and $\sum_{k:I_k\subset I} q_k = 1$. Moreover, with probability at least $1-(\mathcal{D}+1)2^{-r}$, where $\mathcal{D}$ is the degeneracy of the most degenerate eigenvalue with energy inside I, we have

$$\sum_{k:I_k\subset I} q_k(S(\omega_{\delta_k}) - S(\tilde{\omega}_{\delta_k})) \leq r, \tag{63}$$

where c_1 is a system-size independent constant.

Proof. We have

$$\rho_\infty^{UHU^\dagger} = \sum_\nu UP_\nu U^\dagger \rho U P_\nu U^\dagger. \tag{64}$$

Notice that if $E_\nu \in I_k$, then P_ν is a projector in $\mathcal{W}_k$, and since the unitaries U preserve the spaces $\mathcal{W}_k$, $UP_\nu U^\dagger$ is also a projector in $\mathcal{W}_k$. Let W_k be the set of indices ν such that $E_\nu \in I_k$. Then

$$\rho_\infty^{UHU^\dagger} = \sum_{k:I_k\subset I}\sum_{\nu\in W_k} UP_\nu U^\dagger\rho UP_\nu U^\dagger + \sum_{k:I_k\not\subset I}\sum_{\nu\in W_k} UP_\nu U^\dagger\rho UP_\nu U^\dagger =: p_\Delta \sum_{k:I_k\subset I} \underbrace{\frac{p_k}{p_\Delta}}_{=:q_k} \tilde{\omega}_{\delta_k} + (1-p_\Delta)\rho_{\mathrm{tail}} \tag{65}$$

with

$$p_\Delta = \mathrm{tr}\left(\Pi_\Delta \rho_\infty^{UHU^\dagger}\right), \tag{66}$$

$$p_k = \sum_{\nu\in W_k} \mathrm{tr}\left(\rho UP_\nu U^\dagger\right) = \mathrm{tr}\,(\rho\Pi_k), \tag{67}$$

$$\tilde{\omega}_{\delta_k} = \frac{1}{p_k}\sum_{\nu\in W_k} UP_\nu U^\dagger\rho UP_\nu U^\dagger, \tag{68}$$

$$\rho_{\mathrm{tail}} = \frac{1}{1-p_\Delta}\sum_{k:I_k\not\subset I}\sum_{\nu\in W_k} UP_\nu U^\dagger\rho UP_\nu U^\dagger. \tag{69}$$

To proceed, we start by bounding the entropy. We have

$$S(\tilde{\omega}_{\delta_k}) \geq S_2(\tilde{\omega}_{\delta_k}) = -\log\left(\mathrm{tr}\left(\tilde{\omega}_{\delta_k}^2\right)\right) \tag{70}$$

and

$$\mathrm{tr}\left(\tilde{\omega}_{\delta_k}^2\right) = \frac{1}{p_k^2}\sum_{\nu\in W_k}\mathrm{tr}\left((UP_\nu U^\dagger\rho)^2\right). \tag{71}$$

Recall that U is of the form $\bigoplus_k U_k$ where each U_k is drawn from the Haar measure on $\mathcal{W}_k$. Then a standard computation yields

$$\begin{aligned}\mathbb{E}_{U\sim\mathcal{E}(\delta)}\left(\mathrm{tr}\left(\tilde{\omega}_{\delta_k}^2\right)\right) &= \frac{1}{p_k^2}\sum_{\nu\in W_k}\frac{\mathrm{tr}\,(P_\nu)}{d_k^2-1}\left[\mathrm{tr}\left((\Pi_k\rho)^2\right)\left(\mathrm{tr}\,(P_\nu) - \frac{1}{d_k}\right) + \mathrm{tr}\,(\Pi_k\rho)^2\left(1 - \frac{\mathrm{tr}\,(P_\nu)}{d_k}\right)\right] \\ &\leq \frac{\mathrm{tr}\,(\Pi_k\rho)^2}{p_k^2}\sum_{\nu\in W_k}\frac{\mathrm{tr}\,(P_\nu)\,(1+\mathrm{tr}\,(P_\nu))}{d_k(d_k+1)} \leq \frac{1}{d_k+1}(1+\max{}_\nu \mathrm{tr}\,(P_\nu)) =: \frac{D_k+1}{d_k+1},\end{aligned} \tag{72}$$

where D_k is the degeneracy of the maximally degenerate eigenvalue inside W_k. By Markov's inequality

$$\mathrm{Pr}_{U\sim\mathcal{E}(\delta)}\Big[\log(d_k) - S(\tilde{\omega}_{\delta_k}) \geq r\Big] \leq \mathrm{Pr}_{U\sim\mathcal{E}(\delta)}\Big[\mathrm{tr}\left(\tilde{\omega}_{\delta_k}^2\right) \geq \frac{1}{d_k}2^r\Big] \leq \frac{(D_k+1)d_k}{d_k+1}2^{-r} \leq (D_k+1)2^{-r}. \tag{73}$$

Noticing that $\log(d_k) = S(\omega_{\delta_k})$ yields the result. Finally, we need to bound p_Δ. Notice that

$$p_\Delta = \mathrm{tr}\left(\Pi_\Delta \rho_\infty^{UHU^\dagger}\right) = \sum_{k:I_k\subset I}\sum_{\nu\in W_k}\mathrm{tr}\left(\rho U P_\nu U^\dagger\right) = \mathrm{tr}\left(\Pi_\Delta\rho\right). \tag{74}$$

We now wish to bound p_Δ using Theorem 3. By construction if $E_\nu \notin I$, then

$$|\mathrm{tr}\left(\rho H\right) - E_\nu| \geq \Delta \geq \omega(\sqrt{N}). \tag{75}$$

Then, by Theorem 3, we have

$$p_\Delta \geq 1 - \mathrm{e}^{-g_0\frac{\Delta^2}{N}} \tag{76}$$

if ρ is a product state and

$$p_\Delta \geq 1 - \mathrm{e}^{-g_\xi\left(\frac{\Delta^2}{N}\right)^{\frac{1}{D+1}}} \tag{77}$$

if ρ has exponential decay of correlations. $\square$

Now the following version of Theorem 2 in the main text follows from Theorem 2 and Lemma 4.

Theorem 4 (Typical thermalization)**.** *Let H be a k-local Hamiltonian and ρ be a product state. Let $g_\beta(H)$ be the Gibbs state of H at inverse temperature β such that $|\mathrm{tr}\left(g_\beta(H)H\right) - \mathrm{tr}\left(\rho H\right)| \leq \sigma$. Assume $g_\beta(H)$ has exponential decay of correlations and $\sigma\zeta_N \leq R\ln(N)^{2D}N^{-\kappa}$ for $B,\kappa\geq 0$ constants. For any constant $\alpha\in[0,1)$, if $\delta = 3\sqrt{2\pi}N^{\frac{1-\alpha}{D+1}-\kappa}$, then with probability at least $1-(\mathcal{D}+1)\exp(-c_2N^{\frac{1-\alpha}{D+1}})$ drawing U at random from $\mathcal{E}(\delta)$, we have*

$$D_l(\rho_\infty^{UHU^\dagger}, g_\beta(H)) \leq C_2\,N^{-\gamma_2\alpha} + N^{-\gamma_3(1-\alpha)}\ln(N)^{2D}, \tag{78}$$

where $c_2, C_2, \gamma_2, \gamma_3$ are system-size independent constants.

Proof. Let Δ such that

$$\mathrm{e}^{\Delta^2/\sigma^2} = \frac{N^{\frac{1-\alpha}{D+1}}}{R\ln(N)^{2D}}. \tag{79}$$

By applying Lemma 4 with $r := N^{\frac{1-\alpha}{D+1}}$ we have that

$$\rho_\infty^{UHU^\dagger} = p_\Delta\tau + (1-p_\Delta)\rho_{\mathrm{tail}} \tag{80}$$

is approximately GmE satisfying the conditions of Theorem 2, hence we have

$$D_l(\rho_\infty^{UHU^\dagger}, g_\beta(H)) \leq C_2\,N^{-\gamma_2\alpha} + 2(1-p_\Delta) \leq C_2\,N^{-\gamma_2\alpha} + N^{-\gamma_3(1-\alpha)}\ln(N)^{2D}. \tag{81}$$

$\square$

A similar theorem is immediate for states with exponential decay of correlations, but in this case

$$1 - p_\Delta \leq \exp\left(-\Omega\left(\ln\left(\frac{N^{\frac{1-\alpha}{D+1}}}{\ln(N)^{2D}}\right)\right)^{\frac{1}{D+1}}\right). \tag{82}$$

Equilibration also immediately follows from Lemma 4, as a matter of fact for any observable A in the case of non-degenerate energy gaps, that is if $E_\mu - E_\nu = E_{\mu'} - E_{\nu'}$ implies $(\mu,\nu) = (\mu',\nu')$. We have [5]

$$\begin{aligned}\Delta A_\infty &:= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\,\mathrm{tr}\left(A(\rho^{UHU^\dagger}(t) - \rho_\infty^{UHU^\dagger})\right)^2 \leq \|A\|_\infty^2\mathrm{tr}\left(\left(\rho_\infty^{UHU^\dagger}\right)^2\right) \leq \|A\|_\infty^2\Big((1-p_\Delta)^2 + \sum_{k:I_k\subset I}q_k^2\mathrm{tr}\left(\tilde{\omega}_{\delta_k}^2\right)\Big) \\ &\leq \|A\|_\infty^2\Big((1-p_\Delta)^2 + 2^r\sum_{k:I_k\subset I}q_k^2\frac{1}{d_k}\Big) \leq \|A\|_\infty^2\Big((1-p_\Delta)^2 + 2^r\max{}_k\frac{1}{d_k}\Big).\end{aligned} \tag{83}$$

The condition of non-degenerate gaps is expected to hold generically for interacting systems. Then for example taking $r = \min_k \log(\sqrt{d_k})$ we have that with probability at least $1-(\mathcal{D}+1)\max_k \frac{1}{\sqrt{d_k}}$, $\Delta A_\infty \leq \|A\|_\infty^2 \Big((1-p_\Delta)^2 + 2^r \max_k \frac{1}{\sqrt{d_k}}\Big)$. Since d_k is expected to be exponentially small, the system equilibrates. In addition, we have the following statement.

Theorem 5 (Short time evolution)**.** *For any state ρ and $U \sim \mathcal{E}(\delta)$ and any $t \geq 0$*

$$\|\mathrm{e}^{-iHt}\rho\mathrm{e}^{iHt} - \mathrm{e}^{-iUHU^\dagger t}\rho\mathrm{e}^{iUHU^\dagger t}\|_1 \leq 2t\delta. \tag{84}$$

Proof. Let $V = \mathrm{e}^{-iHt}$ and $\tilde{V} = \mathrm{e}^{-iUHU^\dagger t}$. Then

$$\|V\rho V^\dagger - \tilde{V}\rho\tilde{V}^\dagger\|_1 = \|\tilde{V}^\dagger V\rho V^\dagger\tilde{V} - \rho\|_1 = \|\tilde{V}^\dagger V[\rho, V^\dagger\tilde{V}]\|_1 = \|[\rho, V^\dagger\tilde{V} - \mathbb{1}]\|_1 \leq 2\|\rho\|_1\|V - \tilde{V}\|_\infty \tag{85}$$

where we have used Hölder's inequality and standard properties of the trace norm. Let $X = UHU^\dagger - H$, then

$$\|V - \tilde{V}\|_\infty = \|\mathrm{e}^{iHt} - \mathrm{e}^{iHt+iXt}\|_\infty. \tag{86}$$

Now let the function f be $f(s) = \mathrm{e}^{iHt+iXts}$, then

$$\|V - \tilde{V}\|_\infty = \|f(1) - f(0)\|_\infty \leq \int_0^1 \mathrm{d}s\|f'(s)\|_\infty. \tag{87}$$

Then, using

$$f'(s) = it\int_0^1 \mathrm{d}r\, \mathrm{e}^{r(iHt+iXts)}X\mathrm{e}^{(1-r)(iHt+iXts)}, \tag{88}$$

we get

$$\|f'(s)\|_\infty \leq t\|X\|_\infty \tag{89}$$

and hence by Lemma 3

$$\|V - \tilde{V}\|_\infty \leq t\|X\|_\infty = t\|H - UHU^\dagger\|_\infty \leq t\delta, \tag{90}$$

together with $\|\rho\|_1 = 1$ this proves the result. ☐

Choosing

$$\delta = \Omega(N^{\frac{1-\alpha}{D+1}-\kappa}), \tag{91}$$

if the Berry-Esseen error is bounded by $\tilde{O}(N^{-1/2-\kappa})$, we have, again choosing $\alpha \geq 1-\kappa(D+1)$, that δ is a decreasing function of the system size, and the two dynamics are indistinguishable up to ϵ until a time $\sim \epsilon N^{\kappa-\frac{1-\alpha}{D+1}}$. We now move on to proving that the Gibbs states of H and $UHU^\dagger$ are locally indistinguishable.

Theorem 6 (Local indistinguishability of perturbed Gibbs states)**.** *Let H be a k-local Hamiltonian. Let $g_\beta(H)$ be the Gibbs state of H at an inverse temperature β such that $g_\beta(H)$ has exponential decay of correlations. For any U drawn from $\mathcal{E}(\delta)$ and l, κ, α as in Theorem 2 we have*

$$D_l(g_\beta(H), g_\beta(UHU^\dagger)) \leq C_3\, N^{-\gamma_4\alpha-\gamma_5\kappa}, \tag{92}$$

for system-size independent constants C_3, γ_4, γ_5.

Proof. A simple computation reveals

$$S(g_\beta(UHU^\dagger)||g_\beta(H)) = \beta\mathrm{tr}\left((UHU^\dagger - H)g_\beta(H)\right) \leq \beta\|H - UHU^\dagger\|_\infty \leq \beta\delta \tag{93}$$

where we have used Lemma 3. The result follows from Corollary 1, with $\epsilon = \beta\delta N^{-\frac{1}{D+1}} = \beta N^{-\frac{\alpha}{D+1}-\kappa}$, which ends the proof. ☐

III. RELAXATION DYNAMICS

We previously discussed the dynamics in the short time regime, we now turn to discussing the typical late-time relaxation dynamics to the thermal state following an non-equilibrium initial preparation. We abbreviate in the following $\rho_U(t) := \mathrm{e}^{-iUHU^\dagger t}\rho\mathrm{e}^{iUHU^\dagger t}$. To prove the statement in the main text, we write

$$\langle A\rangle_{\rho_U(t)} - \langle A\rangle_{\rho_\infty^{UHU^\dagger}} = \langle A\rangle_{\rho_U(t)} - \mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right) + \mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right) - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right) + \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right) - \langle A\rangle_{\rho_\infty^{UHU^\dagger}} \tag{94}$$

where the expectation values are taken over $U \sim \mathcal{E}(\delta)$. Hence

$$\left|\langle A\rangle_{\rho_U(t)} - \langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right| \leq \left|\langle A\rangle_{\rho_U(t)} - \mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right)\right| + \left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| + R(t), \tag{95}$$

where we have defined the function R to be

$$R(t) := \left|\mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right) - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right|. \tag{96}$$

We now need to show that the first two terms are small with high probability, and the promised decay of $R(t)$ with the relevant spectral assumptions.

Lemma 5 (Time-evolution concentration)**.** *For a Hamiltonian H, an observable A, and some initial state ρ it holds that*

$$\Pr_{U\sim\mathcal{E}(\delta)}\left[\left|\langle A\rangle_{\rho_U(t)} - \mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right)\right| + \left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| \geq \epsilon\right] \leq 4\sum_k \mathrm{e}^{\frac{-C(d_k-1)\epsilon^2}{\|A\|_\infty^2}}, \tag{97}$$

with constant $C > 0$.

Above, the sum is over the energy windows and d_k is the number of eigenstates in the k-th window. Notice that while d_k might become smaller towards the edges of the spectrum, it is still expected to grow with the system size, and if the state ρ is peaked around its average (e.g., a state with exponential decay of correlations) its tails at the edges of the spectrum can be cut without affecting the result.

Proof. Via a union bound, we first get

$$\begin{aligned}\Pr_{U\sim\mathcal{E}(\delta)}&\left[\left|\langle A\rangle_{\rho_U(t)} - \mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right)\right| + \left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| \geq \epsilon\right] \\ &\leq \Pr_{U\sim\mathcal{E}(\delta)}\left[\left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| \geq \epsilon/2\right] + \Pr_{U\sim\mathcal{E}(\delta)}\left[\left|\langle A\rangle_{\rho_U(t)} - \mathbb{E}\left(\langle A\rangle_{\rho_U(t)}\right)\right| \geq \epsilon/2\right].\end{aligned} \tag{98}$$

To upper bound the two probabilities we will make use of a variant of Levy's lemma. Levy's lemma gives a probability concentration bound for Lipschitz functions. A function $f : S \to \mathbb{R}$ on the metric space (S, m) comprised of a space S equipped with a metric m is called Lipschitz with Lipschitz-constant L if

$$L := \sup_{x,y\in S} \frac{|f(x) - f(y)|}{m(x,y)} < \infty. \tag{99}$$

Denote $\mathcal{U}(n)$ the unitary group on a vector space of dimension n and consider the metric space (S, m) constructed from a direct product of finitely many metric spaces $(\mathcal{U}(n_k), m_k = \|\cdot\|_2)$. It has norm defined by $m = \sum_k m_k$.

Lemma 6 (Levy's Lemma)**.** *Denote with μ_k the Haar measure on $\mathcal{U}(n_k)$. For any L-Lipschitz function f on (S, m) it holds that*

$$\Pr_{x\sim\bigotimes_k \mu_k}\left[\left|f(x) - \mathbb{E}[f(x)]\right| \geq \epsilon\right] \leq 2\sum_k \mathrm{e}^{\frac{-c(n_k-1)\epsilon^2}{L^2}}, \tag{100}$$

with constant $c > 0$.

Proof. By Ref. [6, Proposition 1.3 and Theorem 1.7] as well as by applying Ref. [6, Theorem 2.4] to the unitary group, it holds that on the metric space $(\mathcal{U}(n_k), m_k = \|\cdot\|_2)$ equipped with the uniform Haar measure μ_k any L-Lipschitz function f satisfies

$$\Pr_{x\sim\mu_k}\left[\left|f(x) - \mathbb{E}[f(x)]\right| \geq \epsilon\right] \leq 2\mathrm{e}^{\frac{-c(n_k-1)\epsilon^2}{L^2}}, \tag{101}$$

for some constant $c > 0$. Subsequently applying [6, Proposition 1.11] proves Lemma 6. ☐

All we have to do is to compute the Lipschitz constants L_f and L_g of $f = \langle A\rangle_{\rho_U(t)}$ and $g = \langle A\rangle_{\rho_\infty^{UHU^\dagger}}$, respectively, with respect to the metric $m(U,V) = \sum_k \|U_k - V_k\|_2$, where the sum runs over all binnings in our ensemble $\mathcal{E}(\delta)$. We have

$$\begin{aligned} |\langle A\rangle_{\rho_U(t)} - \langle A\rangle_{\rho_V(t)}| &= |\mathrm{tr}\left(A\mathrm{e}^{-\mathrm{i}tUHU^\dagger}\rho\mathrm{e}^{\mathrm{i}tUHU^\dagger}\right) - \mathrm{tr}\left(A\mathrm{e}^{-\mathrm{i}tVHV^\dagger}\rho\mathrm{e}^{\mathrm{i}tVHV^\dagger}\right)| \\ &=: |\mathrm{tr}\left(A_U\mathrm{e}^{-\mathrm{i}tH}\rho_U\mathrm{e}^{\mathrm{i}tH}\right) - \mathrm{tr}\left(A_V\mathrm{e}^{-\mathrm{i}tH}\rho_V\mathrm{e}^{\mathrm{i}tH}\right)| \\ &\le |\mathrm{tr}\left(A_U\mathrm{e}^{-\mathrm{i}tH}(\rho_U - \rho_V)\mathrm{e}^{\mathrm{i}tH}\right)| + |\mathrm{tr}\left((A_V - A_U)\mathrm{e}^{-\mathrm{i}tH}\rho_V\mathrm{e}^{\mathrm{i}tH}\right)| \\ &\le \|A\|_\infty\|\rho_U - \rho_V\|_1 + \|A_U - A_V\|_\infty\|\rho\|_1. \end{aligned} \tag{102}$$

Moreover, for any Hermitian operator X and $p \ge 1$,

$$\|UXU^\dagger - VXV^\dagger\|_p \le \|UXU^\dagger - UXV^\dagger\|_p + \|UXV^\dagger - VXV^\dagger\|_p = 2\|(U-V)X\|_p \tag{103}$$

and hence

$$\begin{aligned} |\langle A\rangle_{\rho_U(t)} - \langle A\rangle_{\rho_V(t)}| &\le 2\|A\|_\infty(\|(U-V)\rho\|_1 + \|U-V\|_\infty) \\ &\le 4\|A\|_\infty\|U-V\|_\infty \\ &\le 4\|A\|_\infty\sum_k\|U_k - V_k\|_2 \end{aligned} \tag{104}$$

resulting in

$$L_f \le 4\|A\|_\infty. \tag{105}$$

For $g = \langle A\rangle_{\rho_\infty^{UHU^\dagger}}$ we get by existence of the limit

$$\langle A\rangle_{\rho_\infty^{UHU^\dagger}} = \mathrm{tr}\left(A\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\rho_U(t)\right) = \lim_{T\to\infty}\frac{1}{T}\int_0^T\mathrm{d}t\langle A\rangle_{\rho_U(t)}. \tag{106}$$

Moreover, by existence of the limit and continuity of the absolute value

$$\begin{aligned} \left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \langle A\rangle_{\rho_\infty^{VHV^\dagger}}\right| &= \left|\lim_{T\to\infty}\frac{1}{T}\int_0^T\mathrm{d}t\Big(\langle A\rangle_{\rho_U(t)} - \langle A\rangle_{\rho_V(t)}\Big)\right| \\ &\le \lim_{T\to\infty}\frac{1}{T}\int_0^T\mathrm{d}t\Big|\langle A\rangle_{\rho_U(t)} - \langle A\rangle_{\rho_V(t)}\Big|. \end{aligned} \tag{107}$$

Consequently we get

$$L_g \le 4\|A\|_\infty. \tag{108}$$

The statement of Lemma 5 follows. □

We now move on to bounding $R(t)$. For an operator A we denote $A_{k,j} := \Pi_k A \Pi_j$ and $A_k = A_{k,k}$, where Π_k projects onto $\mathcal{W}_k$. We have

$$\langle A\rangle_{\rho_U(t)} = \sum_{k,j}\mathrm{tr}\left(A_{j,k}U_k\mathrm{e}^{-iHt}U_k^\dagger\rho_{k,j}U_j\mathrm{e}^{-iHt}U_j^\dagger\right) \tag{109}$$

and

$$\langle A\rangle_{\rho_\infty^{UHU^\dagger}} = \sum_k\sum_{\nu\in W_k}\langle\nu|U^\dagger\rho U|\nu\rangle\langle\nu|U^\dagger AU|\nu\rangle \tag{110}$$

where W_k is the set of all ν such that $E_\nu \in I_k$. In the following, we denote

$$\phi_k(t) := \frac{\mathrm{tr}\left(\Pi_k\mathrm{e}^{-iHt}\right)}{|W_k|} \tag{111}$$

and we abbreviate $d_k := |W_k|$. Simple computations reveal for the expectation value

$$\mathbb{E}\left(\langle A\rangle_{\rho_{U(t)}}\right) = \sum_{k\neq j}\phi_k(t)\phi_j^*(t)\mathrm{tr}\left(A_{j,k}\rho_{k,j}\right) + \sum_k \frac{|\phi_k(t)|^2 d_k^2 - 1}{d_k^2-1}\mathrm{tr}\left(A_k\rho_k\right) + \frac{d_k}{d_k^2-1}(1-|\phi_k(t)|^2)\mathrm{tr}\left(A_k\right)\mathrm{tr}\left(\rho_k\right) \tag{112}$$

as well as

$$\mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right) = \sum_k \frac{|\overline{\phi}_k|^2 d_k^2 - 1}{d_k^2-1}\mathrm{tr}\left(A_k\rho_k\right) + \frac{d_k}{d_k^2-1}(1-|\overline{\phi}_k|^2)\mathrm{tr}\left(A_k\right)\mathrm{tr}\left(\rho_k\right). \tag{113}$$

Here, we have defined

$$|\overline{\phi}_k|^2 := \lim_{T\to\infty}\frac{1}{T}\int_0^T |\phi_k(t)|^2 \mathrm{d}t = \frac{1}{d_k^2}\sum_{\nu\in W_k}\mathrm{tr}\left(P_\nu\right)^2. \tag{114}$$

Then

$$\begin{aligned}\mathbb{E}\left(\langle A\rangle_{\rho_{U(t)}}\right) - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right) = \sum_{k\neq j}\phi_k(t)\phi_j^*(t)\mathrm{tr}\left(A_{j,k}\rho_{k,j}\right) + \\ \sum_k \frac{1}{d_k+1}\left[\mathrm{tr}\left(A_k\rho_k\right)\left(d_k^2\frac{|\phi_k(t)|^2-|\overline{\phi}_k|^2}{d_k-1}\right) + \mathrm{tr}\left(\rho_k\right)\mathrm{tr}\left(A_k\right)\left(d_k\frac{|\overline{\phi}_k|^2-|\phi_k(t)|^2}{d_k-1}\right)\right].\end{aligned} \tag{115}$$

Simplifying this expression, we get

$$\mathbb{E}\left(\langle A\rangle_{\rho_{U(t)}}\right) - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right) = \sum_{k\neq j}\phi_k(t)\phi_j^*(t)\mathrm{tr}\left(A_{j,k}\rho_{k,j}\right) + \sum_k\frac{d_k}{d_k+1}F_k(t)\left(\mathrm{tr}\left(A_k\rho_k\right) - \frac{\mathrm{tr}\left(A_k\right)}{d_k}\mathrm{tr}\left(\rho_k\right)\right) \tag{116}$$

where we have defined

$$F_k(t) := d_k\frac{|\phi_k(t)|^2-|\overline{\phi}_k|^2}{d_k-1}. \tag{117}$$

This gives

$$R(t) \leq \|A\|_\infty K^2 \max{}_{k,j}\phi_k(t)\phi_j^*(t) + \max{}_k F_k(t)\sum_k\left|\mathrm{tr}\left(A_k\rho_k\right) - \frac{\mathrm{tr}\left(A_k\right)}{d_k}\mathrm{tr}\left(\rho_k\right)\right| \tag{118}$$

where we have used

$$\mathrm{tr}\left(A_{j,k}\rho_{k,j}\right) = \mathrm{tr}\left(A\Pi_k\rho\Pi_j\right) \leq \|A\|_\infty\|\Pi_k\rho\Pi_j\|_1 \leq \|A\|_\infty. \tag{119}$$

By the same reasoning, we can conclude that

$$\mathrm{tr}\left(A_k\rho_k\right) \leq \|A\|_\infty\|\rho_k\|_1 \tag{120}$$

and furthermore $\mathrm{tr}\left(A_k\right)/d_k \leq \|A\|_\infty$. This allows us to bound

$$\sum_k\left|\mathrm{tr}\left(A_k\rho_k\right) - \frac{\mathrm{tr}\left(A_k\right)}{d_k}\mathrm{tr}\left(\rho_k\right)\right| \leq 2\|A\|_\infty. \tag{121}$$

Altogether, we can conclude that

$$R(t) \leq \|A\|_\infty\left(K^2\max{}_{k,j}|\phi_k(t)\phi_j^*(t)| + 2\max{}_k F_k(t)\right). \tag{122}$$

To make more progress, we need to be able to compute and discuss the functions ϕ_k. Mathematically speaking, we make a mild spectral assumption expected to be valid in natural quantum many-body systems. We assume that the above characteristic functions decay appropriately.

Assumption 1 (Spectral assumption)**.** *For some $r > 0$ and $t \leq \mathrm{poly}(N)$, we have*

$$|\phi_k(t)| \leq \frac{1}{t^r}. \tag{123}$$

In this case,

$$R(t) \leq O\left(\|A\|_\infty \frac{K^2}{t^{2r}}\right) \leq O\left(\|A\|_\infty \frac{N^2}{\delta^2 t^{2r}}\right). \tag{124}$$

Such a behavior can be derived, for example, by assuming a physically plausible slowly varying density of states inside the each energy window I_k: consider the function $E_k(x) : [0,1] \to \mathbb{R}$ defined such that for $x \in \left[\frac{\nu}{d_k}, \frac{\nu+1}{d_k}\right)$, $E_k(x) = E_\nu$. Then, we have

$$\phi_k(t) = \int_0^1 \mathrm{d}x \mathrm{e}^{itE_k(x)}. \tag{125}$$

Now, define a linear interpolation of E_k: for $x \in \left[\frac{\nu}{d_k}, \frac{\nu+1}{d_k}\right)$

$$\tilde{E}_k(x) = d_k \Delta_\nu x + (\nu+1)E_\nu - \nu E_{\nu+1}, \tag{126}$$

where $\Delta_\nu = E_{\nu+1} - E_\nu$. $\tilde{E}_k(x)$ is piece-wise linear and satisfies $|\tilde{E}_k(x) - E_k(x)| \leq \max\limits_{\nu \in W_k} \Delta_\nu$. In particular, defining

$$\tilde{\phi}_k(t) = \int_0^1 \mathrm{d}x \mathrm{e}^{it\tilde{E}_k(x)}, \tag{127}$$

we have $|\phi_k(t) - \tilde{\phi}_k(t)| \leq \max\limits_{\nu \in W_k} \Delta_\nu$. Notice that $\tilde{E}_k(x)$ is strictly increasing hence invertible, and its inverse is continuous and differentiable almost everywhere. By the change of variable $y = \tilde{E}_k(x)$ we get

$$\tilde{\phi}_k(t) = \int_{e_k}^{e_{k+1}} \mathrm{d}y \mathrm{e}^{ity} \rho_k(y), \tag{128}$$

where $\rho_k(y) := \frac{\mathrm{d}}{\mathrm{d}y} \tilde{E}_k^{-1}(y)$ is the density of states, and satisfies $\rho_k(y) = \frac{1}{d_k \Delta_\nu}$ for $y \in [E_\nu, E_{\nu+1}]$. We take once again a piecewise linear interpolation of $\rho_k(y)$: for $y \in [E_\nu, E_{\nu+1}]$

$$\tilde{\rho}_k(y) = \frac{y - E_\nu}{d_k \Delta_\nu} \left(\frac{1}{\Delta_{\nu+1}} - \frac{1}{\Delta_\nu}\right) + \frac{1}{d_k \Delta_\nu} \tag{129}$$

is piecewise linear and satisfies

$$|\rho_k(y) - \tilde{\rho}_k(y)| \leq \max_{\nu \in W_k} \frac{1}{d_k} \left|\frac{1}{\Delta_{\nu+1}} - \frac{1}{\Delta_\nu}\right|. \tag{130}$$

Hence, defining

$$\tilde{\tilde{\phi}}_k(t) = \int_{e_k}^{e_{k+1}} \mathrm{d}y \mathrm{e}^{ity} \tilde{\rho}_k(y), \tag{131}$$

we have $|\tilde{\tilde{\phi}}_k(t) - \tilde{\phi}_k(t)| \leq \max_{\nu \in W_k} \frac{1}{d_k} \left|\frac{1}{\Delta_{\nu+1}} - \frac{1}{\Delta_\nu}\right|$, and furthermore, integrating by parts gives

$$\tilde{\tilde{\phi}}_k(t) = \frac{1}{it} \int_{e_k}^{e_{k+1}} \mathrm{d}y \tilde{\rho}_k(y) \frac{\mathrm{d}}{\mathrm{d}y} \mathrm{e}^{ity} = \frac{1}{it} \left(\mathrm{e}^{ite_{k+1}} \tilde{\rho}_k(e_{k+1}) - \mathrm{e}^{ite_k} \tilde{\rho}_k(e_k)\right) + \frac{1}{it} \int_{e_k}^{e_{k+1}} \mathrm{d}y \tilde{\rho}_k'(y) \mathrm{e}^{ity}. \tag{132}$$

Then

$$\begin{aligned} |\tilde{\tilde{\phi}}_k(t)| &\leq \frac{1}{t} \left|\left(\mathrm{e}^{ite_{k+1}} \tilde{\rho}_k(e_{k+1}) - \mathrm{e}^{ite_k} \tilde{\rho}_k(e_k)\right)\right| + \frac{1}{t} \left|\int_{e_k}^{e_{k+1}} \mathrm{d}y \tilde{\rho}_k'(y) \mathrm{e}^{ity}\right| \\ &\leq \frac{1}{t} \left(2 + \max_{\nu \in W_k} \frac{1}{d_k} \left|\frac{1}{\Delta_{\nu+1}} - \frac{1}{\Delta_\nu}\right|\right). \end{aligned} \tag{133}$$

Overall, we get

$$|\phi_k(t)| \leq \frac{1}{t} + O\left(\max_{\nu \in W_k} \Delta_\nu\right) + O\left(\max_{\nu \in W_k} \frac{1}{d_k} \left|\frac{1}{\Delta_{\nu+1}} - \frac{1}{\Delta_\nu}\right|\right). \tag{134}$$

This shows that our assumption holds with $r = 1$ on time scales such that $1/t$ is large with respect to the spectral gaps and the difference of nearby spectral gaps. This is also compatible with natural expectations of spectral gaps following a Wigner-Dyson distribution, which due to level repulsion favors an almost uniform distribution of the spectral gaps.

IV. TRANSLATIONAL INVARIANCE

We will now turn to discuss notions of translational invariance. More generally, we will show that if the original Hamiltonian has a certain symmetry $T : [H, T] = 0$, then the equilibrium state $\rho_\infty^{UHU^\dagger}$ with high probability has the same symmetry in an approximate form. Applying this to the translation symmetry yields the claim in the main text. We will first prove the following.

Lemma 7 (Auxiliary lemma)**.** *Let A be an observable such that $\operatorname{tr}(\Pi_i A) = 0$ for all energy windows I_i. Let $\epsilon \geq \sum_k \Omega\left(\frac{D_k}{d_k}\right) \|\rho_k\|_1$. Then with probability at least*

$$1 - \sum_k \mathrm{e}^{-\frac{C(d_k-1)\epsilon^2}{\|A\|_\infty^2}} \tag{135}$$

for U drawn from $\mathcal{E}(\delta)$ and a constant $C > 0$, we have

$$|\operatorname{tr}\left(\rho_\infty^{UHU^\dagger} A\right)| \leq \epsilon. \tag{136}$$

Proof. Since $\operatorname{tr}(A_k) = 0$, from Eq. (113), we can infer that

$$\mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right) = \sum_k \frac{|\overline{\phi}_k|^2 d_k^2 - 1}{d_k^2 - 1} \operatorname{tr}(A_k \rho_k) \tag{137}$$

with

$$|\overline{\phi}_k|^2 = \lim_{T\to\infty} \int_0^T |\phi_k(t)|^2 \mathrm{d}t = \frac{1}{d_k^2} \sum_{\nu \in W_k} \operatorname{tr}(P_\nu)^2 \leq \frac{D_k}{d_k}. \tag{138}$$

Hence, we can bound

$$\left|\mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| \leq \sum_k O\left(\frac{D_k}{d_k}\right) \|\rho_k\|_1. \tag{139}$$

Assuming that $\epsilon \geq \sum_k \Omega\left(\frac{D_k}{d_k}\right) \|\rho_k\|_1$ and using Levy's Lemma (Lemma 6) as well as the Lipschitz function computation in the previous section we get

$$\begin{aligned}\Pr_{U\sim\mathcal{E}(\delta)}\left[\left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right| > \epsilon\right] &\leq \Pr_{U\sim\mathcal{E}(\delta)}\left[\left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| > \epsilon - \left|\mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right|\right] \\ &\leq \Pr_{U\sim\mathcal{E}(\delta)}\left[\left|\langle A\rangle_{\rho_\infty^{UHU^\dagger}} - \mathbb{E}\left(\langle A\rangle_{\rho_\infty^{UHU^\dagger}}\right)\right| > \frac{\epsilon}{2}\right] \leq \sum_k \mathrm{e}^{-\frac{c(d_k-1)(\epsilon/2)^2}{16\|A\|_\infty^2}}.\end{aligned} \tag{140}$$

□

As commented earlier, d_k might be smaller at the edge of the spectrum, but it is still expected to be polynomial in the system size, and the tails at the edge of the spectrum may be cut if the state has exponential decay of correlations. Then, we find the following corollary.

Corollary 2 (Preservation of symmetries)**.** *Suppose the Hamiltonian has no degenerate eigenvalues and let T be a unitary which commutes with the Hamiltonian. Let A be an observable, let $\epsilon \geq \sum_k \Omega\left(\frac{1}{d_k}\right) \|\rho_k\|_1$, then with probability at least*

$$1 - \sum_k \mathrm{e}^{-\frac{C(d_k-1)\epsilon^2}{\|A\|_\infty^2}} \tag{141}$$

for U drawn from $\mathcal{E}(\delta)$ and a constant $C > 0$, we have that

$$|\operatorname{tr}\left((A - TAT^\dagger)\rho_\infty^{UHU^\dagger}\right)| \leq \epsilon. \tag{142}$$

Proof. Since H and T commute, T commutes with the projectors onto the eigenspaces of H, and in particular with the projectors Π_i. We then have

$$\mathrm{tr}\left(\Pi_i T A T^\dagger\right) = \mathrm{tr}\left(\Pi_i A\right). \tag{143}$$

We conclude by applying Lemma 7 to the observable $TAT^\dagger - A$. $\square$

As a corollary, we get the statement in eq. (17) in the main text. Let A be a local observable and call A_C the observable O acting on the hypercube $C \in \mathcal{C}_l$. We have for any $C' \in \mathcal{C}_l$

$$\begin{aligned}\left|\mathrm{tr}\left(A_{C'}(g_\beta(H) - \rho_\infty^{UHU^\dagger})\right)\right| &\leq \frac{1}{|\mathcal{C}_l|}\sum_{C\in\mathcal{C}_l}\left|\mathrm{tr}\left(A_C(g_\beta(H) - \rho_\infty^{UHU^\dagger})\right)\right| + \left|\mathrm{tr}\left((A_{C'} - A_C)(g_\beta(H) - \rho_\infty^{UHU^\dagger})\right)\right| \\ &\leq D_l(\rho_\infty^{UHU^\dagger}, g_\beta(H)) + \left|\mathrm{tr}\left((A_{C'} - A_C)\rho_\infty^{UHU^\dagger}\right)\right|.\end{aligned} \tag{144}$$

where we used that $g_\beta(H)$ is translation invariant. Since $\mathrm{tr}\left((A'_C - A_C)\Pi_k\right) = 0$ for all k, as H is translation invariant, by Lemma 7 we have that with high probability

$$\left|\mathrm{tr}\left((A_{C'} - A_C)\rho_\infty^{UHU^\dagger}\right)\right| \leq \epsilon. \tag{145}$$

[1] F. G. S. L. Brandao and M. Cramer, "Equivalence of statistical mechanical ensembles for non-critical quantum systems," (2015), arXiv:1502.03263.

[2] A. M. Alhambra, PRX Quantum **4**, 040201 (2023).

[3] A. Anshu, New J. Phys. **18**, 083011.

[4] A. A. Mele, "Introduction to haar measure tools in quantum information: A beginner's tutorial," (2024), arXiv:2307.08956.

[5] A. J. Short, New J. Phys. **13**, 053009 (2011).

[6] M. Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs, Vol. 89 (American Mathematical Society, 2005) optissn: 0076-5376, 2331-7159.