Entanglement-ergodic quantum systems equilibrate exponentially well

H. Wilming,[1] M. Goihl,[2] I. Roth,[2] and J. Eisert[2]

[1]*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*
[2]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

One of the outstanding problems in non-equilibrium physics is to precisely understand when and how physically relevant observables in many-body systems equilibrate under unitary time evolution. General equilibration results show that equilibration is generic provided that the initial state has overlap with sufficiently many energy levels. But results not referring to typicality which show that natural initial states actually fulfill this condition are lacking. In this work, we present stringent results for equilibration for systems in which Rényi entanglement entropies in energy eigenstates with finite energy density are extensive for at least some, not necessarily connected, sub-system. Our results reverse the logic of common arguments, in that we derive equilibration from a weak condition akin to the eigenstate thermalization hypothesis, which is usually attributed to thermalization in systems that are assumed to equilibrate in the first place. We put the findings into the context of studies of many-body localization and many-body scars.

Over recent years the study of the relaxation to equilibrium of complex many-body systems has attracted great attention. This interest can be motivated from at least two points of view. From a foundational viewpoint, it is desirable to understand how statistical equilibrium ensembles emerge within the framework of unitary quantum mechanics – without introducing any external probability measures. It is then necessary to first explain how systems undergoing unitary evolution attain equilibrium at all. A key ingredient to explain this behavior has been found to be the dynamical build-up of entanglement from low-entangled initial states and results showing equilibration under quite general conditions have been derived [1–12]. The increase of entanglement over time is a generic feature of complex quantum systems and leads to an increase of the entropy of subsystems over time reminiscent to the second law of thermodynamics.

From a more concrete perspective, the recent interest in the study of non-equilibrium dynamics is motivated by the fact that such dynamics can now be realized in well-controlled experiments, for example in ion traps or optical lattices [13–18]. Moreover, the discovery of many-body localized systems [19], which equilibrate [20] but fail to thermalize [18], shows that there remains much to be understood about the equilibration behavior of complex quantum systems. Despite the great progress in understanding the equilibration behavior of many-body systems, rigorous results showing that systems with natural initial states equilibrate to high precision based on concrete physical properties have been lacking.

In this article we aim to fill this gap, by taking a new perspective to the problem. To do this, we carefully reconsider the entanglement content of energy eigenstates in complex, interacting many-body systems and devise a working-definition of "entanglement-ergodic" systems whose energy eigenstates at finite energy density have a sufficient amount of entanglement between suitable subsystems. The condition we propose is very weak – yet we show that generically such systems equilibrate to exponential precision in the volume of the system if the initial state is given by a product state with finite, non-zero energy density. Commonly, one assumes equilibration and invokes the *eigenstate thermalization hypothesis*

(ETH) [21–25] to make thermalization plausible. Here, we stringently derive equilibration from a highly plausible condition similar to, but we believe much weaker than the eigenstate thermalization hypothesis.

The main ingredient of our proof is a careful discussion of Rényi entanglement entropies in energy eigenstates with finite energy density. Combining this insight with the strongly peaked energy distribution of weakly correlated states and the monotonicity of Rényi entropies allows us to prove that experimentally accessible initial states are well smeared out over the energy spectrum, which implies high-precision equilibration for generic interacting Hamiltonians.

Formal setting. We consider local Hamiltonians

$$H_\Lambda = \sum_{x \in \Lambda} h_x \qquad (1)$$

on a regular lattice Λ in ν spatial dimensions with $N := |\Lambda|$ lattice sites. The Hilbert space is $\otimes_{x \in \Lambda} \mathcal{H}_x$ with $\dim \mathcal{H}_x = d$. Since we will be talking about the scaling of quantities with the lattice size, H_Λ should be seen as a sequence of Hamiltonians, which is, for example, given by a family of translational invariant system on larger and larger square lattices with periodic boundary conditions. We call H_Λ l-local if the diameter of the support of each h_x is at most l, strictly local if it is l-local with l independent of the system size N and uniformly bounded if $\|h_x\| \leq h$ for all $x \in \Lambda$ independent of the system size. Since we are mostly interested in energy densities instead of total energies later, we make the unusual choice to label eigenvectors of the Hamiltonian by their energy *densities* e_i as $|e_i\rangle_\Lambda = |E_i/N\rangle_\Lambda$, with $i = 1, \ldots, d^N$ and E_i being the eigenvalue of H_Λ corresponding to $|e_i\rangle_\Lambda$. We always assume that the ground-state has vanishing total energy. From now on, we will often drop the subscript Λ from states and Hamiltonians to simplify the notation.

Equilibration in closed systems. We now briefly review general equilibration results that we will be using in the following. Consider any bounded observable A and any initial state ρ. Denote the time evolved states by $\rho(t)$. We want to study whether the expectation value $\langle A(t) \rangle_\rho = \text{Tr}(\rho(t)A)$ equilibrates over time. In a finite system, perfect equilibration,

in the sense that $\langle A(t)\rangle$ becomes static for all times after the equilibration process, is impossible due to recurrences. However, it is perfectly possible that this value is very close to a stationary value $\overline{A}$ for most of the time, with rare deviations. The value $\overline{A}$ is then necessarily the infinite time average

$$\overline{A} := \overline{\langle A(t)\rangle} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \text{Tr}(\rho(t)A) = \text{Tr}(\omega A), \quad (2)$$

where ω denotes the time average $\overline{\rho(t)}$, which is again a valid density matrix. In the case where the Hamiltonian has no degenerate *energy differences* $G_{(i,j)} = E_i - E_j$, it has been shown that the time-averaged fluctuations around the equilibrium value are bounded by [5]

$$\text{Var}(A, H, \rho) := \overline{(\langle A(t)\rangle - \overline{A})^2} \leq \|A\|^2 \, \mathrm{e}^{-S_2(\omega)}, \quad (3)$$

where S_α denotes the Rényi-α entropy

$$S_\alpha(\rho) := \frac{1}{1-\alpha} \log\left(\text{Tr}(\rho^\alpha)\right). \quad (4)$$

We note for later use that in the limit $\alpha \to 1$ the von-Neumann entropy is recovered and that Rényi entropies are monotonically decreasing in α. The condition of non-degenerate energy differences is generically fulfilled in interacting systems [26]. However, generalizations of the above result also exist if this condition is weakly violated [8]. It is also possible to show $\text{Var}(A, H, \rho) \leq 3\|A\|^2 \, \mathrm{e}^{-S_\infty(\omega')}$, where ω' is the same operator as ω but with its largest eigenvalue replaced by zero [2], which sometimes gives a stronger bound. In particular, it also incorporates the case of energy eigenstates, which are always fully equilibrated. It is important to stress, however, that the bound Eq. (3) does not lead to implications on the time it takes to observe equilibration.

Similarly to results in terms of bounded observables, one can also bound the distance of a local reduced state on a subsystem S from its time average as [3]

$$\overline{\|\text{Tr}_{S^c}(\rho(t)) - \text{Tr}_{S^c}(\omega)\|_1} \leq 2d_S \mathrm{e}^{-S_2(\omega)/2}. \quad (5)$$

Here, $\|.\|_1$ denotes the trace norm, which bounds the difference in expectation value of all normalized observables as

$$\|\rho - \sigma\|_1 = \max_{A, \|A\|=1} |\text{Tr}(\rho A) - \text{Tr}(\sigma A)|. \quad (6)$$

Roughly speaking, the role of the norm of the observables in the previous bound is here taken by the dimension of the subsystem. We conclude that a large Rényi-2 entropy of the time-averaged state is a sufficient condition for a generic, closed quantum system to equilibrate eventually. Unfortunately, however, there are few general and rigorous results which show that natural initial states lead to time-averaged states whose energy distribution has a large Rényi-2 entropy [27, 28]. Ref. [28] makes it highly plausible from an operational point of view that states that can be prepared in experiments have large effective dimension, but does not show it

for concrete states. Ref. [27] shows that states with finite correlation length have Rényi-2 entropy of at least roughly the order $\log(N)$. While this formally leads to equilibration as $N \to \infty$, it is insufficient to obtain a finite entropy density at equilibrium (which is crucial from a thermodynamic point of view) and requires very large system sizes to explain equilibration. In the following, we show that in systems which have a sufficiently large amount of entanglement in energy eigenstates with finite energy density, finite entropy density and hence exponentially good equilibration in the system size follows for initial product states.

Entanglement and Rényi entropies. Here, we are interested in ergodic, non-integrable systems. Contrary to the case of classical mechanics, there is no generally agreed upon definition of what it exactly means for a quantum system to be ergodic or non-integrable [29]. In recent years, it has been argued that a general characteristic of systems which can be considered ergodic is that energy eigenstates fulfill a so-called *volume law* in terms of their entanglement content. The condition we propose is inspired from this observation, but is much weaker. It is therefore useful to discuss volume laws before stating our condition. Consider an energy eigenvector with finite energy density $|e\rangle$ and denote by $\rho_A(e)$ the reduced density matrix on some (contiguous) subsystem A which is smaller than one half of the total system, but still contains a finite fraction of the total system. A *volume law* means that the entanglement measured by a Rényi entropy S_α grows like the volume of A as

$$S_\alpha(\rho_A(e)) \sim |A|. \quad (7)$$

A natural question to ask is for which value of α this relation is supposed to hold. We now argue that this relation is a meaningful criterion only if $\alpha > 1$. This might come as a surprise since it is common to measure entanglement in terms of the von-Neumann entropy S_1. This is due to the fact that the von-Neumann entropy describes the fraction of EPR-pairs that can be distilled from asymptotically many copies of a state by local operations and classical communication [30]. However, the examples presented in the following proposition show that a volume law in terms of von-Neumann entropy is not a very useful criterion to determine whether a state of a many-body system deviates strongly from an unentangled state.

Proposition 1 (Counterexample). *For any $1 > \epsilon > 0$ there exist state vectors $|\Psi^\epsilon\rangle$ on Λ with the following properties:*

1. *$|\Psi^\epsilon\rangle$ has overlap exponentially close to $1-\epsilon$ in N with a product state vector $|\Psi\rangle$.*

2. *$|\Psi^\epsilon\rangle$ fulfills a* volume law *in the von Neumann entropy: There exist regions A with $|A| = N/2$ such that $S_1(\rho_A) \simeq \frac{\epsilon}{2}\log(d)N$.*

3. *All Rényi entropies with $\alpha > 1$ are bounded by a constant in the system size, $S_\alpha(\rho_A) \leq \text{const}.$*

The state vectors that fulfill these condition are simply of the form $|\Psi^\epsilon\rangle \propto \sqrt{1-\epsilon}\,|\Psi\rangle + \sqrt{\epsilon}\,|\Omega\rangle$, where $|\Omega\rangle$ is maximally entangled between A and its complement. The proof of

these properties is given in Section B in the Supplemental Material. Given Properties 1 and 3, it cannot reasonably be said that the amount of entanglement in $|\Psi^\epsilon\rangle$ grows volume-like – even though it fulfills a volume law in terms of von-Neumann entropy. At the same time it suggests that we should require a volume law in terms of some Rényi entropy with $\alpha > 1$. Importantly, the inequality (proven in the Supplemental Material, Section A)

$$S_\alpha(\rho) \geq S_\infty(\rho) \geq \frac{\beta-1}{\beta} S_\beta(\rho), \quad \forall \alpha \geq 0, \qquad (8)$$

which holds for any $\beta > 1$, shows that all the Rényi entropies with $\alpha > 1$ have the same scaling behaviour. It therefore does not matter which one we consider and in the following we therefore mostly consider the case $\alpha = 2$.

Recently there has been an increasing amount of numerical results and theoretical arguments that show that energy eigenstates of generic non-integrable quantum systems with finite energy density have this property (see, for example, Refs. [31–37]). These results fit well to, and indeed are partly motivated by, the observation that certain properties of complex, strongly interacting systems at finite energy density can be well described by assuming that their Hamiltonians are *random matrices* [38–43], despite the Hamiltonians being local and thus belonging to a set of measure zero. It is well known that Haar random quantum states have extensive Rényi entanglement entropy on bipartite systems with very high probability [44–48]. Since eigenstates of random Hamiltonians are distributed according to the Haar-measure, we expect that eigenstates of random Hamiltonians fulfill a volume law in terms of all Rényi entropies with very high probability [49].

Entanglement-ergodicity. In essence, in the following we will show that an extensive amount of entanglement entropy in energy eigenstates for Rényi-entropies with $\alpha > 1$ is a sufficient criterion for exponentially good equilibration. Importantly, however, we will significantly weaken the assumption from volume laws – where the subsystem under consideration is assumed be a contiguous region – by allowing the subsystem to be almost completely arbitrary. The only property that we demand from the subsystem is that it includes a finite fraction of the total system. In particular, it need not be connected and further may be chosen differently for every energy eigenstate. For example, it may have a fractal-like shape (up to the lattice spacing), or consist of a sub-lattice of spins that are far away from each other when compared to some natural length scale of the system. In the following, we call this (much weaker) form of a volume law a *weak volume law*. In particular, even states that have a finite correlation length, such as *matrix product states (MPS)*, can be expected to generically follow a weak volume law – unless they are product states themselves.

We will now give a formal definition of what we demand from an entanglement-ergodic system. Notably, this definition is perfectly compatible with the mindset of the ETH that argues basically that eigenstates of local Hamiltonians in the bulk of the spectrum should be locally indistinguishable from Gibbs states due to their entanglement.

Definition 1 (Entanglement-ergodicity). *We call a sequence of systems of increasing system size* entanglement-ergodic, *if there exists a system size N_0 and a function $g : \mathbb{R} \to [0,\infty)$, such that for all system sizes $N \geq N_0$ it holds that*

1. *(Weak volume law) For every eigenvector $|e_i\rangle_\Lambda$ there exists some subsystem A_Λ such that the reduced state $\rho_{A_\Lambda}(e_i) := \mathrm{Tr}_{A_\Lambda^c}(|e_i\rangle\langle e_i|_\Lambda)$ fulfills*

$$S_2(\rho_{A_\Lambda}(e_i)) \geq g(e_i)N. \qquad (9)$$

2. *The function g is sufficiently well-behaved: It is Lipschitz continuous and positive for non-extremal energy densities, i.e., $g(e) > 0$ for $0 < e < e_{\max}$.*

Some remarks on this definition are in order: i) As shown by (12), we could have replaced the Rényi-2 entropy with any Rényi entropy with $\alpha > 1$ and would have obtained an equivalent definition. ii) We only require that there is a *lower bound* on the entropy in the region A_Λ as a function of the energy density and not that that the entropy is given by the function $e \mapsto g(e)$. In particular, we do not require that eigenstates with the same energy density also have the same Rényi entanglement entropy and allow for the possibility that states with vanishing energy density do not have a large amount of entanglement. For example, the ground state of the system may be a product state. Similarly, we impose Lipschitz continuity of the function g for simplicity and concreteness. Similar conclusions as in the following can be reached by imposing different regularity assumptions on g. iii) It would be perfectly fine for all what follows to allow for additional negative but sub-leading terms, e.g. in $O(\sqrt{N})$. These terms would only change the sub-leading behavior of our final results and may encode non-asymptotic information about the entropy in eigenstates. For simplicity, we omit such terms here.

iv) We emphasize again that we only demand a weak volume law, i.e., that for any system size and any eigenstate there exists *some* finite fraction of the total system A_Λ whose entropy is sufficiently large. It is not required that this holds for all such subsystems nor is it required that this subsystem has any particular shape. For example, it could consist of every 10-th site of the lattice. Nevertheless, in generic, strongly interacting systems, we expect that choosing A_Λ simply as one half of the system is sufficient.

v) Importantly, the weak assumption on the subsystem A_Λ allows even MPS-like eigenstates, which fulfill an *area law* [50] for *contiguous* regions, to be entanglement-ergodic. As a concrete example, we prove the following statement in the Supplemental Material, Section C: *States that are prepared by a translationally invariant, finite-depth, local quantum circuit and are not a product state have extensive in Rényi-2 entropy on a finite fraction of the system.* Similarly, we expect an analogous result to hold more generally for generic MPS. A detailed discussion on the application of our framework to generic MPS is provided in Section C of the Supplemental Material. These results show that also systems featuring

many-body localization [19], whose eigenstates are expected to be approximable by matrix product states [51, 52], may fall within the framework of entanglement ergodicity. It is well known that such systems equilibrate, but fail to also thermalize. Thus even if we refer to the key property as entanglement-ergodicity, it is a significantly weaker condition than what is commonly understood as ergodicity in current literature.

vi) Similarly, we expect the notion of entanglement-ergodicity to be weaker than the ETH: if one assumes that the ETH applies to some subsystems containing a finite fraction of the total system, then we strongly believe that the ETH implies entanglement-ergodicity and hence equilibration with high precision (see Supplemental Material, Section G).

vii) With stronger assumptions on the regions A_Λ, we can extend the applicability of the definition of entanglement ergodicity to states related by quasi-local unitaries (those generated by time-evolution under local Hamiltonians): *For systems in which the regions A_Λ have an asymptotically vanishing surface-to-volume ratio, entanglement-ergodicity is stable under quasi-local unitaries.* The precise meaning and formal proof of this statement is formulated in Section E of the Supplements. Intuitively it follows by observing that time evolution under a local Hamiltonian for a finite time can only decrease the entropy of a sub-region by an amount that is proportional to its boundary, since the entropy has to "flow" out of the subregion through its boundary.

Consequences of entanglement-ergodicity. Let us now discuss the consequences of entanglement-ergodicity. We first state a result on the diagonal entropy in entanglement-ergodic systems, to then turn to the implication that entanglement-ergodic systems equilibrate exponentially well.

Theorem 2 (Diagonal entropy in entanglement-ergodic systems)**.** *Consider an entanglement-ergodic system with strictly local, uniformly bounded Hamiltonian. Then for any energy density $e > 0$ there exists a constant $k(e) > 0$ and a system-size $N_0(e)$ such that for all system-sizes $N > N_0(e)$ and for all product-states $|\Psi\rangle_\Lambda$ with energy density e, we have*

$$S_\alpha(\omega_\Lambda) \geq k(e)N, \tag{10}$$

where ω_Λ is the time-average of $|\Psi\rangle\langle\Psi|_\Lambda$.

As a direct consequence of this result we obtain from (3) the following bounds on equilibration.

Corollary 3 (Equilibration in entanglement-ergodic systems)**.** *Under the same conditions as in Theorem 2 and the additional assumption of non-degenerate energy gaps of the Hamiltonian, there exists a constant $k(e) > 0$ such that*

$$\mathrm{Var}\left(A, H_\Lambda, |\Psi\rangle\langle\Psi|_\Lambda\right) \leq \|A\|^2 \, \mathrm{e}^{-k(e)N}. \tag{11}$$

Similar bounds hold for the reduced state on a small subsystem, as implied by (5). The proof of Theorem 2 is given in Section D of the Supplemental Material. It relies on recognizing that an extensive amount of Rényi entanglement entropy implies that a state has exponentially small overlap with

all product states and to combine this statement with a recent central limit-type theorem for the energy distribution in product states [53]. The premises of Theorem 2 require the initial state to be a product states. In the Supplements (Section F), we further extend our results to the case where the initial state is prepared from a product state by a quasi-local unitary under the additional assumption that the regions A_Λ have vanishing surface-to-volume ratio. We obtain equilibration bounds scaling as $O(\exp(-N^\beta))$ for some $0 < \beta < 1$.

Quantum many-body scars. Recently, it has been observed that sets of atypical energy eigenstates with small amount of entanglement may rarely show up even at finite energy density in non-integrable, kinetically constrained many-body systems – a phenomenon dubbed *"quantum many-body scars"*, which leads to exceedingly slow equilibration with long-lived oscillations from certain initial product states [54–58]. Even a toy-model with complete absence of equilibration for certain initial product states has been constructed [59]. This complete absence of equilibration can be explained due to a $\mathrm{SU}(2)$ symmetry in the subspace of many-body scars [59], which also emerges approximately in more realistic Hamiltonians and breaks the assumption of non-degenerate energy gaps. Furthermore, many-body scars may have large squared overlap with the initial product state (of order $1/N$), which implies that such systems also violate entanglement-ergodicity within the subspace of many-body scars.

Conclusion. We carefully formalized a notion of ergodicity based on extensive Rényi entanglement entropies in energy eigenstates and showed that this notion suffices to prove exponentially precise equilibration for Hamiltonians with non-degenerate energy gaps. Our condition is quite weak and we expect it to be fulfilled for generic interacting systems. The notion of ergodicity we introduced is connected to the eigenstate thermalization hypothesis (ETH), which asserts that local reduced states of energy eigenstates with finite energy density already resemble the reduced state of a corresponding Gibbs state. We thereby introduce a new perspective to the study of non-equilibrium quantum systems: We do not have to assume equilibration for systems to become apparently stationary and then turn to the ETH to show thermalization. Instead, we show that equilibration already follows from a weak ETH-like assumption. In fact, in this case the ETH implies much more, since our definition does in general not imply that the system also *thermalizes* (this apparent shortcoming is necessary when formulating a criterion that may also apply to many-body localized systems). Here, we did not discuss the time scales for relaxation to equilibrium, but were interested in the precision of equilibration after arbitrarily long times. While some progress in understanding equilibration time-scales has been made recently, both in integrable [60–69] and generic, non-integrable systems [8, 70–77], finding rigorous arguments bounding equilibration time scales from reasonable assumptions remains an outstanding open problem.

[1] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998).
[2] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008), arXiv:0810.3092.
[3] N. Linden, S. Popescu, A. Short, and A. Winter, Phys. Rev. E **79**, 61103 (2009), arXiv:0812.2385.
[4] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Proc. Roy. Soc. A **466**, 3203 (2010).
[5] A. J. Short, New J. Phys. **13**, 053009 (2011).
[6] P. Reimann, Phys. Scr. **86**, 58512 (2012), arXiv:1210.5821.
[7] P. Reimann and M. Kastner, New J. Phys. **14**, 43020 (2012), arXiv:1202.2768.
[8] A. J. Short and T. C. Farrelly, New J. Phys. **14**, 013063 (2012).
[9] L. Masanes, A. J. Roncaglia, and A. Acín, Phys. Rev. E **87**, 32137 (2013), arXiv:1108.0374.
[10] C. Gogolin and J. Eisert, Rep. Prog. Phys. **79**, 56001 (2016), arXiv:1503.07538.
[11] J. Eisert, M. Friesdorf, and C. Gogolin, Nature Phys. **11**, 124 (2015), arXiv:1408.5148.
[12] B. N. Balz and P. Reimann, Phys. Rev. Lett. **118**, 190601 (2017).
[13] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
[14] I. Bloch, J. Dalibard, and S. Nascimbene, Nature Phys. **8**, 267 (2012).
[15] U. Schneider, L. Hackermüller, J. P. Ronzheimer, S. Will, S. Braun, T. Best, I. Bloch, E. Demler, S. Mandt, D. Rasch, and A. Rosch, Nature Phys. **8**, 213 (2012).
[16] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Nature Phys. **8**, 325 (2012).
[17] S. Braun, M. Friesdorf, S. S. Hodgman, M. Schreiber, J. P. p. Ronzheimer, A. Riera, M. del Rey, I. Bloch, J. Eisert, and U. Schneider, Proc. Natl. Ac. Sc. **112**, 3641 (2015).
[18] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Science **349**, 842 (2015).
[19] R. Nandkishore and D. A. Huse, Ann. Rev. Cond. Matter Phys. **6**, 15 (2015).
[20] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. B **90**, 174302 (2014).
[21] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991).
[22] M. Srednicki, Phys. Rev. E **50**, 888 (1994).
[23] M. Rigol, V. Dunjko, and M. Olshanii, Nature **452**, 854 (2008).
[24] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Rev. Mod. Phys. **83**, 863 (2011).
[25] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, Adv. Phys. **65**, 239 (2016).
[26] In case of global symmetries or super-selection sectors, each sector should be considered separately.

[27] T. Farrelly, F. G. Brandão, and M. Cramer, Phys. Rev. Lett. **118**, 140601 (2017).
[28] R. Gallego, H. Wilming, J. Eisert, and C. Gogolin, Phys. Rev. A **98**, 022135 (2018).
[29] J.-S. Caux and J. Mossel, J. Stat. Mech. , P02023 (2011).
[30] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[31] J. R. Garrison and T. Grover, Phys. Rev. X **8**, 021026 (2018).
[32] Y. O. Nakagawa, M. Watanabe, H. Fujita, and S. Sugiura, Nat. Commun. **9** (2018), 10.1038/s41467-018-03883-9.
[33] T.-C. Lu and T. Grover, Phys. Rev. E **99** (2019), 10.1103/physreve.99.032111.
[34] Y. Huang, Nucl. Phys. B **938**, 594 (2019).
[35] L. Vidmar, L. Hackl, E. Bianchi, and M. Rigol, Phys. Rev. Lett. **119**, 020601 (2017).
[36] L. Vidmar and M. Rigol, Phys. Rev. Lett. **119**, 220603 (2017).
[37] J. Goold, C. Gogolin, S. R. Clark, J. Eisert, A. Scardicchio, and A. Silva, Phys. Rev. B **92**, 180202 (2015).
[38] M. V. Berry, Proc. Roy. Soc. A **400**, 229 (1985).
[39] T. Prosen, Phys. Rev. E **60**, 3949 (1999).
[40] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, Phys. Rev. Lett. **93**, 014103 (2004).
[41] C. Kollath, G. Roux, G. Biroli, and A. M. Läuchli, J. Stat. Mech. **2010**, P08011 (2010).
[42] J. P. Keating, N. Linden, and H. J. Wells, Commun. Math. Phys. **338**, 81 (2015).
[43] P. Kos, M. Ljubotina, and T. Prosen, Phys. Rev. X **8** (2018), 10.1103/physrevx.8.021062.
[44] E. Lubkin, J. Math. Phys. **19**, 1028 (1978).
[45] S. Lloyd and H. Pagels, Ann. Phys. **188**, 186 (1988).
[46] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[47] C. Nadal, S. N. Majumdar, and M. Vergassola, Phys. Rev. Lett. **104**, 110501 (2010).
[48] C. Nadal, S. N. Majumdar, and M. Vergassola, Journal of Statistical Physics **142**, 403 (2011).
[49] While it is immediate that each eigenstate of a random Hamiltonian fulfills a volume law with high probability, the distribution of all the basis states is not independent. It is thus not immediate that the probability that all d^N eigenstates fulfill a volume law is also arbitrarily close to unity for large systems. However, the probability for each of them not to follow a volume law is exponentially small in d^N according to Ref. [47, 48], while there are only d^N state, which indicates that even this latter probability is very large.
[50] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[51] B. Bauer and C. Nayak, J. Stat. Mech. **2013**, P09005 (2013).
[52] M. Friesdorf, A. H. Werner, W. Brown, V. B. Scholz, and J. Eisert, Phys. Rev. Lett. **114**, 170505 (2015).
[53] A. Anshu, New J. Phys. **18**, 083011 (2016).
[54] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Nature **551**, 579 (2017).
[55] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Nat. Phys. **14**, 745 (2018).
[56] S. Moudgalya, S. Rachel, B. A. Bernevig, and N. Regnault, Phys. Rev. B **98** (2018), 10.1103/physrevb.98.235155.
[57] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Phys. Rev. B **98** (2018), 10.1103/physrevb.98.155134.
[58] S. Moudgalya, N. Regnault, and B. A. Bernevig, Phys. Rev. B **98** (2018), 10.1103/physrevb.98.235156.
[59] S. Choi, C. J. Turner, H. Pichler, W. W. Ho, A. A. Michailidis, Z. Papi, M. Serbyn, M. D. Lukin, and D. A. Abanin, "Emergent su(2) dynamics and perfect quantum many-body scars," (2018),

1812.05561v1.

[60] P. Calabrese and J. Cardy, Phys. Rev. Lett. **96**, 136801 (2006).

[61] P. Calabrese and J.-S. Caux, J. Stat. Mech. **2007**, P08032 (2007).

[62] V. Eisler and I. Peschel, J. Stat. Mech: Theory Exp. **2007**, P06005 (2007).

[63] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, Phys. Rev. Lett. **98**, 050405 (2007).

[64] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, Phys. Rev. Lett. **100**, 030602 (2008).

[65] M. Cramer, A. Flesch, I. P. McCulloch, U. Schollwöck, and J. Eisert, Phys. Rev. Lett. **101**, 063001 (2008).

[66] T. Barthel and U. Schollwöck, Phys. Rev. Lett. **100**, 100601 (2008).

[67] A. Flesch, M. Cramer, I. P. McCulloch, U. Schollwoeck, and J. Eisert, Phys. Rev. A **78**, 033608 (2008).

[68] P. Calabrese, F. H. L. Essler, and M. Fagotti, Phys. Rev. Lett. **106**, 227203 (2011).

[69] J.-S. Caux and F. H. L. Essler, Phys. Rev. Lett. **110**, 257203 (2013).

[70] S. Goldstein, T. Hara, and H. Tasaki, Phys. Rev. Lett. **111**, 140401 (2013).

[71] A. S. L. Malabarba, L. P. García-Pintos, N. Linden, T. C. Farrelly, and A. J. Short, Phys. Rev. E **90**, 012121 (2014), arXiv:1402.1093v1.

[72] L. P. García-Pintos, N. Linden, A. S. Malabarba, A. J. Short, and A. Winter, Phys. Rev. X **7**, 031027 (2017).

[73] S. Goldstein, T. Hara, and H. Tasaki, New J. Phys. **17**, 045002 (2015).

[74] T. Farrelly, New J. Phys. **18**, 073014 (2016).

[75] P. Reimann, Nature Comm. **7**, 10821 (2016).

[76] T. R. de Oliveira, C. Charalambous, D. Jonathan, M. Lewenstein, and A. Riera, New J. Phys. **20**, 033032 (2018), arXiv:1704.06646.

[77] H. Wilming, M. Goihl, C. Krumnow, and J. Eisert, "Towards local equilibration in closed interacting quantum many-body systems," (2017), 1704.06291.

[78] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).

[79] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, 2000).

[80] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Phys. Rev. Lett. **100**, 070502 (2008).

[81] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).

[82] B. Nachtergaele, Commun. Math. Phys. **175**, 565 (1996).

[83] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quant. Inf. Comp. **5**&**6**, 401 (2006).

[84] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quant. Info. Comput. **8**, 650 (2008).

[85] X. Chen, Z.-C. Gu, and X.-G. Wen, Phys. Rev. B **82**, 155138 (2010).

[86] K. Van Acoleyen, M. Mariën, and F. Verstraete, Phys. Rev. Lett. **111**, 170501 (2013).

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

Entanglement-ergodic quantum systems equilibrate exponentially well - Supplemental Material

Section A: Proof of Equation (12) in main text and a Lemma on overlaps with product states

In this section, we provide the derivation of the inequality

$$S_\alpha(\rho) \geq S_\infty(\rho) \geq \frac{\beta-1}{\beta} S_\beta(\rho), \quad \forall \alpha \geq 0, \beta > 1. \tag{12}$$

Furthermore, we derive a key technical lemma that we will use in multiple of the subsequent arguments.

Proof of Equation (12). Let $\{q_j\}$ be the eigenvalues of ρ and let q_0 be the largest one (in case of degeneracies, pick any of the largest ones). Since $q_0^{\tilde\alpha} \leq \sum_j q_j^{\tilde\alpha}$, we have, for any $\tilde\alpha > 1$,

$$\frac{\tilde\alpha}{\tilde\alpha-1} S_\infty(\rho) = -\frac{\tilde\alpha}{\tilde\alpha-1} \log(q_0) = \frac{1}{1-\tilde\alpha} \log(q_0^{\tilde\alpha}) \tag{13}$$

$$\geq \frac{1}{1-\tilde\alpha} \log\left(\sum_j q_j^{\tilde\alpha}\right) = S_{\tilde\alpha}(\rho). \tag{14}$$

Since the Rényi entropies are monotonically decreasing in α, this proves the claim. $\square$

We now derive a helpful result concerning the overlap of product states with arbitrary pure states. Among others, our proof of Theorem 2 on the equilibration of entanglement ergodic systems will involke this lemma.

Lemma 4 (Overlap of product states with arbitrary pure states). *Let $|\Phi\rangle$ be any pure state vector on a lattice, A be a fixed subregion and σ_A be the corresponding reduced quantum state. Let $\alpha > 1$. Then*

$$|\langle\Phi|\Psi\rangle|^2 \leq e^{-\frac{\alpha-1}{\alpha} S_\alpha(\sigma_A)} \tag{15}$$

for any product vector $|\Psi\rangle = |\Psi_A\rangle \otimes |\Psi_{A^c}\rangle$.

Proof. To prove this Lemma, we make use of well-known properties of the fidelity

$$F(\rho, \sigma) := \mathrm{Tr}\left((\rho^{1/2}\sigma\rho^{1/2})^{1/2}\right) \tag{16}$$

between two quantum states. In the case of two pure states, this quantity is simply given by their overlap. Moreover, the fidelity can only increase when we trace out a subsystem, which follows from Uhlmann's theorem [78, 79] as

$$F(\rho, \sigma) \leq F(\rho_A, \sigma_A). \tag{17}$$

The reduced state of the product state vector $|\Psi\rangle$ is again pure. We denote this state vector by $|\Psi_A\rangle$ and would like to remind the reader that σ_A is the reduced state corresponding to $|\Omega\rangle$. We then have

$$|\langle\Psi|\Omega\rangle|^2 \leq F(|\Psi_A\rangle\langle\Psi_A|, \sigma_A)^2 \tag{18}$$

$$= \langle\Psi_A|\sigma_A|\Psi_A\rangle \leq \|\sigma_A\|_\infty = e^{-S_\infty(\sigma_A)}. \tag{19}$$

The claim then follows from (12). $\square$

Section B: Entropy calculations for the example

Here, we show that $|\Psi^\epsilon\rangle$ fulfills the properties claimed in Proposition 1. We will make use of Lemma 4 of Section A. In particular, for $\alpha = \infty$ the lemma states that $|\langle\Phi|\psi\rangle|^2 \leq \exp[-S_\infty(\sigma_A)]$ with $S_\infty(\sigma_A) = -\log\|\sigma_A\|$.

First let us determine the normalization of $|\Psi^\epsilon\rangle$:

$$\left\|\sqrt{1-\epsilon}\,|\Psi\rangle + \sqrt{\epsilon}\,||\Omega\rangle\right\|_2^2 = 1 + \sqrt{(1-\epsilon)\epsilon}(\langle\Psi|\Omega\rangle + c.c.). \tag{20}$$

For the maximally entangled state $|\Omega\rangle$ we have $S_\infty\left(\mathrm{Tr}_{A^c}\left(|\Omega\rangle\langle\Omega|\right)\right) = \log d_A$, where $d_A := d^{|A|}$ denotes the Hilbert-space dimension of the region A. Since $|\Psi\rangle$ is a product state, Lemma 4 yields (for $\alpha = \infty$)

$$|\langle\Psi|\Omega\rangle|^2 \leq \mathrm{e}^{-\log(d)|A|}. \tag{21}$$

Hence, the vector $\sqrt{1-\epsilon}\,|\Psi\rangle + \sqrt{\epsilon}\,|\Omega\rangle$ is normalized up to an error that is exponentially small in A. Since we are interested in scaling relations, we will in the following therefore simply assume

$$|\Psi^\epsilon\rangle = \sqrt{1-\epsilon}\,|\Psi\rangle + \sqrt{\epsilon}\,|\Omega\rangle. \tag{22}$$

One may explicitly check in the calculations below that the error introduced by this assumption is irrelevant for the scaling statements that we are making.

Let us now show that the state vector $|\Psi^\epsilon\rangle$ fulfills a volume law in terms of von-Neumann entropy, but not in terms of any Rényi entropy with $\alpha > 1$. Since $|\Psi\rangle$ is a product, its marginals on the region A and its complement A^c are pure and we denote the corresponding state vectors as $|\Psi_A\rangle$ and $|\Psi_{A^c}\rangle$, respectively. It is useful to introduce the state vector $|\Phi\rangle$ in the Hilbert space of region A given by

$$|\Phi\rangle := \frac{(\mathbb{1} \otimes \langle\Psi_{A^c}|)\,|\Omega\rangle}{\|(\mathbb{1} \otimes \langle\Psi_{A^c}|)\,|\Omega\rangle\|_2}. \tag{23}$$

Due to Lemma 4, the norm $\delta := \|(\mathbb{1} \otimes \langle\Psi_{A^c}|)\,|\Omega\rangle\|_2$ can be upper bounded as

$$\delta^2 = \max_{|\chi_A\rangle}|(\langle\chi_A|\otimes\langle\Psi_{A^c}|)\,|\Omega\rangle|^2 \leq \mathrm{e}^{-\log(d_A)} = \mathrm{e}^{-\log(d)|A|}. \tag{24}$$

With these ingredients, let us compute the reduced density matrix of the state $|\Psi^\epsilon\rangle$ on region A as

$$\mathrm{Tr}_{A^c}(|\Psi^\epsilon\rangle\langle\Psi^\epsilon|) = (1-\epsilon)\,|\Psi_A\rangle\langle\Psi_A| + \epsilon\frac{\mathbb{1}}{d_A} + \sqrt{\epsilon(1-\epsilon)}\mathrm{Tr}_{A^c}(|\Psi\rangle\langle\Omega| + \mathrm{h.c.})$$

$$= (1-\epsilon)\,|\Psi_A\rangle\langle\Psi_A| + \epsilon\frac{\mathbb{1}}{d_A} + \sqrt{\epsilon(1-\epsilon)}\delta(|\Psi_A\rangle\langle\Phi| + \mathrm{h.c.}), \tag{25}$$

where the second line follows by taking the partial trace in terms of a basis that includes $|\Psi_{A^c}\rangle$ as one of its elements. For small $\delta > 0$, the spectrum of this density matrix is well-approximated by the spectrum of the state $\xi_A := (1-\epsilon)\,|\Psi_A\rangle\langle\Psi_A| + \epsilon\mathbb{1}/d_A$, since the perturbation proportional to δ has support on a two-dimensional subspace spanned by Ψ_A and the part of Φ that is perpendicular to Ψ_A. This perturbation thus only affects two spectral values of ξ_A. Since δ is exponentially small in the size of A, we simply neglect this perturbation in what follows and work with the state ξ_A. The von-Neumann entropy of the state ξ_A can easily be calculated as

$$S_1(\xi_A) = -(1-\epsilon+\frac{\epsilon}{d_A})\log(1-\epsilon+\frac{\epsilon}{d_A}) + (d_A-1)\frac{\epsilon}{d_A}\log\left(\frac{\epsilon}{d_A}\right)$$

$$\geq \frac{d_A-1}{d_A}\epsilon\log(d_A) \approx \epsilon\log(d)|A|, \tag{26}$$

where the last approximation is exponentially good in $|A|$. We thus find a volume law in terms of the von-Neumann entropy. Let us now calculate the Rényi entropy for $\alpha > 1$. In this case we can again use the inequality

$$S_\alpha(\xi_A) \leq \frac{\alpha}{\alpha-1}S_\infty(\xi_A), \quad \alpha > 1. \tag{27}$$

Since the largest eigenvalue of ξ_A is given by $(1-\epsilon+\frac{\epsilon}{d_A})$, we then have

$$S_\alpha(\xi_A) \leq \frac{\alpha}{\alpha-1}\log\left(\frac{1}{1-\epsilon+\frac{\epsilon}{d_A}}\right) \tag{28}$$

$$\leq \frac{\alpha}{\alpha-1}\log\left(\frac{1}{1-\epsilon}\right), \tag{29}$$

where we have made use of the fact that $\log(1/x)$ is monotonically decreasing in x. We thus find that all Rényi entropies for $\alpha > 1$ are upper bounded by a constant independent of the size of A.

Section C: Extensive entropy for states prepared by finite-depth circuits and MPS

In this appendix, we show that a state vector $|\Psi\rangle$ that is prepared by a translationally invariant, finite-depth, local quantum circuit has either extensive Rényi-2 entropy on a suitable sub-system or is a product state itself. By a finite-depth, local quantum circuit we mean a unitary operator U that can be written as a concatenation of D unitaries which consists of a tensor product of unitary operators, each of which acts on spins that are separated by a lattice distance at most k and we assume that D does not depend on the system size. By translationally invariance, we simply mean that the final state is invariant under shifts of the underlying lattice by some finite lattice vector. We assume this to simplify the discussion, but expect similar results to hold for generic such states. We note that states prepared by a local, finite depth quantum circuit are a special sub-class of MPS. We discuss the case of general MPS further below.

The crucial property that we will make use of is that U has a strict "light-cone" for the spreading of correlations. This implies that spins that are separated by a distance more than $2kD+1$ are uncorrelated. Let us hence consider as subsystem A a sub-lattice $\tilde{\Lambda}$ of spins that are separated by at least a distance $2kD + 1$. Then the reduced state on A is a product state:

$$\rho_A = \bigotimes_{x\in\tilde{\Lambda}} \rho_x. \tag{30}$$

By translational invariance, we furthermore ensure that all the ρ_x are identical, i.e. $\rho_x = \varrho$. Since the Rényi entropies are additive over tensor-products, we thus find that

$$S_2(\rho_A) = |\tilde{\Lambda}|S_2(\varrho) \sim NS_2(\varrho). \tag{31}$$

Note that by assumption ϱ does not depend on N. Hence, if the state has a non-extensive amount of entropy it must have $S_2(\varrho) = 0$ and in consequence the state ϱ must be pure. By repeating this argument for different sub-lattices $\tilde{\Lambda}$ we conclude that every spin is in a pure state and hence $|\Psi\rangle$ is a product state. In conclusion, $|\Psi\rangle$ has either extensive Rényi-2 entropy or is a product state.

Let us now turn to the case of generic matrix product states. For simplicity, we again consider here the case of translationally invariant MPS, which take the form

$$|\mathrm{MPS}[A]_N\rangle := \frac{1}{\sqrt{Z}} \sum_{i_1,...i_N=1}^{d} \mathrm{Tr}(A_{i_1} \cdots A_{i_N}) \, |i_1,\ldots,i_N\rangle, \tag{32}$$

where the $\{A_i\}_{i=1}^{d}$ are a set of $D \times D$ matrices (D in this case called *bond dimension* and is related to, but not the same as the D above). In the following we assume that the A_i and the bond dimension are constant, i.e., independent of N. Generic MPS are so-called *injective* MPS, which essentially means that the matrices A_i cannot be brought into a common block-diagonal form by a similarity transformation. In this case the MPS has a finite-correlation length. It is useful to think of injective MPS as generalizations from the above setting of local quantum circuits to the case of quasi-local quantum circuits. Indeed, for injective MPS it can be shown (using the results of Ref. [80]) that the reduced state on a sub-lattice $\tilde{\Lambda}$ in which the spins are separated by a distance l is exponentially close to a product state,

$$\left\| \rho_{\tilde{\Lambda}} - \bigotimes_{x\in\tilde{\Lambda}} \rho_x \right\|_1 \leq |\tilde{\Lambda}|Ce^{-l/\xi}, \tag{33}$$

for some constants $C, \xi > 0$. Thus, at least if l grows weakly with the system size as $\log(N)$, this strongly suggests that the Rényi entropies are extensive on $\tilde{\Lambda}$. This in turn leads to a Rényi entanglement entropy of order $N/\log(N)$ and, correspondingly, to equilibration with a precision of order $\exp(-N/\log(N))$. However, we have not been able to give a formal proof of this statement so far.

Let us remark, however, that for the proof of Theorem 2 it suffices to show that energy eigenstates with finite energy density have exponentially small overlap with product states. We will now prove that generic, translationally invariant MPS fulfill this criterion. Thus exponentially good equilibration also follows if the eigenstates of the Hamiltonian are generic MPS with sufficiently strong entanglement. While we assume translational invariance here, we expect that similar results also hold (with high probability) in the case where the local MPS-tensors are drawn at random.

Lemma 5 (Exponentially small overlap of MPS with product states)**.** *Consider translationally invariant, injective MPSs* $|\mathrm{MPS}[A]_N\rangle$ *with constant bond dimension. Then either* $|\mathrm{MPS}[A]_N\rangle$ *is a product or there exists a constant* $\kappa > 0$ *such that*

$$|\langle\phi|^{\otimes N} \, |\mathrm{MPS}[A]_N\rangle| \leq e^{-\kappa N} \tag{34}$$

for all $|\phi\rangle \in \mathbb{C}^d$.

Before giving the proof of this Lemma, let us note that we have chosen the product state here to be translationally invariant for convenience. One would arrive at the same conclusion if the product state in question has higher periodicity by coarse-graining the lattice. For example, it could be a charge-density wave of the form $|\uparrow, \downarrow, \uparrow, \downarrow, \ldots\rangle$. All that is required is that the state vector can be written as $|\psi\rangle^{N/m}$ for some m corresponding to a coarse graining on the lattice. Then the MPS in question is either a product state on this coarse grained lattice or has exponentially small overlap with all product states on this coarse grained lattice.

Proof. The main ingredient for the proof is that an injective MPS is always the unique ground-state of a local, frustration free Hamiltonian [81–84]. This implies that for each lattice site x there exists a local projector P_x on a finite neighborhood V_x of the lattice sites $x \in \Lambda$ such that

$$P_x \otimes \mathbb{1}_{\Lambda \setminus V_x} |\mathrm{MPS}[A]_N\rangle = |\mathrm{MPS}[A]_N\rangle. \tag{35}$$

Furthermore, $|\mathrm{MPS}[A]_N\rangle$ is the only vector state fulfilling this condition for all x. Let us choose a sub-lattice $\tilde{\Lambda}$ so that the neighborhoods V_x do not overlap. Then the operators $P_x \otimes \mathbb{1}_{\Lambda \setminus V_x}$ commute. In the following, we will omit the identity for brevity of notation and simply write P_x instead of $P_x \otimes \mathbb{1}_{\Lambda \setminus V_x}$. Let $|\Phi\rangle = |\phi\rangle^{\otimes N}$. Then we have

$$|\langle\Phi|\mathrm{MPS}[A]_N\rangle|^2 = \langle\Phi| \prod_{x \in \tilde{\Lambda}} P_x |\mathrm{MPS}[A]_N\rangle\langle\mathrm{MPS}[A]_N| \prod_{y \in \tilde{\Lambda}} P_y|\Phi\rangle \tag{36}$$

$$\leq \langle\Phi| \prod_{x \in \tilde{\Lambda}} P_x|\Phi\rangle \tag{37}$$

$$= \prod_x \langle\Phi_{V_x}|P_x|\Phi_{V_x}\rangle = \mathrm{e}^{-c|\tilde{\Lambda}|}, \tag{38}$$

where we have defined $|\Phi_{V_x}\rangle := \otimes_{x \in V_x} |\phi\rangle$ and implicitly defined $c > 0$. In the first equality we used (35), in the second line we used that the projector onto the MPS is positive and in the last line we used that $|\Phi\rangle$ is a product state and that the P_x are translations of each other since the MPS under question is translationally invariant. Now suppose that $c = 0$, i.e.,

$$\langle\Phi_{V_x}|P_x|\Phi_{V_x}\rangle = 1. \tag{39}$$

In this case we deduce that $P_x|\Phi\rangle = |\Phi\rangle$ for all x. However, by assumption $|\mathrm{MPS}[A]_N\rangle$ is the unique such state vector and we get $|\mathrm{MPS}[A]_N\rangle = |\Phi\rangle$. Now consider

$$\sup_{|\phi\rangle}\langle\Phi|P_X|\Phi\rangle := \mathrm{e}^{-\tilde{\kappa}}, \tag{40}$$

where again $|\Phi\rangle = |\phi\rangle^{\otimes N}$. Since the the set of pure states is compact, the supremum is in fact achieved. But by assumption it cannot be equal to 1, therefore $\tilde{\kappa} > 0$. This step completes the proof by identifying $\kappa > 0$ as the constant that fulfills $\kappa N = \tilde{\kappa}|\tilde{\Lambda}|$. $\qquad\square$

While we have above only considered the case of matrix-product states, we note that the proof applies to higher-dimensional tensor network states such as *projected entangled pair states* (PEPS), since the only property that we used is that the state in question is a unique, translationally invariant ground state of a local, frustration-free Hamiltonian.

Section D: Proof of Theorem 2

Here we provide the proof of the main result. Before we spell out the details, let us explain the basic structure of the proof. Due to the monotonicity of the Rényi entropies in α it suffice to derive a lower bound for S_∞ in order to get the statement of Theorem 2 for all values of $\alpha \geq 0$. In general, there exists a basis of energy eigenstates that diagonalizes the time averaged state ω_Λ. The eigenvalues of ω_Λ are given by the absolute-squared overlap of the initial state with these energy eigenstates. Therefore, an exponential upper bound on the maximal overlap of the initial product state vector $|\Psi\rangle_\Lambda$ with an energy eigenstate implies a lower bound on S_∞. Let us denote the mean energy density of $|\Psi\rangle_\Lambda$ by e. We establish a bound on the overlaps by two separate arguments – one applying to eigenstates whose energy density deviates from e by *at most* δ and one applying to eigenstates with energy density that deviates from e by *at least* δ. Taken together we thus bound the overlaps with *all* energy eigenstates for a given system size $|\Lambda|$, which establishes our final bound. The constant δ is a small constant that we choose in the course of the proof.

Before we continue, note that entanglement-ergodicity is defined for *sequences* of Hamiltonians of increasing system sizes, whereas the statement of Theorem 2 applies to every system size larger than $N_0(e)$, which will be identified below.

As stated above, we first show that $|\Psi\rangle_\Lambda$ has little overlap with energy eigenstates that differ by a large amount in energy. This is due to the fact that the energy distribution of product states is sharply peaked around its mean. More precisely, we make use of the following Theorem due to Anshu [53], whose formulation we slightly adapt to the current notation:

Theorem 6 (Concentration bound [53]). *Let H_Λ be a strictly local, uniformly bounded Hamiltonian. Then there exists a constant $m > 0$ independent of the system size such that for any product state $|\Psi\rangle_\Lambda$ with energy-density e and any $\delta > \sqrt{2/(mN)}$, we have*

$$\langle\Psi|\,\Pi_{(-\infty,N(e-\delta)]}\,|\Psi\rangle \leq \mathrm{e}^{-m\delta^2 N}, \tag{41}$$

$$\langle\Psi|\,\Pi_{[N(e+\delta),\infty)}\,|\Psi\rangle \leq \mathrm{e}^{-m\delta^2 N}, \tag{42}$$

where Π_I denotes the projection onto the subspace spanned by energy eigenstates with energies in the interval I.

In particular, this theorem implies that for any $\delta > \sqrt{2/(mN)}$, a product state vector $|\Psi\rangle_\Lambda$ has exponentially small overlap with energy eigenvectors $|e'\rangle_\Lambda$ whose energy density e' deviates by more than δ from its mean energy e,

$$|\langle\Psi|e'\rangle_\Lambda|^2 \leq \mathrm{e}^{-m\delta^2 N} \text{ if } |e' - e| \geq \delta > \sqrt{2/(mN)} \tag{43}$$

It remains to bound the overlap of $|\Psi\rangle_\Lambda$, which has energy density e, with eigenstates with energy density within the interval $[e-\delta, e+\delta]$. Here, the main ingredient is the fact that product states have exponentially small overlap with energy eigenstates that fulfill a weak volume law. By the definition of entanglement-ergodic systems, for sufficiently large systems and for any energy eigenstate state $\rho_\Lambda(e') = |e'\rangle\langle e'|_\Lambda$ with energy density e' there exists a region A_Λ such that $S_2(\rho_{A_\Lambda}(e')) \geq g(e')N$ provided that $N \geq N_0$ (where N_0 is given by entanglement-ergodicity). Using Lemma 4, we find that the overlap of a product state with eigenstates of energy density e' is hence bounded as

$$|\langle\Psi|e'\rangle_\Lambda|^2 \leq \mathrm{e}^{-\frac{1}{2}g(e')N}. \tag{44}$$

We now use the regularity of g. Let $K > 0$ be its Lipschitz-constant. Then we know that $g(e') \geq g(e) - K\delta$ for all $e' \in [e-\delta, e+\delta]$. We choose

$$N_0(e) := \max\left\{N_0, \frac{8K^2}{mg(e)^2}\right\}, \quad \delta = \frac{g(e)}{2K}. \tag{45}$$

This ensures that $\delta > \sqrt{2/(mN)}$ for all $N > N_0(e)$. Then we obtain

$$|\langle\psi|e'\rangle_\lambda|^2 \leq \mathrm{e}^{-\frac{1}{4}g(e)N}, \quad e' \in [e-\delta, e+\delta]. \tag{46}$$

In summary, we find that *all* eigenstates have exponentially small overlap with $|\Psi\rangle_\Lambda$, either because their energy density differs substantially from e or because they fulfill a weak volume law. Taking

$$k(e) = \frac{1}{4}g(e)\min\{1, m\,g(e)/K^2\}, \tag{47}$$

we thus find $|\langle\psi|e_i\rangle_\lambda|^2 \leq \exp(-k(e)N)$ for all energy eigenstates $|e_i\rangle_\Lambda$. Therefore

$$S_\alpha(\omega_\Lambda) \geq S_\infty(\omega_\Lambda) \geq k(e)N, \tag{48}$$

which completes the proof.

Section E: Stability of entanglement-ergodicity under quasi-local unitaries

In this section we prove the stability of entanglement-ergodicity under time evolution generated by a local Hamiltonian. More precisely, we establish the following result:

Proposition 2 (Stability under quasi-local unitaries). *Let H_Λ be an entanglement-ergodic system where the regions A_Λ have asymptotically vanishing surface-to-volume ratio and let U'_Λ be a quasi-local sequence of unitaries. Then the system with Hamiltonian $U'_\Lambda H_\Lambda U'^\dagger_\Lambda$ is entanglement-ergodic.*

Here, by a quasi-local unitary, we mean the following: Consider a sequence of unitaries U'_Λ. We call the sequence U'_Λ quasi-local if there exists a sequence of l-local Hamiltonians $H'_\Lambda = \sum_x h'_x$ with $\|h'_x\| \leq 1$ and a finite time T such that

$$U'_\Lambda = \mathrm{e}^{-\mathrm{i}H'_\Lambda t_\Lambda}, \quad t_\Lambda \leq T. \tag{49}$$

Examples of quasi-local unitaries are given by finite-depth local quantum circuits, in particular ones that reflect deformations of a quantum state within the same quantum phase [85]. This result above follows from the following Lemma. It can be seen as the Rényi-2 analogue of the result on entanglement rates for the von Neumann entropy laid out in Ref. [86].

Lemma 7 (Rényi-2 entangling rate). *Let ρ_{AB} be a bipartite state that evolves under a Hamiltonian $H = H_A + H_B + V$ with $H_{A,B}$ acting on systems A, B, respectively. Decompose V into a Hermitian operator basis as $V = \sum_{i,j=1}^{v} c_{i,j} A_i \otimes B_j$ with $\|A_i\| = \|B_j\| = 1$ and $A_i^\dagger = A_i, B_j^\dagger = B_j$ for all i, j. We define the Rényi-2 entangling rate of the subsystem A as*

$$\Gamma(A, H) := \left. \frac{\mathrm{d}S_2(\rho_A)}{\mathrm{d}t} \right|_{t=0} = 2\mathrm{i} \frac{\mathrm{Tr}(\rho_A \mathrm{Tr}_B([V, \rho_{AB}]))}{\mathrm{Tr}(\rho_A^2)}. \tag{50}$$

Then

$$|\Gamma(A, H)| \leq 4 \|\mathbb{C}\|_1 = 4 \sum_{i,j} |c_{i,j}|. \tag{51}$$

Before providing the proof of the Lemma, let us discuss the application to a strictly local many-body systems. In this case $H = \sum_x h_x$ with $\|h_x\| \leq 1$ and where the support of each term h_x has diameter bounded by some constant. Choose a region A and decompose $H = H_A + H_{A^c} + H_{\partial A}$, where $H_{\partial A}$ contains only those terms h_x that have support both on A and A^c. Clearly, the number of these terms scales like the boundary of A. We then have

$$\Gamma(A, H) = 2\mathrm{i} \frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} = 2\mathrm{i} \left[\frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H_{\partial A}, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} + \frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H_A, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} + \frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H_{A^c}, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} \right] \tag{52}$$

$$= 2\mathrm{i} \left[\frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H_{\partial A}, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} + \frac{\mathrm{Tr}\left(\rho_A[H_A, \rho_A]\right)}{\mathrm{Tr}\left(\rho_A^2\right)} + \frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H_{A^c}, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} \right] \tag{53}$$

$$= 2\mathrm{i} \frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([H_{\partial A}, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} = \sum_{x \in \partial A} 2\mathrm{i} \frac{\mathrm{Tr}\left(\rho_A \mathrm{Tr}_{A^c}([h_x, \rho])\right)}{\mathrm{Tr}\left(\rho_A^2\right)} = \sum_{x \in \partial A} \Gamma(A, h_x), \tag{54}$$

where $x \in \partial A$ denotes those x for which the support of h_x is neither fully contained in A nor in A^c. The third line follows from cyclicity of the trace for the second term and by choosing the eigenbasis of H_{A^c} to trace-out system A^c for the third term in the second line. We can now use the lemma to bound $|\Gamma(A, h_x)| \leq 4 \|\mathbf{C}_x\|_1$, where $\mathbf{C}$ is the vector of coefficients $c_{i,j}$ associated to the Hamiltonian term h_x. We hence obtain

$$|\Gamma(A, H)| \leq 4 |\partial A| \max_{x \in \Lambda} \|\mathbf{C}_x\|_1. \tag{55}$$

Since we assume that the Hamiltonian is strictly local and uniformly bounded, there is a maximum value of $\|\mathbf{C}_x\|_1$ independent of the system size. Integrating over time we obtain

$$|S_2(\rho_A(t)) - S_2(\rho_A(0))| \leq 4t |\partial A| \max_{x \in \Lambda} \|\mathbf{C}_x\|_1. \tag{56}$$

Thus, time evolution under a local Hamiltonian for a fixed time t (independent of the system size) cannot change the volume law of the Rényi-2 entropy as long as the surface area of A is asymptotically vanishingly small in comparison with its volume. Using the equivalence of volume laws for different values of $\alpha > 1$, we thus find that an entanglement-ergodic system remains entanglement-ergodic (possibly with a renormalized function g) under such time evolution. Clearly, the same is true if a finite number of such time evolutions U'_Λ is concatenated as long as this number does not grow with the size of the lattice.

Proof of Lemma 7. First we write the entangling rate as (using the same method as above)

$$\Gamma(A, H) = 2\mathrm{i} \frac{\mathrm{Tr}(\rho_A \mathrm{Tr}_B([V, \rho_{AB}]))}{\mathrm{Tr}(\rho_A^2)}. \tag{57}$$

We now use $\mathrm{Tr}_B([A \otimes B, \rho_{AB}]) = [A, \mathrm{Tr}_B(\rho_{AB}B)]$, which follows from the cyclicity of the trace. This allows us to write

$$|\mathrm{Tr}(\rho_A \mathrm{Tr}_B([V, \rho_{AB}]))| \leq \sum_{i,j} |c_{i,j}| |\mathrm{Tr}(\rho_A[A_i, \mathrm{Tr}_B(\rho_{AB}B_j)])| \tag{58}$$

$$\leq 2 \sum_{i,j} |c_{i,j}| \left[\mathrm{Tr}(\rho_A^2 A_i^2) \mathrm{Tr}(\mathrm{Tr}_B(\rho_{AB}B_j)^\dagger \mathrm{Tr}_B(\rho_{AB}B_j)) \right]^{1/2}, \tag{59}$$

where we have used the Cauchy-Schwarz inequality. We can now use the fact that $\|A_i\| = 1$ to obtain

$$|\mathrm{Tr}(\rho_A \mathrm{Tr}_B([V, \rho_{AB}]))| \leq 2 \sum_{i,j} |c_{i,j}| \left[\mathrm{Tr}(\rho_A^2) \mathrm{Tr}(\mathrm{Tr}_B(\rho_{AB}B_j)^\dagger \mathrm{Tr}_B(\rho_{AB}B_j)) \right]^{1/2}. \tag{60}$$

Let us now consider the term $\mathrm{Tr}\left(\mathrm{Tr}_B(\rho_{AB}B_j)^\dagger \mathrm{Tr}_B(\rho_{AB}B_j)\right)$. Let us decompose B_j into its eigenbasis and write

$$\mathrm{Tr}_B(\rho_{AB}B_j) = \sum_\alpha b_j^\alpha \mathrm{Tr}_B(\rho_{AB}P_j^\alpha), \tag{61}$$

where P_j^α are the eigenprojectors of B_j and b_j^α its eigenvalues. Upper bounding each b_j^α by the maximum one, we then find the operator inequality

$$\mathrm{Tr}_B(\rho_{AB}B_j)^\dagger \mathrm{Tr}_B(\rho_{AB}B_j) = \sum_{\alpha,\beta} b_j^\alpha b_j^\beta \mathrm{Tr}_B(\rho_{AB}P_j^\alpha) \mathrm{Tr}_B(\rho_{AB}P_j^\beta) \tag{62}$$

$$\leq \|B_j\|^2 \, \mathrm{Tr}_B(\rho_{AB})^\dagger \mathrm{Tr}_B(\rho_{AB}) \tag{63}$$

$$= \|B_j\|^2 \, \rho_A^2. \tag{64}$$

Since $\|B_j\| = 1$, we thus finally obtain

$$|\mathrm{Tr}(\rho_A \mathrm{Tr}_B([V, \rho_{AB}]))| \leq 2 \sum_{i,j} |c_{i,j}| \mathrm{Tr}(\rho_A^2) = 2 \|\mathbf{C}\|_1 \mathrm{Tr}(\rho_A^2). \tag{65}$$

We thus obtain for the entangling rate

$$|\Gamma(A, H)| \leq 4 \|\mathbf{C}\|_1. \tag{66}$$

This step completes the proof. $\qquad\square$

The above lemma shows stability of entanglement-ergodic systems under arbitrary quasi-local unitaries. We would like to highlight that stability under finite-depth quantum circuits follows immediately from the fact that finite-depth quantum circuits can only change the Rényi entropy of some region by an amount proportional to the area of the boundary of the region. This is due to the fact that a finite-depth quantum circuit can only (de-)correlate degrees of freedom that are within a finite distance from each other.

Section F: Quasi-local initial states

Here, we discuss the case of initial state vectors $|\phi_\Lambda\rangle = U'_\Lambda |\Psi\rangle_\Lambda$, where U'_Λ is an arbitrary quasi-local unitary and $|\Psi\rangle_\Lambda$ is a product vector. In this case, we can show that

$$S_\alpha(\omega) \geq k(e) N^{\frac{1}{\nu+1}}, \tag{67}$$

which in turn implies equilibration with precision of order $\exp(-N^{\frac{1}{\nu+1}})$. To see this, first note that due to stability of entanglement-ergodicity under quasi-local unitaries almost all the steps of the proof of Theorem 2 transfer directly. In particular the overlap of $|\Psi\rangle_\Lambda$ with energy eigenstates of energy density $e' \in [e - \delta, e + \delta]$ is exponentially small in N due to entanglement-ergodicity. The only step which cannot be transferred is the argument that the overlap with energy eigenstates whose energy density differs by more than δ from e is exponentially small. Indeed, the best general result for quasi-local states known to us, from Ref. [53], shows that this overlap is only sub-exponentially small. More precisely, the following theorem holds:

Theorem 8 ([53]). *Let H be a strictly local Hamiltonian on a ν-dimensional lattice with N sites, let ρ be a quantum state with correlation length $\xi > 0$ and $\langle H \rangle_\rho = \text{tr}(H\rho)$ be the average energy of ρ. Furthermore denote by Π_X the projector onto all energy eigenstates with energy in the set $X \subset \mathbb{R}$. For $a \geq (2^{O(\nu)}/N\xi)^{1/2}$, it holds that*

$$\text{tr}(\rho\Pi_{[\langle H \rangle_\rho + Na, \infty]}) \leq O(\xi) \exp\left(-\frac{(Na^2\xi)^{1/(\nu+1)}}{O(1)\nu\xi} \right) \tag{68}$$

and

$$\text{tr}(\rho\Pi_{[0, \langle H \rangle_\rho - Na]}) \leq O(\xi) \exp\left(-\frac{(Na^2\xi)^{1/(\nu+1)}}{O(1)\nu\xi} \right). \tag{69}$$

Note that this theorem not only bounds the overlap with individual energy eigenstates, as is needed to prove our result, but in fact the cumulative overlap with *all* energy eigenstates whose energy density deviates from e. Since these are the vast majority of all states in the exponentially large Hilbert-space, it seems highly plausible that despite the fact that the cumulative overlap with all these states only scales sub-exponentially, the overlap with each individual such energy eigenstate is exponentially small, which would imply exponentially good equilibration also for quasi-local initial states. We thus conjecture that (possibly under mild additional assumptions) it is possible to prove exponentially good equilibration in entanglement-ergodic systems for arbitrary initial states with a finite correlation length.

Section G: Entanglement ergodicity and eigenstate thermalization

In this section we argue that entanglement ergodicity follows from the eigenstate thermalization hypothesis, provided that one assumes that the latter applies to subsystems that contain a finite fraction of the total system. While we believe that our observations make this statement highly plausible, we leave the formal proof as an open problem. We thus assume that for any energy density e, the reduced state of some subsystem A is given by the reduced state of the Gibbs state with inverse temperature $\beta(e)$ corresponding to the energy density e,

$$\text{Tr}_{A^c}\left(|e\rangle\langle e|\right) = \text{Tr}_{A^c}\left(\frac{e^{-\beta(e)H}}{Z_{\beta(e)}} \right). \tag{70}$$

We further only consider energy densities for which the thermal state has a finite correlation length and for simplicity assume that the system is translationally invariant. In this case the thermal state of the total system at inverse temperature β has a finite equilibrium free energy density $f(\beta)$. Let us take the convention that the ground-state of the Hamiltonian has energy zero. In this convention, the total free energy $F_\beta = Nf(\beta) = -\frac{1}{\beta}\log(Z_\beta)$ is negative. Furthermore the ground-state probability $p(\beta)$ is given by

$$p(\beta) = 1/Z = e^{\beta F_\beta} \tag{71}$$

and hence the Rényi-∞ entropy of the total system is given by

$$S_\infty = -\beta F_\beta, \tag{72}$$

which is a positive, extensive quantity, since the free energy is extensive. We thus correspondingly expect that to leading order, i.e., up to boundary terms, the corresponding entropy of a contiguous subsystem A is given by

$$S_\infty(A) = -\beta|A|f(\beta(e)), \tag{73}$$

which then implies entanglement-ergodicity. This is further supported by the fact that the reduced state on region A is well approximated by the reduced state of the Gibbs state of the Hamiltonian H restricted to A plus a "buffer" region around it – a phenomenon known as locality of temperature [87].