Commun. Math. Phys. 340, 499–561 (2015)
Digital Object Identifier (DOI) 10.1007/s00220-015-2473-y

Communications in
Mathematical
Physics

Thermalization and Canonical Typicality in Translation-Invariant Quantum Lattice Systems

Markus P. Müller[1,2,3], **Emily Adlam**[4], **Lluís Masanes**[5,6], **Nathan Wiebe**[7,8,9]

[1] Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 19, 69120 Heidelberg, Germany
E-mail: m.mueller@thphys.uni-heidelberg.de
[2] Department of Applied Mathematics, Department of Philosophy, University of Western Ontario, 1151 Richmond Street, London, ON N6A 5B7, Canada
[3] Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada
[4] Centre for Quantum Information and Foundations, DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, UK
[5] H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, UK
[6] Department of Physics and Astronomy, University College London, London WC1E 6BT, UK
[7] Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052, USA
[8] Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada
[9] Department of Combinatorics and Opt., University of Waterloo, Waterloo, ON, N2L 3G1, Canada

Received: 4 August 2014 / Accepted: 7 July 2015
Published online: 10 September 2015 – © Springer-Verlag Berlin Heidelberg 2015

Abstract: It has previously been suggested that small subsystems of closed quantum systems thermalize under some assumptions; however, this has been rigorously shown so far only for systems with very weak interaction between subsystems. In this work, we give rigorous analytic results on thermalization for translation-invariant quantum lattice systems with finite-range interaction of arbitrary strength, in all cases where there is a unique equilibrium state at the corresponding temperature. We clarify the physical picture by showing that subsystems relax towards the reduction of the global Gibbs state, not the local Gibbs state, if the initial state has close to maximal population entropy and certain non-degeneracy conditions on the spectrum are satisfied. Moreover, we show that almost all pure states with support on a small energy window are locally thermal in the sense of canonical typicality. We derive our results from a statement on equivalence of ensembles, generalizing earlier results by Lima, and give numerical and analytic finite-size bounds, relating the Ising model to the finite de Finetti theorem. Furthermore, we prove that global energy eigenstates are locally close to diagonal in the local energy eigenbasis, which constitutes a part of the eigenstate thermalization hypothesis that is valid regardless of the integrability of the model.

Contents

1. Introduction

How do closed quantum systems thermalize? The last few years have seen a resurgence of interest in this old question, motivated by new experimental [1] and numerical [2] methods, relying on new ideas and methods from quantum information theory [3–10]. Clearly, closed quantum systems in any given pure initial state cannot literally thermalize: unitary time evolution enforces that the global state remains pure and will never become thermal, unless there is at least a tiny interaction with some environment. However, small subsystems of closed quantum systems can equilibrate in a certain sense, as entanglement between the subsystem and its remainder will lead to locally mixed states, and one may hope that these will in many cases resemble the ensembles of statistical physics.

Along these lines, it was suggested in [3] that typical pure quantum states in many-body systems resemble thermal states on small subsystems due to entanglement, a property called "canonical typicality". However, no rigorous mathematical formulation of this was given in [3]. Almost at the same time, it was rigorously proven in [4] that typical pure quantum states in subspaces of bipartite Hilbert spaces are locally close to some equilibrium state. However, this equilibrium state is not thermal in general. This raises the question of what conditions are needed to ensure that the local equilibrium state will be thermal, i.e., a Gibbs state.

In addition to these kinematical results, there has been major progress in understanding how closed quantum systems equilibrate dynamically [5–9]. Regarding the emergence of the Gibbs state, the situation is similar to the kinematical case: the subsystems approach some equilibrium state (for most times in some time interval), which is however not thermal in general. The question is thus the same: under what conditions will the equilibrium state be thermal?

Important progress on this question was made in [10]: a rigorous bound on the distance $\mathcal{D}$ between the local equilibrium state and a thermal state was established. This result has two drawbacks, however. First, the given bound is rather cumbersome, which is due to the great generality of considering arbitrary Hamiltonians. Second, and more importantly, the upper bound on the distance $\mathcal{D}$ grows with the operator norm of the interaction Hamiltonian which couples the subsystem to its surroundings. Thus, the bound becomes trivial as soon as the boundary of the subsystem becomes moderately large, or the interaction becomes strong.

In this work, we give rigorous analytic proofs of dynamical and kinematic formulations of thermalization for interactions of finite range, but arbitrary strength. By restricting to the special case of translation-invariant lattice systems as in Fig. 1, we are able to prove the common belief that small subsystems are indeed close to thermal, under various natural conditions on the spectrum and the initial state that depend on the specific setup and boundary conditions. Our work also clarifies how thermalization should generally be formalized by showing that the resulting state will in general not be the

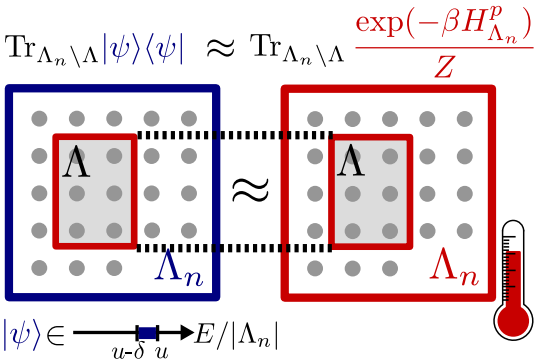

Fig. 1. Canonical typicality. A rectangular lattice Λ_n evolves according to a translation-invariant finite-range interaction Hamiltonian $H^p_{\Lambda_n}$, where "p" is for periodic boundary conditions (the case of arbitrary boundary conditions is treated in the Sect. 3). If $|\psi\rangle$ is a generic state occupying only energies E with $u-\delta \leq E/|\Lambda_n| \leq u$, then small subsystems $\Lambda \subset \Lambda_n$ will, for large n, behave as if the full system was in a Gibbs state of the corresponding temperature, for all possible measurements in the subsystem. Dynamically, the same will be true for $|\psi(t)\rangle$ for most times t if the initial state $|\psi(0)\rangle$ has close to maximal population entropy, and the spectrum satisfies certain non-degeneracy conditions

local Gibbs state; rather, it is the reduction of the global system's Gibbs state. This identification is made clear from the fact that the expected distance between the local reduced state and the thermal state goes to zero in the thermodynamic limit. In contrast, we show that boundary effects cause the local Gibbs state in general to remain distinct from the thermal state even in the thermodynamic limit. This shows why earlier work led to bounds on the distance that necessarily grow with the interaction strength.

We are further able to provide concrete finite-size bounds, rather than asymptotic bounds, for two cases of interest. We give tight analytic bounds for the distance between the reduction of a typical global pure state and the local Gibbs state in the non-interacting case, which already turns out to be a non-trivial problem, and we give numerical finite-size estimates for interacting models in one lattice dimension. Building on the results by Low [26], we also show that the kinematical result on canonical typicality remains true if global pure states are not drawn with respect to the unitarily invariant measure (which is hard to implement), but according to an approximation of this measure (an "8-design") that can be sampled efficiently.

Finally, we address the question of whether the given thermalization results can hold even on the level of single energy eigenstates, as conjectured in the eigenstate thermalization hypothesis (ETH) [33,34]. In a nutshell, the ETH claims that global energy eigenstates are locally close to a thermal state. It is easy to see that the ETH cannot be true for all the models that we consider, and that additional assumptions (along the lines of nonintegrability) are needed. However, we prove a result that constitutes a part of the ETH that is true for all models with finite interaction range: global energy eigenstates are locally close to diagonal in the local energy eigenbasis. We hope that this result (proven via Lieb-Robinson bounds) may serve as a first step towards a complete resolution of the ETH in future work.

2. Summary of the Main Results

We provide a self-contained summary of the main results of this paper in this section, focusing on periodic boundary conditions. The case of arbitrary boundary conditions

will be treated in Sect. 3. While the detailed definitions will be given in Sect. 3 (and are close to [13]), here we describe the setup and notation in a less formal way.

Our work considers the thermalization of interacting d-dimensional systems in a cubic or rectangular lattice in ν spatial dimensions. These spins are constrained to interact with each other via finite-range translationally invariant Hamiltonians with arbitrary boundary conditions. Although these restrictions are stringent, many models relevant to condensed matter physics, such as the Ising and Heisenberg models, satisfy these requirements.

We introduce the following notation to describe the lattice. We define the set of lattice sites to be $\Lambda := [\lambda_1, \mu_1] \times \cdots \times [\lambda_\nu, \mu_\nu]$, where $[\lambda, \mu] \subset \mathbb{Z}$ denotes the interval of integers between λ and $\mu \geq \lambda$. In particular, we consider sequences of regions $\Lambda_1 \subset \Lambda_2 \subset \Lambda_3 \ldots$ that converge to the full infinite lattice $\mathbb{Z}^\nu$; for example, we may have the sequence of hypercubes $\Lambda_n = [-n, n]^\nu$. The physical interpretation is that a region Λ_n describes the actual physical system in the laboratory, and a subregion $\Lambda \subset \Lambda_n$ describes a small subsystem, cf. Fig. 1. The number of sites in a region Λ is denoted $|\Lambda|$. The "particles" located at each of these sites carry a d-dimensional Hilbert space $\mathbb{C}^d$.

Time evolution in Λ_n is determined by a Hamiltonian $H_{\Lambda_n}^{BC}$ where the superscript explicitly denotes the type of boundary conditions that the Hamiltonian satisfies. The choice of Hamiltonian is subject to some conditions defined as follows. To every finite region $X \subset \mathbb{Z}^\nu$, we associate a self-adjoint operator h_X, and define the Hamiltonian with open boundary conditions to be $H_\Lambda := \sum_{X \subset \Lambda} h_X$. We assume translation-invariance, i.e. h_{X+y} equals h_X (translated to the corresponding lattice sites), and finite-range of interaction, i.e. there is some $r < \infty$ such that $h_X = 0$ whenever the diameter of X is larger than r. In the following, we will exclude the case that the map $X \mapsto h_X$ is, up to physical equivalence [13], everywhere identically zero. As a simple example in one dimension, the Heisenberg model $H_{[1,n]} = -J \sum_{i=1}^{n-1} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} - h \sum_{i=1}^{n} \sigma_i^Z$, with Pauli matrices $\boldsymbol{\sigma} = (\sigma^X, \sigma^Y, \sigma^Z)$, fits into this framework, if we define h_X as $-h\sigma_i^Z$ if $X = \{i\}$ for some integer i, as $-J\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1}$ if $X = \{i, i+1\}$, and as zero for all other X.

The Hamiltonian with open boundary conditions, H_{Λ_n}, can be augmented with additional non-translationally invariant terms on the boundary of Λ_n to obtain some $H_{\Lambda_n}^{BC}$. The case of periodic boundary conditions is of particular importance to the remainder of the discussion and we denote such Hamiltonians by $H_{\Lambda_n}^{p}$. More general boundary conditions are also permitted. The only assumption will be that $\| H_{\Lambda_n}^{BC} - H_{\Lambda_n} \|_\infty / |\Lambda_n| \to 0$ as $n \to \infty$, where $\| \cdot \|_\infty$ is the operator norm. That is, the boundary terms only contribute a vanishing energy density.

While we aim at statements for *finite* regions Λ_n, the thermodynamic limit $n \to \infty$ becomes important as a proof tool and an indicator of phase transitions [13,14]. We make extensive use of the following properties, which characterize the system's behavior in the thermodynamic limit. States ω on the infinite lattice $\mathbb{Z}^\nu$ are given by families of density matrices $(\omega_\Lambda)_{\Lambda \subset \mathbb{Z}^\nu \text{finite}}$, with $\omega_\Lambda = \mathrm{Tr}_{\Lambda' \setminus \Lambda}\, \omega_{\Lambda'}$ if $\Lambda \subseteq \Lambda'$. Translation-invariant states ω on $\mathbb{Z}^\nu$ have entropy density $s(\omega) := \lim_{n\to\infty} \frac{1}{|\Lambda_n|} S(\omega_{\Lambda_n})$, with $S(\rho) = -\mathrm{tr}(\rho \log \rho)$ the von Neumann entropy, and energy density $u(\omega) := \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\omega_{\Lambda_n} H_{\Lambda_n})$. A characteristic quantity for any given model and $\beta > 0$ is the equilibrium Helmholtz free energy density $f_{\mathrm{th}}(\beta) := (-1/\beta) \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \mathrm{tr} \exp(-\beta H_{\Lambda_n})$. It holds

$$f_{\mathrm{th}}(\beta) = \inf\{f(\omega) \mid \omega \text{ translation-invariant state}\},$$

where $f(\omega) := u(\omega) - s(\omega)/\beta$ is the Helmholtz free energy density [13] of state ω. For any finite region Λ, the Gibbs state at inverse temperature β is $\gamma_\Lambda^{BC}(\beta) :=$

$\exp(-\beta H_\Lambda^{BC})/Z$, with Z the partition function. Gibbs states on the infinite lattice can be defined in several different equivalent ways; here we use a variational principle: a translation-invariant state ω on the infinite lattice is by definition a Gibbs state at inverse temperature β if it minimizes the free energy density, i.e. if $f(\omega) = f_{\text{th}}(\beta)$. This definition is equivalent to the well-known KMS condition [16].

For every inverse temperature β, there is at least one Gibbs state ω_β on the infinite lattice; however, the possibility of finite-temperature phase transitions implies that there may be more than one Gibbs state at the same β. Consequently, we say that *there is a unique equilibrium state around inverse temperature* β if there is a small interval around β such that for all β' in that interval, there is only one Gibbs state at inverse temperature β'. This is true, for example, if β is smaller than some model-dependent critical inverse temperature [18], and it is true for all β if the lattice dimension is $\nu = 1$ [17]. A given energy density value u will be called *thermal* if it is strictly larger than the ground state energy density $u_{\min}$, and strictly smaller than the infinite-temperature energy density $u_{\max}$. These are given by $u_{\min} = \lim_{n\to\infty} \lambda_{\min}(H_{\Lambda_n})/|\Lambda_n|$ with $\lambda_{\min}$ the smallest eigenvalue, and $u_{\max} := \lim_{n\to\infty} \operatorname{tr}(H_{\Lambda_n})/(|\Lambda_n| d^{|\Lambda_n|})$. If u is thermal, then there is exactly one positive inverse temperature $\beta \equiv \beta(u)$ such that the energy density $u(\omega_\beta)$ of the corresponding Gibbs state ω_β equals u [13].

2.1. Canonical typicality. As suggested in [3], we show that the Gibbs state arises in translation-invariant quantum lattice systems due to entanglement between small subsystems and the remainder. Consider any model with a given thermal energy density u such that there is a unique equilibrium state around the corresponding inverse temperature $\beta = \beta(u)$. For $\delta > 0$, define the *microcanonical subspace*

$$T_n^p := \operatorname{span}\{|E\rangle \mid u - \delta \le E/|\Lambda_n| \le u\}, \tag{1}$$

where $H_{\Lambda_n}^p |E\rangle = E|E\rangle$ denotes the periodic boundary condition energy eigenstates on the global region Λ_n. Choose any pure state $|\psi\rangle \in T_n^p$ at random according to the unitarily invariant measure. Then, with high probability, this state will locally in $\Lambda \subset \Lambda_n$ be very close to the reduction of the global Gibbs state, as depicted in Fig. 1:

Theorem 1 (Summary of Theorem 25). *Fix $\delta > 0$ and u thermal. Then for every $\varepsilon \ge 0$, the probability p that a state $|\psi\rangle \in T_n^p$ sampled according to the unitarily invariant measure satisfies*

$$\left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda} |\psi\rangle\langle\psi| - \operatorname{Tr}_{\Lambda_n\setminus\Lambda} \frac{\exp(-\beta H_{\Lambda_n}^p)}{Z} \right\|_1 \ge \varepsilon + \Delta_{n,\Lambda}$$

is doubly-exponentially small in the lattice size $|\Lambda_n|$; that is, $p \le \exp\left(-\varepsilon^2 \exp(|\Lambda_n| s + o(|\Lambda_n|))\right)$, where $s = s(\omega_\beta)$ is the entropy density of the corresponding Gibbs state, and $\Delta_{n,\Lambda}$ is a sequence of positive real numbers with $\lim_{n\to\infty} \Delta_{n,\Lambda} = 0$ for every fixed Λ. Here, β can either be set equal to $\beta(u)$ as defined above, or equal to the solution of $\operatorname{tr}(H_{\Lambda_n}^p \gamma_{\Lambda_n}^p(\beta))/|\Lambda_n| = u$ (which depends on n).

As illustrated in Fig. 1, in the limit of large n, almost all pure states $|\psi\rangle$ in an energy window subspace will be locally almost indistinguishable from the Gibbs state at the corresponding temperature, since the one-norm distance $\|\rho - \sigma\|_1 = 2 \max_{P=P^\dagger=P^2} |\operatorname{tr}(\rho P) - \operatorname{tr}(\sigma P)|$ being small means that ρ and σ give similar expectation value for all possible measurements. The theorem does not say how quickly $\Delta_{n,\Lambda}$

tends to zero with increasing n; we will come back to the question of finite-size estimates later. Earlier work [3,10] attempted to prove that $\mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,|\psi\rangle\langle\psi|$ is arbitrarily close to the local Gibbs state $\gamma_\Lambda(\beta) = \exp(-\beta H_\Lambda)/Z$. However, this can only be true if the interaction across the boundary of Λ is very weak [10]; in particular, the given upper bound on the distance grows with the boundary of Λ and is thus interesting only if Λ is small or if the lattice is one-dimensional. Our theorem shows that in general the local Gibbs state has to be replaced by the reduction of the global Gibbs state to obtain arbitrary closeness in the thermodynamics limit, unless one considers models that are fine-tuned such that the local Gibbs state agrees with the reduction of the global Gibbs state.

Before we turn to the proof, we note that the unitarily invariant (Haar) measure in Theorem 1 can be replaced by a more physically realistic measure, namely an η-approximate t-design [25,27], for $t = 8$ and $\eta = \exp(-|\Lambda_n|s + o(|\Lambda_n|))$. Here $o(\cdot)$ is (asymptotic) Landau notation, where $f_n = o(g_n)$ means that f_n/g_n converges to zero in the limit $n \to \infty$. Such t-designs are approximations to the Haar measure that can be efficiently generated in a time which is polynomial in the lattice size $|\Lambda_n|$. It follows from the results of Low [26] that Theorem 1 remains valid, however with a probability value that is only exponentially (not doubly-exponentially) small in the lattice site – see Theorem 28.

To prove Theorem 1, we invoke the results of [4], which tell us that $\mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,|\psi\rangle\langle\psi|$ is with high probability close to $\mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\tau_n$, where τ_n is the uniformly mixed state on T_n^p. We obtain Theorem 1 directly, with all constants, if we set $\Delta_{n,\Lambda}$ up to corrections of order $\exp\left(-\frac{1}{2}|\Lambda_n|s + o(|\Lambda_n|)\right)$ (cf. Eq. (36)) equal to

$$\delta_{n,\Lambda} := \left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\tau_n - \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\frac{\exp(-\beta H^p_{\Lambda_n})}{Z} \right\|_1 . \tag{2}$$

It remains to prove that $\delta_{n,\Lambda} \to 0$ as $n \to \infty$. However, τ_n is nothing but the microcanonical ensemble, and the statement left to prove is that its predictions on small subsystems Λ are equivalent to those of the canonical ensemble in the thermodynamic limit. Thus, we are naturally led to study the problem of equivalence of ensembles in our setting.

2.2. Equivalence of ensembles. To state our result, note that we can regard Λ_n as a torus, by identifying $\mu_i + 1$ in the interval $[\lambda_i, \mu_i]$ with λ_i; this way, we can define periodic translations of Λ_n as those of the resulting torus. A state τ_n on Λ_n will be called Λ_n-translation-invariant if it is invariant with respect to all periodic translations of Λ_n. Using this notion, our main technical result on equivalence of ensembles reads as follows:

Theorem 2 (Summary of Theorem 10). *Suppose that $(\tau_n)_{n\in\mathbb{N}}$ is any sequence of Λ_n-translation-invariant states on Λ_n, and $\beta > 0$ such that there is a unique equilibrium state around inverse temperature β. If*

$$\limsup_{n\to\infty} \frac{1}{|\Lambda_n|}\left(\mathrm{tr}(\tau_n H^{BC}_{\Lambda_n}) - S(\tau_n)/\beta\right) \leq f_{\mathrm{th}}(\beta) \tag{3}$$

for some choice of boundary conditions BC, then

$$\lim_{n\to\infty} \left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\tau_n - \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\frac{\exp(-\beta_n H^p_{\Lambda_n})}{Z_n} \right\|_1 = 0, \tag{4}$$

where we may set β_n either equal to the fixed value β, or equal to the solution of $\mathrm{tr}(H^p_{\Lambda_n}\gamma^p_{\Lambda_n}(\beta_n))/|\Lambda_n| = u(\beta)$.

Theorem 2 implies Theorem 1: If τ_n is the microcanonical ensemble, i.e. maximal mixture on T^p_n, then $\mathrm{tr}(\tau_n H^p_{\Lambda_n})/|\Lambda_n| \leq u$ by construction, and $S(\tau_n) = \log\dim(T^p_n) = s|\Lambda_n| + o(|\Lambda_n|)$ according to [13, Theorem IV.2.14] (as [13] does not provide a proof, we reproduce the proof in Lemma 11 below). Since $u - s/\beta = f_{\mathrm{th}}(\beta)$, (3) holds, which shows equivalence to the canonical ensemble, $\lim_{n\to\infty}\delta_{n,\Lambda} = 0$, and establishes Theorem 1.

The crucial property of the microcanonical subspace T^p_n used in this proof is its dimensionality (which is close to maximal given its energy density u), namely $\lim_{n\to\infty}(1/|\Lambda_n|)\log\dim T^p_n = s$. It follows from Lemma 11 that this property is satisfied if the width $\delta > 0$ in the definition of the microcanonical subspace (1) is constant in n, which corresponds to an extensive energy uncertainty. In general, one can also choose an n-dependent width $\delta \equiv \delta_n$; as long as δ_n tends to zero slowly enough, the necessary limit identity will still hold. Unfortunately, giving a concrete expression for a possible choice of δ_n amounts to proving a generalization of Lemma 11 for "small" microcanonical subspaces, and we do not currently have such a generalization.

However, in the special case of the non-interacting Ising model described in Sect. 2.4 below, it is easy to see via standard inequalities (like the ones used in the proof of Theorem 37) that one can choose $\delta_n \geq c(\log n)/n$, with $c > 0$ some constant depending on u. It is therefore plausible to expect that a comparable scaling of δ_n might be possible also in the interacting case.

We now sketch the proof of Theorem 2. We first show that $(\tau_n)_{n\in\mathbb{N}}$ has at least one limit point ω as a state on the infinite lattice. Since every τ_n is Λ_n-translation-invariant, ω is translation-invariant, and (3) implies that $f(\omega) = f_{\mathrm{th}}(\beta)$. Thus, ω is the unique Gibbs state ω_β, and so

$$\lim_{n\to\infty} \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\tau_n = (\omega_\beta)_\Lambda. \tag{5}$$

Consider the special case where τ_n equals the local Gibbs state $\gamma_n := \gamma^p_{\Lambda_n}(\beta)$ which appears in (4). Every γ_n is Λ_n-translation-invariant and minimizes the free energy locally, hence

$$\mathrm{tr}(\gamma_n H^p_{\Lambda_n}) - S(\gamma_n)/\beta \leq \mathrm{tr}((\omega_\beta)_{\Lambda_n} H^p_{\Lambda_n}) - S((\omega_\beta)_{\Lambda_n})/\beta \overset{n\to\infty}{\longrightarrow} f_{\mathrm{th}}(\beta),$$

which shows that (3) is satisfied for $\tau_n = \gamma_n$. Consequently $\lim_{n\to\infty}\mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\gamma_n = (\omega_\beta)_\Lambda$, and combining this with (5) proves the theorem.

This proof strategy has been pioneered by Lima [11,12]; however, our result is more general. In particular, we allow a more general set of possible interactions, and permit $\beta_n \neq \beta$ to be determined from the finite region Λ_n.

2.3. Dynamical thermalization. It has been shown in [6–8] that subsystems of closed quantum systems equilibrate, subject to some conditions on the initial state and spectrum. In general, the equilibrium state depends on the initial state, and is not thermal unless additional conditions are met [10]. However, for translation-invariant systems, we can say more. Consider any initial state $\rho_0^{(n)}$ on Λ_n, pure or mixed. The index n indicates that the state is chosen to be a function of the lattice size n. We can think of a simple dependence such as $\rho_0^{(n)} = \rho_0^{\otimes\Lambda_n}$ for some fixed (single-site) state ρ_0 on $\mathbb{C}^d$; however, the only technical condition we need to assume is that the density of the

inner energy $U_n := \mathrm{tr}(\rho_0^{(n)} H_{\Lambda_n}^p)$ converges to some well-defined thermal energy density $u := \lim_{n\to\infty} U_n/|\Lambda_n|$.

The state evolves unitarily under the Hamiltonian $H_{\Lambda_n}^p$, i.e. $\rho^{(n)}(t) = \exp(-itH_{\Lambda_n}^p)\,\rho_0^{(n)}\exp(itH_{\Lambda_n}^p)$. We can define the *population entropy* $\bar{S}(\rho_0^{(n)})$ as follows. From the spectral decomposition $H_{\Lambda_n}^p = \sum_i E_i \pi_i$, compute the weights $\lambda_i := \mathrm{tr}(\rho_0^{(n)}\pi_i)$, and set $\bar{S}(\rho_0^{(n)}) := -\sum_i \lambda_i \log \lambda_i$. Similarly, there is an inverse temperature β_n corresponding to $\rho_0^{(n)}$, defined by $\mathrm{tr}(H_{\Lambda_n}^p \gamma_{\Lambda_n}^p(\beta_n)) = U_n$. Denote the time average by $\langle\cdot\rangle$, i.e. $\rho_{\mathrm{avg}}^{(n)} := \langle\rho^{(n)}(t)\rangle := \lim_{T\to\infty}(1/T)\int_0^T \rho^{(n)}(t)dt$. Then the actual state at time t is close to $\rho_{\mathrm{avg}}^{(n)}$ for most times t, and this state is close to thermal:

Theorem 3 (Summary of Theorem 31). *If there is a unique equilibrium state around inverse temperature $\beta := \lim_{n\to\infty}\beta_n$, if the (possibly pure) initial state has close to maximal population entropy, in the sense that*

$$\bar{S}(\rho_0^{(n)}) \geq S(\gamma_{\Lambda_n}^p(\beta_n)) - o(|\Lambda_n|), \tag{6}$$

and if each $H_{\Lambda_n}^p$ is non-degenerate (i.e. all eigenspaces are one-dimensional), then unitary time evolution thermalizes the subsystem Λ for most times t:

$$\left\langle \left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\rho^{(n)}(t) - \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\rho_{\mathrm{avg}}^{(n)} \right\|_1 \right\rangle \leq d^{|\Lambda|}\sqrt{D_G}\,\exp\left(-\frac{s(\omega_\beta)^2}{4\log d}|\Lambda_n| + o(|\Lambda_n|)\right), \quad \textit{and} \tag{7}$$

$$\lim_{n\to\infty}\left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\rho_{\mathrm{avg}}^{(n)} - \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\,\frac{\exp(-\beta_n H_{\Lambda_n}^p)}{Z_n} \right\|_1 = 0, \tag{8}$$

where D_G is the gap degeneracy [8] of $H_{\Lambda_n}^p$, defined by $D_G = \max_E |\{(i,j) \mid i \neq j, E_i - E_j = E\}|$, where E_i denotes the eigenvalues of $H_{\Lambda_n}^p$.

In Theorem 33, we generalize this result to the case of arbitrary boundary conditions and degenerate $H_{\Lambda_n}^{BC}$. Unlike (8), which expresses equivalence of the time-averaged state $\rho_{\mathrm{avg}}^{(n)}$ and the thermal state $\gamma_{\Lambda_n}^p(\beta_n)$ for local observables A on Λ, the generalized version shows equivalence of these global states on a different set of observables [8], arising from averaging observables A over translations of Λ. We also show numerically in Sect. 3.5 that the conditions of non-degeneracy of $H_{\Lambda_n}^p$ and $D_G = 1$ are generically satisfied for randomly chosen translation-invariant nearest-neighbor interactions in one lattice dimension.

Our proof of Theorem 3 follows similarly to the proof of the results of [8]. First we have to show that the "effective dimension" $d_{\mathrm{eff}} = e^{S_2(\lambda)}$ is large, with $S_\alpha(\rho) := (\log \mathrm{tr}(\rho^\alpha))/(1-\alpha)$ the α-Rényi entropy. We do this via the inequality $S_2 \geq 2\varepsilon(S - \varepsilon/(1+\varepsilon)S_0)$ for $0 \leq \varepsilon \leq 1$, which we prove from results of [29], establishing (7). From $S(\rho_{\mathrm{avg}}^{(n)}) \geq \bar{S}(\rho_0^{(n)})$, we conclude that $\rho_{\mathrm{avg}}^{(n)} =: \tau_n$ satisfies (3). We then apply Theorem 2 to prove (8).

As an example, if $\rho_0^{(n)}$ is a pure state $|\psi_0^{(n)}\rangle \sim \sum_{u-\delta<E_i/|\Lambda_n|<u}|E_i\rangle$ which is a "flat" uniform superposition of eigenstates $|E_i\rangle$ of $H_{\Lambda_n}^p$, Theorem 3 applies. This recovers results of [10], albeit in a different context.

2.4. Finite-size estimates. Estimates on how large Λ_n has to be to have good agreement with our asymptotic results, in particular bounds on $\delta_{n,\Lambda}$ in (2), are expected to depend strongly on the details of the model, such as distance to phase transitions, correlation lengths etc. [23]. To get some intuition, we now give analytic bounds for the non-interacting Ising model, which already turns out to be a non-trivial problem. For this model, it was already shown in [4] that local reduced states are close to thermal in the sense of Theorem 1; however, no explicit analytic bounds on the distance have been given in [4]. Here we provide tight analytic finite-size bounds.

We set $\Lambda_n = [1, n] \subset \mathbb{Z}^1$, and $H_\Lambda := \sum_{i\in\Lambda} Z_i$, where Z_i is the Pauli Z-matrix at site i. Then the microcanonical state τ_n is permutation-invariant, and the canonical state is a product state, $\gamma_{\Lambda_n}(\beta) = \gamma_\beta^{\otimes n}$, with $\gamma_\beta := \gamma_{\{1\}}(\beta)$ the single-site Gibbs state. We are interested in estimating the distance $\delta_{n,\Lambda}$ in (2). In the case where the energy value of the microcanonical subspace (1) is sharp, i.e. $\delta = 0$, the state τ_n is the uniform mixture over a type class, that is, over the subspace spanned by eigenvectors with a fixed frequency of "spin-up". In this case, it turns out that we can apply the proof of the classical finite de Finetti theorem [22] to obtain

$$\left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\, \tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \frac{4m}{n}, \tag{9}$$

where $m := |\Lambda|$. Thus, in order to maintain a fixed 1-norm distance between the states, the total system size n has to be increased linearly with the size of the subsystem m. As mentioned before Eq. (2), this also upper-bounds the distance $\Delta_{n,\Lambda}$ in Theorem 1 up to corrections exponentially small in the lattice size.

The case of finite energy uncertainty $\delta > 0$ is more difficult to treat. If we assume each of the lattice sites holds a qubit ($d = 2$) and take an appropriate rescaling of the energy then

$$S\left(\gamma_\beta^{\otimes m} \,\middle\|\, \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n\right) \leq \frac{(1-\delta)u}{u-\delta}\cdot\frac{m}{n-m} + \frac{mu\delta}{u-\delta}\left(1+\frac{m}{n-m}\right)$$

whenever $m \leq n(u-\delta)$, with $S(\rho\|\sigma) := \mathrm{tr}(\rho\log\rho - \rho\log\sigma)$ the quantum relative entropy. This claim is formally stated as Lemma 36. For $\delta = 0$ (and $m \ll n$), this inequality is similar to (9) above, but now with the relative entropy as distance measure. We expect it to be tight (i.e. not to allow for significant improvements) in the case $\delta = 0$, since it is well-known that the bound in the classical finite de Finetti theorem, and thus (9), cannot be significantly improved. However, this inequality on the relative entropy has the drawback that it is only interesting as long as $\delta \lesssim 1/m$. The question arises how n has to be scaled with growing subsystem size m in order to achieve a fixed distance $\delta > 0$ (for $\delta = 0$, we have seen that n has to be increased linearly with m). In Theorem 37, we settle this question up to a correction term of the order $\log n$: under some conditions on the variables, we show that

$$\left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \frac{2\delta}{n\sqrt{u}} + \sqrt{\frac{m}{n-m}\left(1+\frac{4\log n}{\log\frac{1-u}{u}}\right)}.$$

This inequality is not tight in general (as one sees by comparing with (9) for $\delta = 0$), but it shows that n has to be increased only slightly superlinearly with m in order to achieve a fixed 1-norm distance also in the case $\delta > 0$. We leave it as an open question whether the $\log n$ term can be removed.

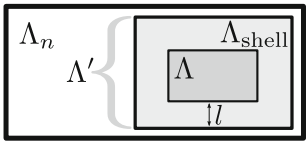

Fig. 2. Subregions of the whole lattice Λ_n. We enlarge Λ by setting $\Lambda' = \Lambda \cup \Lambda_{\text{shell}}$, where Λ_{shell} contains all sites outside of Λ which have distance l or less to Λ. The number of terms of H_{Λ_n} that have support on both Λ' and $\Lambda_n \backslash \Lambda'$ is denoted A, which quantifies the size of the boundary area of Λ'

In order to get some intuition for what happens in the *interacting* case, we numerically study random nearest-neighbor interactions in one lattice dimension in Sect. 3.5. It turns out that the behavior that we have shown analytically for non-interacting models remains approximately valid also in the interacting case (as far as one can tell for the small lattice sizes $n \leq 11$ that are numerically tractable), see in particular Fig. 6. However, we leave it open whether a similar behavior remains valid in lattice dimensions $\nu \geq 2$, where finite-temperature phase transitions become relevant.

2.5. Towards eigenstate thermalization. The question whether some of the results above can be strengthened to hold for *individual energy eigenstates* is known as the eigenstate thermalization hypothesis (ETH) [33,34]. For example, consider our result on dynamical thermalization, Theorem 3. For this result to hold, Eq. (6) must be satisfied, which says that the initial state populates a large number of energy levels.

The question arises whether this assumption can be dropped. In the most extreme case, we could have an energy eigenstate $|E\rangle$ as the initial state, i.e. $\rho_0 = |E\rangle\langle E|$. (This notation does not assume non-degeneracy of the spectrum; $|E\rangle$ is an arbitrary pure state in the eigenspace corresponding to energy E.) Energy eigenstates do not evolve, such that $\rho^{(n)}(t) = \rho_0$ is constant in time. Thus $\rho^{(n)}(t)$ is close to thermal for most times t if and only if the reduced state $\text{Tr}_{\Lambda_n \backslash \Lambda} |E\rangle\langle E|$ is close to thermal.

To formulate eigenstate thermalization in more detail, consider the setup in Fig. 2. We have argued above that one should not expect that the local marginals of random global pure states $|\psi\rangle$ are close to a local Gibbs state, due to boundary effects (which led us to consider the reduction of the global Gibbs state instead). More generally, to take boundary effects into account, we can enlarge the subregion Λ by a shell of width l; if l is large enough, one would expect that

$$\text{Tr}_{\Lambda_n \backslash \Lambda} |E\rangle\langle E| \approx \text{Tr}_{\Lambda_n \backslash \Lambda} \gamma_{\Lambda_n}(\beta) \approx \text{Tr}_{\Lambda_{\text{shell}}} \gamma_{\Lambda'}(\beta). \tag{10}$$

It is immediately clear that a statement like this cannot literally be true for all eigenstates $|E\rangle$ of all models that we consider: the non-interacting Ising model, where some eigenstates are product states (and thus marginals are pure and not thermal), is a counterexample.

However, we can prove a weaker version of this statement which is true for all eigenstates of all translation-invariant models with finite range interaction: there is a state ω_E on Λ' such that $\text{Tr}_{\Lambda_n \backslash \Lambda} |E\rangle\langle E| \approx \text{Tr}_{\Lambda_{\text{shell}}} \omega_E$, where ω_E partially resembles a thermal state. That is, ω_E does not necessarily have Boltzmann weights on its diagonal (as one would expect from the thermal state $\gamma_{\Lambda'}(\beta)$), but its off-diagonal elements are close to zero, as they are for the thermal state.

We formulate and prove this result by applying a version of the Lieb-Robinson bound [30–32]: for models with finite-range interaction, it states that there are constants

$c, C, v > 0$ such that for all operators X and Y supported on finite regions $\mathcal{X}, \mathcal{Y}$ of distance Δ, it holds $\|[X(t), Y]\|_\infty \leq C\,\|X\|_\infty \|Y\|_\infty \min\{|\mathcal{X}|, |\mathcal{Y}|\}\, e^{-c[\Delta - v|t|]}$, where $X(t) = e^{iH_{\Lambda_n}t} X e^{-iH_{\Lambda_n}t}$. The constants also appear in the following theorem, where we assume in particular that the Hamiltonian only has interactions between sites of distance r or less.

Theorem 4 (Summary of Theorem 38). *There is a state ω_E on Λ' such that*

$$\left\| \mathrm{Tr}_{\Lambda_{\mathrm{shell}}}(\omega_E) - \mathrm{Tr}_{\Lambda_n \setminus \Lambda} |E\rangle\langle E| \right\|_1 \leq \kappa \cdot e^{-c(l-r)/2}, \tag{11}$$

where $\kappa = 2AJ(CA+2)\sqrt{\frac{l-r}{8cv^2}}$ and $J = \max_X \|h_X\|_\infty$, which is close to diagonal in the eigenbasis $\{|e\rangle\}$ of $H_{\Lambda'}$, i.e.

$$|\langle e_1|\omega_E|e_2\rangle| \leq e^{-(l-r)(e_1-e_2)^2/(8cv^2)}. \tag{12}$$

This result does not assume translation-invariance; finite range of interaction is sufficient for its validity. The ETH corresponds to the claim that the theorem holds for the particular choice $\omega_E = \gamma_{\Lambda'}(\beta)$. As discussed above, the ETH cannot be true in general for all eigenstates of all models we consider; intuitively, some additional assumptions, possibly along the lines of nonintegrability, are needed.

Even though the mathematical details of the proof are cumbersome, it has a simple physical interpretation. We define ω_E by evolving $\mathrm{Tr}_{\Lambda_n \setminus \Lambda'} |E\rangle\langle E|$ according to $H_{\Lambda'}$ and averaging the result over small t; concretely, $\omega_E := \int_{-\infty}^{\infty} dt\, g(t)\, e^{-iH_{\Lambda'}t}\, \mathrm{Tr}_{\Lambda_n \setminus \Lambda'} |E\rangle\langle E|\, e^{iH_{\Lambda'}t}$, with $g(t)$ some Gaussian. The Lieb-Robinson bound guarantees finite speed of information transmission, such that the result will within Λ still look very much as if the initial state $|E\rangle\langle E|$ evolved according to the full Hamiltonian H_{Λ_n}, if the shell is large enough. Since $|E\rangle\langle E|$ is stationary, this leads to (11). On the other hand, interaction across the boundary of Λ' will decohere the state $\mathrm{Tr}_{\Lambda_n \setminus \Lambda'} |E\rangle\langle E|$; in particular, coherences corresponding to energy levels e_1, e_2 with large $|e_1 - e_2|$ will be suppressed, which yields (12).

In Theorems 1, 2 and 3, we quantify thermalization by the distance to the marginal of the global Gibbs state, that is, $\mathrm{Tr}_{\Lambda_n \setminus \Lambda} \exp(-\beta_n H^p_{\Lambda_n})/Z_n$. This contrasts with the above setup (10), where one starts with the thermal state in a subregion Λ' containing a shell around Λ. In fact, it is easy to see that Theorems 1, 2 and 3 also hold if we replace the thermal state on all of Λ_n by the thermal state on $\Lambda' := \Lambda_{n'} \supseteq \Lambda$ for $n' \ll n$, as long as $n' \to \infty$ with $n \to \infty$; that is, if we consider the distance to $\mathrm{Tr}_{\Lambda_{n'} \setminus \Lambda} \exp(-\beta_{n'} H^p_{\Lambda_{n'}})/Z_{n'}$ instead. The drawback, however, is that we do not have any non-trivial bounds for fixed finite n' (in contrast to Theorem 4), which renders the formulation of the first three theorems in terms of the more general shell setup mathematically equivalent to their current formulation. Finding non-trivial finite-size bounds, in particular on the size of the shell, remains as an interesting open problem.

3. Proofs of the Main Results

All results are formulated in two versions, namely for periodic and for arbitrary boundary conditions (BC). The main results can be found in the following places:

- *Equivalence of ensembles* Main technical statement is Theorem 10 (periodic BC), with Example 13 giving the standard formulation in comparing the microcanonical

and the canonical ensemble. The corresponding formulations for arbitrary BC are given in Theorem 20 and Example 21. Note that the norm $\|\cdot\|_{\{m\}}$ appearing there can be replaced by $\|\cdot\|_{[m]}$, measuring the difference of expectation values on *m-block averaged observables* only, due to Lemma 24.

- *Canonical typicality* For periodic BC, the main result is Theorem 25, with a derandomized version in terms of 8-designs given in Theorem 28. The corresponding formulations for arbitrary BC are given in Theorem 26 and Theorem 29.
- *Dynamical thermalization* For periodic BC, the main result is Theorem 31, and for arbitrary BC it is Theorem 33.
- *Finite-size bounds without interaction* Lemma 34 relates the Ising model and equivalence of ensembles for sharp energy eigenspaces ($\delta = 0$) and local Hilbert space dimension $d = 2$ to the finite de Finetti theorem. Lemma 36 is a new derivation (compared to Diaconis and Freedman [22]) in terms of the relative entropy. The main result is Theorem 37, proving that the scaling (bath size increasing linearly with system size to achieve fixed 1-norm error) remains basically valid also for $\delta > 0$.
- *Numerical results* They are given in Sect. 3.5, confirming that some assumptions from the main theorems (on degeneracy of spectra etc.) are generically satisfied. Moreover, they show that the qualitative finite-size scaling that has been proven analytically for non-interacting systems seems to remain valid for interacting systems with periodic BC, at least for lattice dimension $\nu = 1$.
- *Eigenstate thermalization* The main result is Theorem 38, showing that energy eigenstates are locally "weakly diagonal". Note that this result does not assume translation-invariance (only finite range of interaction).

The notation is specified in Sect. 3.1 below. In comparison to Sect. 2, statements about the minimization of the Helmholtz free energy density $f(\omega) := u(\omega) - s(\omega)/\beta$ are replaced by statements about the maximization of $-\beta\, f(\omega) = s(\omega) - \beta\, u(\omega)$ (following mathematical physics tradition), which has to be compared with the "pressure"

$$p(\beta, \Phi) = -\beta\, f_{\text{th}}(\beta) \qquad (\beta > 0).$$

This has the advantage that $p(\beta, \Phi)$ is also defined for $\beta = 0$, i.e. infinite temperature. Moreover, convexity of $\beta \mapsto p(\beta, \Phi)$ will play a crucial role. Similarly, to conform with mathematical physics literature, we will write $\Phi(X)$ instead of h_X, and the map Φ will be called an "interaction". Furthermore, we will assume that the small subsystem Λ equals Λ_m for some fixed m, which is no loss of generality.

3.1. Equivalence of ensembles. We start by fixing some notation. We consider a ν-dimensional quantum lattice system, with local Hilbert space dimension d. To every $x \in \mathbb{Z}^\nu$, we associate a local algebra of observables $\mathcal{A}_x$, which is a copy of $M_d(\mathbb{C})$, the algebra of complex $d \times d$ matrices. For every finite region $\Lambda \subset \mathbb{Z}^\nu$, we have the local observable algebra $\mathcal{A}_\Lambda := \bigotimes_{x\in\Lambda} \mathcal{A}_x$. For every $y \in \mathbb{Z}^\nu$, there is a translation automorphism γ_y, mapping observables A in a region Λ, i.e. $A \in \mathcal{A}_\Lambda$, to the corresponding observable $\gamma_y(A)$ in the translated region $\Lambda + y$, i.e. $\gamma_y(A) \in \mathcal{A}_{\Lambda+y}$.

To every finite region $X \subset \mathbb{Z}^\nu$, we associate an *interaction* $\Phi(X)$, which is a self-adjoint operator in $\mathcal{A}_X$, describing the interaction of the spins in region Λ. For finite $\Lambda \subset \mathbb{Z}^\nu$, the *local Hamiltonian* H_Λ is

$$H_\Lambda := \sum_{X\subset\Lambda} \Phi(X).$$

We assume that our interaction has finite range, i.e. that $\Phi(X) = 0$ whenever the diameter of X is larger than r for some fixed $r \in \mathbb{N}$. Furthermore, we assume translation-invariance, which can be expressed as

$$\Phi(X + y) = \gamma_y(\Phi(X)) \qquad \text{for all } X \subset \mathbb{Z}^\nu \text{ finite}, y \in \mathbb{Z}^\nu.$$

We can also define an observable algebra $\mathcal{A}_\infty$ for the infinite lattice $\mathbb{Z}^\nu$ by a suitable limit procedure, called the *quasi-local algebra*, see [13] for details. The (operator) norm on $\mathcal{A}_\infty$ will be denoted $\|\cdot\|_\infty$. A *state* ω on $\mathcal{A}_\infty$ is a positive linear functional with $\omega(\mathbf{1}) = 1$. States are automatically weak *-continuous. A state ω is *translation-invariant* if $\omega(\gamma_y(A)) = \omega(A)$ for all $A \in \mathcal{A}_\infty$ and $y \in \mathbb{Z}^\nu$ (it is sufficient to demand this for all $A \in \mathcal{A}_\Lambda$ for all finite regions Λ). If $\Lambda \in \mathbb{Z}^\nu$ is finite, there is a density matrix $\omega_\Lambda \in \mathcal{A}_\Lambda$ such that

$$\mathrm{tr}(\omega_\Lambda A) = \omega(A) \qquad \text{for all } A \in \mathcal{A}_\Lambda.$$

This yields the following consistency condition: if $\Lambda \subset \Lambda'$ and Λ' is finite, then $\omega_\Lambda = \mathrm{Tr}_{\Lambda' \setminus \Lambda} \omega_{\Lambda'}$. Conversely, every consistent family of density matrices defines a state on $\mathcal{A}_\infty$.

For translation-invariant states, the following definitions are crucial. To state them, we consider sequences of boxes (that is, hyperrectangles) $(\Lambda_n)_{n\in\mathbb{N}}$ with $\Lambda_n \subset \Lambda_{n+1}$ and with the property that for every $x \in \mathbb{Z}^\nu$ there is some $n \in \mathbb{N}$ with $x \in \Lambda_n$. Unless specified otherwise, all sequences of regions Λ_n in the following will be assumed to have these properties.

All logarithms are in base e, i.e. $\log(\exp(x)) = x$.

Definition 5. Let ω be a translation-invariant state on $\mathcal{A}_\infty$. Then the following expressions exist:

- Energy density: $u(\omega) := \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\omega_{\Lambda_n} H_{\Lambda_n})$,
- entropy density: $s(\omega) := -\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\omega_{\Lambda_n} \log \omega_{\Lambda_n})$.

Moreover, there is the state-independent quantity *pressure*

$$p(\beta, \Phi) := \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \mathrm{tr} \exp(-\beta H_{\Lambda_n})$$

for all $\beta \geq 0$. It satisfies

$$p(\beta, \Phi) = \sup\{s(\varphi) - \beta\, u(\varphi) \mid \varphi \text{ is any translation-invariant state on } \mathcal{A}_\infty\}. \qquad (13)$$

See [13] for more details. In the following, we consider Gibbs state on the infinite lattice. They are defined by any one of the following equivalent conditions.

Definition 6. Let ω be a translation-invariant state on the quasi-local algebra $\mathcal{A}_\infty$ over $\mathbb{Z}^\nu$, with translation-invariant finite-range interaction Φ, and let $\beta > 0$. Then the following conditions are equivalent:

- *Variational principle* it holds $p(\beta, \Phi) = s(\omega) - \beta\, u(\omega)$, which is the maximal possible value according to (13).
- *KMS condition* at inverse temperature β (see [16],
- *Gibbs condition* at inverse temperature β (see also [16]).

If ω satisfies one of these equivalent conditions, we will call ω a *Gibbs state* at inverse temperature β. We say that Gibbs states are unique around inverse temperature β for a given interaction Φ if there is an open interval containing β such that for every β' in this interval, there is a unique (only one) Gibbs state at inverse temperature β'.

Since we do not use the KMS and the Gibbs conditions, we do not explain them in detail here. We refer the reader to [13,16].

For what follows, we need to extend the notion of translation-invariance to finite regions. This is done in the obvious way. Let $\mathcal{A}_n := \mathcal{A}_{\Lambda_n}$, with Λ_n a sequence of boxes tending to infinity as $n \to \infty$ in the sense specified above. Call an observable $A \in \mathcal{A}_n$ Λ_n*-translation-invariant* if it is translation-invariant with respect to periodic translations of Λ_n; that is, translations in which we regard Λ_n as a torus. In more detail, write Λ_n as the product of intervals

$$\Lambda_n = [\lambda_1, \mu_1] \times [\lambda_2, \mu_2] \times \cdots \times [\lambda_\nu, \mu_\nu],$$

where $\lambda_i, \mu_i \in \mathbb{Z}$, $\lambda_i \leq \mu_i$. The statement that Λ_n tends to infinity means that all $\lambda_i \to -\infty$ and all $\mu_i \to +\infty$ as $n \to \infty$. Define ν independent translations $(T_j)_{j=1,\ldots,\nu}$ for $x \in \Lambda_n$ by

$$T_j(x) \equiv T_j(x_1, \ldots, x_\nu) = (x_1, \ldots, x_{j-1}, x_j \oplus 1, x_{j+1}, \ldots, x_\nu),$$

where

$$x_j \oplus 1 = \begin{cases} x_j + 1 & \text{if } x_j + 1 \leq \mu_j, \\ \lambda_j & \text{otherwise.} \end{cases}$$

We can interpret T_j as a unitary operator, translating the computational basis vectors, constructed from the translation automorphisms γ_y. An observable A will be called Λ_n-translation-invariant if $T_j A T_j^\dagger = A$ for all $j = 1, \ldots, \nu$.

We can also formalize this definition somewhat differently. Denote by $\mathbf{T}(\Lambda_n)$ the set of all periodic translations of Λ_n into itself; in other words, regard Λ_n as a torus, and $\mathbf{T}(\Lambda_n)$ as the set of translations on the torus. These are arbitrary compositions of translations T_j. If $\alpha \in \mathbb{Z}^\nu$, then the periodic translation by vector α will be denoted $T_\alpha \in \mathbf{T}(\Lambda_n)$; it equals $T_\alpha = \bigcirc_{j=1}^{\nu} T_j^{\alpha_j}$, where the circle denotes composition and the T_j are mutually commuting. Then an observable A is Λ_n-translation-invariant if and only if $TAT^\dagger = A$ for all $T \in \mathbf{T}(\Lambda_n)$.

So far, we have defined H_Λ for finite regions Λ by summing up all interaction terms that are fully contained in Λ. This is usually called the Hamiltonian with *open boundary conditions*. Alternatively, one can consider periodic or other, more general boundary conditions. We use the following definition.

Definition 7 *(Periodic and arbitrary boundary conditions).* Let Φ be any finite-range translation-invariant interaction. A region $\Lambda \subset \mathbb{Z}^\nu$ is called *large enough* if for every region X with $\Phi(X) \neq \emptyset$, there is $y \in \mathbb{Z}^\nu$ such that the translation $X + y$ is contained in Λ.

A *choice of boundary conditions* is a map that assigns to every large enough, finite set $\Lambda \subset \mathbb{Z}^\nu$ a Hamiltonian H_Λ^{BC} such that

$$\lim_{n\to\infty} \frac{\left\| H_{\Lambda_n}^{BC} - H_{\Lambda_n} \right\|_\infty}{|\Lambda_n|} = 0 \tag{14}$$

for every sequence of boxes $(\Lambda_n)_{n\in\mathbb{N}}$ that tends to infinity in the sense specified above.

A particularly important example of choice of boundary conditions is given by *periodic boundary conditions*, with corresponding Hamiltonians denoted H_Λ^p. Following [13], we define it as

$$H_\Lambda^p := {\sum_{X\cap\Lambda\neq\emptyset}}' T_{-\alpha}\left[\gamma_\alpha(\Phi(X))\right]T_{-\alpha}^\dagger,$$

where $\alpha \in \mathbb{Z}^\nu$ denotes any vector that translates X into Λ, i.e. $X+\alpha \in \Lambda$, and the prime on the sum indicates that regions X, X' with $X_i' = X_i + n_i a_i$ for all i, where a_i is the ith sidelength of the boxes Λ, and $n_i \in \mathbb{Z}$, are not included twice in the sum, but only once (i.e. only X or X' will be included).

The fact that we are demanding that regions Λ are large enough implies that our definition of H_Λ^p agrees with both of what Simon [13] calls $H_\Lambda^{p,1}$ and $H_\Lambda^{p,2}$. To see that H_Λ^p satisfies (14), denote by $\partial\Lambda$ the discrete boundary of Λ, that is

$$\partial\Lambda := \left\{x \in \Lambda \mid \exists y \in \mathbb{Z}^\nu \setminus \Lambda : \ \operatorname{dist}(x,y) \leq 1\right\}, \tag{15}$$

where $\operatorname{dist}(x,y) := \max_i |y_i - x_i|$. Suppose $x \in \mathbb{Z}^\nu$ is any point. Since Φ has finite range and is translation-invariant, there is some finite integer $\kappa \in \mathbb{N}$ equal to the number of finite regions X that contain x and have $\Phi(X) \neq 0$. This number is the same for every $x \in \mathbb{Z}^\nu$. Also, $\|\Phi\| := \max_X \|\Phi(X)\|_\infty$ is finite. Thus

$$\left\|H_{\Lambda_n}^p - H_{\Lambda_n}\right\|_\infty \leq \sum_{X\cap\Lambda_n\neq\emptyset,\ X\not\subset\Lambda_n} \|\Phi(X)\|_\infty \leq \sum_{x\in\partial\Lambda_n} \kappa\|\Phi\| = \kappa\|\Phi\|\,|\partial\Lambda_n|.$$

Since $|\partial\Lambda_n|/|\Lambda_n|$ tends to zero for $n \to \infty$, this proves (14). We can write $H_{\Lambda_n}^p$ in an alternative form. Given Λ_n, denote by $X_1, \ldots, X_N$ subsets of Λ_n with the property that no X_i is a periodic translation of any other X_j, and such that all subsets of Λ_n can be generated by periodically translating some X_i. For example, if $\Lambda = 0, 1, 2, 3$ on a one-dimensional lattice, then $X_1 = \{0\}$, $X_2 = \{0, 1\}$, $X_3 = \{0, 2\}$, $X_4 = \{0, 1, 2\}$ and $X_5 = \{0, 1, 2, 3\}$ is a possible choice of those sets. Then we have

$$H_{\Lambda_n}^p = \sum_{i=1}^N \sum_{T\in\mathbf{T}(\Lambda_n)} T\Phi(X_i)T^\dagger,$$

and from the representation it becomes clear that H_Λ^p is Λ_n-translation-invariant.

Note that we do not consider what Simon calls "external boundary conditions".

In the following, we will frequently use that the energy density does not depend on the choice of boundary conditions; that is, if $(\tau_n)_{n\in\mathbb{N}}$ is an arbitrary sequence of states on $\mathcal{A}_n$, then

$$\lim_{n\to\infty}\left(\frac{\operatorname{tr}(\tau_n H_{\Lambda_n})}{|\Lambda_n|} - \frac{\operatorname{tr}(\tau_n H_{\Lambda_n}^{BC})}{|\Lambda_n|}\right) = 0.$$

This is because $|\operatorname{tr}(\tau_n H_{\Lambda_n}) - \operatorname{tr}(\tau_n H_{\Lambda_n}^{BC})| \leq \|H_{\Lambda_n} - H_{\Lambda_n}^{BC}\|_\infty$.

Lemma 8. *Let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of density matrices in $\mathcal{A}_n$ such that every τ_n is Λ_n-translation-invariant. For every $m \in \mathbb{N}$, consider the sequence of states $(\rho_n^{(m)})_{n\in\mathbb{N}} \in \mathcal{A}_m$, defined for $n \geq m$ by*

$$\rho_n^{(m)} := \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n.$$

Define $L^{(m)}$ as the set of all limit points of the sequence $(\rho_n^{(m)})_{n\in\mathbb{N}}$, and L as the set of all possible sequences $(\sigma_m)_{m\in\mathbb{N}}$ with $\sigma_m \in L^{(m)}$ and $\sigma_{m-1} = \mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} \sigma_m$. Then L is not empty, and every element of L defines a translation-invariant state on the quasi-local algebra. Additionally, if $\beta \geq 0$ is such that

$$\liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \left(S(\tau_n) - \beta\, \mathrm{tr}(\tau_n H_{\Lambda_n})\right) \geq p(\beta, \Phi), \tag{16}$$

then every state $\omega \in L$ is a Gibbs state at inverse temperature β, and we have equality in (16). Furthermore, if L contains only a single element ω_β, then $\lim_{n\to\infty} \frac{1}{|\Lambda_n|}\mathrm{tr}(\tau_n H_{\Lambda_n}) = u(\omega_\beta)$ and $\lim_{n\to\infty} \frac{1}{|\Lambda_n|} S(\tau_n) = s(\omega_\beta)$.

Proof. First, we observe that every element of $L^{(m)}$ generates an element of $L^{(m-1)}$ by taking the partial trace over $\Lambda_m \setminus \Lambda_{m-1}$; that is,

$$\mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} L^{(m)} \subseteq L^{(m-1)}. \tag{17}$$

Similarly, suppose that $\rho \in L^{(m-1)}$. By definition, this means that there is a strictly increasing sequence of natural numbers $(n_k)_{k\in\mathbb{N}}$ such that $\rho_{n_k}^{(m-1)} \overset{k\to\infty}{\longrightarrow} \rho$. Now consider the sequence $\rho_{n_k}^{(m)}$; since m is fixed, it is a bounded sequence on a finite-dimensional vector space. By Bolzano-Weierstraß, it must have at least one limit point $\bar\rho$. Since $\rho_{n_k}^{(m-1)} = \mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} \rho_{n_k}^{(m)}$, we obtain $\rho = \mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} \bar\rho$. We have thus proven that

$$\text{for every } \rho \in L^{(m-1)}, \text{ there is } \bar\rho \in L^{(m)} \text{ such that } \rho = \mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} \bar\rho. \tag{18}$$

Furthermore, by Bolzano-Weierstraß, $L^{(1)}$ is non-empty. Combining the properties (17) and (18), we obtain $\mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} L^{(m)} = L^{(m-1)}$ as an equality between non-empty sets. This is sketched in Fig. 3, where we plot elements of $L^{(m)}$ as dots, with an edge connecting two dots if the left element (in $L^{(m-1)}$) is the partial trace of the right one (in $L^{(m)}$). Wandering from left to the right, no path will lead to a dead end; furthermore, every point can be reached this way by starting with some element in $L^{(1)}$. Thus, there is at least one path that starts with some element $\sigma_1 \in L^{(1)}$ and extends to infinity – that is, a sequence $(\sigma_m)_{m\in\mathbb{N}}$ with $\sigma_{m-1} = \mathrm{Tr}_{\Lambda_m\setminus\Lambda_{m-1}} \sigma_m$. L is the set of all these paths and hence not empty. Every $\omega \in L$ can be interpreted as a state: for any finite region $\Lambda \subset \mathbb{Z}^\nu$, take the smallest n such that $\Lambda \subset \Lambda_n$, and set $\omega_\Lambda := \mathrm{Tr}_{\Lambda_n\setminus\Lambda} \omega_m$. This defines a consistent family of density matrices, hence a state on the quasi-local algebra.

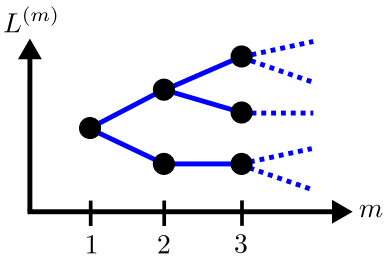

Fig. 3. Schematics of the sequence $L^{(m)}$ of limit points, as defined in the proof of Lemma 8. Note that $|L^{(1)}| > 1$ is possible

Now let $\omega \in L$ be any state. We claim that ω is translation-invariant. It will be sufficient to show the invariance equation $\omega(\gamma_y(A)) = \omega(A)$ for observables $0 \leq A \leq \mathbf{1}$ and translations γ_{δ_j}, where $\delta_j = (0, \ldots, 0, \underbrace{1}_{j}, 0, \ldots, 0)$. So let $\Lambda \subset \mathbb{Z}^\nu$ be finite, $A \in \mathcal{A}_\Lambda$ an observable with $0 \leq A \leq \mathbf{1}$, and $j \in \{1, \ldots, d\}$; set $\gamma := \gamma_{\delta_j}$. Choose m large enough such that $\Lambda \subseteq \Lambda_m$ and $\Lambda + \delta_j \subseteq \Lambda_m$. Let $\varepsilon > 0$ be arbitrary. Since $\omega_{\Lambda_m} \in L^{(m)}$, there is some $n \geq m$ such that

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \omega_{\Lambda_m} \right\|_1 < \varepsilon.$$

The effect of the translation of the observable A in the region Λ_n can be written

$$\gamma(A) \otimes \mathbf{1}_{\Lambda_n \setminus (\Lambda + \delta_j)} = T_j (A \otimes \mathbf{1}_{\Lambda_n \setminus \Lambda}) T_j^\dagger,$$

where $T_j \in \mathcal{A}_{\Lambda_n}$ is the unitary translation operator in Λ_n as defined shortly before Definition 7. Using this, we see that there are two real numbers Δ, $\Delta' \in \mathbb{R}$ with $|\Delta| < 2\varepsilon$, $|\Delta'| < 2\varepsilon$ such that

$$\begin{aligned}
\omega(\gamma(A)) &= \mathrm{tr}\left[\omega_{\Lambda_m}\left(\gamma(A) \otimes \mathbf{1}_{\Lambda_m \setminus (\Lambda + \delta_j)}\right)\right] \\
&= \mathrm{tr}\left[\left(\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n\right)\left(\gamma(A) \otimes \mathbf{1}_{\Lambda_m \setminus (\Lambda + \delta_j)}\right)\right] + \Delta \\
&= \mathrm{tr}\left[\tau_n \left(\gamma(A) \otimes \mathbf{1}_{\Lambda_n \setminus (\Lambda + \delta_j)}\right)\right] + \Delta = \mathrm{tr}\left[\tau_n T_j (A \otimes \mathbf{1}_{\Lambda_n \setminus \Lambda}) T_j^\dagger\right] + \Delta \\
&= \mathrm{tr}\left[T_j^\dagger \tau_n T_j (A \otimes \mathbf{1}_{\Lambda_n \setminus \Lambda})\right] + \Delta = \mathrm{tr}\left[\tau_n (A \otimes \mathbf{1}_{\Lambda_n \setminus \Lambda})\right] + \Delta \\
&= \mathrm{tr}\left[\left(\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n\right)(A \otimes \mathbf{1}_{\Lambda_m \setminus \Lambda})\right] + \Delta = \mathrm{tr}\left[\omega_{\Lambda_m}(A \otimes \mathbf{1}_{\Lambda_m \setminus \Lambda})\right] + \Delta + \Delta' \\
&= \omega(A) + \Delta + \Delta'.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this proves translation-invariance of ω.

In particular, every $\omega \in L$ has a well-defined entropy rate $s(\omega)$; what can we say about it? Fix $m \in \mathbb{N}$, and let $\omega_m := \omega_{\Lambda_m}$. Remember that $\rho_n^{(m)} = \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n$. Since $\omega_m \in L^{(m)}$, there exists a sequence $(n_k)_{k \in \mathbb{N}}$ such that

$$\rho_{n_k}^{(m)} \overset{k \to \infty}{\longrightarrow} \omega_m.$$

Fix $k \in \mathbb{N}$. We now decompose Λ_{n_k} into a disjoint union of boxes, where each box is a translate of Λ_m (where we consider translations as in the notion of Λ_{n_k}-translation invariance – that is, we regard Λ_{n_k} as a torus). In general, this cannot be done perfectly, but there will be some remaining part of Λ_{n_k} not covered by a translate of Λ_m. To spell out the details, let $a_m^{(1)}, \ldots, a_m^{(\nu)}$ denote the sidelengths of the box Λ_m, and $a_{n_k}^{(1)}, \ldots, a_{n_k}^{(\nu)}$ the sidelengths of Λ_{n_k}. Write

$$a_{n_k}^{(i)} = \ell_i \cdot a_m^{(i)} + j_i, \qquad \text{where } 0 \leq j_i < a_m^{(i)}.$$

Clearly, all ℓ_i tend to infinity for $k \to \infty$ on fixed m. Let $N_k := \ell_1 \cdot \ell_2 \cdot \ldots \cdot \ell_\nu$, then there are N_k translates $\Lambda_m^{(1)}, \Lambda_m^{(2)}, \ldots, \Lambda_m^{(N_k)}$ of Λ_m and the remainder $\Lambda_{\mathrm{rem}} \subset \Lambda_{N_k}$, all of them pairwise disjoint, such that

$$\Lambda_{n_k} = \bigcup_{i=1}^{N_k} \Lambda_m^{(i)} \cup \Lambda_{\mathrm{rem}}.$$

We have $|\Lambda_{\mathrm{rem}}| = |\Lambda_{n_k}| - N_k \cdot a_m^{(1)} \cdot \ldots \cdot a_m^{(\nu)}$, hence

$$\begin{aligned}\frac{|\Lambda_{\mathrm{rem}}|}{N_k} &= \frac{|\Lambda_{n_k}|}{N_k} - a_m^{(1)} \cdot \ldots \cdot a_m^{(\nu)} = \frac{a_{n_k}^{(1)} \cdot a_{n_k}^{(2)} \cdot \ldots \cdot a_{n_k}^{(\nu)}}{\ell_1 \ell_2 \ldots \ell_\nu} - a_m^{(1)} a_m^{(2)} \ldots a_m^{(\nu)} \\ &< \frac{(\ell_1+1) a_m^{(1)} \cdot \ldots \cdot (\ell_\nu+1) a_m^{(\nu)}}{\ell_1 \ell_2 \ldots \ell_\nu} - a_m^{(1)} a_m^{(2)} \ldots a_m^{(\nu)} \\ &= |\Lambda_m| \cdot \left[\left(1+\frac{1}{\ell_1}\right) \cdots \left(1+\frac{1}{\ell_\nu}\right) - 1 \right] \overset{k\to\infty}{\longrightarrow} 0.\end{aligned}$$

As a consequence, we also obtain

$$\lim_{k\to\infty} \frac{|\Lambda_{n_k}|}{N_k} = \lim_{k\to\infty} \left(|\Lambda_m| + \frac{|\Lambda_{\mathrm{rem}}|}{N_k} \right) = |\Lambda_m|.$$

Since τ_{n_k} is Λ_{n_k}-translation-invariant, its marginals on all the boxes $\Lambda_m^{(i)}$ are equal, that is, equal to $\rho_{n_k}^{(m)}$. Due to subadditivity of von Neumann entropy S, we have

$$S(\tau_{n_k}) \leq N_k S(\rho_{n_k}^{(m)}) + S(\tau_{\Lambda_{\mathrm{rem}}}) \leq N_k S(\rho_{n_k}^{(m)}) + |\Lambda_{\mathrm{rem}}| \cdot \log d,$$

where d is the single-site Hilbert space dimension. Thus, we obtain

$$\begin{aligned}S(\omega_m) &= \lim_{k\to\infty} S(\rho_{n_k}^{(m)}) \\ &\geq \limsup_{k\to\infty} \frac{1}{N_k} \left[S(\tau_{n_k}) - |\Lambda_{\mathrm{rem}}| \cdot \log d \right] \\ &= \limsup_{k\to\infty} \frac{1}{N_k} S(\tau_{n_k}) = |\Lambda_m| \limsup_{k\to\infty} \frac{1}{|\Lambda_{n_k}|} S(\tau_{n_k}). \end{aligned} \tag{19}$$

Furthermore, we can estimate the energy expectation value of ω_m as follows. Define a Hamiltonian $H_{\Lambda_{n_k}}^{(m)}$ on Λ_{n_k} by "switching off" all interaction terms that are not fully contained in one of the $\Lambda_m^{(i)}$, that is,

$$H_{\Lambda_{n_k}}^{(m)} := \sum_{i=1}^{N_k} \sum_{X \subset \Lambda_m^{(i)}} \Phi(X).$$

We can estimate the norm difference of $H_{\Lambda_{n_k}}^{(m)}$ and $H_{\Lambda_{n_k}}$ as follows. All missing terms are either fully contained in Λ_{rem}, or act across the boundary of some $\Lambda_m^{(i)}$. With the boundary $\partial \Lambda_m^{(i)}$ as defined in (15), we obtain $\lim_{m\to\infty} |\partial \Lambda_m| / |\Lambda_m| = 0$, and due to finite-range of the interaction Φ, there are constants $c_1, c_2 > 0$ such that

$$\left\| H_{\Lambda_{n_k}} - H_{\Lambda_{n_k}}^{(m)} \right\|_\infty \leq c_1 |\Lambda_{\mathrm{rem}}| + c_2 \sum_{i=1}^{N_k} |\partial \Lambda_m^{(i)}| = c_1 |\Lambda_{\mathrm{rem}}| + c_2 N_k |\partial \Lambda_m|.$$

By construction and translation-invariance of Φ, we have

$$\mathrm{tr}\left(\tau_{n_k} H_{\Lambda_{n_k}}^{(m)} \right) = N_k \, \mathrm{tr}\left(\rho_{n_k}^{(m)} H_{\Lambda_m} \right).$$

Combining these identities, we get

$$\begin{aligned}\frac{1}{|\Lambda_m|}\mathrm{tr}(\omega_m H_{\Lambda_m}) &= \frac{1}{|\Lambda_m|}\lim_{k\to\infty}\mathrm{tr}\left(\rho^{(m)}_{n_k} H_{\Lambda_m}\right) = \frac{1}{|\Lambda_m|}\lim_{k\to\infty}\frac{1}{N_k}\mathrm{tr}\left(\tau_{n_k} H^{(m)}_{\Lambda_{n_k}}\right)\\ &\le \frac{1}{|\Lambda_m|}\limsup_{k\to\infty}\frac{1}{N_k}\left(\mathrm{tr}(\tau_{n_k} H_{\Lambda_{n_k}}) + c_1|\Lambda_{\mathrm{rem}}| + c_2 N_k|\partial\Lambda_m|\right)\\ &= \frac{c_2|\partial\Lambda_m|}{|\Lambda_m|} + \limsup_{k\to\infty}\frac{1}{|\Lambda_{n_k}|}\mathrm{tr}\left(\tau_{n_k} H_{\Lambda_{n_k}}\right). \end{aligned} \tag{20}$$

Since $\liminf(a_n + b_n) \le \liminf a_n + \limsup b_n$, we obtain

$$\begin{aligned}\frac{1}{|\Lambda_m|}\left(S(\omega_m) - \beta\,\mathrm{tr}(\omega_m H_{\Lambda_m})\right) &\ge \limsup_{k\to\infty}\frac{1}{|\Lambda_{n_k}|}S(\tau_{n_K})\\ &\quad -\beta\limsup_{k\to\infty}\frac{1}{|\Lambda_{n_k}|}\mathrm{tr}\left(\tau_{n_k} H_{\Lambda_{n_k}}\right) - \beta c_2\frac{|\partial\Lambda_m|}{|\Lambda_m|}\\ &\ge \liminf_{k\to\infty}\frac{1}{|\Lambda_{n_k}|}\left(S(\tau_{n_k}) - \beta\,\mathrm{tr}(\tau_{n_k} H_{\Lambda_{n_k}})\right) - \beta c_2\frac{|\partial\Lambda_m|}{|\Lambda_m|}\\ &\ge \liminf_{n\to\infty}\frac{1}{|\Lambda_n|}\left(S(\tau_n) - \beta\,\mathrm{tr}(\tau_n H_{\Lambda_n})\right) - \beta c_2\frac{|\partial\Lambda_m|}{|\Lambda_m|}. \end{aligned}$$

Taking the limit $m \to \infty$ finally shows that $s(\omega) - \beta\, u(\omega) \ge p(\beta, \Phi)$. Since ω is translation-invariant, we must have equality, and ω must be a Gibbs state.

Now suppose that L contains only a single element, then so does $L^{(m)}$; hence $\rho^{(m)}_n$ converges for $n \to \infty$, and we can choose the convergent subsequence to be $n_k = k$. Repeating the calculation of (20) with inequality in both directions yields

$$\begin{aligned}\liminf_{k\to\infty}\frac{1}{|\Lambda_k|}\mathrm{tr}(\tau_k H_{\Lambda_k}) &\ge \frac{1}{|\Lambda_m|}\mathrm{tr}(\omega_m H_{\Lambda_m}) - \frac{c_2|\partial\Lambda_m|}{|\Lambda_m|},\\ \limsup_{k\to\infty}\frac{1}{|\Lambda_k|}\mathrm{tr}(\tau_k H_{\Lambda_k}) &\le \frac{1}{|\Lambda_m|}\mathrm{tr}(\omega_m H_{\Lambda_m}) + \frac{c_2|\partial\Lambda_m|}{|\Lambda_m|}, \end{aligned}$$

By taking the limit $m \to \infty$ of the right-hand side, we obtain $\lim_{n\to\infty}\frac{1}{|\Lambda_n|}\mathrm{tr}(\tau_n H_{\Lambda_n}) = u(\omega_\beta)$. Then it follows directly from (16) that $\liminf_{n\to\infty}\frac{1}{|\Lambda_n|}S(\tau_n) = s(\omega_\beta)$. Furthermore, (19) shows that $\limsup_{n\to\infty}\frac{1}{|\Lambda_n|}S(\tau_n) \le s(\omega_\beta)$, hence $\lim_{n\to\infty}\frac{1}{|\Lambda_n|}S(\tau_n) = s(\omega_\beta)$. □

We can always define a *maximally mixed state* ω on the quasi-local algebra $\mathcal{A}_\infty$, by defining its local density matrix for finite $\Lambda \subset \mathbb{Z}^\nu$ as $\omega_\Lambda := \mathbf{1}_\Lambda / d^{|\Lambda|}$. It is easy to check that this is a consistent family of density matrices, defining a translation-invariant state on $\mathcal{A}_\infty$. According to Definition 5, its energy density exists; it is $u(\omega) = \lim_{n\to\infty}\mathrm{tr}(H_{\Lambda_n})/d^{|\Lambda_n|}$. This fact will be used in the following lemma. To state that lemma, we have to assume that the interaction Φ does not vanish – and, in addition, that it is not *physically equivalent* to zero. An example would be an interaction in one dimension (i.e. $\nu = 1$) with $\Phi(\{1,2\}) = -\Phi(\{1\}) \otimes \mathbf{1}_2$, such that the resulting Hamiltonian is zero up to boundary terms. For a formal definition of physical equivalence and a further example see [13]. Note also that Φ is physically equivalent to zero if and only if $p(\beta, \Phi) = \log d$ for all $\beta \ge 0$, which is the same value as for $\Phi = 0$.

Lemma 9. *Let Φ be an interaction which is not physically equivalent to zero, with ground state energy density $u_{\min}(\Phi) = \lim_{n\to\infty} \lambda_{\min}(H_{\Lambda_n})/|\Lambda_n| = -\lim_{\beta\to\infty} \beta^{-1} p(\beta, \Phi)$ and infinite temperature energy density $u_{\max}(\Phi) := \lim_{n\to\infty} \mathrm{tr}(H_{\Lambda_n})/(|\Lambda_n| d^{|\Lambda_n|})$. Then, for every $u \in (u_{\min}(\Phi), u_{\max}(\Phi)]$, there exists a unique $\beta \equiv \beta(u) \geq 0$ such that there is at least one Gibbs state ω at inverse temperature β with energy density $u(\omega) = u$. Its entropy density is $s(\omega) = s(u) := p(\beta(u), \Phi) + u\,\beta(u)$, and this is the maximal possible entropy density of any translation-invariant state with energy density u.*

Proof. These statements are proven in [13]; uniqueness of $\beta(u)$ can be seen as follows. If Φ is not physically equivalent to zero, then the function $\beta \mapsto p(\beta, \Phi)$ is strictly convex, see [13, p. 349 and Theorem II.1.5]. Consider any translation-invariant state ω; it defines an affine-linear map $\beta \mapsto s(\omega) - \beta\, u(\omega) =: \ell_\omega(\beta)$. According to (13), the line ℓ_ω lies completely on or below of the graph of p; that is, $\ell_\omega(\beta) \leq p(\beta, \Phi)$ for all β. According to Definition 6, it is a Gibbs state if and only if ℓ_ω touches the graph of p; that is, if there is some β such that $\ell_\omega(\beta) = p(\beta)$. If we are given some value of u, then every translation-invariant state with this energy density has a corresponding line ℓ_ω with slope $(-u)$. Consider all those lines. Then only one of them can touch the graph of p, and it can do so in only one point, due to the strict convexity of p. The β-value of the unique touching point is then $\beta(u)$. □

Now we have all ingredients to prove our main theorem on the equivalence of ensembles.

Theorem 10 (Equivalence of ensembles). *Let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of Λ_n-translation-invariant states on $\mathcal{A}_n$, let $\beta \geq 0$, and let Φ be a translation-invariant finite-range interaction which is not physically equivalent to zero, and for which there is a unique Gibbs state ω_β at inverse temperature β. Suppose that*

$$\liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \left(S(\tau_n) - \beta\,\mathrm{tr}(\tau_n H_{\Lambda_n})\right) \geq p(\beta, \Phi),$$

then we have equality in this expression, and

$$\lim_{n\to\infty} \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m} \tau_n = (\omega_\beta)_{\Lambda_m} \tag{21}$$

for every $m \in \mathbb{N}$. Furthermore, we have

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} S(\tau_n) = s(\omega_\beta), \qquad \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\tau_n H_{\Lambda_n}) = u(\omega_\beta),$$

and

$$\lim_{n\to\infty} \left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m} \tau_n - \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right\|_1 = 0, \tag{22}$$

where $Z_n = \mathrm{tr}(\exp(-\beta H^p_{\Lambda_n}))$, and $H^p_{\Lambda_n}$ is the Hamiltonian on Λ_n with periodic boundary conditions. If the lattice dimension is $\nu = 1$, then $H^p_{\Lambda_n}$ in (22) can be replaced by H_{Λ_n}, the Hamiltonian with open boundary conditions. Furthermore, if Gibbs states are unique around inverse temperature β, define β_n^{BC} as the solution of the equation $\frac{1}{|\Lambda_n|} \mathrm{tr}\left(H^{BC}_{\Lambda_n} \frac{\exp(-\beta_n^{BC} H^p_{\Lambda_n})}{Z_n} \right) = u_n$, where BC denotes an arbitrary fixed choice of boundary conditions, and $(u_n)_{n\in\mathbb{N}}$ is an arbitrary sequence with $\lim_{n\to\infty} u_n = u(\omega_\beta)$. Then $\lim_{n\to\infty} \beta_n^{BC} = \beta$, and

$$\lim_{n\to\infty} \left\| \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n - \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta_n^{BC} H^p_{\Lambda_n})}{Z_n} \right\|_1 = 0. \tag{23}$$

Proof. Set $\rho_n^{(m)} := \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n$, and define $L^{(m)}$ and L exactly as in the statement of Lemma 8. Since there is only one Gibbs state ω_β at inverse temperature β, Lemma 8 implies that $L = \{\omega_\beta\}$, and so $L^{(m)} = (\omega_\beta)_{\Lambda_m}$ for all $m \in \mathbb{N}$. In other words, for every m, the state $(\omega_\beta)_{\Lambda_m}$ is the unique limit point of the sequence $(\rho_n^{(m)})_{n\in\mathbb{N}}$, and thus the limit of this sequence. This proves the first identity. To infer the second identity, Eq. (22), either apply [13, Theorem IV.2.12], or note that $\tau_n' := \exp(-\beta H^p_{\Lambda_n})/Z_n$ maximizes the functional $\rho \mapsto S(\rho) - \beta\,\mathrm{tr}(H^p_{\Lambda_n}\rho)$, thus

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\tau_n') - \beta\,\mathrm{tr}(\tau_n' H_{\Lambda_n})\right) &= \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\tau_n') - \beta\,\mathrm{tr}(\tau_n' H^p_{\Lambda_n})\right) \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S((\omega_\beta)_{\Lambda_n}) - \beta\,\mathrm{tr}((\omega_\beta)_{\Lambda_n} H^p_{\Lambda_n})\right) \\
&= s(\omega_\beta) - \beta\, u(\omega_\beta) = p(\beta,\Phi).
\end{aligned}$$

Thus $\lim_{n\to\infty}(\mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n - \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n') = (\omega_\beta)_{\Lambda_m} - (\omega_\beta)_{\Lambda_m} = 0$. Note that this also shows that $\lim_{n\to\infty}\frac{1}{|\Lambda_n|}\mathrm{tr}(\tau_n' H_{\Lambda_n}) = u(\omega_\beta)$. In the case of lattice dimension $\nu = 1$, apply the fact that in this case, the local Gibbs state $\exp(-\beta H_{\Lambda_n})/Z_n$ weakly converges to the unique global Gibbs state in the limit $n\to\infty$, as shown in [15].

It remains to prove (23). To this end, use the notation $Z_n(\beta) := \mathrm{tr}(\exp(-\beta H^p_{\Lambda_n}))$, and $\tau_n' := \exp(-\beta_n^{BC} H^p_{\Lambda_n})/Z_n(\beta_n^{BC})$. First we have to show that β_n^{BC} is well-defined for n large enough and that it is a bounded sequence. Set $\rho(\beta') := \exp(-\beta' H^p_{\Lambda_n})/Z_n(\beta')$ for $\beta' \geq 0$. Choose $\beta_0, \beta_1 \in \mathbb{R}$ such that $0 < \beta_0 < \beta < \beta_1$, and such that the Gibbs states at inverse temperatures β_0 and β_1 are unique. Then the previous results show that $\lim_{n\to\infty}\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H^{BC}_{\Lambda_n}\rho(\beta_i)\right) = u_i$ for $i = 0, 1$, where $u_i := u(\omega_{\beta_i})$. It follows $u_0 > u > u_1$, and thus for n large enough, we have $\frac{1}{|\Lambda_n|}\mathrm{tr}(H^{BC}_{\Lambda_n}\rho(\beta_0)) > u_n > \frac{1}{|\Lambda_n|}\mathrm{tr}(H^{BC}_{\Lambda_n}\rho(\beta_1))$, so $\beta_0 < \beta_n^{BC} < \beta_1$ for n large enough; in particular, a solution β_n^{BC} can be found in the interval (β_0, β_1). Moreover, since β_0 and β_1 can be chosen arbitrarily close to β, this proves that $\lim_{n\to\infty}\beta_n^{BC} = \beta$. Direct calculation shows that $S(\tau_n') = \log Z_n(\beta_n^{BC}) + \beta_n^{BC} u_n |\Lambda_n|$, thus

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\tau_n') - \beta\,\mathrm{tr}(\tau_n' H_{\Lambda_n})\right) \\
&= \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(\log Z_n(\beta_n^{BC}) + \beta_n^{BC}\, u_n|\Lambda_n| - \beta\,\mathrm{tr}(\tau_n' H^{BC}_{\Lambda_n})\right) \\
&= \liminf_{n\to\infty} \left((\beta_n^{BC} - \beta)u_n + \frac{1}{|\Lambda_n|}\log Z_n(\beta_n^{BC})\right) \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(\log Z_n(\beta) - |\log Z_n(\beta_n^{BC}) - \log Z_n(\beta)|\right) \\
&\geq p(\beta,\Phi) - \limsup_{k\to\infty} \frac{|\beta_n^{BC} - \beta|\,\|H^p_{\Lambda_n}\|_\infty}{|\Lambda_n|} = p(\beta,\Phi),
\end{aligned}$$

where we have used that $|\log Z_n(\beta_n^{BC}) - \log Z_n(\beta)| \leq |\beta_n^{BC} - \beta| \cdot \|H^p_{\Lambda_n}\|_\infty$, see [13, Lemma II.2.2Q]. This shows that $\lim_{n\to\infty} \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m} \tau'_n = (\omega_\beta)_{\Lambda_m}$. Combining this with (21) proves (23). □

In order to obtain some concrete instances of this equivalence of ensembles result, we need a series of lemmas. The first one is given in [13, Theorem IV.2.14], though with typos; see also [11,12], and for newer results on equivalence of ensembles, see [19]. Since the lemma is crucial for our paper, we give the proof for completeness, translating the proof of [13, Theorem III.4.15] to the quantum case.

Lemma 11. *Suppose that Φ is any finite-range translation-invariant interaction, not physically equivalent to zero. Then, for all $u \in (u_{\min}(\Phi), u_{\max}(\Phi))$, we have*

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left|\left\{\textit{eigenvalues of } H_{\Lambda_n} \leq u \cdot |\Lambda_n|\right\}\right| = s(u),$$

where $s(u)$ is defined in Lemma 9.

Proof. We transfer the classical proof of [13, Thm. III.4.15] to the quantum case (with slight modifications and simplifications, using notation established earlier). Define

$$N_{\Lambda_n}(u) := \left|\left\{\text{eigenvalues of } H_{\Lambda_n} \leq u \cdot |\Lambda_n|\right\}\right|$$
$$\bar{s}(u) := \limsup_{n\to\infty} \frac{1}{|\Lambda_n|} \log N_{\Lambda_n}, \qquad \underline{s}(u) := \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \log N_{\Lambda_n}.$$

Denote the eigenvalues of H_{Λ_n} by E_i, and $Z := \mathrm{tr}(\exp(-\beta(u)H_{\Lambda_n}))$, then

$$1 \geq \frac{1}{Z} \sum_{E_i\,:\, E_i/|\Lambda_n| \leq u} e^{-\beta(u)E_i} \geq \frac{1}{Z} N_{\Lambda_n} e^{-\beta(u)u|\Lambda_n|}.$$

Taking logarithms, we obtain $\frac{1}{|\Lambda_n|} \log N_{\Lambda_n} \leq \frac{1}{|\Lambda_n|} \log Z + \beta(u)u \overset{n\to\infty}{\longrightarrow} s(u)$, hence

$$\bar{s}(u) \leq s(u). \tag{24}$$

The converse inequality is more involved. Fix $u_1 \leq u_2$, $\delta > 0$, and $0 < \lambda < 1$. Use the notation of the proof of Lemma 8, where we have split Λ_{n_k} into disjoint regions $\Lambda_m^{(i)}$, $i = 1, \ldots, N_k$, and Λ_{rem}. Set $n_k = k$. Denote by $|E_1\rangle, \ldots, |E_M\rangle$ mutually orthonormal eigenvectors of H_{Λ_m} with energy density less than or equal to u_1, and $|E'_1\rangle, \ldots, |E'_N\rangle$ mutually orthonormal eigenvectors of H_{Λ_m} with energy density less than or equal to $u_2 - \delta$, where $M := N_{\Lambda_m}(u_1)$ and $N := N_{\Lambda_m}(u_2 - \delta)$. Set $i := (i_1, \ldots, i_{N_k})$, where $i_1, \ldots, i_{\lfloor \lambda N_k \rfloor} \in \{1, \ldots, M\}$, and $i_{\lfloor \lambda N_k \rfloor + 1}, \ldots, i_{N_k} \in \{1, \ldots, N\}$. For every possible choice of i, define

$$|\psi_i\rangle := \bigotimes_{l=1}^{\lfloor \lambda N_k \rfloor} |E_{i_l}\rangle_{\Lambda_m^{(l)}} \otimes \bigotimes_{l=\lfloor \lambda N_k \rfloor + 1}^{N_k} |E'_{i_l}\rangle_{\Lambda_m^{(l)}} \otimes |0\rangle_{\Lambda_{\mathrm{rem}}},$$

where $|0\rangle_{\Lambda_{\mathrm{rem}}}$ is an arbitrary pure state on Λ_{rem}. Then we have

$$\begin{aligned}
\frac{\langle\psi_i|H_{\Lambda_k}|\psi_i\rangle}{|\Lambda_k|} &\leq \frac{\langle\psi_i|H^{(m)}_{\Lambda_k}|\psi_i\rangle + \|H_{\Lambda_k} - H^{(m)}_{\Lambda_k}\|_\infty}{|\Lambda_k|} \\
&\leq \frac{\lfloor \lambda N_k \rfloor}{|\Lambda_k|} |\Lambda_m| u_1 + \frac{N_k - \lfloor \lambda N_k \rfloor}{|\Lambda_k|} |\Lambda_m| (u_2 - \delta) \\
&\quad + c_1 \frac{|\Lambda_{\mathrm{rem}}|}{|\Lambda_k|} + c_2 \frac{N_k}{|\Lambda_k|} |\partial \Lambda_m|.
\end{aligned}$$

If k and m are large enough (while $k \gg m$), the right-hand side is less than $u' := \lambda u_1 + (1-\lambda)u_2$. Furthermore, if $i \neq i'$ then $|\psi_i\rangle \perp |\psi_{i'}\rangle$, thus $N_{\Lambda_k}(u') \geq |\{|\psi_i\rangle\}| = M^{\lfloor \lambda N_k \rfloor} N^{N_k - \lfloor \lambda N_k \rfloor}$. Taking logarithms, we obtain

$$\frac{1}{|\Lambda_k|} \log N_{\Lambda_k}\left(u'\right) \geq \frac{1}{|\Lambda_k|} \left(\lfloor \lambda N_k \rfloor \log N_{\Lambda_m}(u_1) + (N_k - \lfloor \lambda N_k \rfloor) \log N_{\Lambda_m}(u_2 - \delta) \right).$$

Since $\lim_{k\to\infty} N_k/|\Lambda_k| = 1/|\Lambda_m|$, this yields

$$\underline{s}\left(u'\right) \geq \frac{\lambda}{|\Lambda_m|} \log N_{\Lambda_m}(u_1) + \frac{1-\lambda}{|\Lambda_m|} \log N_{\Lambda_m}(u_2 - \delta),$$

and thus

$$\underline{s}\left(\lambda u_1 + (1-\lambda)u_2\right) \geq \lambda \underline{s}(u_1) + (1-\lambda)\underline{s}(u_2 - \delta). \tag{25}$$

Now consider a fixed value of u, and set $\beta := \beta(u)$. We use the elementary inequalities for $a \leq b$:

$$\limsup_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_{i:\, E_i/|\Lambda_k| \in [a,b]} e^{-\beta E_i} \leq -\beta a + \bar{s}(b), \tag{26}$$

$$\liminf_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_{i:\, E_i/|\Lambda_k| \in [a,b]} e^{-\beta E_i} \leq -\beta a + \underline{s}(b), \tag{27}$$

where the E_i are now the eigenvalues of H_{Λ_k}. Suppose that β is a point of differentiability of $p(\cdot, \Phi)$ such that (due to strict convexity) $s(u') - \beta u' < p(\beta, \Phi)$ for all $u' \neq u$. Let $\alpha > 0$ such that $s(u') - \beta u' \leq p(\beta, \Phi) - \alpha$ for all u' with $|u' - u| > \delta$. Choose $\varepsilon > 0$ such that $\varepsilon\beta \leq \alpha/2$. Now we decompose the energy density interval into a disjoint union

$$(u_{\min}(\Phi), u_{\max}(\Phi)] \setminus (u-\delta, u+\delta) = \bigcup_{j=1}^{n-1} I_j, \qquad \text{where } I_j = (a_j, b_j]$$
$$\text{with } |b_j - a_j| \leq \varepsilon; \;\; I_n := (u_{\max}(\Phi), \infty).$$

Due to (26), we have $\limsup_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_{i:\, E_i/|\Lambda_k| \in I_j} e^{-\beta E_i} \leq \beta a_j + \bar{s}(b_j)$, and due to (24), we obtain for $j \leq n-1$

$$\bar{s}(b_j) - \beta a_j \leq s(b_j) - \beta b_j + \beta(b_j - a_j) \leq s(b_j) - \beta b_j + \frac{\alpha}{2} \leq p(\beta, \Phi) - \frac{\alpha}{2},$$

and for $j = n$, we get

$$\bar{s}(b_j) - \beta a_j = \log d - \beta u_{\max}(\Phi) = s(a_n) - \beta a_n \leq p(\beta, \Phi) - \alpha.$$

Thus

$$\begin{aligned}
\limsup_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_{i:\, |E_i/|\Lambda_k| - u| \geq \delta} e^{-\beta E_i} &= \limsup_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_{j=1}^{n} \sum_{i:\, E_i/|\Lambda_k| \in I_j} e^{-\beta E_i} \\
&\leq \limsup_{k\to\infty} \frac{1}{|\Lambda_k|} \left(\log n + \max_j \sum_{i:\, E_i/|\Lambda_k| \in I_j} e^{-\beta E_i} \right) \\
&\leq p(\beta, \Phi) - \frac{1}{2}\alpha.
\end{aligned}$$

But since $\lim_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_i e^{-\beta E_i} = p(\beta, \Phi)$ by definition of the pressure, we obtain

$$\liminf_{k\to\infty} \frac{1}{|\Lambda_k|} \log \sum_{i:|E_i/|\Lambda_k|-u|\leq\delta} e^{-\beta E_i} \geq p(\beta, \Phi).$$

Comparing this to (27) yields $p(\beta, \Phi) \leq -\beta(u-\delta) + \underline{s}(u+\delta)$, hence

$$\underline{s}(u+\delta) \geq \lim_{\delta\to 0} \underline{s}(u+\delta) \geq \lim_{\delta\to 0} p(\beta, \Phi) + \beta(u-\delta) = p(\beta, \Phi) + \beta u = s(u).$$

This finally shows that

$$\underline{s}(u+\delta) \geq s(u) \qquad \text{for all } \delta > 0, \text{ if } \beta(u) \text{ is a point of differentiability of } \beta \mapsto p(\beta, \Phi).$$

Since $\beta \mapsto p(\beta, \Phi)$ is strictly convex, the right and left derivatives D^+p and D^-p exist everywhere, and the set $B := \{\beta > 0 \mid (D^+p)(\beta) \neq (D^-)(p)(\beta)\}$ is countable. Furthermore, the set $A := \{u \in (u_{\min}(\Phi), u_{\max}(\Phi)) \mid \beta(u) \in B\}$ is a countable union of closed intervals. If u is any value such that there is a sequence $(u_n)_{n\in\mathbb{N}}$, $u_n \leq u$, with $\lim_{n\to\infty} u_n = u$ and $u_n \notin A$, then $\underline{s}(u) \geq s(u_n)$, and due to continuity of s, we get $\underline{s}(u) \geq s(u)$. We get this inequality for all $u \notin A$ and the left-hand endpoints of intervals in A.

Finally, let $[u_0, u_1] \subset A$ be an isolated closed interval and $u \in (u_0, u_1]$. Then for every $\varepsilon \geq 0$ there is $\lambda_\varepsilon \in (0, 1)$ with $u = \lambda_\varepsilon(u_1 + \varepsilon) + (1-\lambda_\varepsilon)u_0$. Then, for every $\varepsilon > 0$ small enough such that $u_1 + \varepsilon \notin A$, and $\delta > 0$ small enough such that $u_0 - \delta \notin A$, we get due to (25)

$$\begin{aligned}\underline{s}(u) &= \underline{s}(\lambda_\varepsilon(u_1+\varepsilon) + (1-\lambda_\varepsilon)u_0) \geq \lambda_\varepsilon \underline{s}(u_1+\varepsilon) + (1-\lambda_\varepsilon)\underline{s}(u_0-\delta) \\ &\geq \lambda_\varepsilon s(u_1+\varepsilon) + (1-\lambda_\varepsilon)s(u_0-\delta).\end{aligned}$$

Since s is continuous, we can first take the limit $\delta \to 0$ and then the limit $\varepsilon \to 0$ to obtain

$$\underline{s}(u) \geq \lambda_0 s(u_1) + (1-\lambda_0)s(u_0) = s(u),$$

where we have used the fact that s is linear on $[u_0, u_1]$. Together with (24), this completes the proof. □

This lemma only refers to the Hamiltonian H_{Λ_n} corresponding to open boundary conditions. However, we need this in more generality, in particular for the case of periodic boundary conditions.

Lemma 12. *Let $H^{BC}_{\Lambda_n}$ be the Hamiltonians corresponding to an arbitrary choice of boundary conditions in the sense of Definition 7. Then*

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \textit{eigenvalues of } H^{BC}_{\Lambda_n} \leq u \cdot |\Lambda_n| \right\} \right| = s(u).$$

Proof. Define $H^{BC}_{\partial\Lambda_n} := H^{BC}_{\Lambda_n} - H_{\Lambda_n}$. Fix u, and let $\tilde{u} < u$ be arbitrary. If n is large enough, then

$$\tilde{u}|\Lambda_n| + \|H^{BC}_{\partial\Lambda_n}\|_\infty \leq u|\Lambda_n|.$$

Thus, due to Weyl's Perturbation Theorem [20], if $\lambda_1, \ldots, \lambda_k$ are the k smallest eigenvalues of H_{Λ_n}, then $H^{BC}_{\Lambda_n}$ has eigenvalues $\lambda'_i \leq \lambda_i + \|H^{BC}_{\partial\Lambda_n}\|$. Therefore

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \text{eigenvalues of } H^{BC}_{\Lambda_n} \leq u|\Lambda_n| \right\} \right| \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \text{eigenvalues of } H^{BC}_{\Lambda_n} \leq \tilde{u}|\Lambda_n| + \|H^{BC}_{\partial\Lambda_n}\| \right\} \right| \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \text{eigenvalues of } H_{\Lambda_n} \leq \tilde{u}|\Lambda_n|\| \right\} \right| = s(\tilde{u}).
\end{aligned}$$

By continuity of s, since this is true for all $\tilde{u} < u$, the previous inequality is also true if $s(\tilde{u})$ is replaced by $s(u)$. Similarly, if $\tilde{u} > u$ is arbitrary, then

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \text{eigenvalues of } H^{BC}_{\Lambda_n} \leq u|\Lambda_n| \right\} \right| \\
&\leq \limsup_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \text{eigenvalues of } H^{BC}_{\Lambda_n} \leq \tilde{u}|\Lambda_n| - \|H^{BC}_{\partial\Lambda_n}\| \right\} \right| \\
&\leq \limsup_{n\to\infty} \frac{1}{|\Lambda_n|} \log \left| \left\{ \text{eigenvalues of } H_{\Lambda_n} \leq \tilde{u}|\Lambda_n|\| \right\} \right| = s(\tilde{u}).
\end{aligned}$$

This proves the claim. □

As an immediate consequence we obtain the following result.

Example 13 (Microcanonical versus canonical ensemble). The sequence of states $(\tau_n)_{n\in\mathbb{N}}$ which are defined as the maximal mixtures on the microcanonical subspaces

$$T^p_n := \operatorname{span}\left\{ |E\rangle \;\middle|\; H^p_{\Lambda_n}|E\rangle = E|E\rangle,\ \frac{E}{|\Lambda_n|} \in (u-\delta, u) \right\},$$

where $H^p_{\Lambda_n}$ is the Hamiltonian on Λ_n with periodic boundary conditions satisfies the premises of Theorem 10. That is, we obtain equivalence of ensembles in the standard sense:

$$\lim_{n\to\infty} \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \tau_n - \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right\|_1 = 0,$$

where one may either set β equal to $\beta(u)$, the inverse temperature corresponding to energy density u in the thermodynamic limit, or equal to the (n-dependent) solution of $\frac{1}{|\Lambda_n|}\operatorname{tr}\left(H^{BC}_{\Lambda_n} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right) = u$, where BC denotes an arbitrary fixed choice of boundary conditions.

In this example, as well as in Theorem 10, the partial traces cannot be removed: globally, the microcanonical and the canonical ensemble will in general have large one-norm distance. In the example of a non-interacting system of binary spins, the well-known tightness of the classical finite de Finetti theorem provides a proof of this, see Lemma 34.

Furthermore, it is crucial to use the reduction of the global Gibbs state, $\operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \exp(-\beta H^p_{\Lambda_n})/Z_n$, instead of the local Gibbs state, $\exp(-\beta H^p_{\Lambda_m})$. Replacing the former by the latter renders the statement of the theorem false in general. This is rather obvious: the local Gibbs state will in general be different from the reduction of the global one, due to interaction terms across the boundary of Λ_m. This phenomenon will also occur in

Sect. 3.6, where we prove a special case of the "eigenstate thermalization hypothesis" only by taking the boundary terms into account. A concrete counterexample to the naive version of equivalence of ensembles is already given by the classical Ising model, interpreted as a quantum model.

Example 14 (The Ising model). Consider the one-dimensional model on $\Lambda_n := \{-n, \ldots, n\}$

$$H^p_{\Lambda_n} := -J \sum_{i=-n}^{n} Z_i Z_{i+1} - h \sum_{i=-n}^{n} Z_i, \tag{28}$$

identifying $n+1 \equiv -n$. Here, Z_i denotes the Pauli Z-matrix $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ on lattice site i. This model has a unique Gibbs state ω_β (in the thermodynamic limit $n \to \infty$) for all $\beta \geq 0$, see [24]. Fix $m = 0$, and consider the reduction of the global microcanonical state τ_n to $\Lambda_0 = \{0\}$, a single lattice site. Due to Example 13 and Theorem 10, we have

$$\lim_{n\to\infty} \mathrm{Tr}_{\Lambda_n \setminus \Lambda_0} \tau_n = (\omega_\beta)_{\Lambda_0}.$$

On the other hand, using the known formula for the magnetization of the Ising model [24], we have

$$\mathrm{tr}\left((\omega_\beta)_{\Lambda_0} Z\right) = \frac{1}{2n+1} \sum_{i=-n}^{n} \mathrm{tr}\left((\omega_\beta)_{\Lambda_n} Z_i\right) = \frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + \exp(-4\beta J)}}, \tag{29}$$

where the first equality is due to translation-invariance, and the second equality follows from taking the limit $n \to \infty$ and using the well-known result for the magnetization of this model. We can compare this with the local Gibbs state $\omega^{\mathrm{loc}}_\beta$, which is defined as the normalization of $\exp(-\beta H_{\Lambda_0})$. We run into an immediate conceptual problem: how do we define H_{Λ_0}? The most obvious choice is $H_{\Lambda_0} = Z_0$, but we have the freedom to interpret (28) in different ways, by subtracting local terms from $Z_i Z_{i+1}$ and adding them to the Z_i-term. This is exactly the freedom that we encountered before, in the definition of *physical equivalence* that we discussed before Lemma 9. Whatever we define to be H_{Λ_0}, it should be *some* fixed Hamiltonian which can be written in the form $H_{\Lambda_0} = U \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} U^\dagger$, with U unitary and $E_1, E_2 \in \mathbb{R}$ its energy eigenvalues. Our crucial assumption will be that whatever H_{Λ_0} is, it should be independent of β. But then

$$\mathrm{tr}(\omega^{\mathrm{loc}}_\beta Z) = \frac{\mathrm{tr}\left[U \begin{pmatrix} \exp(-\beta E_1) & 0 \\ 0 & \exp(-\beta E_2) \end{pmatrix} U^\dagger Z\right]}{\exp(-\beta E_1) + \exp(-\beta E_2)}.$$

Regarding this as a function $f(\beta)$ for complex $\beta \in \mathbb{C}$, we obtain a function that is holomorphic except for possibly countably many isolated singularities on the imaginary axis (if $E_1 \neq E_2$). This is not true for (29) which is a function with branch cut singularities due to the presence of the square root. This shows that $\omega^{\mathrm{loc}}_\beta \neq (\omega_\beta)_{\Lambda_0}$ at least for some values of $\beta > 0$, no matter how we define H_{Λ_0}. Thus $\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n$ cannot converge to $\omega^{\mathrm{loc}}_\beta$ in the thermodynamics limit where $n \to \infty$.

The standard microcanonical ensemble (mentioned in Example 13 above) is defined as a flat distribution on the energy windows subspace corresponding to the interval $(u-\delta, u)$. However, we can apply Theorem 10 more generally. In order to slightly generalize Example 13, we need another simple lemma:

Lemma 15. *Let $(p_1, \ldots, p_n)$ be discrete probability distribution, and suppose that there exists $M \geq 1$ such that $\frac{p_i}{p_j} \leq M$ for all $i \neq j$. Then its Shannon entropy satisfies $H(p) \geq \log n - \log M$.*

Proof. Let $\ell_i := \log(1/p_i)$, then $\log(p_i/p_j) = \ell_j - \ell_i$, and the condition above implies $|\ell_i - \ell_j| \leq \log M$ for all i, j. Then all ℓ_i lie in the interval $[\ell_{min}, \ell_{max}]$, where $\ell_{min} := \min_i \ell_i$ and $\ell_{max} := \max_i \ell_i$. This interval has size at most $\ell_{max} - \ell_{min} \leq \log M$. Since $\min_i p_i \leq 1/n \leq \max_i p_i$, the quantity $\log n$ must be contained in this interval. Thus $|\ell_i - \log n| \leq \log M$ for all i. It follows that

$$\begin{aligned} |H(p) - \log n| &= \left| \sum_i p_i \log \frac{1}{p_i} - \sum_i p_i \log n \right| \\ &= \left| \sum_i p_i \ell_i - \sum_i p_i \log n \right| \leq \sum_i p_i |\ell_i - \log n| \leq \log M. \end{aligned}$$

□

Now we apply this to prove a generalization of Example 13.

Example 16 (Microcanonical ensemble with given distribution function). Let Φ be an interaction which is not physically equivalent to zero, and let $u_{\min}(\Phi) < u \leq u_{\max}(\Phi)$. Let $f : [u_{\min}(\Phi), u_{\max}(\Phi)] \to \mathbb{R}$ be a bounded nonnegative function such that $f(x) = 0$ for all $x > u$ and such that there exists $\delta > 0$ such that f is continuous and strictly positive on $[u-\delta, u]$, cf. Fig. 4. For every $n \in \mathbb{N}$, let $\{|E_i\rangle\}_i$ be an arbitrary energy eigenbasis of $H^p_{\Lambda_n}$, the Hamiltonian on Λ_n with periodic boundary conditions. Then the set of states defined by

$$\tau_n := \frac{1}{\mathcal{N}} \sum_i f\left(\frac{E_i}{|\Lambda_n|}\right) |E_i\rangle\langle E_i|,$$

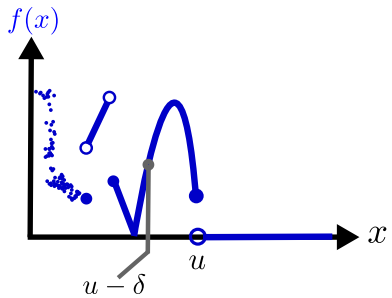

Fig. 4. Spectral density functions that satisfy the premises of Example 16, and yield equivalence of ensembles in the sense that the corresponding microcanonical state locally resembles the canonical state. The non-negative bounded function f must satisfy $f(x) = 0$ for all $x \geq u$, and there must be $\delta > 0$ such that f is continuous and strictly positive on the interval $[u-\delta, u]$. For $x < u - \delta$, f can have all kinds of discontinuities

where $\mathcal{N} := \sum_i f(E_i/|\Lambda_n|)$, satisfies the premises of Theorem 10. That is, this modified microcanonical ensemble resembles locally the canonical ensemble.

Proof. Since $\mathrm{tr}(\tau_n H^p_{\Lambda_n}) \leq u|\Lambda_n|$, we have $\bar{u} := \limsup_{n\to\infty} \frac{1}{|\Lambda_n|}\mathrm{tr}(\tau_n H_{\Lambda_n}) \leq u$. Define $I_n := \{i \mid E_i/|\Lambda_n| \in [u-\delta, u]\}$, then $\tau_n = (1-\lambda_n)\sigma_n + \lambda_n \sigma_n'$, where

$$\sigma_n = \sum_{i\in I_n} \frac{f(E_i/|\Lambda_n|)}{\sum_{j\in I_n} f(E_j/|\Lambda_n|)} |E_i\rangle\langle E_i|, \qquad \sigma_n' = \sum_{i\notin I_n} \frac{f(E_i/|\Lambda_n|)}{\sum_{j\notin I_n} f(E_j/|\Lambda_n|)} |E_i\rangle\langle E_i|,$$

$$\lambda_n = \frac{\sum_{j\notin I_n} f(E_j/|\Lambda_n|)}{\sum_j f(E_j/|\Lambda_n|)}.$$

According to Lemma 12, we have

$$\sum_{j\notin I_n} f(E_j/|\Lambda_n|) \leq \#\{i \mid E_i/|\Lambda_n| < u-\delta\} \cdot \|f\|_\infty = \exp[|\Lambda_n|\, s(u-\delta) + o(|\Lambda_n|)].$$

On the other hand,

$$\sum_j f(E_j/|\Lambda_n|) \geq \sum_{j\in I_n} f(E_j/|\Lambda_n|) \geq \#I_n \cdot \min_{x\in[u-\delta,u]} f(x) = \exp[|\Lambda_n|\, s(u) - o(|\Lambda_n|)].$$

This shows that $\lim_{n\to\infty} \lambda_n = 0$, and concavity of the entropy, i.e. $S(\tau_n) \geq (1-\lambda_n)S(\sigma_n) + \lambda_n S(\sigma_n')$, yields

$$\liminf_{n\to\infty} \frac{1}{|\Lambda_n|} S(\tau_n) \geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} S(\sigma_n).$$

But the eigenvalues of σ_n are $p_i := \dfrac{f(E_i/|\Lambda_n|)}{\sum_{j\in I_n} f(E_j/|\Lambda_n|)}$, such that $\dfrac{p_i}{p_j} = \dfrac{f(E_i/|\Lambda_n|)}{f(E_j/|\Lambda_n|)} \leq \dfrac{b}{a}$, where $a := \min_{x\in[u-\delta,u]} f(x)$ and $b := \max_{x\in[u-\delta,u]} f(x)$. Thus, Lemma 15 shows that

$$S(\sigma_n) \geq \log \#I_n - \log \frac{b}{a} = |\Lambda_n| \cdot s(u) - o(|\Lambda_n|),$$

and so $\underline{s} := \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} S(\tau_n) \geq s(u)$. In summary, we obtain

$$\liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \big(S(\tau_n) - \beta\, \mathrm{tr}(\tau_n H_{\Lambda_n})\big) \geq \underline{s} - \beta\, \bar{u} \geq s(u) - \beta\, u = p(\beta, \Phi).$$

This proves all the premises of Theorem 10. □

Remark. The condition that f has a discontinuity at u (i.e. $f(u) > 0$, but $f(x) = 0$ for all $x > u$) can be relaxed: the statement above will remain valid if $f(u) = 0$ as long as $f(x)$ does not tend to zero too quickly as $x \to u$. However, the question what "too quickly" means mathematically seems to depend on the choice of the model, because it depends on subtle properties of the spectrum of H_Λ, in particular on the number of eigenvalues in certain intervals with diameters of order $o(|\Lambda|)$. In this paper, we only analyze what can be said in full generality from translation-invariance alone, without reference to any details of the model.

The main proof idea used in this subsection—to apply the variational principle (13)—has been pioneered by Lima [11,12]. Our result however is more general:

- It involves more general spectral density functions (the function f in Example 16) instead of only the flat distribution,
- it allows to determine the inverse temperature from the energy density on the finite region Λ_n,
- it allows local lattice site dimensions larger than two, and, most significantly,
- Lima considers only a restricted set of interactions that commute with a particle number operator, see [12, p. 183], and [11, p. 63]. There is no such restriction in this work.

In the remainder of this subsection, we will consider the case of sequences of states $(\tau_n)_{n\in\mathbb{N}}$ that are not necessarily Λ_n-translation-invariant. The simplest example is given by the microcanonical ensembles (in the sense of Example 13) if boundary conditions are not periodic. The proof of Theorem 10 does not work any more, because we cannot guarantee that limit points of this sequence, as states on the quasi-local algebra, are translation-invariant.

However, we can still prove a version of equivalence of ensembles in this case, even though it will be a weaker version. This was already seen by Lima [12]. In a nutshell, we will prove an equivalence of ensemble result for a restricted set of observables. The following definition specifies the class of observables that we will consider.

Definition 17 *(m-block periodically averaged observable).* For $m \leq n$, an operator $A \in \mathcal{A}_n$ will be called an *m-block periodically averaged observable* if there exists $A' \in \mathcal{A}_m$ with $A' = (A')^\dagger$ such that

$$A = \frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{T\in\mathbf{T}(\Lambda_n)} T(A' \otimes \mathbf{1})T^\dagger, \tag{30}$$

where $\mathbf{T}(\Lambda_n)$ denotes all periodic translations of the finite region Λ_n into itself, and the unit observable is supported on $\Lambda_n \setminus \Lambda_m$. Moreover, A will be called an *m-block periodically averaged effect* if there exists $A' \in \mathcal{A}_m$ which satisfies the equation above, and additionally satisfies $0 \leq A' \leq \mathbf{1}$.

Note that m-block periodically averaged observables A on Λ_n are automatically Λ_n-translation-invariant. The notion "effect" refers to the property that they satisfy $0 \leq A \leq \mathbf{1}$ (as inherited from A'), and can thus be interpreted as defining a binary measurement with POVM elements $(A, \mathbf{1} - A)$.

The usual $\|\cdot\|_1$-distance on density matrices (which is twice the trace distance) can be interpreted (up to a factor of two) as the maximal possible difference of probabilities in any binary measurement that is applied to the states:

$$\|\rho - \sigma\|_1 = 2 \max_{0\leq P\leq\mathbf{1}} |\mathrm{tr}(P\rho) - \mathrm{tr}(P\sigma)| .$$

Similarly, we can define a pseudonorm that quantifies the extent to which two states differ in the expectation value of m-block periodically averaged effects: for $m \in \mathbb{N}$ and $M = M^\dagger \in \mathcal{A}_n$ with $n \geq m$, set

$$\|M\|_{\{m\}} := 2 \max \{|\mathrm{tr}(PM)| \mid P \text{ is an } m\text{-block periodically averaged effect on } \Lambda_n\} .$$

As a consequence, $\|\rho - \sigma\|_{\{m\}}$ denotes the maximal difference in probabilities of any measurements described by m-block averaged effects that are performed on ρ resp. σ. It is clear that $0 \leq \|A\|_{\{m\}} \leq \|A\|_1$, and the norm properties $\|\lambda A\|_{\{m\}} = |\lambda| \, \|A\|_{\{m\}}$ for

$\lambda \in \mathbb{R}$ as well as $\|A+B\|_{\{m\}} \leq \|A\|_{\{m\}} + \|B\|_{\{m\}}$ are satisfied. However, $\|A\|_{\{m\}}$ can be zero without A being zero, which shows that $\|\cdot\|_{\{m\}}$ is not a norm.

In the case where we have an m-block periodically averaged observable which does not come from an effect, we have the following inequality:

Lemma 18. *Let A be an m-block periodically averaged observable on Λ_n, coming from an observable $A' \in \mathcal{A}_m$ according to (30). Then for all quantum states ρ, σ on Λ_n, we have*

$$|\mathrm{tr}(\rho A) - \mathrm{tr}(\sigma A)| \leq \|A'\|_\infty \|\rho - \sigma\|_{\{m\}}.$$

Proof. Denote by $\lambda_{\min}$ resp. $\lambda_{\max}$ the smallest resp. largest eigenvalue of A'. If $\lambda_{\max} = \lambda_{\min}$ then there is nothing to prove. Otherwise, set $B' := (\lambda_{\max} - \lambda_{\min})^{-1}(A' - \lambda_{\min}\mathbf{1})$, then $0 \leq B' \leq \mathbf{1}$. Define $B := \frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{T \in \mathbf{T}(\Lambda_n)} T(B' \otimes \mathbf{1})T^\dagger$, then B is an m-block periodically averaged effect, and hence

$$|\mathrm{tr}(\rho B) - \mathrm{tr}(\sigma B)| \leq \frac{1}{2}\|\rho - \sigma\|_{\{m\}}.$$

On the other hand, we have $B = (\lambda_{\max} - \lambda_{\min})^{-1}(A - \lambda_{\min}\mathbf{1})$. Substituting this into the previous inequality, and using that $\lambda_{\max} - \lambda_{\min} \leq 2\|A'\|_\infty$, we obtain the claimed inequality. □

As a preparation, we need a lemma which says that *periodically averaged* local Gibbs states for arbitrary boundary conditions converge to the global Gibbs state if it is unique.

Lemma 19. *Fix any $\beta \geq 0$, and let $H_{\Lambda_n}^{BC}$ be a sequence of Hamiltonians with arbitrary boundary conditions, corresponding to an interaction Φ which is not physically equivalent to zero and which has a unique Gibbs state ω_β at inverse temperature β. Then, for every $m \in \mathbb{N}$,*

$$\lim_{n\to\infty} \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \left(\frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{T \in \mathbf{T}(\Lambda_n)} T \, \frac{\exp(-\beta H_{\Lambda_n}^{BC})}{Z_n} \, T^\dagger \right) = (\omega_\beta)_{\Lambda_m}, \tag{31}$$

where $Z_n = \mathrm{tr}\left(\exp(-\beta H_{\Lambda_n}^{BC})\right)$. Furthermore, if Gibbs states are unique around inverse temperature $\beta > 0$, and if we define β_n^{BC} as the solution of the equation $\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H_{\Lambda_n}^{BC} \frac{\exp(-\beta_n^{BC} H_{\Lambda_n}^{BC})}{Z_n}\right) = u_n$, with $(u_n)_{n\in\mathbb{N}}$ an arbitrary sequence with $\lim_{n\to\infty} u_n = u(\omega_\beta)$, then $\lim_{n\to\infty} \beta_n^{BC} = \beta$, and

$$\lim_{n\to\infty} \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \left(\frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{T \in \mathbf{T}(\Lambda_n)} T \, \frac{\exp(-\beta_n^{BC} H_{\Lambda_n}^{BC})}{Z_n'} \, T^\dagger \right) = (\omega_\beta)_{\Lambda_m},$$

where $Z_n' = \mathrm{tr}\left(\exp(-\beta_n^{BC} H_{\Lambda_n}^{BC})\right)$.

Proof. Set $\rho_n(\beta') := \exp(-\beta' H_{\Lambda_n}^{BC})/Z_n(\beta')$, and $\rho_n := \rho_n(\beta)$. By construction, ρ_n maximizes the functional $\rho \mapsto S(\rho) - \beta\,\mathrm{tr}(H_{\Lambda_n}^{BC}\rho)$. Thus

$$S(\rho_n) - \beta \operatorname{tr}(H^{BC}_{\Lambda_n}\rho_n) \geq S\left((\omega_\beta)_{\Lambda_n}\right) - \beta \operatorname{tr}\left(H^{BC}_{\Lambda_n}(\omega_\beta)_{\Lambda_n}\right).$$

Set $\rho'_n := 1/(|\mathbf{T}(\Lambda_n)|)\sum_{T\in\mathbf{T}(\Lambda_n)} T\rho_n T^\dagger$, then concavity of the entropy implies $S(\rho'_n) \geq S(\rho_n)$. Since $T^\dagger H^p_{\Lambda_n} T = H^p_{\Lambda_n}$ for all $T \in \mathbf{T}(\Lambda_n)$, we have $\operatorname{tr}(\rho'_n H^p_{\Lambda_n}) = \operatorname{tr}(\rho_n H^p_{\Lambda_n})$. We obtain

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\rho'_n) - \beta \operatorname{tr}(\rho'_n H_{\Lambda_n})\right)\\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\rho_n) - \beta \operatorname{tr}(\rho'_n H^p_{\Lambda_n})\right)\\
&= \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\rho_n) - \beta \operatorname{tr}(\rho_n H^p_{\Lambda_n})\right)\\
&= \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\rho_n) - \beta \operatorname{tr}(\rho_n H^{BC}_{\Lambda_n})\right)\\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S\left((\omega_\beta)_{\Lambda_n}\right) - \beta \operatorname{tr}\left(H^{BC}_{\Lambda_n}(\omega_\beta)_{\Lambda_n}\right)\right)\\
&= s(\omega_\beta) - \beta\, u(\omega_\beta) = p(\beta, \Phi).
\end{aligned} \tag{32}$$

Since every ρ'_n is Λ_n-translation-invariant, Theorem 10 proves (31) and also $\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \operatorname{tr}(\rho'_n H_{\Lambda_n}) = u(\omega_\beta)$. Thus

$$u(\omega_\beta) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|}\operatorname{tr}(\rho'_n H^p_{\Lambda_n}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|}\operatorname{tr}(\rho_n H^p_{\Lambda_n}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|}\operatorname{tr}(\rho_n H^{BC}_{\Lambda_n}).$$

Choose $\beta_0, \beta_1 \in \mathbb{R}$ such that $0 < \beta_0 < \beta < \beta_1$, and such that the Gibbs states at inverse temperatures β_0 and β_1 are unique. Then the previous results show that $\lim_{n\to\infty} \frac{1}{|\Lambda_n|}\operatorname{tr}\left(H^{BC}_{\Lambda_n}\rho_n(\beta_i)\right) = u_i$ for $i = 0, 1$, where $u_i := u(\omega_{\beta_i})$. It follows $u_0 > u > u_1$, and thus for n large enough, we have $\frac{1}{|\Lambda_n|}\operatorname{tr}(H^{BC}_{\Lambda_n}\rho_n(\beta_0)) > u_n > \frac{1}{|\Lambda_n|}\operatorname{tr}(H^{BC}_{\Lambda_n}\rho_n(\beta_1))$, so $\beta_0 < \beta^{BC}_n < \beta_1$ for n large enough; in particular, a solution β^{BC}_n can be found in the interval (β_0, β_1). Moreover, since β_0 and β_1 can be chosen arbitrarily close to β, this proves that $\lim_{n\to\infty} \beta^{BC}_n = \beta$. We can then repeat the calculation (32), with β after the minus sign replaced by β^{BC}_n where necessary, ω_β left unchanged, ρ_n replaced by $\rho_n(\beta^{BC}_n)$, and ρ'_n replaced by $\rho'_n(\beta^{BC}_n) := 1/(|\mathbf{T}(\Lambda_n)|)\sum_{T\in\mathbf{T}(\Lambda_n)} T\rho_n(\beta^{BC}_n)T^\dagger$, proving the final claim of the lemma. □

Now we have all the ingredients to prove our main theorem on equivalence of ensembles.

Theorem 20 *(Equivalence of ensembles, non-translation-invariant states). Let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of states on $\mathcal{A}_n$, let $\beta \geq 0$, and let Φ be a translation-invariant finite-range interaction which is not physically equivalent to zero, and for which there is a unique Gibbs state ω_β at inverse temperature β. Suppose that*

$$\liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\tau_n) - \beta \operatorname{tr}(\tau_n H_{\Lambda_n})\right) \geq p(\beta, \Phi),$$

then we have equality in this expression, and

$$\lim_{n\to\infty}\left\|\tau_n - \frac{\exp(-\beta H^{BC}_{\Lambda_n})}{Z_n}\right\|_{\{m\}} = 0, \qquad \textit{as well as} \qquad \lim_{n\to\infty}\frac{1}{|\Lambda_n|}\mathrm{tr}(\tau_n H_{\Lambda_n}) = u(\omega_\beta),$$

where $Z_n = \mathrm{tr}(\exp(-\beta H^{BC}_{\Lambda_n}))$*, and* $H^{BC}_{\Lambda_n}$ *is the Hamiltonian on* Λ_n *corresponding to* Φ *with arbitrary boundary conditions. Furthermore, if Gibbs states are unique around inverse temperature* $\beta > 0$*, we have*

$$\lim_{n\to\infty}\left\|\tau_n - \frac{\exp(-\beta^{BC}_n H^{BC}_{\Lambda_n})}{Z_n}\right\|_{\{m\}} = 0,$$

where β^{BC}_n *is defined as the solution of the equation* $\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H^{BC}_{\Lambda_n}\frac{\exp(-\beta^{BC}_n H^{BC}_{\Lambda_n})}{Z_n}\right) = u_n$*, where* $(u_n)_{n\in\mathbb{N}}$ *is any sequence with* $\lim_{n\to\infty} u_n = u(\omega_\beta)$*.*

Proof. We prove both claims at once, by defining two sequences $(\beta_n)_{n\in\mathbb{N}}$ and $(\beta'_n)_{n\in\mathbb{N}}$, either setting $\beta_n := \beta$ and $\beta'_n := \beta$, *or* setting $\beta_n := \beta^{BC}_n$ and $\beta'_n := \beta^p_n$. Define $\Omega(\sigma) := \frac{1}{|\mathbf{T}(\Lambda_n)|}\sum_{T\in\mathbf{T}(\Lambda_n)} T\sigma T^\dagger$, then it is easy to check that Ω is Hilbert-Schmidt self-adjoint, i.e. $\mathrm{tr}(A\Omega(B)) = \mathrm{tr}(\Omega(A)B)$ for $A = A^\dagger$, $B = B^\dagger$. Furthermore, define $\tau'_n := \Omega(\tau_n)$, then concavity of the entropy implies that $S(\tau'_n) \geq S(\tau_n)$. Since the Hamiltonian with periodic boundary conditions satisfies $T H^p_{\Lambda_n} T^\dagger = H^p_{\Lambda_n}$, we obtain $\mathrm{tr}(\tau'_n H^p_{\Lambda_n}) = \mathrm{tr}(\tau_n H^p_{\Lambda_n})$, and thus

$$\begin{aligned}
\liminf_{n\to\infty}\frac{1}{|\Lambda_n|}\left(S(\tau'_n) - \beta\,\mathrm{tr}(\tau'_n H_{\Lambda_n})\right) &\geq \liminf_{n\to\infty}\frac{1}{|\Lambda_n|}\left(S(\tau_n) - \beta\,\mathrm{tr}(\tau'_n H^p_{\Lambda_n})\right)\\
&= \liminf_{n\to\infty}\frac{1}{|\Lambda_n|}\left(S(\tau_n) - \beta\,\mathrm{tr}(\tau_n H^p_{\Lambda_n})\right)\\
&= \liminf_{n\to\infty}\frac{1}{|\Lambda_n|}\left(S(\tau_n) - \beta\,\mathrm{tr}(\tau_n H_{\Lambda_n})\right)\\
&\geq p(\beta, \Phi).
\end{aligned}$$

Thus, τ'_n satisfies the premises of Theorem 10, and (22) and (23) tell us that

$$\lim_{n\to\infty}\left\|\mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau'_n - \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\frac{\exp(-\beta'_n H^p_{\Lambda_n})}{Z'_n}\right\|_1 = 0,$$

where $Z'_n = \mathrm{tr}\left(\exp(-\beta'_n H^p_{\Lambda_n})\right)$. Now let A be any m-block periodically averaged effect on Λ_n, then it is of the form (30) with $A' \in \mathcal{A}_m$, $0 \leq A' \leq \mathbf{1}$. A simple calculation shows that $\mathrm{tr}(\tau_n A) = \mathrm{tr}(\tau'_n(A'\otimes\mathbf{1}))$, and Λ_n-translation-invariance of $H^p_{\Lambda_n}$ implies that $\mathrm{tr}\left(\frac{\exp(-\beta'_n H^p_{\Lambda_n})}{Z'_n}A\right) = \mathrm{tr}\left(\frac{\exp(-\beta'_n H^p_{\Lambda_n})}{Z'_n}(A'\otimes\mathbf{1})\right)$. Thus

$$\begin{aligned}
&\left\| \tau_n - \frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right\|_{\{m\}} \\
&= 2 \max_A \left| \mathrm{tr}(\tau_n A) - \mathrm{tr}\left(\frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} A \right) \right| \\
&= 2 \max_{A'} \left| \mathrm{tr}\left(A' \,\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n' \right) - \mathrm{tr}\left(A' \,\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right) \right| \\
&\leq \left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n' - \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right\|_1 \overset{n\to\infty}{\longrightarrow} 0. \qquad (33)
\end{aligned}$$

Now we extend this to arbitrary boundary conditions. Let A be any m-block periodically averaged effect, then there exists $A' \in \mathcal{A}_m$ such that $A = \Omega(A' \otimes \mathbf{1})$. Setting $Z_n = \mathrm{tr}(\exp(-\beta_n H^{BC}_{\Lambda_n}))$, we obtain

$$\begin{aligned}
&\left| \mathrm{tr}\left(A \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right) - \mathrm{tr}\left(A \frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right) \right| \\
&= \left| \mathrm{tr}\left(\Omega(A' \otimes \mathbf{1}) \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right) - \mathrm{tr}\left(\Omega(A' \otimes \mathbf{1}) \frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right) \right| \\
&= \left| \mathrm{tr}\left((A' \otimes \mathbf{1}) \Omega\left(\frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right) \right) - \mathrm{tr}\left((A' \otimes \mathbf{1}) \Omega\left(\frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right) \right) \right| \\
&= \left| \mathrm{tr}\left(A' \,\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \Omega\left(\frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right) \right) - \mathrm{tr}\left(A' \,\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \Omega\left(\frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right) \right) \right| \\
&\leq \frac{1}{2} \left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \Omega\left(\frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right) - \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \Omega\left(\frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right) \right\|_1 \overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

for all $m \in \mathbb{N}$ according to Lemma 19. Taking the supremum over all A shows that

$$\lim_{n\to\infty} \left\| \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} - \frac{\exp(-\beta_n' H^p_{\Lambda_n})}{Z_n'} \right\|_{\{m\}} = 0 \qquad \text{for all } m \in \mathbb{N}.$$

Combining this with (33) proves the second claim. Furthermore, Theorem 10 implies that

$$u(\omega_\beta) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\tau_n' H^p_{\Lambda_n}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\tau_n H^p_{\Lambda_n}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\tau_n H_{\Lambda_n}).$$

This completes the proof of the theorem. □

The simplest example application is as follows.

Example 21 (Microcanonical versus canonical ensemble, arbitrary boundary conditions). The sequence of states $(\tau_n)_{n\in\mathbb{N}}$ which are defined as the maximal mixtures on the microcanonical subspaces

$$T^{BC}_n := \mathrm{span}\left\{ |E\rangle \;\middle|\; H^{BC}_{\Lambda_n} |E\rangle = E|E\rangle, \;\; \frac{E}{|\Lambda_n|} \in (u-\delta, u) \right\},$$

where $H_{\Lambda_n}^{BC}$ is a Hamiltonian on Λ_n with arbitrary boundary conditions, satisfies the premises of Theorem 20. That is, if Gibbs states are unique around inverse temperature $\beta := \beta(u)$, we obtain equivalence of ensembles on m-block periodically averaged observables:

$$\lim_{n\to\infty} \left\| \tau_n - \frac{\exp(-\beta H_{\Lambda_n}^{BC})}{Z_n} \right\|_{\{m\}} = 0 \qquad \text{for all } m \in \mathbb{N}.$$

Furthermore, the same result is true if β is defined as the (n-dependent) solution of $\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H_{\Lambda_n}^{BC} \frac{\exp(-\beta H_{\Lambda_n}^{BC})}{Z_n}\right) = u$.

Remark. The choice of boundary conditions in the definition of T_n^{BC} and in the statement of the example need not be identical.

Proof. Apply Lemma 12 and $\frac{1}{|\Lambda_n|}\mathrm{tr}(\tau_n H_{\Lambda_n}^{BC}) \leq u$ to show that $\liminf_{n\to\infty} \frac{1}{|\Lambda_n|}\left(S(\tau_n) - \beta\,\mathrm{tr}(\tau_n H_{\Lambda_n})\right) \geq p(\beta, \Phi)$. □

For non-periodic boundary conditions, it is somewhat unnatural to consider periodically averaged observables. Instead, we may consider m-block averaged observables, where the region Λ_m is translated only inside the boundaries of Λ_n, without considering the periodic extension of the latter.

Definition 22 *(m-block averaged observable).* For $m \leq n$, define $\mathbf{T}(\Lambda_m, \Lambda_n) := \{y \in \mathbb{Z}^\nu \mid \Lambda_m + y \subset \Lambda_n\}$. An operator $A \in \mathcal{A}_n$ will be called an *m-block averaged observable* if there exists $A' \in \mathcal{A}_m$ with $A' = (A')^\dagger$ (resp. *m-block averaged effect* if $0 \leq A' \leq \mathbf{1}$) such that

$$A = \frac{1}{|\mathbf{T}(\Lambda_m, \Lambda_n)|} \sum_{y \in \mathbf{T}(\Lambda_m, \Lambda_n)} \gamma_y(A') \otimes \mathbf{1}, \tag{34}$$

where the unit observable is supported on $\Lambda_n \setminus (\Lambda_m + y)$. Moreover, we define the pseudonorm $\|\cdot\|_{[m]}$ on self-adjoint operators $M \in \mathcal{A}_n$ by

$$\|M\|_{[m]} := 2 \max\{|\mathrm{tr}(PM)| \mid P \text{ is an } m\text{-block averaged effect on } \Lambda_n\}.$$

The following lemma translates Lemma 18 to the pseudonorm $\|\cdot\|_{[m]}$ and also generalizes it.

Lemma 23. *Let A be an m-block averaged observable on Λ_n, coming from an observable $A' \in \mathcal{A}_m$ according to (34). Then for all quantum states ρ, σ on Λ_n, we have*

$$|\mathrm{tr}(\rho A) - \mathrm{tr}(\sigma A)| \leq \|A'\|_\infty \|\rho - \sigma\|_{[m]}.$$

Furthermore, we have $\|A\|_\infty \leq \|A'\|_\infty$; if in addition $A' \geq 0$, then we also have $\|A\|_\infty \geq \frac{1}{|\Lambda_m|}\|A'\|_\infty$. In the special case where $|\Lambda_m| = 1$, we have $\|A\|_\infty = \|A'\|_\infty$ whether or not A' is positive.

Proof. The proof of the first statement is identical to that of Lemma 18 and thus omitted. Clearly, $\|A\|_\infty \leq \|A'\|_\infty$ follows directly from the definition (34) and $\|\gamma_y(A') \otimes \mathbf{1}\|_\infty = \|\gamma_y(A')\|_\infty = \|A'\|_\infty$. Since Λ_m is a box, it can be written $\Lambda_m = [\lambda_1, \mu_1] \times \ldots \times [\lambda_\nu, \mu_\nu]$. Consider two boxes $\Lambda \subset \Lambda_n$ and $\Lambda' \subset \Lambda_n$ which are congruent to Λ_m, i.e. are translations of Λ_m. We call Λ and Λ' *equivalent* if there is a translation y such that

$\Lambda' = \Lambda + y$, which has components $y_i = k_i(\mu_i - \lambda_i)$ with $k_i \in \mathbb{Z}$. In other words, equivalent boxes (which are shaped like Λ_m) do not overlap, and they can tesselate Λ_n (up to sites close to the boundary).

Every equivalence class is uniquely determined by an element $x \in \Lambda_m$, which specifies a box Λ in that equivalence class which is $\Lambda_m + y$, where $x = (\lambda_1, \ldots, \lambda_v) + y$. Thus, the number of equivalence classes is upper-bounded by $|\Lambda_m|$. Now call two translations $y, z \in \mathbf{T}(\Lambda_m, \Lambda_n)$ equivalent if $\Lambda_m + y$ is equivalent to $\Lambda_m + z$ in the sense just specified. There will be N equivalence classes $\mathbf{T}_1, \ldots, \mathbf{T}_N$, where $N \leq |\Lambda_m|$, and $\mathbf{T}(\Lambda_m, \Lambda_n) = \bigcup_{i=1}^{N} \mathbf{T}_i$, which is a disjoint union. Consequently, at least one of them—say, $\mathbf{T}_j$—must have $|\mathbf{T}_j| \geq |\mathbf{T}(\Lambda_m, \Lambda_n)|/N$. For the moment, suppose that A' is a positive-semidefinite matrix. Then there is a state $|\psi\rangle$ on Λ_m such that $\|A'\|_\infty = \langle\psi|A'|\psi\rangle$. We can write $\Lambda_n = \bigcup_{y\in\mathbf{T}_j}(\Lambda_m + y) \cup \Lambda_{rest}$, where unions are disjoint. Now we define a state $|\Psi\rangle$ on Λ_n, by taking the tensor product of copies of $|\psi\rangle$ in the regions $\Lambda_m + y$, and an arbitrary pure reference state $|0\rangle$ on Λ_{rest}. We get

$$\begin{aligned}
\|A\|_\infty \geq \langle\Psi|A|\Psi\rangle &= \frac{1}{|\mathbf{T}(\Lambda_m, \Lambda_n)|} \sum_{i=1}^{N} \sum_{y\in\mathbf{T}_i} \langle\Psi|\gamma_y(A') \otimes \mathbf{1}|\Psi\rangle \\
&\geq \frac{1}{|\mathbf{T}(\Lambda_m, \Lambda_n)|} \sum_{y\in\mathbf{T}_j} \langle\Psi|\gamma_y(A') \otimes \mathbf{1}|\Psi\rangle \\
&= \frac{|\mathbf{T}_j|}{|\mathbf{T}(\Lambda_m, \Lambda_n)|} \langle\psi|A'|\psi\rangle \geq \frac{1}{N}\|A'\|_\infty \geq \frac{1}{|\Lambda_m|}\|A'\|_\infty.
\end{aligned}$$

If $|\Lambda_m| = 1$, choose the single-site state $|\psi\rangle$ such that $\|A'\|_\infty = |\langle\psi|A'|\psi\rangle|$. Let $|\psi^{\otimes\Lambda_n}\rangle$ be the state $|\psi\rangle$, copied onto every lattice site of Λ_n. Then

$$\begin{aligned}
\|A\|_\infty &\geq \left|\langle\psi^{\otimes\Lambda_n}|A|\psi^{\otimes\Lambda_n}\rangle\right| \\
&= \left|\frac{1}{|\mathbf{T}(\Lambda_m, \Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \langle\psi^{\otimes\Lambda_n}|\gamma_y(A') \otimes \mathbf{1}|\psi^{\otimes\Lambda_n}\rangle\right| \\
&= |\langle\psi|A'|\psi\rangle| = \|A'\|_\infty.
\end{aligned}$$

The claim follows. □

Asymptotically, that is for large n, the pseudonorms $\|\cdot\|_{\{m\}}$ and $\|\cdot\|_{[m]}$ are equivalent. This is the statement of the following lemma. Thus, our equivalence of ensemble results in Theorem 20 and Example 21 remain valid of the former pseudonorm is replaced by the latter. This yields a more natural physical interpretation of our results.

Lemma 24 (Equivalence of both averaging methods). *For every $m \leq n$ and all states ρ, σ on Λ_n, we have*

$$\left|\, \|\rho - \sigma\|_{\{m\}} - \|\rho - \sigma\|_{[m]} \right| \leq 8|\Lambda_m| \cdot \frac{|\partial\Lambda_n|}{|\Lambda_n|}$$

which tends to zero for fixed m as $n \to \infty$.

Proof. Define the completely positive map $\Phi : \mathcal{A}_m \to \mathcal{A}_n$ by setting $\Phi(A')$ as the right-hand side of (30). Similarly, define the completely positive map $\Phi' : \mathcal{A}_m \to \mathcal{A}_n$

by setting $\Phi'(A')$ as the right-hand side of (34). Note that $\Phi(\mathbf{1}) = \mathbf{1} = \Phi'(\mathbf{1})$. Then $\|M\|_{\{m\}} = 2\max_{0\le A'\le \mathbf{1}} |\mathrm{tr}(M\Phi(A'))|$ and $\|M\|_{[m]} = 2\max_{0\le A'\le \mathbf{1}} |\mathrm{tr}(M\Phi'(A'))|$. If M is traceless (as is the case for $M = \rho - \sigma$), then $\mathrm{tr}(M\Phi(\mathbf{1} - A')) = -\mathrm{tr}(M\Phi(A'))$ and similarly for Φ', and the absolute values under the maxima can be removed. Thus

$$\begin{aligned}\big|\, \|M\|_{\{m\}} - \|M\|_{[m]}\big| &= 2\left|\max_{0\le A'\le \mathbf{1}} \mathrm{tr}(M\Phi(A')) - \max_{0\le A'\le \mathbf{1}} \mathrm{tr}(M\Phi'(A'))\right| \\ &\le 2\max_{0\le A'\le \mathbf{1}} \left|\mathrm{tr}(M\Phi(A')) - \mathrm{tr}(M\Phi'(A'))\right| \\ &\le 2\|M\|_1 \max_{0\le A'\le \mathbf{1}} \left\|\Phi(A') - \Phi'(A')\right\|_\infty .\end{aligned}$$

To compare Φ and Φ', we note that we can interpret every translation $y \in \mathbf{T}(\Lambda_m, \Lambda_n)$ as a periodic translation $T \in \mathbf{T}(\Lambda_n)$ such that $\gamma_y(A')\otimes\mathbf{1} = T(A'\otimes\mathbf{1})T^\dagger$ for every $A' \in \mathcal{A}_m$; this is an equality of observables on Λ_n. In this sense, we can write $\mathbf{T}(\Lambda_m, \Lambda_n) \subset \mathbf{T}(\Lambda_n)$. A simple application of the triangle inequality and $\|A'\|_\infty \le 1$ gives

$$\begin{aligned}&\left\|\Phi(A') - \Phi'(A')\right\|_\infty \\ &\le \left\| \frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{T\in\mathbf{T}(\Lambda_n)} T(A'\otimes\mathbf{1})T^\dagger - \frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \gamma_y(A')\otimes\mathbf{1} \right\|_\infty \\ &\quad + \left\| \frac{1}{|\mathbf{T}(\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \gamma_y(A')\otimes\mathbf{1} - \frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \gamma_y(A')\otimes\mathbf{1} \right\|_\infty \\ &= \frac{1}{|\mathbf{T}(\Lambda_n)|} \left\| \sum_{T\in\mathbf{T}(\Lambda_n)\setminus\mathbf{T}(\Lambda_m,\Lambda_n)} T(A'\otimes\mathbf{1})T^\dagger \right\|_\infty \\ &\quad + \left(\frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} - \frac{1}{|\mathbf{T}(\Lambda_n)|} \right) \left\| \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \gamma_y(A')\otimes\mathbf{1} \right\|_\infty \\ &\le 2\frac{|\mathbf{T}(\Lambda_n)| - |\mathbf{T}(\Lambda_m,\Lambda_n)|}{|\mathbf{T}(\Lambda_n)|}.\end{aligned}$$

Estimating this expression is a matter of simple lattice geometry. First, it is easy to see that $|\mathbf{T}(\Lambda_n)| = |\Lambda_n|$, the number of sites in the region. Consider any translation $T \in \mathbf{T}(\Lambda_n) \setminus \mathbf{T}(\Lambda_m, \Lambda_n)$. It translates Λ_m periodically inside Λ_n, but *not* in a way such that the same is achieved by a non-periodic translation γ_y with $y \in \mathbb{Z}^\nu$. Instead, the corresponding y-translation will map Λ_m partially inside and partially outside of Λ_n. That is, there must be some intersection of $y + \Lambda_m$ with the boundary of Λ_n defined in (15). However, for every given boundary point $x \in \partial\Lambda_n$, there are only $|\Lambda_m|$ many translations y such that $x \in \Lambda_m + y$. Hence

$$|\mathbf{T}(\Lambda_n)| - |\mathbf{T}(\Lambda_m, \Lambda_n)| = |\mathbf{T}(\Lambda_n) \setminus \mathbf{T}(\Lambda_m, \Lambda_n)| \le |\partial\Lambda_n| \cdot |\Lambda_m|.$$

Combining the previous inequalities, and using that $\|\rho - \sigma\|_1 \le 2$, completes the proof. □

3.2. Canonical typicality. With the results of the previous subsection, in particular Examples 13 and 21, it is easy to prove a general result on canonical typicality for translation-invariant quantum systems.

Theorem 25 (Canonical typicality, periodic boundary conditions). *Let Φ be any translation-invariant finite-range interaction, not physically equivalent to zero, with corresponding periodic boundary condition Hamiltonians $H^p_{\Lambda_n}$, let $u_{\min}(\Phi) < u \leq u_{\max}(\Phi)$ and $\delta > 0$. Suppose that there is a unique infinite-volume Gibbs state ω_β at inverse temperature $\beta \equiv \beta(u)$. Consider the microcanonical subspace*

$$T^p_n := \operatorname{span}\left\{ |E\rangle \;\middle|\; H^p_{\Lambda_n}|E\rangle = E|E\rangle,\ \frac{E}{|\Lambda_n|} \in (u-\delta, u) \right\}.$$

If $|\psi\rangle \in T^p_n$ is a random pure state, then for every $m \in \mathbb{N}$ there is a sequence of positive real numbers $(\Delta_{m,n})_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \Delta_{m,n} = 0$, such that

$$\begin{aligned} &\operatorname{Prob}\left\{ \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} |\psi\rangle\langle\psi| - \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right\|_1 \geq \Delta_{m,n} + \varepsilon \right\} \\ &\quad \leq \exp\Big(-\varepsilon^2 \exp(|\Lambda_n|\, s(u) + o(|\Lambda_n|)) \Big) \end{aligned}$$

for every $\varepsilon \geq 0$. Furthermore, if Gibbs states are unique around inverse temperature $\beta > 0$, then the same result is true if β is chosen as the (n-dependent) solution of $\frac{1}{|\Lambda_n|} \operatorname{tr}\left(H^{BC}_{\Lambda_n} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right) = u$, where BC denotes an arbitrary fixed choice of boundary conditions.

Proof. It follows from [4, Theorem 1] that

$$\operatorname{Prob}\left\{ \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} |\psi\rangle\langle\psi| - \Omega_{m,n} \right\|_1 \geq \varepsilon + \frac{d^{|\Lambda_m|}}{\sqrt{|T^p_n|}} \right\} \leq 2\exp\left(-\frac{|T^p_n|\varepsilon^2}{18\pi^3} \right)$$

for all $\varepsilon \geq 0$, where $|T^p_n|$ denotes the dimension of the subspace T^p_n, and $\Omega_{m,n} := \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \tau_n$, with τ_n the maximally mixed state on T^p_n. Set

$$\delta_{m,n} := \left\| \Omega_{m,n} - \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right\|_1, \tag{35}$$

then Example 13 resp. Theorem 10 imply that $\lim_{n\to\infty} \delta_{m,n} = 0$. Thus, the previous statements imply

$$\begin{aligned} &\operatorname{Prob}\left\{ \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} |\psi\rangle\langle\psi| - \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n} \right\|_1 \geq \varepsilon + \frac{d^{|\Lambda_m|}}{\sqrt{|T^p_n|}} + \delta_{m,n} \right\} \\ &\quad \leq 2\exp\left(-\frac{|T^p_n|\varepsilon^2}{18\pi^3} \right). \end{aligned}$$

Furthermore, according to Lemma 12, we have $|T^p_n| = \exp[|\Lambda_n|\, s(u) + o(|\Lambda_n|)]$. Setting

$$\Delta_{m,n} := \delta_{m,n} + d^{|\Lambda_m|}/\sqrt{|T^p_n|} \tag{36}$$

completes the proof of the theorem. □

Example 14 shows again that we cannot in general replace the restriction of the global Gibbs state, $\mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\exp(-\beta H^p_{\Lambda_n})/Z_n$, with the local Gibbs state, $\exp(-\beta H^{BC}_{\Lambda_m})/Z_m$, no matter what boundary conditions we choose for $H^{BC}_{\Lambda_m}$.

Similarly as for our equivalence of ensembles result, we can prove an analogue of this theorem in the case of arbitrary boundary conditions by replacing $\|\cdot\|_1$ by $\|\cdot\|_{\{m\}}$.

Theorem 26 (Canonical typicality, arbitrary boundary conditions). *Let Φ be any translation-invariant finite-range interaction, not physically equivalent to zero, with corresponding arbitrary boundary condition Hamiltonians $H^{BC}_{\Lambda_n}$, let $u_{\min}(\Phi) < u \leq u_{\max}(\Phi)$ and $\delta > 0$. Suppose that there is a unique infinite-volume Gibbs state ω_β at inverse temperature $\beta \equiv \beta(u)$. Consider the microcanonical subspace*

$$T^{BC}_n := \mathrm{span}\left\{|E\rangle \;\middle|\; H^{BC}_{\Lambda_n}|E\rangle = E|E\rangle,\ \frac{E}{|\Lambda_n|} \in (u-\delta, u)\right\}.$$

If $|\psi\rangle \in T^{BC}_n$ is a random pure state, then for every $m \in \mathbb{N}$ there is a sequence of positive real numbers $(\Delta_{m,n})_{n\in\mathbb{N}}$ with $\lim_{n\to\infty}\Delta_{m,n} = 0$, such that

$$\begin{aligned}\mathrm{Prob}&\left\{\left\||\psi\rangle\langle\psi| - \frac{\exp(-\beta H^{BC}_{\Lambda_n})}{Z_n}\right\|_{[m]} \geq \Delta_{m,n} + \varepsilon\right\}\\ &\leq \exp\left(-\varepsilon^2\exp(|\Lambda_n|\, s(u) + o(|\Lambda_n|))\right)\end{aligned}$$

for every $\varepsilon \geq 0$. Furthermore, if Gibbs states are unique around inverse temperature $\beta > 0$, then the same result is true if β is chosen as the (n-dependent) solution of $\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H^{BC}_{\Lambda_n}\frac{\exp(-\beta H^{BC}_{\Lambda_n})}{Z_n}\right) = u$.

Proof. Denote by τ^{BC}_n the maximally mixed state on T^{BC}_n. Suppose that $\eta \geq 0$ is any real number such that

$$\left\||\psi\rangle\langle\psi| - \tau^{BC}_n\right\|_{[m]} \geq \eta. \tag{37}$$

By definition, this means that there exists some observable $A' \in \mathcal{A}_m$ such that

$$2\left|\frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|}\sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)}\left(\langle\psi|\gamma_y(A')\otimes\mathbf{1}|\psi\rangle - \mathrm{tr}(\tau^{BC}_n\gamma_y(A')\otimes\mathbf{1})\right)\right| \geq \eta,$$

and thus, there must be some $y \in \mathbf{T}(\Lambda_m,\Lambda_n)$ such that

$$2\left|\langle\psi|\gamma_y(A')\otimes\mathbf{1}|\psi\rangle - \mathrm{tr}(\tau^{BC}_n\gamma_y(A')\otimes\mathbf{1})\right| \geq \eta.$$

Let $\Lambda := \Lambda_m + y$, then $|\Lambda| = |\Lambda_m|$, $\Lambda \subset \Lambda_n$, and

$$\left\|\mathrm{Tr}_{\Lambda_n\setminus\Lambda}|\psi\rangle\langle\psi| - \mathrm{Tr}_{\Lambda_n\setminus\Lambda}\tau^{BC}_n\right\|_1 \geq \eta.$$

Now consider the case $\eta = \varepsilon + d^{|\Lambda_m|}/\sqrt{|T^{BC}_n|}$. According to [4, Theorem 1], the probability that the previous inequality holds on Haar-random choice of $|\psi\rangle$ is upper-bounded by $2\exp\left(-|T^{BC}_n|\varepsilon^2/(18\pi^3)\right)$. Thus

$$\mathrm{Prob}\left\{\left\||\psi\rangle\langle\psi| - \tau^{BC}_n\right\|_{[m]} \geq \varepsilon + \frac{d^{|\Lambda_m|}}{\sqrt{|T^p_n|}}\right\} \leq 2\exp\left(-\frac{|T^{BC}_n|\varepsilon^2}{18\pi^3}\right).$$

Now set

$$\delta_{m,n} := \left\| \tau_n^{BC} - \frac{\exp(-\beta H_{\Lambda_n}^{BC})}{Z_n} \right\|_{[m]},$$

and set $\Delta_{m,n} := \delta_{m,n} + d^{|\Lambda_m|}/\sqrt{|T_n^{BC}|}$. Example 21 and Lemma 24 show that $\lim_{n\to\infty} \delta_{m,n} = 0 = \lim_{n\to\infty} \Delta_{m,n}$, and arguing as in the proof of Theorem 25 completes the proof. □

Drawing a pure state $|\psi\rangle$ according to the Haar measure is a process that cannot be achieved efficiently in practice, as parameter counting shows. Thus, it is also to be expected that no process in nature really produces a Haar-random state. However, what *can* be achieved efficiently—for example, by application of random local unitaries [25]—are approximations to the Haar measure known as (approximate) *unitary t-designs*. As shown in [26], they give a way to "derandomize" results like the canonical typicality theorems above.

There are different definitions of what is called an ε-approximate k-design ν; they all have in common that the computational effort of sampling from them scales polynomially in $\log \varepsilon$ and $\log d$, where d is the underlying Hilbert space dimension.

Here, we use the definition from [26]. It utilizes the notion of a *balanced monomial of degree k* of a matrix U, which is a monomial in the components of U and $U^\dagger$ which contains the same number (k) of conjugated as unconjugated elements. For example, $U_{ij}U_{pq}^*$ is a balanced monomial of degree 1.

Definition 27 *(Approximate design).* A measure ν on the unitary group $U(d)$ is called an ε-approximate (unitary) k-design, if for all balanced monomials M of degree less than or equal to k, we have

$$|\mathbb{E}_{U\sim\nu} M(U) - \mathbb{E}_{U\sim\mu_H} M(U)| \leq \frac{\varepsilon}{d^k},$$

where $\mathbb{E}_{U\sim\mu}$ denotes the expectation with respect to a measure μ, and μ_H is the Haar measure.

We now use Theorem 1.4 in [26] to prove a derandomized version of canonical typicality. Note that the theorem in [26] uses as an implicit additional assumption that k is an integer-multiple of 8.

Theorem 28 (Canonical typicality, periodic boundary conditions, derandomized version). *Let Φ be any translation-invariant finite-range interaction, not physically equivalent to zero, with corresponding periodic boundary condition Hamiltonians $H_{\Lambda_n}^p$, let $u_{\min}(\Phi) < u \leq u_{\max}(\Phi)$ and $\Delta > 0$. Suppose that there is a unique infinite-volume Gibbs state ω_β at inverse temperature $\beta \equiv \beta(u)$. Consider the microcanonical subspace*

$$T_n^p := \operatorname{span}\left\{ |E\rangle \;\middle|\; H_{\Lambda_n}^p |E\rangle = E|E\rangle,\ \frac{E}{|\Lambda_n|} \in (u-\delta, u) \right\}.$$

Choose a state $|\psi\rangle$ at random from T_n^p by choosing a unitary from an ε-approximate 8-design and applying it to a fixed initial pure state, where $\varepsilon = \exp(-|\Lambda_n|s(u) + o(|\Lambda_n|))$. Then for every $m \in \mathbb{N}$ large enough such that $d^{|\Lambda_m|} \geq 14$, there is a sequence of positive real numbers $(\delta_{m,n})_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \delta_{m,n} = 0$, such that

$$\begin{aligned}&\mathrm{Prob}_\nu\left\{\left\|\mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}|\psi\rangle\langle\psi| - \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n}\right\|_1 \geq \delta_{m,n}+\kappa\right\}\\ &\leq \frac{d^{3|\Lambda_m|}}{\kappa^2}\exp\big(-|\Lambda_n|s(u)+o(|\Lambda_n|)\big)\end{aligned}$$

for all $\kappa > 0$. *Furthermore, if Gibbs states are unique around inverse temperature* $\beta > 0$, *then the same result is true if* β *is chosen as the (n-dependent) solution of* $\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H^{BC}_{\Lambda_n}\frac{\exp(-\beta H^p_{\Lambda_n})}{Z_n}\right) = u$, *where BC denotes an arbitrary fixed choice of boundary conditions.*

Proof. Let τ_n be the maximally mixed state on T^p_n, and set $\varepsilon := 6d^{3|\Lambda_m|}/|T^p_n|$. Due to [26, Theorem 1.4], we have

$$\mathrm{Prob}_\nu\left\{\left\|\mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}|\psi\rangle\langle\psi| - \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n\right\|_1 \geq \kappa\right\} \leq \frac{24d^{3|\Lambda_m|}}{|T^p_n|\kappa^2} \tag{38}$$

for all $\kappa > 0$. Define $\delta_{m,n}$ as in (35), use Example 13 and absorb the factor 24 into the $\exp(o|\Lambda_n|)$-term. □

One still has concentration on the thermal state; however, in contrast to the Haar measure result in Theorem 25, the concentration is now exponential in the number of sites $|\Lambda_n|$, not doubly-exponential. This behavior is more in line with standard expectations on physical systems in statistical mechanics.

It is now clear how Theorem 26 can be derandomized, by imitating the proof of Theorem 26 in conjunction with the T^{BC}_n-analogue of (38) and the inequality $\|\cdot\|_{[m]} \leq \|\cdot\|_1$. We omit the details.

Theorem 29 (Canonical typicality, arbitrary boundary conditions, derandomized version). *Let* Φ *be any translation-invariant finite-range interaction with corresponding arbitrary boundary condition Hamiltonians* $H^{BC}_{\Lambda_n}$, *let* $u_{\min}(\Phi) < u \leq u_{\max}(\Phi)$ *and* $\delta > 0$. *Suppose that there is a unique infinite-volume Gibbs state* ω_β *at inverse temperature* $\beta \equiv \beta(u)$. *Consider the microcanonical subspace*

$$T^{BC}_n := \mathrm{span}\left\{|E\rangle \;\middle|\; H^{BC}_{\Lambda_n}|E\rangle = E|E\rangle,\ \frac{E}{|\Lambda_n|}\in(u-\delta,u)\right\}.$$

Choose a state $|\psi\rangle$ *at random from* T^{BC}_n *by choosing a unitary from an* ε*-approximate* 8*-design and applying it to a fixed initial pure state, where* $\varepsilon = \exp(-|\Lambda_n|s(u)+o(|\Lambda_n|))$. *Then for every* $m \in \mathbb{N}$ *large enough such that* $d^{|\Lambda_m|} \geq 14$, *there is a sequence of positive real numbers* $(\delta_{m,n})_{n\in\mathbb{N}}$ *with* $\lim_{n\to\infty}\delta_{m,n} = 0$, *such that*

$$\begin{aligned}&\mathrm{Prob}_\nu\left\{\left\||\psi\rangle\langle\psi| - \frac{\exp(-\beta H^{BC}_{\Lambda_n})}{Z_n}\right\|_{[m]} \geq \delta_{m,n}+\kappa\right\}\\ &\leq \frac{d^{3|\Lambda_m|}}{\kappa^2}\exp\big(-|\Lambda_n|s(u)+o(|\Lambda_n|)\big)\end{aligned}$$

for all $\kappa > 0$. *Furthermore, if Gibbs states are unique around inverse temperature* $\beta > 0$, *then the same result is true if* β *is chosen as the (n-dependent) solution of* $\frac{1}{|\Lambda_n|}\mathrm{tr}\left(H^{BC}_{\Lambda_n}\frac{\exp(-\beta H^{BC}_{\Lambda_n})}{Z_n}\right) = u$.

Since the effort of sampling from an ε-approximate 8-design ν scales polynomially in $\log \varepsilon$ and the logarithm of the Hilbert space dimension, we obtain that sampling from ν in the theorems above amounts to an effort that grows only polynomially in $|\Lambda_n|$, i.e. the particle number.

3.3. Dynamical thermalization. We can apply the previous results to obtain statements about dynamical thermalization, using the results of [8] which are elaborations of earlier results in [6,7]. However, for the technicalities, we need to relate the von Neumann entropy with the Rényi entropy of order two. For $\alpha > 0$ with $\alpha \neq 1$ and density matrices ρ, we define [28]

$$S_\alpha(\rho) := \frac{1}{1-\alpha} \log \operatorname{tr}(\rho^\alpha),$$

and the limit $\alpha \to 1$ recovers von Neumann entropy, $S_1(\rho) := S(\rho) = -\operatorname{tr}(\rho \log \rho)$, and the limit $\alpha \to 0$ yields $S_0(\rho) := \log \operatorname{rank}(\rho)$. If $\alpha \leq \alpha'$ then $S_\alpha \geq S_{\alpha'}$. In fact, we will use Rényi entropy only for classical probability vectors $\lambda = (\lambda_1, \ldots, \lambda_N)$, and write sloppily $S_\alpha(\lambda)$ for classical Rényi entropy, which is the same as the quantum Rényi entropy of the diagonal matrix with entries λ_i. We use some inequalities and insights from [29] to show the following:

Lemma 30. *For every $0 \leq \varepsilon \leq 1$, we have $S_2(\rho) \geq 2\varepsilon \left(S(\rho) - \frac{\varepsilon}{1+\varepsilon} S_0(\rho) \right) \geq 2\varepsilon (S(\rho) - \varepsilon S_0(\rho))$.*

Proof. As shown in [29], we have $\frac{\partial}{\partial q} \frac{q-1}{q} S_q \geq 0$, hence $\frac{q-1}{q} S_q \leq \frac{1}{2} S_2$ for all $q \in [1, 2]$. Since the function $q \mapsto S_q$ is convex, the value of S_q lies on or above the line $g(x) := S_0 - (S_0 - S_1)x$ that connects S_0 and S_1, i.e. $S_q \geq g(q) = S_0 - (S_0 - S_1)q$. We get

$$S_2 \geq \frac{2(q-1)}{q} S_q \geq \frac{2(q-1)}{q} \left(S_0 - (S_0 - S_1)q \right) = 2(q-1)S_1 - \frac{2(q-1)^2}{q} S_0.$$

Setting $q =: 1 + \varepsilon$ proves the claim. □

Following [8], for any Hamiltonian H, we define its *gap degeneracy* by

$$D_G(H) := \max_E \left| \{(i, j) \mid i \neq j, \ E_i - E_j = E\} \right|,$$

where the E_i denotes the (energy) eigenvalues of H. Using Theorem 3 of [8], we can easily show the following.

Theorem 31 (Thermalization, periodic boundary conditions). *Let Φ be a translation-invariant finite-range interaction which is not physically equivalent to zero, and $(\rho_0^{(n)})_{n \in \mathbb{N}}$ any sequence of initial states on Λ_n which have energy expectation value of $U_n := \operatorname{tr}(\rho_0^{(n)} H_{\Lambda_n}^p)$ with density $U_n/|\Lambda_n|$ converging to some value $u \in (u_{\min}(\Phi), u_{\max}(\Phi))$ as $n \to \infty$.*

Suppose that the initial states have close to maximal "population entropy" in the following sense. Define $\bar{S}(\rho_0^{(n)}) := S(\lambda_1, \ldots, \lambda_N)$, where S is Shannon entropy, and $\lambda_i := \operatorname{tr}(\rho_0^{(n)} \pi_i)$ is the probability that the i-th energy level is populated, where $H_{\Lambda_n}^p =: \sum_{i=1}^N E_i \pi_i$ is the spectral decomposition. Furthermore, suppose that either $H_{\Lambda_n}^p$ is non-degenerate, or that every $\pi_i \rho_0^{(n)} \pi_i$ is Λ_n-translation-invariant. Then, determine the corresponding inverse temperature β_n for which

$$\operatorname{tr}(H^p_{\Lambda_n}\gamma^p_{\Lambda_n}(\beta_n)) = U_n, \qquad \text{where } \gamma^p_{\Lambda_n}(\beta_n) := \frac{\exp(-\beta_n H^p_{\Lambda_n})}{Z_n}.$$

If the initial states have close to maximal population entropy in the sense that

$$\bar{S}(\rho_0^{(n)}) \geq S(\gamma^p_{\Lambda_n}(\beta_n)) - o(|\Lambda_n|), \tag{39}$$

then unitary time evolution $\rho^{(n)}(t) := \exp(-itH^p_{\Lambda_n})\rho_0^{(n)}\exp(itH^p_{\Lambda_n})$ *thermalizes the subsystem* Λ_m *for most times* t*:*

$$\begin{aligned}
&\left\langle \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \rho^{(n)}(t) - \left\langle \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \rho^{(n)}(t) \right\rangle \right\|_1 \right\rangle \\
&\quad \leq d^{|\Lambda_m|}\sqrt{D_G(H^p_{\Lambda_n})}\exp\left(-\frac{s(\omega_\beta)^2}{4\log d}|\Lambda_n| + o(|\Lambda_n|)\right), \qquad \text{and} \\
&\lim_{n\to\infty} \left\| \left\langle \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \rho^{(n)}(t) \right\rangle - \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta_n H^p_{\Lambda_n})}{Z_n} \right\|_1 = 0,
\end{aligned}$$

where $Z_n = \operatorname{tr}(\exp(-\beta_n H^p_{\Lambda_n}))$*, and* $\langle\cdot\rangle$ *denotes the average over all times* $t \geq 0$*. Furthermore, in this statement,* β_n *can be replaced by* $\beta := \beta(u)$*.*

Remark. If $H^p_{\Lambda_n}$ is non-degenerate, we have $\bar{S}(\rho_0^{(n)}) = S(\bar{\rho}_0^{(n)})$, where $\bar{\rho}_0^{(n)} := \sum_i \pi_i \rho_0^{(n)} \pi_i$ is the dephased initial state. Furthermore, we can summarize the result by saying that

$$\left\langle \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \rho^{(n)}(t) - \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \frac{\exp(-\beta_n H^p_{\Lambda_n})}{Z_n} \right\|_1 \right\rangle \xrightarrow{n\to\infty} 0$$

as long as the gap degeneracy D_G grows at most subexponentially with $|\Lambda_n|$. However, the more detailed formulation above contains more information: while the difference to the Gibbs state may tend to zero polynomially in $|\Lambda_n|$, the result shows strong equilibration of time evolution indicated by a trace distance which goes to zero exponentially in $|\Lambda_n|$.

Proof. According to [8, Theorem 3 resp. 25], we have

$$\left\langle \left\| \operatorname{Tr}_{\Lambda_n\setminus\Lambda_m} \rho^{(n)}(t) - (\bar{\rho}_0^{(n)})_{\Lambda_m} \right\|_1 \right\rangle \leq d^{|\Lambda_m|}\sqrt{\frac{D_G(H^p_{\Lambda_n})}{d_{\text{eff}}}}, \tag{40}$$

where $d_{\text{eff}}^{-1} = \sum_i \lambda_i^2$, thus $d_{\text{eff}} = \exp(S_2(\lambda_1, \ldots, \lambda_N))$, and $\bar{\rho}_0^{(n)} = \langle \rho^{(n)}(t)\rangle = \sum_{i=1}^N \pi_i \rho_0^{(n)} \pi_i$. If $H^p_{\Lambda_n}$ is non-degenerate, then every $\pi_i \rho_0^{(n)} \pi_i$ is a real multiple of π_i and thus Λ_n-translation-invariant. Thus, the conditions of the lemma ensure that $\bar{\rho}_0^{(n)}$ is Λ_n-translation-invariant. Since the $\pi_i \rho_0^{(n)} \pi_i/\lambda_i$ for $\lambda_i \neq 0$ are density matrices with mutually orthogonal supports, we have

$$\begin{aligned}
S(\bar{\rho}_0^{(n)}) &= S\left(\sum_{i:\lambda_i\neq 0} \lambda_i \frac{\pi_i \rho_0^{(n)} \pi_i}{\lambda_i}\right) \\
&= S(\lambda_1, \ldots, \lambda_N) + \sum_{i:\lambda_i\neq 0} \lambda_i S\left(\frac{\pi_i \rho_0^{(n)} \pi_i}{\lambda_i}\right) \geq S(\lambda_1, \ldots, \lambda_N) = \bar{S}(\rho_0^{(n)}).
\end{aligned}$$

Note that $\mathrm{tr}(\rho_0^{(n)} H^p_{\Lambda_n}) = \mathrm{tr}(\bar\rho_0^{(n)} H^p_{\Lambda_n})$. Furthermore, Theorem 10 shows that $\lim_{n\to\infty} \beta_n = \beta := \beta(u)$, and $\rho(\beta_n)$ maximizes the functional $\rho \mapsto S(\rho) - \beta_n \mathrm{tr}(\rho H^p_{\Lambda_n})$. Thus

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \Big(S(\bar\rho_0^{(n)}) - \beta\, \mathrm{tr}(\bar\rho_0^{(n)} H_{\Lambda_n}) \Big) \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \Big(\bar S(\rho_0^{(n)}) - \beta\, \mathrm{tr}(\bar\rho_0^{(n)} H^p_{\Lambda_n}) \Big) \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \Big(S(\gamma^p_{\Lambda_n}(\beta_n)) - \beta\, \mathrm{tr}(\rho_0^{(n)} H^p_{\Lambda_n}) \Big) \\
&= \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \Big(S(\gamma^p_{\Lambda_n}(\beta_n)) - \beta\, \mathrm{tr}(\gamma^p_{\Lambda_n}(\beta_n) H^p_{\Lambda_n}) \Big) \\
&= \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \Big(S(\gamma^p_{\Lambda_n}(\beta_n)) - \beta_n \mathrm{tr}(\gamma^p_{\Lambda_n}(\beta_n) H^p_{\Lambda_n}) \Big) \\
&\geq \liminf_{n\to\infty} \frac{1}{|\Lambda_n|} \Big(S((\omega_\beta)_{\Lambda_n}) - \beta_n \mathrm{tr}((\omega_\beta)_{\Lambda_n} H^p_{\Lambda_n}) \Big) \\
&= s(\omega_\beta) - \beta\, u(\omega_\beta) = p(\beta, \Phi),
\end{aligned} \tag{41}$$

and Theorem 10 proves that

$$\lim_{n\to\infty} \left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \bar\rho_0^{(n)} - \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \frac{\exp(-\beta_n H^p_{\Lambda_n})}{Z_n} \right\|_1 = 0$$

and $\lim_{n\to\infty} \frac{1}{|\Lambda_n|} S(\bar\rho_0^{(n)}) = s(\omega_\beta)$ as well as $\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\bar\rho_0^{(n)} H_{\Lambda_n}) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathrm{tr}(\rho_0^{(n)} H^p_{\Lambda_n}) = u(\omega_\beta)$. Together with (39), this implies that $\bar S(\rho_0^{(n)}) = s(\omega_\beta)|\Lambda_n| + o(|\Lambda_n|)$. It remains to estimate d_{eff}. This will be done via Lemma 30. Writing $\lambda = (\lambda_1, \ldots, \lambda_N)$ and using that $S_0(\lambda) \leq \log N \leq |\Lambda_n| \log d$ and $S(\lambda) = \bar S(\rho_0^{(n)})$, we obtain $S_2(\lambda) \geq 2\varepsilon(S(\lambda) - \varepsilon|\Lambda_n| \log d) = 2\varepsilon \big((s(\omega_\beta) - \varepsilon \log d)|\Lambda_n| + o(|\Lambda_n|) \big)$ for all $0 \leq \varepsilon \leq 1$. The special case $\varepsilon = s(\omega_\beta)/(2 \log d)$ yields $d_{\mathrm{eff}} = \exp(S_2(\lambda)) \geq \exp\left(\frac{s(\omega_\beta)^2}{2 \log d} |\Lambda_n| + o(|\Lambda_n|) \right)$. □

Here is an example of a suitable sequence of initial states that appeared in work by Riera et al. [10]:

Example 32 ("Flat" pure initial state). Consider pure initial states $\rho_0^{(n)} = |\psi_0^{(n)}\rangle\langle\psi_0^{(n)}|$ which have a flat energy distribution in an energy window, as discussed in [10]. Concretely, denote the energy eigenstates of $H^p_{\Lambda_n}$ by $|E_i\rangle$, fix $\delta > 0$, and set (up to normalization)

$$|\psi_0^{(n)}\rangle \sim \sum_{u-\delta < E_i/|\Lambda_n| < u} |E_i\rangle.$$

If $H^p_{\Lambda_n}$ is non-degenerate, then $\bar S(\rho_0^{(n)})$ is the logarithm of the number of energy levels between densities $u - \delta$ and u, which is $s(u)|\Lambda_n| + o(|\Lambda_n|) = S(\gamma^p_{\Lambda_n}(\beta)) + o(|\Lambda_n|)$ according to Lemma 12. Thus, Theorem 31 proves thermalization of small subsystems. The same conclusion holds if $|\psi_0^{(n)}\rangle$ is not exactly flat, but populates the energy levels as given in Example 16 and Fig. 4.

This example and Theorem 31 (in one formulation) assume that $H^p_{\Lambda_n}$ is non-degenerate. In fact, we show numerically in Sect. 3.5 that generic models of the kind we consider are non-degenerate, despite translation-invariance. Alternatively, we can lift the condition of non-degeneracy or periodic boundary conditions by proving a weaker statement about m-block-averaged observables.

Theorem 33 (Thermalization, arbitrary boundary conditions). *Let Φ be a translation-invariant finite-range interaction which is not physically equivalent to zero, and $(\rho_0^{(n)})_{n\in\mathbb{N}}$ any sequence of initial states on Λ_n which have energy expectation value of $U_n := \mathrm{tr}(\rho_0^{(n)} H^{BC}_{\Lambda_n})$ with density $U_n/|\Lambda_n|$ converging to some value $u \in (u_{\min}(\Phi), u_{\max}(\Phi))$ as $n \to \infty$, where BC denotes an arbitrary fixed choice of boundary conditions.*

Suppose that the initial states have close to maximal "population entropy" in the following sense. Define $\bar{S}(\rho_0^{(n)}) := S(\lambda_1, \ldots, \lambda_N)$, where S is Shannon entropy, and $\lambda_i := \mathrm{tr}(\rho_0^{(n)}\pi_i)$, where $H^{BC}_{\Lambda_n} =: \sum_{i=1}^N E_i\pi_i$ is the spectral decomposition. Then, determine the corresponding inverse temperature β_n for which

$$\mathrm{tr}(H^{BC}_{\Lambda_n}\gamma^{BC}_{\Lambda_n}(\beta_n)) = U_n, \qquad \text{where } \gamma^{BC}_{\Lambda_n}(\beta_n) := \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n}.$$

If the initial states have close to maximal population entropy in the sense that

$$\bar{S}(\rho_0^{(n)}) \geq S(\gamma^{BC}_{\Lambda_n}(\beta_n)) - o(|\Lambda_n|),$$

then unitary time evolution $\rho^{(n)}(t) := \exp(-itH^{BC}_{\Lambda_n})\rho_0^{(n)}\exp(itH^{BC}_{\Lambda_n})$ thermalizes all m-block averaged observables for most times t:

$$\left\langle \left\| \rho^{(n)}(t) - \left\langle \rho^{(n)}(t) \right\rangle \right\|_{[m]} \right\rangle$$
$$\leq d^{|\Lambda_m|}\sqrt{D_G(H^{BC}_{\Lambda_n})}\exp\left(-\frac{s(\omega_\beta)^2}{4\log d}|\Lambda_n| + o(|\Lambda_n|)\right), \qquad \text{and}$$
$$\lim_{n\to\infty}\left\| \left\langle \rho^{(n)}(t) \right\rangle - \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right\|_{[m]} = 0,$$

where $Z_n = \mathrm{tr}(\exp(-\beta_n H^{BC}_{\Lambda_n}))$, and $\langle\cdot\rangle$ denotes the average over all times $t \geq 0$. Furthermore, in this statement, β_n can be replaced by $\beta := \beta(u)$.

Remark. As in the previous theorem, we can summarize the result (at the expense of losing some information) as

$$\lim_{n\to\infty}\left\langle \left\| \rho^{(n)}(t) - \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n} \right\|_{[m]} \right\rangle = 0$$

whenever the gap degeneracy D_G does not grow too quickly with $|\Lambda_n|$. In fact, we can always force D_G to be equal to one—that is, remove degeneracies—by adding appropriate boundary conditions in the sense of Definition 7.

Proof. For any $X = X^\dagger \in \mathcal{A}_n$, we can estimate the $\|\cdot\|_{[m]}$-norm via

$$\begin{aligned}
\|X\|_{[m]} &= 2\max\left\{ \left| \frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \operatorname{tr}\left[X(\gamma_y(A')\otimes\mathbf{1})\right] \right| \;\middle|\; A'\in\mathcal{A}_m,\ 0\le A'\le\mathbf{1} \right\} \\
&\le 2\max\left\{ \frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \left|\operatorname{tr}\left[X(\gamma_y(A')\otimes\mathbf{1})\right]\right| \;\middle|\; A'\in\mathcal{A}_m,\ 0\le A'\le\mathbf{1} \right\} \\
&\le \frac{2}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \max\left\{ \left|\operatorname{tr}\left[X(\gamma_y(A')\otimes\mathbf{1})\right]\right| \;\middle|\; A'\in\mathcal{A}_m,\ 0\le A'\le\mathbf{1}\right\} \\
&= \frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \left\|\operatorname{Tr}_{\Lambda_n\setminus(\Lambda_m+y)} X\right\|_1 .
\end{aligned}$$

Using again the results of [8] in the form (40), setting again $\bar\rho_0^{(n)} := \langle\rho^{(n)}(t)\rangle = \sum_i \pi_i\rho_0^{(n)}\pi_i$, we obtain

$$\begin{aligned}
\left\langle \left\|\rho^{(n)}(t) - \bar\rho_0^{(n)}\right\|_{[m]} \right\rangle &\le \frac{1}{|\mathbf{T}(\Lambda_m,\Lambda_n)|} \\
&\qquad \sum_{y\in\mathbf{T}(\Lambda_m,\Lambda_n)} \left\langle \left\|\operatorname{Tr}_{\Lambda_n\setminus(\Lambda_m+y)}\rho^{(n)}(t) - \operatorname{Tr}_{\Lambda_n\setminus(\Lambda_m+y)}\bar\rho_0^{(n)}\right\|_1 \right\rangle \\
&\le d^{|\Lambda_m|}\sqrt{\frac{D_G(H^{BC}_{\Lambda_n})}{d_{\mathrm{eff}}}},
\end{aligned}$$

where $d_{\mathrm{eff}} = \exp(S_2(\lambda))$. As in the proof of Theorem 31, we have $S(\bar\rho_0^{(n)}) \ge \bar S(\rho_0^{(n)}) = S(\lambda)$, and also $\operatorname{tr}(\rho_0^{(n)}H^{BC}_{\Lambda_n}) = \operatorname{tr}(\bar\rho_0^{(n)}H^{BC}_{\Lambda_n})$. Furthermore, Lemma 19 implies that $\lim_{n\to\infty}\beta_n = \beta := \beta(u)$. Thus, we can repeat the calculation (41) in the proof of Theorem 31, and obtain that $\liminf_{n\to\infty}\frac{1}{|\Lambda_n|}\left(S(\bar\rho_0^{(n)}) - \beta\operatorname{tr}(\bar\rho_0^{(n)}H_{\Lambda_n})\right) \ge p(\beta,\Phi)$. Consequently, Theorem 20 and Lemma 24 imply that

$$\lim_{n\to\infty}\left\|\bar\rho_0^{(n)} - \frac{\exp(-\beta_n H^{BC}_{\Lambda_n})}{Z_n}\right\|_{[m]} = 0 \qquad\text{and}\qquad \lim_{n\to\infty}\frac{1}{|\Lambda_n|}\operatorname{tr}(\bar\rho_0^{(n)}H^{BC}_{\Lambda_n}) = u(\omega_\beta).$$

As in the proof of Theorem 31, it also follows that $\bar S(\rho_0^{(n)}) \ge s(\omega_\beta)|\Lambda_n| + o(|\Lambda_n|) = S(\lambda)$. Repeating the final steps of the proof of Theorem 31 yields the claimed estimate for d_{eff}. □

3.4. Finite-size estimates for systems without interaction. As the most simple special case, consider the non-interacting Hamiltonian

$$H_\Lambda := \sum_{x\in\Lambda} h_x,$$

where $h_x = \gamma_x(h)$ denotes a fixed self-adjoint matrix h sitting on site $x \in \Lambda$. This corresponds to an interaction Φ of the form

$$\Phi(X) = \begin{cases} \gamma_X(h) & \text{if } \#X = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Since there is no interaction, the dimension ν of the lattice $\mathbb{Z}^\nu$ does not play any role; without loss of generality, we may assume that $\nu = 1$. Similarly, we set $\Lambda_m = [1, m] \subset \mathbb{Z}$. Without interaction, the (restriction of the global) Gibbs state becomes the product state

$$\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \frac{\exp(-\beta H_{\Lambda_n})}{Z_n} = \gamma_\beta^{\otimes m},$$

where $\gamma_\beta = \exp(-\beta h)/Z_1$ is the single-site Gibbs state with single-site partition function Z_1. We will now look at equivalence of ensembles – and its finite-size behavior – in this special case. That is, we consider the maximally mixed state τ_n on

$$T_n := \mathrm{span}\left\{ |E\rangle \;\middle|\; \frac{E}{n} \in [u - \delta, u] \right\},$$

where $\delta > 0$ and u will be considered fixed in what follows. On every site, we can choose the local basis such that h is diagonal, denoting the corresponding single-site eigenstates of h by $\{|0\rangle, \ldots, |d-1\rangle\}$. (Recall that d denotes the single-site Hilbert space dimension.) For $0 \leq j \leq d-1$, the eigenvalue corresponding to $|j\rangle$ will be denoted E_j; that is,

$$h|j\rangle = E_j|j\rangle.$$

We may always choose a basis and shift the energy such that $0 = E_0 \leq E_1 \leq \ldots \leq E_{d-1}$, i.e.

$$h = \begin{pmatrix} 0 & & & \\ & E_1 & & \\ & & \ddots & \\ & & & E_{d-1} \end{pmatrix}.$$

Every string $s = s_1 s_2 \ldots s_n$ of length n over the alphabet $\{0, \ldots, d-1\}$ describes an eigenvector $|s\rangle := |s_1\rangle \otimes \ldots \otimes |s_n\rangle$ of H on n sites, where $H|s\rangle = \sum_i h|s_i\rangle = \sum_i E_{s_i}$. Thus, the microcanonical subspace can also be written

$$T_n = \mathrm{span}\left\{ |s\rangle \;\middle|\; s \in \{0, 1, \ldots, d-1\}^n, \;\; \frac{1}{n}\sum_{i=1}^n E_{s_i} \in [u - \delta, u] \right\}.$$

Our goal is to estimate the difference

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1. \tag{42}$$

Since all relevant operators compute, we can restrict to the probability distributions on the diagonal; we have a purely classical problem. Our first observation is that a tight upper bound on this expression is known in the special case $\delta = 0$ and $d = 2$; it has been obtained in proofs of the finite classical de Finetti Theorem [22].

Lemma 34. *In the case of a perfectly sharp microcanonical subspace, i.e. $\delta = 0$, and of qubit systems, i.e. $d = 2$, we have*

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \frac{4m}{n},$$

assuming that the energy density u is chosen such that the corresponding microcanonical subspace T_n is not empty.

Proof. We have $h = \begin{pmatrix} 0 & \\ & E_1 \end{pmatrix}$, and so $\rho_\beta = \begin{pmatrix} 1 & \\ & e^{-\beta E_1} \end{pmatrix} \cdot \frac{1}{1+e^{-\beta E_1}}$. The inverse temperature β is determined by $\mathrm{tr}(\gamma_\beta h) = u$. In this case, $u = E_1 \cdot p_1$, where p_1 is the relative frequencies of 1's in the strings s with $|s\rangle \in T_n$. This equation implies $\gamma_\beta = \begin{pmatrix} 1-p_1 & \\ & p_1 \end{pmatrix}$, with classical probability distribution $P_\beta := (1-p_1, p_1)$ on the diagonal. If we denote by Q the classical probability distribution on $\{0,1\}^m$ determined by the diagonal elements of $\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n$, we have

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 = \left\| Q - P_\beta^{\otimes m} \right\|_1,$$

where $\| \cdot \|_1$ on the right-hand side denotes the variation distance of two probability distributions:

$$\|P - Q\|_1 = \sum_{i=1}^{2^m} |P_i - Q_i| = 2 \max_{A \subseteq \{1,\ldots,2^m\}} |P(A) - Q(A)|.$$

Consider an urn U with n balls, where $p_1 \cdot n$ of them are marked by a "1" and all others marked by a "0". Then $P_\beta^{\otimes m}$ describes the distribution obtained by m draws from U with replacement, whereas Q described the distribution obtained by m draws from U *without* replacement, where in both cases the order of the results is taken into account. These distributions are considered in [22] in the proof a finite version of the classical de Finetti theorem. The main result then follows from Theorem (4) in [22]. □

For $d \geq 3$, even if $\delta = 0$, the results of [22] do not directly yield an upper bound on expression (42). This is for two reasons. First, the typical subspace T_n will in general not be spanned by a single type class, but by several ones. For example, consider the case $d = 3$ with energies $E_0 = 0$, $E_1 = 1$, and $E_2 = 2$. Fixing the energy density to $u = 2/3$ yields the microcanonical subspace

$$T_3 = \mathrm{span}\,\{|011\rangle, |101\rangle, |110\rangle, |002\rangle, |020\rangle, |200\rangle\}\,.$$

This is a disjoint union of two type classes. While $T_4 = \{0\}$ and $T_5 = \{0\}$, we have

$$T_6 = \mathrm{span}\,\{|000022\rangle, \ldots, |000112\rangle, \ldots, |001111\rangle, \ldots\}\,,$$

where the dots denotes all permutations. This is a union of three type classes. Then the results in [22] do not prove directly that $\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n$ is close to a product state, but that it is close to a *convex combination* of product states, resembling the de Finetti theorem.

In this particular example, it can be checked numerically that the qualitative behavior of Lemma 34 remains true: n needs to be increased linearly with m in order to achieve a fixed one-norm distance error. The inverse temperature turns out to be $\beta = \log[(1+\sqrt{33})/4]$, and $\gamma_\beta = \mathrm{diag}\left((15-\sqrt{33})/18, (\sqrt{33}-3)/9, (9-\sqrt{33})/18\right)$.

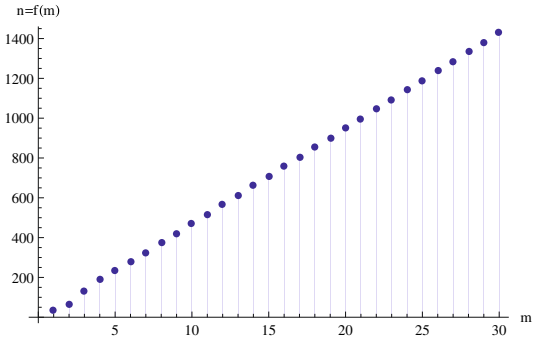

Fig. 5. Minimal number of sites n to guarantee that a subsystem of given size m is ε-close to the local Gibbs state, where $\varepsilon = 1/100$, energy density $u = 2/3$, local Hilbert space dimension $d = 3$, and energy levels $E_0 = 0$, $E_1 = 1$ and $E_2 = 2$. In this case, the microcanonical subspace of width $\delta = 0$ is spanned by more than one type class. It can be seen that the size of the "bath" has to be increased linearly with the size of the subsystem

The subspace T_n is non-trivial whenever n is a multiple of 3. Define the function $f : \mathbb{N} \to \mathbb{N}$ by

$$f(m) := \text{ smallest possible } n \in \mathbb{N} \text{ such that } \left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \frac{1}{100}.$$

This function is evaluated numerically in Fig. 5. It can be seen that $n = f(m)$ increases linearly with m.

It turns out that for $\delta = 0$, the previous example is atypical in the sense that *generic* energy windows usually lead to microcanonical subspaces T_n that contain only a single type class. This can be characterized as in the following lemma. We use the standard terminology to call a set of real numbers $E_1, \ldots, E_{d-1}$ *rationally dependent* if there are rational numbers $\lambda_1, \ldots, \lambda_{d-1} \in \mathbb{Q}$, not all of them zero, such that $\sum_{i=1}^{d-1} \lambda_i E_i = 0$, and otherwise *rationally independent*.

Lemma 35. *Suppose that $\delta = 0$. Then, all non-trivial microcanonical subspaces $T_n \neq \{0\}$, for all n and u, are spanned by a single type class if and only if the energies $E_1, \ldots, E_{d-1}$ are rationally independent.*

Proof. We denote type classes as follows:

$$T(k_0, \ldots, k_{d-1}) := \left\{ s \in \{0, \ldots, d-1\}^n \mid \#\{i : s_i = j\} = k_j \text{ for all } j \right\},$$

that is, the set of all strings that have k_0 zeroes, k_1 ones, and so on. All strings s in the same type class have the same energy $\langle s|H|s\rangle =: E_s = \sum_i E_{s_i} = \sum_{j=0}^{d-1} k_j E_j$. Thus, the microcanonical subspace T_n must be a disjoint union of (spans of) type classes.

Suppose the energies are rationally independent, and suppose that $\mathrm{span}\, T(k_0, \ldots, k_{d-1}) \subset T_n$ and at the same time $\mathrm{span}\, T(k'_0, \ldots, k'_{d-1}) \subset T_n$. Then

$$u \cdot n = k_1 E_1 + \ldots + k_{d-1} E_{d-1} = k'_1 E_1 + \ldots + k'_{d-1} E_{d-1}.$$

Thus

$$\underbrace{(k_1 - k'_1)}_{\in \mathbb{Z}} E_1 + \cdots + \underbrace{(k_{d-1} - k'_{d-1})}_{\in \mathbb{Z}} E_{d-1} = 0,$$

and rational independence implies that $k_j = k'_j$ for all j, so T_n is the span of a single type class.

Conversely, suppose that $E_1, \ldots, E_{d-1}$ are rationally dependent. That is, there are $k_1, \ldots, k_{d-1} \in \mathbb{Z}$, not all $k_i = 0$, such that

$$k_1 E_1 + \ldots + k_{d-1} E_{d-1} = 0.$$

There is at least one energy E_i with $E_i > 0$, so this equation can only be satisfied if $\min_i k_i =: k_j < 0$ and $\max_i k_i > 0$. Set

$$E := -k_j(E_1 + \cdots + E_{d-1}) = (k_1 - k_j)E_1 + \cdots + (k_{d-1} - k_j)E_{d-1}$$

Then all $k'_i := k_i - k_j \geq 0$ are integers, and they cannot all be zero. Choose any $n \in \mathbb{N}$ with

$$n \geq \max\{(d-1)|k_j|, (k_1 - k_j) + (k_2 - k_j) + \cdots + (k_{d-1} - k_j)\}.$$

Set $k_0 := n - (d-1)|k_j| \geq 0$ and $k'_0 := n - [(k_1 - k_j) + \cdots + (k_{d-1} - k_j)] \geq 0$, and set the energy density to $u := E/n$. Then we have

$$\begin{aligned} \operatorname{span} T(k_0, |k_j|, |k_j|, \ldots, |k_j|) &\subseteq T_n, \\ \operatorname{span} T(k'_0, k_1 - k_j, k_2 - k_j, \ldots, k_{d-1} - k_j) &\subseteq T_n. \end{aligned}$$

Thus, T_n is spanned by at least two different type classes. □

There is a second reason why the results in [22] cannot directly be used if $d \geq 3$, even in the case where $\delta = 0$ *and* assuming the rational independence of the energies. It follows from [22] that in this case

$$\left\| \operatorname{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma^{\otimes m} \right\|_1 \leq \frac{2dm}{n}; \tag{43}$$

however, the state γ is in general *not* equal to γ_β for any β. Instead, γ is the single-site density matrix with the symbols' relative frequencies in the type class as eigenvalues, and this is in general not a thermal state.

As a simple example, consider the case $d = 3$, with single-site Hamiltonian $h = \operatorname{diag}\left(0, 1, \sqrt{2}\right)$ and energy density $u = (2 + \sqrt{2})/6$. If n is a multiple of 6, then T_n contains all basis vectors $|s\rangle$ with strings $s \in \{0, 1\}^n$ that have $n/2$ zeroes, $n/3$ ones and $n/6$ twos. Then the γ appearing in (43) is $\gamma = \operatorname{diag}(1/2, 1/3, 1/6)$, and there does not exist any β such that $\gamma = \gamma_\beta$.

In the following, we will generalize the result of Lemma 34, by showing that also in the case of a microcanonical subspace of width $\delta > 0$, the qualitative behavior of Fig. 5 remains true, at least in the case $d = 2$, i.e. in the qubit case. First, we prove a lemma which shows this for $\delta = 0$ or δ depending on n and approaching zero fast enough. Later, we will extend the result to arbitrary fixed $\delta > 0$ by some large deviations argument.

Lemma 36. *Consider the case of qubits, i.e. $d = 2$, and shift the energies such that $E_0 = 0$ and $E_1 = 1$. If τ_n is the maximal mixture on the non-trivial microcanonical subspace corresponding to the energy interval $n \cdot [u - \delta, u]$, with $0 \leq \delta < u \leq \frac{1}{2}$, and γ_β is the single-site Gibbs state with corresponding inverse temperature β, then we have for subsystems of size $m \leq n(u - \delta)$,*

$$S\left(\gamma_\beta^{\otimes m} \,\middle\|\, \operatorname{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n\right) \leq \frac{(1-\delta)u}{u-\delta} \cdot \frac{m}{n-m} + \frac{mu\delta}{u-\delta}\left(1 + \frac{m}{n-m}\right),$$

where S denotes the quantum relative entropy (with logarithm in base e). In particular, if $\delta = 0$, the relative entropy is upper-bounded by $m/(n-m)$, and the Pinsker inequality yields

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \sqrt{\frac{1}{2} \cdot \frac{m}{n-m}} \qquad (\textit{special case } \delta = 0).$$

Proof. As explained above, the calculation is classical: we can regard $\gamma_\beta^{\otimes m}$ as a classical probability distribution on the binary strings of length m, given by

$$P_\beta^{\otimes m}(x) = u^k (1-u)^{m-k}, \qquad \text{where } k \text{ is the number of ones in } x.$$

From elementary combinatorics, the marginal distribution Q is given by

$$Q(x) = (\dim T_n)^{-1} \cdot \sum_{\ell \in [\lceil n(u-\delta) \rceil, \lfloor nu \rfloor], \ \ell \geq k} \binom{n-m}{\ell - k},$$

and the numerator counts all possible ways to complete x to a string of length n which has ℓ ones such that the energy is in the suitable interval. Since $k \leq m \leq n(u-\delta) \leq \lceil n(u-\delta) \rceil$, the condition $\ell \geq k$ is automatically satisfied for all ℓ in the summation interval; hence this condition can be removed from the specification of the sum. The dimension of the microcanonical subspace is given by

$$\dim T_n = \sum_{\ell = \lceil n(u-\delta) \rceil}^{\lfloor nu \rfloor} \binom{n}{\ell}. \tag{44}$$

Thus, the quantum relative entropy S can be written in terms of the classical relative entropy H,

$$\begin{aligned} S\left(\gamma_\beta^{\otimes m} \,\middle\|\, \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n\right) &= H\left(P_\beta^{\otimes m} \,\middle\|\, Q\right) \\ &= \sum_{x \in \{0,1\}^m} P_\beta^{\otimes m}(x) \left(\log P_\beta^{\otimes m}(x) - \log Q(x)\right) \\ &= \sum_{k=0}^{m} \binom{m}{k} u^k (1-u)^{m-k} \left(\log P_k - \log Q_k\right), \end{aligned} \tag{45}$$

where

$$\begin{aligned} P_k &= u^k (1-u)^{m-k}, \\ Q_k &= (\dim T_n)^{-1} \cdot \sum_{\ell = \lceil n(u-\delta) \rceil}^{\lfloor nu \rfloor} \binom{n-m}{\ell - k}. \end{aligned} \tag{46}$$

Using that $(n-m)! = n!/[(n-m+1)(n-m+2)\ldots n]$ and similar identities for $(\ell - k)!$ and $[n - \ell - (m-k)]!$, we obtain

$$Q_k = (\dim T_n)^{-1} \cdot \sum_{\ell = \lceil n(u-\delta) \rceil}^{\lfloor nu \rfloor} \frac{n! \prod_{j=0}^{k-1} (\ell - j) \prod_{j=0}^{m-k-1} (n - \ell - j)}{\prod_{j=0}^{m-1} (n-j)\, \ell! (n-\ell)!}$$

In order to eliminate all ℓ-variables from all products, we substitute the inequalities

$$\begin{aligned} \ell - j &\geq \lceil n(u-\delta)\rceil - j, \\ n - \ell - j &\geq n - \lfloor nu\rfloor - j \end{aligned}$$

and obtain

$$Q_k \geq (\dim T_n)^{-1} \frac{\prod_{j=0}^{k-1}\left(\lceil n(u-\delta)\rceil - j\right)\prod_{j=0}^{m-k-1}\left(n-\lfloor nu\rfloor - j\right)}{\prod_{j=0}^{m-1}(n-j)} \sum_{\ell=\lceil n(u-\delta)\rceil}^{\lfloor nu\rfloor}\binom{n}{\ell}.$$

Thus, the sum on the right-hand side exactly cancels the factor $(\dim T_n)^{-1}$ according to (44), and we obtain

$$\begin{aligned} \log Q_k &\geq \sum_{j=0}^{k-1}\log\left(\lceil n(u-\delta)\rceil - j\right) + \sum_{j=0}^{m-k-1}\log\left(n-\lfloor nu\rfloor - j\right) - \sum_{j=0}^{m-1}\log(n-j) \\ &= \sum_{j=0}^{k-1}\log\frac{\lceil n(u-\delta)\rceil - j}{n-j-m+k} + \sum_{j=0}^{m-k-1}\log\frac{n-\lfloor nu\rfloor - j}{n-j}. \end{aligned}$$

It is easy to check that the addends in both sums are (negative and) decreasing functions in j; thus, we can lower-bound the sums by integrals:

$$\begin{aligned} \log Q_k &\geq \int_0^k \log\frac{\lceil n(u-\delta)\rceil - j}{n-j-m+k}\,dj + \int_0^{m-k}\log\frac{n-\lfloor nu\rfloor - j}{n-j}\,dj \\ &= \lceil n(u-\delta)\rceil\log\lceil n(u-\delta)\rceil - \lceil n(u-\delta)\rceil\log\left(\lceil n(u-\delta)\rceil - k\right) \\ &\quad + k\log\left(\lceil n(u-\delta)\rceil - k\right) \\ &\quad + (n-m)\log(n-m) + (n-\lfloor nu\rfloor)\log\left(n-\lfloor nu\rfloor\right) \\ &\quad - (n-\lfloor nu\rfloor)\log\left(n-\lfloor nu\rfloor - m + k\right) \\ &\quad + (m-k)\log\left(n-\lfloor nu\rfloor - m + k\right) - n\log n. \end{aligned}$$

The right-hand side contains the expressions $f\left(\lceil n(u-\delta)\rceil\right)$ and $g\left(n-\lfloor nu\rfloor\right)$, where $f(x) := x\log x - x\log(x-k) + k\log(x-k)$ and $g(x) := x\log x - x\log(x-m+k) + (m-k)\log(x-m+k)$. It is easy to check that f and g are both increasing in the relevant intervals, thus we have $f\left(\lceil n(u-\delta)\rceil\right) \geq f\left(n(u-\delta)\right)$ and $g\left(n-\lfloor nu\rfloor\right) \geq g(n-nu)$, and all the floors and ceilings in the inequality above can be dropped.

Due to (46), we have $\log P_k = k\log u + (m-k)\log(1-u)$, thus

$$\begin{aligned} \log P_k - \log Q_k \leq\ & k\log u + (m-k)\log(1-u) - n(u-\delta)\log[n(u-\delta)] \\ & + n(u-\delta)\log[n(u-\delta)-k] \\ & - k\log[n(u-\delta)-k] - (n-m)\log(n-m) \\ & - n(1-u)\log[n(1-u)] \\ & + n(1-u)\log[n(1-u)-m+k] - (m-k) \\ & \log[n(1-u)-m+k] + n\log n. \end{aligned}$$

The largest contribution to the sum in (45) will be those k where $k \approx mu$. This motivates the definition $\varepsilon_k := k - mu$ (despite the name, this can be a negative number). Replacing all k by $mu + \varepsilon_k$ yields

$$\begin{aligned}\log P_k - \log Q_k \leq\ & [n-m-nu+mu+\varepsilon_k]\log\left(1+\frac{\varepsilon_k}{(1-u)(n-m)}\right)\\ & -(mu+\varepsilon_k)\log\left(1-\frac{\delta}{u}\right)\\ & +(nu-mu-\varepsilon_k-n\delta)\log\left(1-\frac{\varepsilon_k+m\delta}{(u-\delta)(n-m)}\right)\\ & -n\delta\log\left(1-\frac{m}{n}\right).\end{aligned}$$

All real numbers $x > -1$ satisfy $x/(1+x) \leq \log(1+x) \leq x$. Thus

$$\begin{aligned}\log P_k - \log Q_k \leq\ & [(n-m)(1-u)+\varepsilon_k]\cdot\frac{\varepsilon_k}{(1-u)(n-m)} + (mu+\varepsilon_k)\frac{\delta/u}{1-\delta/u}\\ & +[n(u-\delta)-mu-\varepsilon_k]\left(-\frac{\varepsilon_k+m\delta}{(u-\delta)(n-m)}\right)+n\delta\frac{m/n}{1-m/n}.\end{aligned} \tag{47}$$

We have the following three equations for the Binomial distribution:

$$\sum_{k=0}^{m}\binom{m}{k}u^k(1-u)^{m-k} = 1, \tag{48}$$

$$\sum_{k=0}^{m}\binom{m}{k}u^k(1-u)^{m-k}\varepsilon_k = 0, \tag{49}$$

$$\sum_{k=0}^{m}\binom{m}{k}u^k(1-u)^{m-k}\varepsilon_k^2 = mu(1-u), \tag{50}$$

where (48) is simply the normalization of the Binomial distribution, (50) is its variance, and (49) follows from its expectation value. Thus, when substituting (47) into the expression (45) for the relative entropy, we can drop all terms linear in ε_k. We obtain

$$\begin{aligned}S\left(\gamma_\beta^{\otimes m}\ \|\ \mathrm{Tr}_{\Lambda_n\setminus\Lambda_m}\tau_n\right) \leq\ & \sum_{k=0}^{m}\binom{m}{k}u^k(1-u)^{m-k}\\ & \left[\frac{\varepsilon_k^2}{(1-u)(n-m)}+\frac{m\delta}{1-\delta/u}+n\delta\frac{m}{n-m}\right.\\ & \left.+\frac{\varepsilon_k^2}{(u-\delta)(n-m)}-\frac{m\delta}{(u-\delta)(n-m)}\,(n(u-\delta)-mu)\right]\\ =\ & \frac{(1-\delta)u}{u-\delta}\cdot\frac{m}{n-m}+\frac{mu\delta}{u-\delta}\left(1+\frac{m}{n-m}\right).\end{aligned}$$

This proves the claim. □

Theorem 37. *Consider the case of qubits, i.e. $d = 2$, and shift the energies such that $E_0 = 0$ and $E_1 = 1$. Suppose that τ_n is the maximal mixture on the non-trivial microcanonical subspace corresponding to the energy interval $n\cdot[u-\delta, u]$, with*

$0 \leq \delta < u < \frac{1}{2}$, and γ_β is the single-site Gibbs state with corresponding inverse temperature β. If the size of the subsystem m is large enough such that $\frac{20}{m} \log \frac{m}{u} \leq \log \frac{1-u}{u}$, and at the same time $5 \leq m \leq n(u-\delta)$, then we have

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \frac{2\delta}{n\sqrt{u}} + \sqrt{\frac{m}{n-m}\left(1 + \frac{4 \log n}{\log \frac{1-u}{u}}\right)}.$$

Proof. We start by introducing some notation. For arbitrary subsets $S \subseteq [0, u]$ define τ_n^S to be the maximally mixed state on the subspace

$$T_n^S := \mathrm{span}\left\{ |s\rangle \;\middle|\; s \in \{0,1\}^n, \;\; \frac{1}{n}\sum_{i=1}^n E_{s_i} \in S \right\}.$$

As before, we set $T_n := T_n^{[u-\delta,u]}$ and $\tau_n := \tau_n^{[u-\delta,u]}$. Moreover, define

$$\mu_n^S := \frac{\dim T_n^S}{\dim T_n};$$

then, if we write $[u-\delta, u]$ as any disjoint union of two sets S and T, the microcanonical state can be written as a convex combination, $\tau_n = \mu_n^S \tau_n^S + \mu_n^T \tau_n^T$. In the following, $(\alpha_n)_{n\in\mathbb{N}}$ will be any sequence of positive real numbers tending to zero, satisfying $1/n < \alpha_n < u - 1/n$, to be specified later. We start with the identity

$$\tau_n = \mu_n^{[u-\delta,u-\alpha_n)} \tau_n^{[u-\delta,u-\alpha_n)} + \mu_n^{[u-\alpha_n,u]} \tau_n^{[u-\alpha_n,u]}.$$

Due to convexity, the Pinsker inequality, and Lemma 36, we have

$$\begin{aligned}
\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 &\leq \mu_n^{[u-\delta,u-\alpha_n)} \left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n^{[u-\delta,u-\alpha_n)} - \gamma_\beta^{\otimes m} \right\|_1 \\
&\quad + \mu_n^{[u-\alpha_n,u]} \left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n^{[u-\alpha_n,u]} - \gamma_\beta^{\otimes m} \right\|_1 \\
&\leq 2\mu_n^{[u-\delta,u-\alpha_n)} + \sqrt{\frac{1}{2} S\left(\gamma_\beta^{\otimes m} \,\middle\|\, \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n^{[u-\alpha_n,u]}\right)} \\
&\leq 2\mu_n^{[u-\delta,u-\alpha_n)} \\
&\quad + \sqrt{\frac{1}{2}} \cdot \sqrt{\frac{(1-\alpha_n)u}{u-\alpha_n} \cdot \frac{m}{n-m} + \frac{mu\alpha_n}{u-\alpha_n}\left(1 + \frac{m}{n-m}\right)}.
\end{aligned} \tag{51}$$

Let $\tilde{u}$ be the largest $p \in [u-\delta, u]$ with the property that $T_n^{\{p\}} \neq \{0\}$; it is given by the equation $\lfloor u \cdot n \rfloor = \tilde{u} \cdot n$. Then we can upper-bound the measure $\mu_n^{[u-\delta,u-\alpha_n)}$ in the following way:

$$\begin{aligned}
\mu_n^{[u-\delta,u-\alpha_n)} &= \frac{\sum_{p\in[u-\delta,u-\alpha_n)} \dim T_n^{\{p\}}}{\sum_{p\in[u-\delta,u]} \dim T_n^{\{p\}}} \\
&\leq \frac{\sum_{j\in\mathbb{N}:\, j/n\in[u-\delta,u-\alpha_n)} \binom{n}{j}}{\dim T_n^{\{\tilde{u}\}}}
\end{aligned}$$

$$\leq \frac{\#\{j \in \mathbb{N} : j/n \in [u-\delta, u-\alpha_n)\} \cdot \binom{n}{\lfloor n(u-\alpha_n)\rfloor}}{\binom{n}{n\tilde{u}}}$$
$$\leq \frac{n\delta \binom{n}{\lfloor n(u-\alpha_n)\rfloor}}{\binom{n}{\lfloor nu\rfloor}}. \tag{52}$$

The Binomial coefficients can be estimated by using Lemma 17.5.1 in [21]: For $0 < p < 1$ such that np is an integer, we have

$$\frac{1}{\sqrt{8np(1-p)}} \leq \binom{n}{np} e^{-nH(p)} \leq \frac{1}{\sqrt{\pi np(1-p)}},$$

where $H(p) = -p\log p - (1-p)\log(1-p)$ is the binary entropy function. Substituting this into (52), defining p by $np = \lfloor n(u-\alpha_n)\rfloor$, and using that $p \leq u - \alpha_n$ as well as $u \geq \tilde{u} \geq u - 1/n$, we obtain

$$\mu_n^{[u-\delta, u-\alpha_n)} \leq n\delta \binom{n}{np}\binom{n}{n\tilde{u}}^{-1} \leq n\delta \sqrt{\frac{\tilde{u}(1-\tilde{u})}{p(1-p)}} e^{n[H(u-\alpha_n) - H(u-1/n)]} \tag{53}$$

(note that $u > 1/n$ due to $n \geq m/(u-\delta) > m/u \geq 1/u$). Since the binary entropy function H is concave in the interval $[0, 1/2]$, we have

$$H(u-\alpha_n) \leq H\left(u - \frac{1}{n}\right) - H'\left(u - \frac{1}{n}\right) \cdot \left(\alpha_n - \frac{1}{n}\right)$$
$$\Rightarrow \quad H(u-\alpha_n) - H\left(u - \frac{1}{n}\right) \leq -\left(\alpha_n - \frac{1}{n}\right) \log \frac{1 - (u - 1/n)}{u - 1/n}.$$

Substituting this and $\tilde{u}(1-\tilde{u}) \leq 1/4$ as well as $1/\sqrt{p(1-p)} \leq \sqrt{2/p}$ and $p \geq u - \alpha_n - 1/n$ into (53), we get

$$\mu_n^{[u-\delta, u-\alpha_n)} \leq \frac{1}{2} n\delta \sqrt{\frac{2}{u - \alpha_n - 1/n}} \, (c_n)^{-n\left(\alpha_n - \frac{1}{n}\right)}, \qquad \text{where } c_n = \left(\frac{1 - \left(u - \frac{1}{n}\right)}{u - \frac{1}{n}}\right).$$

Now set

$$\alpha_n := \frac{1}{n} + \frac{2\log n}{n \log c_n} = \mathcal{O}\left(\frac{\log n}{n}\right). \tag{54}$$

Since $n > m/u$, this is less than $u - 1/n$ as necessary if m is large enough; it turns out that $m \geq 5$ and $(20/m)\log(m/u) \leq \log((1-u)/u)$ gives in fact $\alpha_n < u/2 - 1/n$. This yields $(c_n)^{-n\left(\alpha_n - \frac{1}{n}\right)} = n^{-2}$, and so

$$\mu_n^{[u-\delta, u-\alpha_n)} \leq \frac{\delta}{n\sqrt{u}}.$$

Substituting this and $u - \alpha_n > u/2$ as well as $m/(n-m) \leq 1$ into (51) yields

$$\left\| \mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} \tau_n - \gamma_\beta^{\otimes m} \right\|_1 \leq \frac{2\delta}{n\sqrt{u}} + \frac{1}{\sqrt{2}} \sqrt{2\frac{m}{n-m} + 4m\alpha_n}.$$

Then the claim follows by substituting (54) and $\log c_n = \log(1 - u + 1/n) - \log(u - 1/n) > \log(1-u) - \log u$. □

This mainly recovers the result depicted in Fig. 5, where the size of the "bath", $n-m$, has to be increased linearly with the size of the subsystem, m, to achieve a fixed error. In this theorem, for $\delta > 0$, the $(\log n)$-term contributes a small correction to this behavior, and n has to be increased slightly super-linearly with m.

3.5. Numerical results on finite-size behavior in one dimension. Here we provide numerical examples that not only show that random local Hamiltonians satisfy our requirements for canonical typicality and dynamical thermalization, but also that the replacing the global Gibbs state with the local Gibbs state does not give the correct statistics. This emphasizes that entanglement is key to understanding why closed quantum systems can conform to thermodynamic predictions. The class of Hamiltonians that we consider are random 2-local Hamiltonians acting on n qubits on a line with periodic boundary conditions:

$$H^p_{\Lambda_n} = \sum_{i=1}^{n} \left(H_0^{(i)} + H_{\text{int}}^{(i,i+1 \bmod n)} \right), \tag{55}$$

where the onsite term is of the form for constants a_1, a_2 and a_3,

$$H_0^{(i)} = a_1 \sigma_x^{(i)} + a_2 \sigma_y^{(i)} + a_3 \sigma_z^{(i)}, \tag{56}$$

and the interaction term takes the form, for constants $b_{1,1}, b_{1,2}, \ldots, b_{3,3}$,

$$H_{\text{int}}^{(i,j)} = b_{1,1} \sigma_x^{(i)} \sigma_x^{(j)} + b_{1,2} \sigma_x^{(i)} \sigma_y^{(j)} + \cdots + b_{3,3} \sigma_z^{(i)} \sigma_z^{(j)}. \tag{57}$$

The constants a_i and $b_{i,j}$ are chosen randomly according to a Gaussian distribution with zero mean and unit variance. For ease of comparison, each random translationally invariant Hamiltonian is re-normalized to have unit norm. Note that one-dimensional translation-invariant systems with finite-range interaction do not exhibit finite temperature phase transitions.

The numerical experiments begin by drawing a random Hamiltonian $H^p_{\Lambda_n}$ for a fixed value of β and energy window δ. The first step is to compute the energy density u using $u = \frac{1}{|\Lambda_n|} \text{Tr}\left(\gamma^p_{\Lambda_n} H^p_{\Lambda_n} \right)$ where $\gamma^p_{\Lambda_n} = \exp(-\beta H^p_{\Lambda_n})/Z$ is the thermal state that results from the choice of β. The Hamiltonian is then diagonalized and all energy eigenvectors within the window $(u-\delta, u)$ are found. A random state $|\psi\rangle$ is then constructed out of the span of these vectors, and then we compute $\left\| \text{Tr}_{\Lambda_n \setminus \Lambda_m} |\psi\rangle\langle\psi| - \text{Tr}_{\Lambda_n \setminus \Lambda_m} \frac{\exp(-\beta H^p_{\Lambda_n})}{Z} \right\|_1$, as per Theorem 25. We take the subsystem to consist of a single qubit, i.e. $m = 1$, and the bath contain $n-1$ qubits in all these examples. This process is repeated for many such random Hamiltonians and we compute the mean and the standard deviation of these distances, which allows us to see whether the correspondence predicted by Theorem 25 is typical for this ensemble of random local Hamiltonians.

The data in Fig. 6 shows that the distance between the reduced density matrix of the pure state and the Gibbs state shrinks as n increases, roughly as $\mathcal{O}(1/n)$. The error bars (representing the standard deviation of the discrepancy with the canonical state) also shrink as n increases, illustrating that almost all such random translationally invariant 2–local Hamiltonians agree with the predictions of Theorem 25 and in turn that there is a strong correspondence between the subsystem traces of the global Gibbs state and $|\psi\rangle\langle\psi|$.

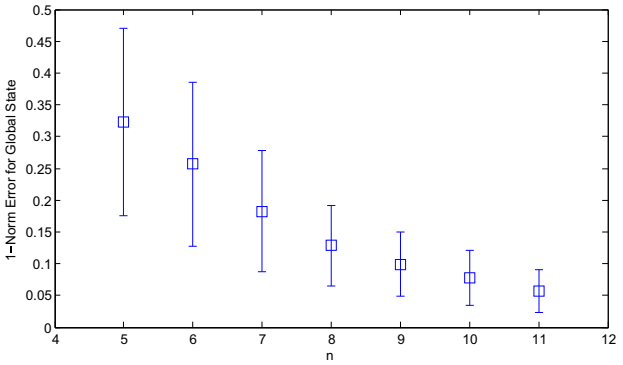

Fig. 6. 1-norm difference between the reduced density operator and the reduced global Gibbs state for a system of n qubits. The *squares* represent the ensemble means and the *error bars* give the standard deviations of the differences between the Gibbs state and the subsystem trace. The data was collected for $\beta = 0.1$ and $\delta = 0.02n$ and 400 random Hamiltonians were considered for each n

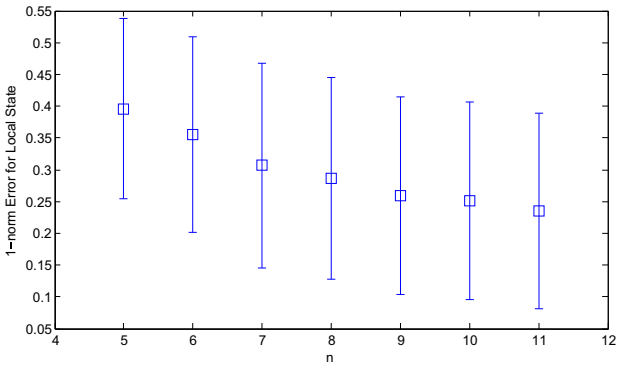

Fig. 7. 1-norm difference between the reduced density operator and the reduced local Gibbs state for a system of n qubits. The *squares* and *error bars* are defined identically to those in Fig. 6. The data was collected for $\beta = 0.1$ and $\delta = 0.02n$ and 400 random Hamiltonians were considered for each n

On the other hand, Fig. 7 shows that substituting the local Gibbs state for the subsystem trace of the global Gibbs state causes this correspondence to break down. In particular, we see no clear evidence that the ensemble mean of the differences between $\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} |\psi\rangle\langle\psi|$ and the local Gibbs state approaches zero as n increases; more tellingly, the standard deviation of the differences does not seem to decrease with n. These results suggest that even as n increases, $\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} |\psi\rangle\langle\psi|$ remains distinct from the local Gibbs state. Thus the correspondence suggested by Theorem 25 is correct and the naïve correspondence between the local Gibbs state and $\mathrm{Tr}_{\Lambda_n \setminus \Lambda_m} |\psi\rangle\langle\psi|$ is incorrect.

Regarding dynamical thermalization, there are two caveats that we need to check in order to justify the applicability of Theorem 31. First, we need to ensure that almost all Hamiltonians drawn from this random ensemble are non-degenerate, in order to ensure thermalization for arbitrary initial states with maximal population entropy. Figure 8 shows that the probability of small eigenvalue gaps is suppressed, hence Hamiltonians that are typical of the random local Hamiltonian ensemble will be non-degenerate. Second, we need to show that the gap degeneracy $D_G(H^p_{\Lambda_n})$ is not too large. Figure 9 shows that, with high probability, the eigenvalue gaps between any two energy levels

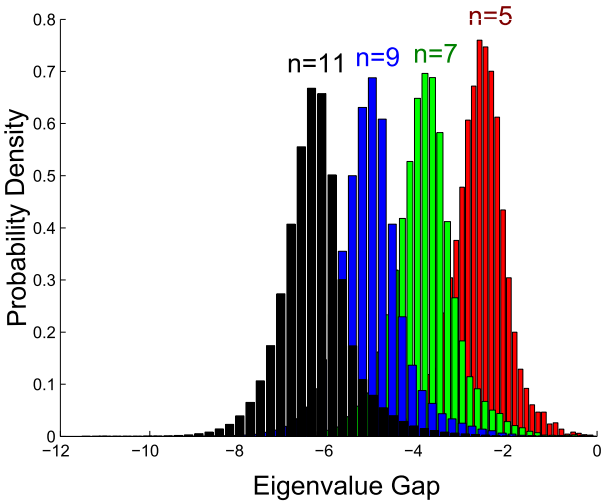

Fig. 8. Probability density of eigenvalue gaps for random Hamiltonians with $n = 5, 7, 9$ and 11 qubits. The x-axis is $\log_{10}(\text{gap})$ for 100 random Hamiltonians. No degenerate eigenvalues were ever detected in this sample within numerical error

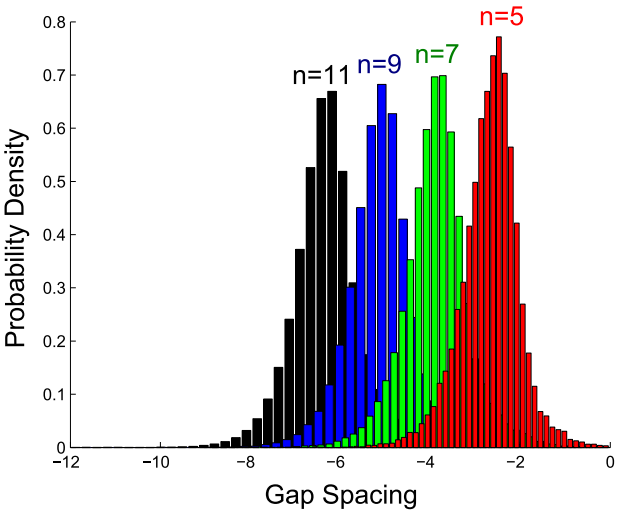

Fig. 9. Probability density of for the eigenvalue gap spacings for random Hamiltonians with $n = 5, 7, 9$ and 11 qubits. The x-axis is $\log_{10}(\text{gap}(\text{gap}))$ for 100 random Hamiltonians. No degenerate eigenvalue gaps were ever detected in this sample within numerical error

will be distinct from any other such gap in the system, hence $D_G(H^p_{\Lambda_n}) = 1$ with high probability.

These results illustrate the application of our results to a wide range of physically realistic random 2-local Hamiltonians. It is further reasonable to expect that broad classes of physically realistic closed quantum systems will agree with the canonical distribution, illuminating the mechanism by which thermodynamics emerges for macroscopic closed quantum systems.

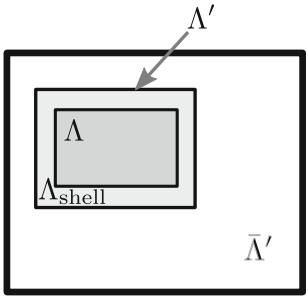

Fig. 10. Subdivision of the whole lattice, $\Lambda_{\text{lattice}} = \Lambda_n$, into regions as used in this subsection. We have $\Lambda' = \Lambda \cap \Lambda_{\text{shell}}$

3.6. Local diagonality of energy eigenstates. A strong sense in which the eigenstates of a local Hamiltonian H could thermalize is that their reduced density matrix of a region Λ (much smaller than the full lattice $\Lambda_{\text{lattice}} = \Lambda_n$) is approximately equal to a thermal state in that region,

$$\text{Tr}_{\bar{\Lambda}} |E\rangle\langle E| \approx \frac{e^{-\beta H_\Lambda}}{\text{tr}\, e^{-\beta H_\Lambda}}, \tag{58}$$

where $\text{Tr}_{\bar{\Lambda}}$ denotes trace on the Hilbert space associated to the complementary region $\bar{\Lambda} = \Lambda_{\text{lattice}} \backslash \Lambda$, and H_Λ is the sum of all terms of H which are fully contained in the region Λ. The inverse temperature β should be chosen such that $\langle E|H_\Lambda|E\rangle = \text{tr}(H_\Lambda e^{-\beta H_\Lambda})/\text{tr}\, e^{-\beta H_\Lambda}$ holds.

A possible concern is that the Hamiltonian H_Λ has open boundary conditions, hence we expect boundary effects in the eigenstates of H_Λ which are not present in $\text{Tr}_{\bar{\Lambda}} |E\rangle\langle E|$; and this makes unlikely that relation (58) holds. A way to get rid of the boundary effects is by defining a slightly larger region Λ' which includes a shell of width l around Λ, cf. Fig. 10; that is

$$\Lambda' := \{x \in \Lambda_{\text{lattice}} : \exists y \in \Lambda : \text{dist}(x, y) \leq l\}. \tag{59}$$

If instead of (58) we consider the thermal state in Λ' and trace out the shell $\Lambda_{\text{shell}} := \Lambda' \setminus \Lambda$, then the approximate equality

$$\text{Tr}_{\bar{\Lambda}} |E\rangle\langle E| \approx \text{Tr}_{\Lambda_{\text{shell}}} \left(\frac{e^{-\beta H_{\Lambda'}}}{\text{tr}\, e^{-\beta H_{\Lambda'}}} \right) \tag{60}$$

is more likely to hold in generic systems, because by tracing out the shell we may eliminate the boundary effects of the eigenvectors of $H_{\Lambda'}$. (As before, we denote by $H_{\Lambda'}$ the sum of all terms in H which are fully contained in Λ'.)

It is expected that the relation (60) holds for generic local Hamiltonians, but not for all local Hamiltonians. For example, consider the translational-invariant quantum Ising Hamiltonian in one dimension that we analyzed in Sect. 3.4. This is a Hamiltonian without interaction terms, such as $H_{\Lambda_{\text{shell}}} = \sum_{i=1}^{n} h_i$ for $\Lambda_{\text{shell}} = [1, n]$, with constant single-site terms h_i. If, for example, $h_i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, the computational basis vectors $|E\rangle = |x_1 x_2 \ldots x_n\rangle$ with $x_i \in \{0, 1\}$ are energy eigenstates. Even for those eigenstates that correspond to finite energies $E > 0$ with corresponding inverse temperature $\beta < \infty$, the local reduced state on $\Lambda = [1, m]$, $m \ll n$, is $\text{Tr}_{\bar{\Lambda}} |E\rangle\langle E| = |x_1 \ldots x_m\rangle\langle x_1 \ldots x_m|$.

This is a pure state, far away from any thermal state of temperature β. Thus, (60) does not hold for the Ising model.

In summary, extra conditions are necessary for (60) to hold. Folk wisdom tells us that such conditions could be along the lines of non-integrability, although this is not yet a clear and mathematically well-defined concept within quantum theory. In this work, we follow a different approach: instead of looking for additional conditions, we relax the statement (60). One way to do this is by noticing that the state $e^{-\beta H_{\Lambda'}}/\mathrm{tr}\, e^{-\beta H_{\Lambda'}}$ is diagonal in the eigenbasis of $H_{\Lambda'}$. Our weakened statement is informally the following:

For any eigenvalue E of H there is a density matrix ω_E defined in the extended region Λ' which is weakly diagonal in the eigenbasis of $H_{\Lambda'}$ and satisfies

$$\mathrm{Tr}_{\bar{\Lambda}} |E\rangle\langle E| \;\approx\; \mathrm{Tr}_{\Lambda_{\mathrm{shell}}} \omega_E .$$

The meaning of weakly diagonal will be made precise in the statement of the theorem below. But before, let us specify the type of systems that we are considering. Exactly as explained at the beginning of Sect. 3.1, we consider local Hamiltonians on a cubic lattice, with a finite-dimensional Hilbert space at each site. By local we mean that the Hamiltonian H has finite interaction range r. This means that if we write it as

$$H = \sum_{\mathcal{X} \subseteq \Lambda_{\mathrm{lattice}}} \Phi(\mathcal{X}),$$

where $\Phi(\mathcal{X})$ has only support on the region $\mathcal{X}$, then for any region $\mathcal{X} \subseteq \Lambda_{\mathrm{lattice}}$ such that $\mathrm{diam}\, \mathcal{X} := \max_{x,x' \in \mathcal{X}} \mathrm{dist}(x, x') > r$ we have $\Phi(\mathcal{X}) = 0$ (the definition of dist is given in (15)). However, in contrast to the previous subsection, we do not need to assume that the interaction is translation-invariant. This type of Hamiltonian satisfies a Lieb-Robinson bound [30,31] (see [32] for a simpler proof). That is, let X, Y be two matrices acting non-trivially in the regions $\mathcal{X}, \mathcal{Y} \subseteq \Lambda_{\mathrm{lattice}}$ which are separated by a distance $\mathrm{dist}(\mathcal{X}, \mathcal{Y})$, and let $X(t) = e^{iHt} X e^{-iHt}$. There are positive constants C, c, v such that

$$\|[X(t), Y]\|_\infty \;\leq\; C\, \|X\|_\infty \|Y\|_\infty \min\{|\mathcal{X}|, |\mathcal{Y}|\}\, e^{-c[\mathrm{dist}(\mathcal{X},\mathcal{Y}) - v|t|]}. \tag{61}$$

The constants C, c, v only depend on coarse features of the lattice and the Hamiltonian, like the interaction length, and the $\|\cdot\|_\infty$-norm of the local terms in the Hamiltonian. The constant v is called the Lieb-Robinson velocity, and it is an upper-bound for the speed at which information travels through the lattice.

Theorem 38 (Weak local diagonality). *Let Φ be any finite-range interaction (not necessarily translation-invariant), let $\Lambda \subseteq \Lambda_{\mathrm{lattice}}$ be any region of the lattice, and let $\Lambda' \subseteq \Lambda_{\mathrm{lattice}}$ be the set of points at distance not larger than l from Λ, as defined in (59). Define the regions $\Lambda_{\mathrm{shell}} = \Lambda' \setminus \Lambda$ and the complements $\bar{\Lambda} = \Lambda_{\mathrm{lattice}} \setminus \Lambda$ and $\bar{\Lambda}' = \Lambda_{\mathrm{lattice}} \setminus \Lambda'$. Let H be a local Hamiltonian as defined above, with finite interaction range $r \leq l$. For each eigenvector $|E\rangle$ of the Hamiltonian H we define the state ω_E in the region Λ' as*

$$\omega_E := \int_{-\infty}^{\infty} dt\, g(t)\, e^{-iH_{\Lambda'}t}\, \mathrm{Tr}_{\bar{\Lambda}'}(|E\rangle\langle E|)\, e^{iH_{\Lambda'}t},$$

where $g(t) = (2\pi\sigma^2)^{-1/2}\, e^{-t^2/(2\sigma^2)}$ and $\sigma^2 = (l - r)/(4cv^2)$. The state ω_E is weakly diagonal in the eigenbasis of $H_{\Lambda'}$, denoted $|e\rangle$, in the sense that

$$|\langle e_1|\omega_E|e_2\rangle| \leq e^{-(l-r)(e_1-e_2)^2/(8cv^2)}. \tag{62}$$

The state ω_E is almost indistinguishable from $|E\rangle\langle E|$ inside the region Λ, that is

$$\left\|\mathrm{Tr}_{\Lambda_{\mathrm{shell}}}(\omega_E) - \mathrm{Tr}_{\bar{\Lambda}}(|E\rangle\langle E|)\right\|_1 \leq \frac{2}{\sqrt{2\pi}} A J \sigma (C A + 2) e^{-c(l-r)/2}, \tag{63}$$

where A is the number of subsets X with $\Phi(X) \neq 0$ that have non-empty intersection with both Λ' and $\bar{\Lambda}'$.

Note that the number A quantifies the size of the boundary of Λ'; so for a three-dimensional lattice, A is an area. Also, we stress the fact that closeness in $\|\cdot\|_1$-norm is a very strong feature, and it really implies that the two states in the left-hand side of (63) are almost indistinguishable. The right-hand side of (63) can be made small by choosing the thickness of the shell to be

$$l \gtrsim \frac{6}{c} \log A + r.$$

Still, for large regions Λ, the relative volume of the shell $lA/|\Lambda|$ vanishes.

If the local dimension is d, then the dimension of the Hilbert space associated to the region Λ' is $d^{|\Lambda'|}$. Hence, the expected size of the entries of ω_E is of the order of $d^{-|\Lambda'|}$, which is very small. This may rise the concern that bound (62) is trivial. To see that this is not the case, we note that the largest entry of ω_E is at least $d^{-|\Lambda'|}$. Also, since $H_{\Lambda'}$ is a local Hamiltonian, the range of energies is $\Delta e \sim J|\Lambda'|$. This implies that the exponent of (62) is proportional to $|\Lambda'|^2$, while the exponent of the largest entry is proportional to $|\Lambda'|$, which is much smaller. In summary, for large enough regions $|\Lambda|$, the bound (62) is non-trivial. It is a consequence of the locality of interactions as expressed by the Lieb-Robinson bound.

Proof. Using the fact that the $|e\rangle$ are the eigenvectors of $H_{\Lambda'}$ we obtain

$$\begin{aligned}\langle e_1|\omega_E|e_2\rangle &= \int dt\, g(t)\, e^{-i(e_1-e_2)t}\, \langle e_1|\mathrm{Tr}_{\bar{\Lambda}'}(|E\rangle\langle E|)|e_2\rangle \\ &= e^{-(e_1-e_2)^2\sigma^2/2}\langle e_1|\mathrm{Tr}_{\bar{\Lambda}'}(|E\rangle\langle E|)|e_2\rangle,\end{aligned}$$

which implies (62). Using the triangle inequality for the norm $\|\cdot\|_1$ we obtain

$$\begin{aligned}&\left\|\mathrm{Tr}_{\Lambda_{\mathrm{shell}}}(\omega_E) - \mathrm{Tr}_{\bar{\Lambda}}(|E\rangle\langle E|)\right\|_1 \\ &= \left\|\int dt\, g(t)\, \mathrm{Tr}_{\bar{\Lambda}}\left(e^{-iH_{\Lambda'}t}|E\rangle\langle E|e^{iH_{\Lambda'}t} - |E\rangle\langle E|\right)\right\|_1 \\ &\leq \int dt\, g(t)\, \left\|\mathrm{Tr}_{\bar{\Lambda}}\left(e^{-iH_{\Lambda'}t}|E\rangle\langle E|e^{iH_{\Lambda'}t} - |E\rangle\langle E|\right)\right\|_1 .\end{aligned} \tag{64}$$

Next, we use the identity $\|Y\|_1 = \max_X |\mathrm{Tr}(XY)|$, where the maximum is over all Hermitian matrices X which satisfy $-\mathbf{1} \leq X \leq \mathbf{1}$. Since we apply this to an observable on Λ, it follows that X is fully supported on Λ. We also use the fact that $e^{iHt}|E\rangle\langle E|e^{-iHt} = |E\rangle\langle E|$ for any t, obtaining

$$\begin{aligned}&\left\|\mathrm{Tr}_{\bar{\Lambda}}\left(e^{-iH_{\Lambda'}t}|E\rangle\langle E|e^{iH_{\Lambda'}t} - |E\rangle\langle E|\right)\right\|_1 \\ &= \max_X \left|\mathrm{Tr}\left[X\left(e^{-iH_{\Lambda'}t}e^{iHt}|E\rangle\langle E|e^{-iHt}e^{iH_{\Lambda'}t} - |E\rangle\langle E|\right)\right]\right| \\ &= \max_X \left|\langle E|e^{-iHt}e^{iH_{\Lambda'}t}Xe^{-iH_{\Lambda'}t}e^{iHt} - X|E\rangle\right| .\end{aligned} \tag{65}$$

Now we use the inequality $|\langle\alpha|Y|\beta\rangle| \leq \|Y\|_\infty$ for any pair of unit vectors $|\alpha\rangle$, $|\beta\rangle$. Also, we use the fact that $[X, H_{\bar{\Lambda}'}] = [H_{\Lambda'}, H_{\bar{\Lambda}'}] = 0$, and define $H_A := H - H_{\bar{\Lambda}'} - H_{\Lambda'}$. We obtain

$$\left|\langle E|e^{-iHt}e^{iH_{\Lambda'}t}Xe^{-iH_{\Lambda'}t}e^{iHt} - X|E\rangle\right| \leq \left\|e^{-iHt}e^{i(H-H_A)t}Xe^{-i(H-H_A)t}e^{iHt} - X\right\|_\infty \tag{66}$$

Next, we use the matrix identity $M(t) - M(0) = \int_0^t dt_1 \frac{\partial}{\partial t_1} M(t_1)$, the triangle inequality, and the unitary invariance of the operator norm, $\|e^{-iHt_1}Ye^{iHt_1}\|_\infty = \|Y\|_\infty$. If $t \geq 0$ then

$$\begin{aligned}
&\left\|e^{-iHt}e^{i(H-H_A)t}Xe^{-i(H-H_A)t}e^{iHt} - X\right\|_\infty \\
&= \left\|\int_0^t dt_1 \frac{\partial}{\partial t_1}\left(e^{-iHt_1}e^{i(H-H_A)t_1}Xe^{-i(H-H_A)t_1}e^{iHt_1}\right)\right\|_\infty \\
&\leq \int_0^{|t|} dt_1 \left\|\left[H_A, e^{iH_{\Lambda'}t_1}Xe^{-iH_{\Lambda'}t_1}\right]\right\|_\infty .
\end{aligned}$$

If $t < 0$, then the substitution $t_2 := -t_1$ in the integral yields

$$\begin{aligned}
&\left\|e^{-iHt}e^{i(H-H_A)t}Xe^{-i(H-H_A)t}e^{iHt} - X\right\|_\infty \\
&= \left\|\int_0^{|t|} dt_2 \frac{\partial}{\partial t_2}\left(e^{iHt_2}e^{-i(H-H_A)t_2}Xe^{i(H-H_A)t_2}e^{-iHt_2}\right)\right\|_\infty \\
&\leq \int_0^{|t|} dt_2 \left\|\left[H_A, e^{-iH_{\Lambda'}t_2}Xe^{iH_{\Lambda'}t_2}\right]\right\|_\infty .
\end{aligned}$$

In both cases, we can apply the Lieb-Robinson bound to the two regions $\mathcal{X} = \Lambda$ and $\mathcal{Y}$ the support region of H_A (covering the boundary of Λ' and of $\bar{\Lambda}'$). For all $t \in \mathbb{R}$, we get

$$\left\|\left[H_A, e^{iH_{\Lambda'}t}Xe^{-iH_{\Lambda'}t}\right]\right\|_\infty \leq \|H_A\|_\infty \min\left\{2, CA\, e^{-c(l-r)+cv|t|}\right\},$$

which implies

$$\begin{aligned}
&\left\|e^{-iHt}e^{i(H-H_A)t}Xe^{-i(H-H_A)t}e^{iHt} - X\right\|_\infty \\
&\leq \|H_A\|_\infty \min\left\{\int_0^{|t|} dt_1 \cdot 2,\ \int_0^{|t|} dt_1\, CAe^{-c(l-r)+cv|t_1|}\right\} \\
&\leq \|H_A\|_\infty \min\left\{2|t|,\ CA|t|e^{-c(l-r)+cv|t|}\right\}.
\end{aligned}$$

Combining this with (64)–(66), and dividing the integration (64) into two intervals, we get for $t_0 \geq 0$

$$\begin{aligned}
&\left\|\mathrm{Tr}_{\Lambda_{\mathrm{shell}}}(\omega_E) - \mathrm{Tr}_{\bar{\Lambda}}(|E\rangle\langle E|)\right\|_1 \\
&\leq 2\|H_A\|_\infty \int_0^\infty dt\, g(t) \min\{2t,\ CAt\, e^{-c(l-r)+cvt}\} \\
&\leq 2\|H_A\|_\infty \left(\int_0^{t_0} dt\, g(t)\, CAt\, e^{-c(l-r)+cvt} + \int_{t_0}^\infty dt\, g(t)\, 2t\right)
\end{aligned}$$

$$\leq 2\|H_A\|_\infty \left(CA\, e^{-c(l-r)+cvt_0} \int_0^\infty dt\, g(t)t + \frac{\sigma}{\sqrt{2\pi}} 2\, e^{-t_0^2/(2\sigma^2)} \right)$$
$$\leq 2\|H_A\|_\infty \left(\frac{\sigma}{\sqrt{2\pi}} CAe^{-c(l-r)+cvt_0} + \frac{\sigma}{\sqrt{2\pi}} 2\, e^{-t_0^2/(2\sigma^2)} \right).$$

Now choose $t_0 := (l-r)/(2v)$ such that $-c(l-r) + cvt_0 = -t_0^2/(2\sigma^2)$, and use $\sigma^2 = (l-r)/(4cv^2)$. Furthermore,

$$H_A = H_{\Lambda_{\text{lattice}}} - H_{\bar{\Lambda}'} - H_{\Lambda'} = \sum_{X \subset \Lambda_{\text{lattice}}:\, X \cap \Lambda' \neq \emptyset \text{ and } X \cap \bar{\Lambda}' \neq \emptyset} \Phi(X),$$

such that $\|H_A\|_\infty \leq A\, J$, where $J = \max_{X \subset \mathbb{Z}^\nu} \|\Phi(X)\|_\infty$, and A is the number of subsets X with $\Phi(X) \neq 0$ that have non-empty intersection with both Λ' and $\bar{\Lambda}'$. □

4. Conclusions

Our work provides a significant step towards a rigorous understanding of how closed quantum systems thermalize. Our key innovations come from combining methods from quantum information theory and from more traditional mathematical physics techniques to address the problem. Through this approach, we find that small subsystems of closed translation-invariant quantum systems with finite-range interaction thermalize, in the sense that they relax towards the reduction of the global Gibbs state. In doing so, we not only provide a rigorous explanation for how a wide class of physically significant Hamiltonians thermalize, but also show that the correct correspondence is with a reduction of the global system's Gibbs state, not its local Gibbs state.

This work opens a number of interesting avenues for future work. One open problem is to obtain more explicit finite-size bounds, but these may well depend on details of the specific model or interaction. Similarly, an interesting open question is whether ω_E in Theorem 4 has Boltzmann weights on its diagonal. However, rigorously answering this question in the affirmative, and thus proving a complete version of the eigenstate thermalization hypothesis, seems to require additional assumptions along the lines of nonintegrability. Thus, one may hope that attempts to prove the ETH for quantum lattice systems will also lead to a better understanding and rigorous mathematical definition of the notion of integrability in the quantum case. We further believe that the methodology we provide will lead to further applications to be discovered in the future. In particular, it may turn out that giving finite versions of asymptotic mathematical physics results will prove to be as promising as using asymptotic results to prove statements on finite systems, which was the approach taken in this paper.

Acknowledgement. MM would like to thank Jens Eisert, Joe Emerson, Patrick Hayden, and Sandu Popescu for discussions in early stages of this project, and Oscar Dahlsten for comments. Research at Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI. LM acknowledges support from the EU ERC Advanced Grant NLST (PHYS RQ8784), EU Qessence project, EPSRC and the Templeton Foundation. This work was partially supported by the COST Action MP1209.

References

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

2. Bañuls, M.C., Cirac, J.I., Hastings, M.B.: Strong and weak thermalization of infinite nonintegrable quantum systems. Phys. Rev. Lett. **106**, 050405 (2011)
3. Goldstein, S., Lebowitz, J.L., Tumulka, R., Zanghi, N.: Canonical typicality. Phys. Rev. Lett. **96**, 050403 (2006)
4. Popescu, S., Short, A.J., Winter, A.: Entanglement and the foundations of statistical mechanics. Nat. Phys. **2**, 754 (2006)
5. Reimann, P.: Foundation of statistical mechanics under experimentally realistic conditions. Phys. Rev. Lett. **101**, 190403 (2008)
6. Linden, N., Popescu, S., Short, A.J., Winter, A.: Quantum mechanical evolution towards thermal equilibrium. Phys. Rev. E **79**, 061103 (2009)
7. Short, A.J.: Equilibration of quantum systems and subsystems. New J. Phys. **13**(5), 053009 (2011)
8. Short, A.J., Farrelly, T.C.: Quantum equilibration in finite time. New J. Phys. **14**, 013063 (2012)
9. Ududec, C., Wiebe, N., Emerson, J.: Information-theoretic equilibration: the appearance of irreversibility under complex quantum dynamics. Phys. Rev. Lett. **111**, 080403 (2013)
10. Riera, A., Gogolin, C., Eisert, J.: Thermalization in nature and on a quantum computer. Phys. Rev. Lett. **108**, 080802 (2012)
11. Lima, R.: Equivalence of ensembles in quantum lattice systems. Annales de l'I. H.P. **15**(1), 61–68 (1971)
12. Lima, R.: Equivalence of ensembles in quantum lattice systems: states. Commun. Math. Phys. **24**, 180–192 (1972)
13. Simon, B.: The Statistical Mechanics of Lattice Gases, Vol. 1. Princeton University Press, Princeton (1993)
14. Bratteli, O., Robinson, D.W.: Operator Algebras and Quantum Statistical Mechanics I and II. Springer, New York (2002)
15. Araki, H.: Gibbs states of a one dimensional quantum lattice. Commun. Math. Phys. **14**, 120–157 (1969)
16. Araki, H.: On the equivalence of the KMS condition and the variational principle for quantum lattice systems. Commun. Math. Phys. **38**, 1–10 (1974)
17. Araki, H.: On uniqueness of KMS states of one-dimensional quantum lattice systems. Commun. Math. Phys. **44**, 1–7 (1975)
18. Kliesch, M., Gogolin, C., Kastoryano, M.J., Riera, A., Eisert, J.: Locality of temperature. Phys. Rev. X **4**, 031019 (2014). arXiv:1309.0816
19. De Roeck, W., Maes, C., Netočný, K.: Quantum macrostates, equivalence of ensembles and an H-theorem. J. Math. Phys. **47**, 073303 (2006)
20. Bhatia, R.:Perturbation bounds for matrix eigenvalues. SIAM Class. Appl. Math. (2007)
21. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (2006)
22. Diaconis, P., Freedman, D.: Finite exchangeable sequences. Ann. Probab. **8**(4), 745–764 (1980)
23. Deserno, M.: Microcanonical and canonical two-dimensional Ising model: an example. http://www.cmu.edu/biolphys/deserno/pdf/microcan
24. Huang, K.: Statistical Mechanics. John Wiley & Sons, New York (1987)
25. Brandão, F.G.S.L., Harrow, A.W., Horodecki, M.: Local random quantum circuits are approximate polynomial-designs. arXiv:1208.0692
26. Low, R.A.: Large deviation bounds for k-designs. In: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 465, pp. 3289–3308. The Royal Society, London (2009)
27. Ambainis, A., Emerson, J.: Quantum t-designs: t-wise independence in the quantum world. In: 49th Annual IEEE Symposium on Foundations of Computer Science, pp. 813-822 (2008)
28. Petz, D.: Quantum Information Theory and Quantum Statistics. Springer, Berlin-Heidelberg (2008)
29. Zyczkowski, K.: Rényi extrapolation of Shannon entropy. Open Sys. Inf. Dyn. **10**, 297–310 (2003)
30. Lieb, E.H., Robinson, D.W.: The finite group velocity of quantum spin systems. Commun. Math. Phys. **28**, 251–257 (1972)
31. Bravyi, S., Hastings, M.B., Verstraete, F.: Lieb-Robinson bounds and the generation of correlations and topological quantum order. Phys. Rev. Lett. **97**, 050401 (2006)
32. Masanes, L.: Area law for the entropy of low-energy states. Phys. Rev. A **80**, 052104 (2009)
33. Deutsch, J.M.: Quantum statistical mechanics in a closed system. Phys. Rev. A **43**, 2046 (1991)
34. Srednicki, M.: Chaos and quantum thermalization. Phys. Rev. E **50**, 888 (1994)

Communicated by M. M. Wolf