Locality of Temperature

M. Kliesch,[1,*] C. Gogolin,[1] M. J. Kastoryano,[1] A. Riera,[1,2] and J. Eisert[1]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
Am Mühlenberg 1, 14476 Potsdam-Golm, Germany*

This work is concerned with thermal quantum states of Hamiltonians on spin- and fermionic-lattice systems with short-range interactions. We provide results leading to a local definition of temperature, thereby extending the notion of "intensivity of temperature" to interacting quantum models. More precisely, we derive a perturbation formula for thermal states. The influence of the perturbation is exactly given in terms of a generalized covariance. For this covariance, we prove exponential clustering of correlations above a universal critical temperature that upper bounds physical critical temperatures such as the Curie temperature. As a corollary, we obtain that above the critical temperature, thermal states are stable against distant Hamiltonian perturbations. Moreover, our results imply that above the critical temperature, local expectation values can be approximated efficiently in the error and the system size.

I. INTRODUCTION

The ongoing miniaturization of devices, with structures reaching the nanoscale, has lead to the development of extremely small thermometers [1,2], some of which are so small that they can only be read out with powerful electron microscopes [3]. Even small thermal machines working in the quantum regime have been suggested [4,5]. In order to understand the working of such devices, it is necessary to formulate a theory of statistical mechanics and thermodynamics at the microscopic and mesoscopic scales. A prerequisite for such a formulation is a good understanding of the limitations of the concept of temperature at small scales.

The problem with assigning locally a temperature to a small subsystem of a globally thermal system is the following: Interactions between the subsystem and its environment that generate correlations can lead to noticeable deviations of the state of the subsystem from a thermal state (see Fig. 1). Hence, given only a subsystem state, there is no canonical way to assign a temperature to the subsystem. We call this the locality-of-temperature problem.

The first steps toward a solution of the locality-of-temperature problem have been taken in Refs. [6–8], and more recently, within the mind-set of quantum information theory, in Ref. [9]. The general locality-of-temperature problem is, however, still open. In this work, we conclusively solve it for spin- and fermionic-lattice systems.

More precisely, we first show that the locality-of-temperature problem is equivalent to a decay of correlations measured by an averaged generalized covariance that precisely captures the response of expectation values to perturbations of the Hamiltonian. We expect the corresponding equality to be useful for applications beyond the scope of this article.

We then provide conditions under which the generalized covariance decays exponentially with the distance, including a detailed analysis of the preasymptotic, and of the finite-size regime. In particular, this exponential decay holds above a universal critical temperature that only depends on the "connectivity" of the underlying graph of the model and is an upper bound on physically relevant critical temperatures such as the Curie temperature.

While, in the low-temperature regime, quantum lattice models exhibit a great diversity of phases, many of which involve the emergence of long-range or topological order [10], in the high-temperature regime, exponential clustering
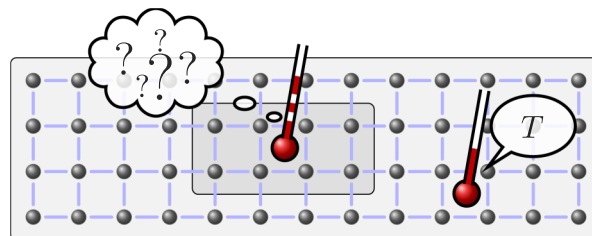

FIG. 1. The locality-of-temperature problem: Subsystems of thermal states are themselves, in general, not in a state with a locally well-defined temperature. Down to which length scale can temperature be an intensive quantity?

*[*]martin.kliesch@gmail.com

of correlations is expected. Our rigorous results help to delineate the boundary between these two regimes. They build upon and go significantly beyond previous results on the clustering of correlations in classical systems [11], for quantum gases [12], i.e., translation-invariant Hamiltonians in the continuum, and cubic lattices [13–15].

Mathematically, we significantly contribute to the problem of whether and under which precise conditions thermal quantum states are stable against distant Hamiltonian perturbations. Stability of thermal states is particularly relevant in the broader scheme of phase transitions in classical and quantum lattice models [14,16] as well as for the foundations of statistical mechanics and the equilibration and thermalization behavior of closed quantum systems [17–25]. In the light of the recent surge of interest in these topics, developing a better understanding of the properties of thermal states has become a timely issue.

A major obstacle to progress on some of the most interesting open questions in this context, such as equilibration time scales in closed quantum systems, is the limited set of mathematical tools available for exploiting the structure of locally interacting Hamiltonians [25]. Our results are among the first that explicitly exploit properties of local Hamiltonians, without being limited to very specific models.

For quantum Monte Carlo simulations [26], our results provide a guideline as to how large the finite system size has to be taken in order to be able to sample from the right partition function and, conversely, to identify observables that are best suited to detect long-range correlations.

In fact, our results are reminiscent of known statements about ground states. If a Hamiltonian has a unique ground state and is gapped, correlations in its ground state cluster exponentially and faraway regions become essentially uncorrelated. This clustering of correlations is rigorously proven using information-theory-inspired methods such as Lieb-Robinson bounds and quasiadiabatic continuation [27–29]. These rigorous results allow for certified algorithms that efficiently approximate ground states of gapped Hamiltonians on classical computers [30]. In the same spirit, we are able to show that an exponential decay of correlations renders thermal states locally efficiently simulatable.

The rest of this paper is structured as follows: In Sec. II, we formulate the precise setting and explain the main results and their implications. In Sec. III, we discuss connections to known results on phase transitions, thermalization in closed quantum systems, and matrix product operator approximations. Then, in Sec. IV, we discuss basic properties of the generalized covariance, explain how our results can be made applicable to finite-range k-body interactions, and state the results for fermionic lattices. We proceed with proving all theorems in Sec. V and conclude in Sec. VI. In the Appendix, we provide a detailed proof of two bounds on truncated cluster expansions, one of which is an important ingredient to the proof of clustering of correlations.

II. SETTING AND MAIN RESULTS

In this section, we introduce the setting, state the locality-of-temperature problem more formally, and state our results.

A. Perturbation formula for thermal states

As the first result, we state a perturbation formula, which is a general statement about the response of the expectation value of an observable in the thermal state, upon changes in the system Hamiltonian. It does not make any reference to the locality structure of the Hamiltonian but turns out to be especially useful when correlations between local observables decay rapidly with distance.

Throughout the paper, we assume the Hilbert space to be finite dimensional [31] and denote the thermal state, or Gibbs state, of a Hamiltonian H at inverse temperature β by

$$g(\beta) := \frac{e^{-\beta H}}{Z(\beta)}, \tag{1}$$

with $Z(\beta) := \mathrm{Tr}(e^{-\beta H})$ being the partition function. If we mean the thermal state or partition function of a different Hamiltonian H', we write $g[H'](\beta)$ or $Z[H'](\beta)$.

We measure correlations by the (generalized) covariance that we define for any two operators A and A', full-rank quantum state ρ, and parameter $\tau \in [0, 1]$ as

$$\mathrm{cov}_\rho^\tau(A, A') := \mathrm{Tr}(\rho^\tau A \rho^{1-\tau} A') - \mathrm{Tr}(\rho A)\mathrm{Tr}(\rho A'). \tag{2}$$

We discuss various properties of this covariance and generalizations to arbitrary-rank quantum states in Sec. IV A.

The generalized covariance appears naturally in our first theorem about the response of expectation values to perturbations. More precisely, when we are given an unperturbed Hamiltonian H_0 and a perturbed Hamiltonian H, then the difference of expectation values in the corresponding thermal states is captured by that covariance.

Theorem 1 (Perturbation formula).—Let H_0 and H be Hamiltonians acting on the same Hilbert space. For $s \in [0, 1]$, define the interpolating Hamiltonian by $H(s) := H_0 + s(H - H_0)$ and denote its thermal state by $g_s := g[H(s)]$. Then,

$$\mathrm{Tr}[A g_0(\beta)] - \mathrm{Tr}[A g(\beta)]$$
$$= \beta \int_0^1 d\tau \int_0^1 ds \, \mathrm{cov}_{g_s(\beta)}^\tau(H - H_0, A) \tag{3}$$

for any operator A.

The proof of the theorem, which is presented in Sec. V A, relies on the fundamental theorem of calculus and Duhamel's formula. We refer to the double integral over

the covariance in Eq. (3) as the averaged (generalized) covariance.

B. Spin-lattice systems

In the remainder of this work, we will be concerned with spin- and fermionic-lattice systems. We will only write out everything for spin systems and then later, in Secs. IV C and V C, explain the necessary modifications for fermionic systems. In the case of spin-lattice systems, the Hilbert space is given by $\mathcal{H} = \bigotimes_{x \in V} \mathcal{H}_x$, where V is called the vertex set and is assumed to be finite. To make the presentation more accessible, many of the following definitions are highlighted in Fig. 2. A local Hamiltonian with interaction (hyper)graph (V, E) is a sum

$$H = \sum_{\lambda \in E} h_\lambda \qquad (4)$$

of local Hamiltonian terms h_λ acting on $\mathcal{H}$. The (hyper) edge set E is the set of supports $\lambda = \mathrm{supp}(h_\lambda) \subset V$ of the local terms h_λ. For any subset of edges $F \subset E$, we denote by $H_F := \sum_{\lambda \in F} h_\lambda$ the Hamiltonian only containing the interactions in F, and for any subsystem $B \subset V$, we define the truncated Hamiltonian to be $H_{\upharpoonright B} := H_{E(B)}$, where $E(B) \subset \{\lambda \in E : \lambda \subset B\}$ is the restricted edge set and we take $H_{\upharpoonright B}$ to be an operator on the Hilbert space $\mathcal{H}_B := \bigotimes_{x \in B} \mathcal{H}_x$.

Given some subsystem $S \subset V$, there are two natural thermal states associated with it.

(i) $g_{\upharpoonright S}(\beta) := g[H_{\upharpoonright S}](\beta)$ denotes the thermal state of S alone, i.e., the thermal state of the truncated Hamiltonian $H_{\upharpoonright S}(\beta)$.

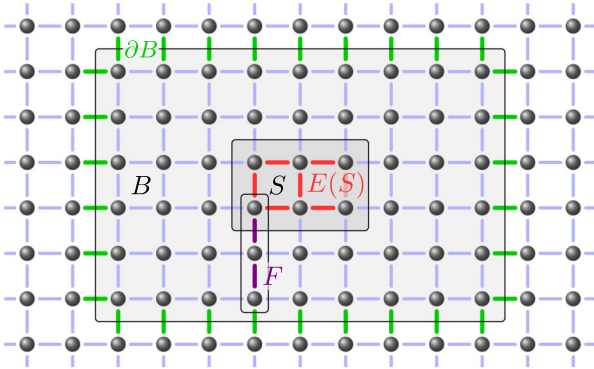(ii) $g^S(\beta) := \mathrm{Tr}_{S^c}[g(\beta)]$ denotes the full thermal state reduced to S.

FIG. 2. A 2D square lattice: The boxes indicate subsystems $S \subset B \subset V$. The edges in S are $E(S)$, boundary edges of B are ∂B, and F is a shortest path connecting S and ∂B; hence, $d(S, \partial B) = |F| = 2$. The set of edges $E(S)$ is an example for an animal of size $|E(S)| = 7$, while ∂B is not connected and hence not an animal.

For a noninteracting Hamiltonian, these two states coincide, but, in general, this is not the case due to correlations between S and its environment. This discrepancy raises the question of how to locally define temperature as an intensive quantity, i.e., the locality-of-temperature problem.

C. Locality of temperature

In order to locally assign a temperature to the subsystem $S \subset V$, it was suggested, e.g., in Ref. [9], to extend S by a buffer region and define the temperature of S via the thermal state of the Hamiltonian truncated outside the extended region B; see Fig. 2. The role of the buffer region B is to remove the boundary effects and the correlations with the rest of the system that are intuitively the reason for the locality-of-temperature problem. Nevertheless, it is not obvious how these correlations should be quantified and how large this buffer region needs to be. We will see shortly that Theorem 1 answers these questions.

By $\partial B \subset E$, we denote the set of boundary edges of B, i.e., the edges having overlap with both B and its complement $B^c := V \backslash B$. Then, by choosing $H_0 = H - H_{\partial B}$ in Theorem 1, using that $g_0 = g_{\upharpoonright B} \otimes g_{\upharpoonright B^c}$, and tracing over B^c, we obtain the following corollary (see also Fig. 3).

Corollary 1 (Truncation formula).—Let H be a local Hamiltonian, let $B \subset V$ be a subsystem, and denote the corresponding boundary Hamiltonian by $H_{\partial B}$ and the interpolating Hamiltonian by $H(s) := H - (1 - s)H_{\partial B}$ with its thermal state $g_s := g[H(s)]$. Then, for any operator $A = A_B \otimes \mathbb{1}_{B^c}$ supported on B,

$$\mathrm{Tr}[A_B g_{\upharpoonright B}(\beta)] - \mathrm{Tr}[A g(\beta)]$$
$$= \beta \int_0^1 d\tau \int_0^1 ds \, \mathrm{cov}^\tau_{g_s(\beta)}(H_{\partial B}, A). \qquad (5)$$

Now, we choose $S \subset B \subset V$ (see Fig. 2). If, for a given inverse temperature β, correlations over the distance between S and ∂B are negligible, then the corollary clearly implies that

$$\mathrm{Tr}[A g(\beta)] \approx \mathrm{Tr}[A_B g_{\upharpoonright B}(\beta)] \qquad (6)$$

for any observable $A_B = A_S \otimes \mathbb{1}_{B \backslash S}$ on S. Also note that such an approximate equality does not hold whenever average correlations over lengths exceeding the distance between S and ∂B are non-negligible.

Hence, we have the following equivalence for the temperature defined via thermal states.

Implication 1 (Locality of temperature).—Temperature is intensive on a given length scale if and only if correlations (measured by the averaged generalized covariance) are negligible compared to $1/\beta$ on that length scale.

In order to fully exploit Corollary 1 it is necessary to bound the generalized covariance, which we will do for high temperatures in the next section.

D. Clustering of correlations at high temperatures

For small temperatures, correlations can be arbitrarily long ranged, as is, e.g., the case for the ferromagnetic Ising model in two or higher dimensions below the Curie temperature. On the other hand, above a universal critical temperature, depending only on a local property of the interaction graph, correlations cluster exponentially, as we will see next. Given the combinatorial nature of parts of the arguments leading to this result, we need additional notation related to edges and vertices of the lattice. Most of the following definitions can be understood intuitively, as is shown in Fig. 2.

We say that two subsystems X, $Y \subset V$ overlap if $X \cap Y \neq \emptyset$, a set $X \subset V$ and a set $F \subset E$ overlap if F contains an edge that overlaps with X, and two sets F, $F' \subset E$ overlap if F overlaps with any of the edges in F'. A subset of edges $F \subset E$ connects X and Y if F contains a sequence of pairwise overlapping edges such that the first overlaps with X and the last overlaps with Y and similarly for the case where X and/or Y are just vertices.

The graph distance on V, and also the induced distance on subsets of V, are denoted by d. The distance $d(X, F)$ of a subset $X \subset V$ and a subset $F \subset E$ is 0 if X and F overlap and otherwise equal to the size of the smallest subset of E that connects X and F. Sometimes, we denote the support of an operator by the operator itself, e.g., for two operators A and A', their distance is $d(A, A') := d(\text{supp}\, A, \text{supp}\, A')$ and $\partial A \subset E$ are the edges across the boundary of $\text{supp}(A)$.

A subset of edges $F \subset E$ that connects all pairs of its elements λ, $\lambda' \in F$ is called connected. Such a connected set F is also called an (edge) animal. The size $|F|$ of an animal F is given by the number of edges contained in F. The results presented here apply to Hamiltonians with interaction graphs (V, E) whose number a_m of lattice animals of size m containing some fixed edge is exponentially bounded. With

$$a_m := \sup_{\lambda \in E} |\{F \subset E \text{ connected}: \lambda \in F, |F| = m\}|, \quad (7)$$

the growth constant α is the smallest constant satisfying

$$a_m \leq \alpha^m. \quad (8)$$

For example, the growth constant of a D-dimensional cubic lattice can be bounded as $\alpha \leq 2De$ (Lemma 2 in Ref. [32]), where e is Euler's number. Moreover, α is finite for any regular lattice [33]. Upper bounds to growth constants for so-called spread-out graphs [32] render our results applicable for the case of bounded-range two-body interactions. By a simple embedding argument, one can also bound the growth constant for the case of local k-body interactions on a regular lattice, which we explain in Sec. IV B in detail.

For any operator A and $p \in [1, \infty]$, we denote by $\|A\|_p$ its Schatten p norm; e.g., $\|A\|_\infty$ is the operator norm and

$\|A\|_1$ is the trace norm of A. We call $J := \max_{\lambda \in E} \|h_\lambda\|_\infty$ the local interaction strength of a local Hamiltonian, as given in Eq. (4).

We are able to provide a universal inverse critical temperature β^*, which is, in particular, independent of the system size, below which correlations decay exponentially with a thermal correlation length $\xi(\beta)$.

Theorem 2 (Clustering of correlations at high temperatures).—Let $g(\beta)$ be the thermal state at inverse temperature β of a local Hamiltonian with finite interaction (hyper) graph (V, E) having growth constant α and local interaction strength J. Define the quantities

$$\beta^* := \ln[(1 + \sqrt{1 + 4/\alpha})/2]/(2J) \quad (9)$$

and

$$\xi(\beta) := |\ln[\alpha e^{2|\beta|J}(e^{2|\beta|J} - 1)]|^{-1}. \quad (10)$$

Then, for every $|\beta| < \beta^*$, parameter $\tau \in [0, 1]$, every two operators A and B with $d(A, B) \geq L_0(\beta, a)$ [given in Eq. (50)], and $a := \min\{|\partial A|, |\partial B|\}$,

$$|\text{cov}^\tau_{g(\beta)}(A, B)| \leq \frac{4a\|A\|_\infty \|B\|_\infty}{\ln(3)(1 - e^{-1/\xi(\beta)})} e^{-d(A,B)/\xi(\beta)}. \quad (11)$$

The proof is given in Sec. V B.

In the following sections, we outline some of the applications of Theorem 2.

E. Universal locality and stability at high temperatures

If one is interested in the state $g^S(\beta)$ of some subsystem S, then one can truncate the Hamiltonian to S extended by some buffer region and obtain the approximation via the thermal state of the truncated Hamiltonian. The following theorem implies that the approximation error is exponentially small in the width of the buffer region.

For any operator ρ, we denote its reduction to a subsystem $S \subset V$ by $\rho^S := \text{Tr}_{S^c}[\rho]$ and note that

$$\|\rho^S\|_1 = \sup\{|\text{Tr}[A\rho]| : \text{supp}(A) = S, \|A\|_\infty = 1\}. \quad (12)$$

Then, as a consequence of Corollary 1 and Theorem 2, we obtain the following corollary.

Corollary 2 (Universal locality at high temperatures).— Let H be a Hamiltonian satisfying the conditions of Theorem 2, let $|\beta| < \beta^*$, and let $S \subset B \subset V$ be subsystems with $d(S, \partial B) \geq L_0(\beta, |\partial S|)$. Then,

$$\|g^S(\beta) - g^S_{\restriction B}(\beta)\|_1 \leq \frac{v|\beta|J}{1 - e^{-1/\xi(\beta)}} e^{-d(S,\partial B)/\xi(\beta)}, \quad (13)$$

where $g^S_{\restriction B}$ denotes the thermal state of B reduced to S and $v := 4|\partial S||\partial B|/\ln(3)$.
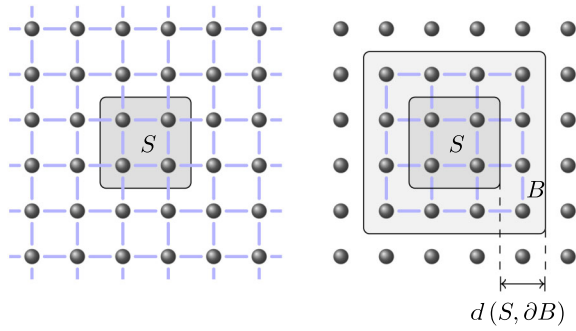
FIG. 3.   The truncation from Corollary 2 and Implication 3: For $\beta < \beta^*$ and $d(S, \partial B) \ll \xi(\beta)$, Corollary 2 implies that $g^S(\beta)$, depicted on the left, and $g^S_{\restriction B}(\beta)$, depicted on the right, are approximately equal.

Similarly, as a corollary of Theorems 1 and 2, we obtain the following implication.

Implication 2 (Stability).—Below the critical inverse temperature β^* [from Eq. (9)], thermal states of local Hamiltonians are exponentially stable against distant, locally bounded perturbations.

F. Efficient approximation

Corollary 2 on the universal locality of thermal states also has the following complexity-theoretic consequence.

Implication 3 (Efficient approximation).—For $|\beta| < \beta^*$, local expectation values can be approximated with a computational cost independent of the system size and bounded polynomially in the reciprocal error.

In this sense, the error bound (see Fig. 4) of Corollary 2 is reminiscent of the quasilocality of dynamics, as, e.g., presented in Ref. [34], which is a consequence of
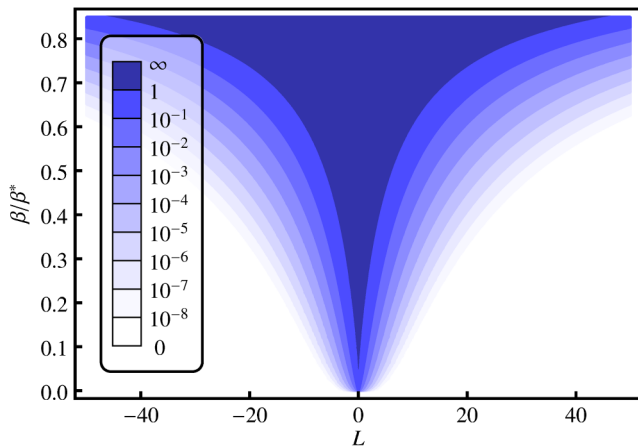

FIG. 4.   One can obtain slightly tighter error bounds in Corollaries 2 and 5 by directly using Eq. (47). The plot shows this bound on the approximation error $\|g^S(\beta) - g^S_{\restriction B}(\beta)\|_1$ for the case of S being a single site on a 2D square lattice as a function of the inverse temperature β in units of the critical temperature and the width of the buffer region L. This bound can be seen as an imaginary-time Lieb-Robinson "cone" with diverging width as $\beta \to \beta^*$.

Lieb-Robinson bounds [35,36]. The quasilocality theorem [34] allows for an approximation of time-evolved local observables by truncating the Hamiltonian in the time-evolution operator at a distance $L > 0$ far away from the spacetime cone of the observable's support and has an approximation error that is exponentially small in L.

G. Fermions

In Ref. [37], it was shown for fermionic systems that two-point functions of observables that are odd polynomials in the fermionic operators decay exponentially with a correlation length proportional to the inverse temperature. Here, we obtain an exponential decay of the covariance above the critical temperature for all operators.

Observation 1 (Fermions).—All results also hold for locally interacting fermions on a lattice. See Theorem 4 and Corollaries 4 and 5 in Sec. IV C for the precise statements.

III. RELATIONS TO KNOWN RESULTS

In this section, we discuss the critical temperature from the clustering theorem, the connection of this work to concepts related to thermalization, and approximations of thermal states with so-called matrix product operators; as a last point, we briefly mention similarities with local topological quantum order.

A. Critical temperatures and phase transitions

Our results show that the quantity β^*, as defined in Eq. (9), provides a potentially coarse but universal and completely general upper bound on physical critical temperatures like the Curie temperature. For the ferromagnetic two-dimensional isotropic Ising model without an external field, our bound yields, for example, $1/(\beta^* J) = 2/\ln[(1 + \sqrt{1 + 1/e})/2] \approx 24.58$, whereas the phase transition between the disordered paramagnetic and the ordered ferromagnetic phases is known to really happen at $1/(\beta_c J) = 2/\ln(1 + \sqrt{2}) \approx 2.27$ [16]. Our universal bound is about an order of magnitude higher than the actual value for this example. To put this discrepancy into perspective, it is worth pointing out that it is a very difficult task to estimate physical critical temperatures—numerically or analytically. In fact, analytic expressions for critical temperatures or even just bounds on their values are known only for very few models.

One of the few known general statements is the Mermin-Wagner-Hohenberg theorem [38]. It states that in certain low-dimensional systems with short-range interactions, there cannot be any phase transition involving the spontaneous breaking of a continuous symmetry at any nonzero temperature. However, such systems can still have a low-temperature phase with quasi-long-range order characterized by power-law-like decaying correlations. Consequently, even for systems covered by the Mermin-Wagner-Hohenberg theorem, our Theorem 2 is nontrivial.

For example, it implies an upper bound on the critical temperature of the Kosterlitz-Thouless transition in the two-dimensional *XY* model [39].

In this work, we have concentrated on the general picture, but it seems likely that refinements of the methods employed and developed here can yield much tighter bounds on critical temperatures if more specific properties of a model are taken into account. At the same time, it remains an open problem to actually find a model with a phase transition with long-range order at the universal highest possible temperature.

B. Foundations of statistical mechanics

The recent years have seen a large number of numerical and experimental (see Ref. [21] for a review) as well as analytical investigations (see, for example, Refs. [17–19,22–24]) of equilibration and thermalization in closed quantum systems. In the focus of these works are the approach to equilibrium or properties of energy eigenstates. The current work complements this body of literature in that it shows fundamental properties of systems in thermal equilibrium. A feature that makes the current work unique is that, contrary to essentially all other works, the results derived here explicitly use the structure of locality interacting systems. (Noteworthy exceptions are Ref. [24] and, albeit in a very special setting, Ref. [18].)

The locality of thermal states is also of interest for recent results [24] on the dynamical thermalization of translation-invariant lattice models: Our Corollary 2 guarantees the existence of a "unique phase" [24] for all temperatures above our critical temperature. Hence, it implies that at sufficiently high temperatures, Theorems 1, 2, and 3 of Ref. [24] are applicable for any translation-invariant Hamiltonian.

There is also an interesting connection of our locality and stability results to the so-called eigenstate-thermalization hypothesis (ETH) [20,21]. The ETH essentially conjectures that the expectation values of certain physically relevant observables (for example, local ones) in energy eigenstates of sufficiently complex Hamiltonians should be very similar to the expectation values in thermal states with the same average energy. Corollary 2 and Implication 2 thus imply that the eigenstates of a Hamiltonian in the center of the spectrum (which correspond to high-temperature thermal states) must, if the Hamiltonian fulfills the ETH, also be locally stable against perturbations of the Hamiltonian. This insight could put constraints on the class of Hamiltonians that fulfills the ETH, provide new insights into the properties of their eigenstates, and open up new ways to test the ETH.

C. MPO approximation of thermal states

Matrix product operators (MPOs) are a certain class of operators that are tractable on classical computers for one-dimensional systems. Therefore, they play an important role in numerical simulations based on so-called tensor networks.

An important ingredient to our proof of Theorem 2 on the clustering of correlations will be a bound on a truncated cluster expansion (Lemma 1). The original result on the cluster expansion (Lemma 2 in the Appendix) is due to Hastings and was first used to approximate thermal states with inverse temperature below $2\beta^*$ by MPOs [40]. This approximation is summarized in the next theorem.

In one spatial dimension, this MPO approximation yields a tensor size bounded polynomially in the system size and the approximation error (see the subsequent corollary). In higher dimensions, however, the MPO approximation yields a tensor size bounded only subexponentially in the system size and is hence not computationally efficient, albeit exponentially cheaper than storing the full density matrix $g(\beta)$. In order to explain the MPO approximation in more detail, we start the discussion with a slightly nonstandard definition of MPOs.

Definition 1 (MPO).—Let $(b[x]^{(j)})_{j=1}^{d^2}$ be a basis for the operators on $\mathcal{H}_x$ and write an arbitrary operator A on $\mathcal{H}$ in the product basis as

$$A = \sum_{k \in [d^2]^V} A_k \bigotimes_{x \in V} b[x]^{(k_x)}, \tag{14}$$

with expansion coefficients $A_k \in \mathbb{C}$ and where $[d^2] := \{1, 2, \ldots, d^2\}$. If the A_k are of the form

$$A_k = \prod_{x \in V} a[x](k), \tag{15}$$

where every $a[x](k)$ only depends on at most r of the $|V|$ indices k_x, then A is called an MPO with tensor size d^{2r}.

Thermal states can be approximated by such MPOs. The following theorem is a consequence of Lemma 2, which we will prove in the Appendix along with Lemma 1.

Theorem 3 (MPO approximation of thermal states [40]).—Let $H = \sum_{\lambda \in E} h_\lambda$ be a local Hamiltonian with finite interaction graph (V, E) having a growth constant α and local interaction strength $J = \max_{\lambda \in E} \|h_\lambda\|_\infty$, and define $b(\beta J) := \alpha e^{|\beta J|}(e^{|\beta J|} - 1)$. Moreover, let β be small enough such that $b(\beta J) < 1$. Then, for each $L \in \mathbb{Z}^+$, there exists a self-adjoint MPO $\rho(\beta, L)$ [given in Eq. (A3)] with tensor size $d^{2N(L)}$, where

$$N(L) := \sup_{x_0 \in V} |\{x \in V : d(x, x_0) < L\}| \tag{16}$$

is the number of vertices within a distance less than L. The approximation error is bounded as

$$\|g(\beta) - \rho(\beta, L)\|_1 \leq \exp\left(|E| \frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1; \tag{17}$$

i.e., for fixed $|\beta J| < b^{-1}(1)$, the trace norm difference scales as $O(|E|e^{-|\ln[b(\beta J)]|L})$ for large enough L.

In particular, the above theorem implies the following corollary.

Corollary 3 (Bound on the tensor size).—Let D be the spatial dimension of the Hamiltonian's interaction graph (V, E), let $n := |E|$ be the system size, and let $\beta < 2\beta^*$ with β^* from Eq. (9). Then, the MPO approximation in Theorem 3 gives rise to a tensor size of the MPO $\rho(\beta, L)$ scaling as

$$\log_d(\text{tensor size}) \leq O[\ln(Cn/\epsilon)^D], \qquad (18)$$

with some β-dependent constant C. In particular, for $D = 1$, the bound on the tensor size scales polynomially with n/ϵ.

Let us consider a one-dimensional system and suppose we are explicitly given the MPO tensors $a'[x]$ [see Eq. (15)] of an approximation to a state ρ and, similarly, an observable A of MPO form with MPO tensors $a[x]$. If the tensor sizes of both MPOs scale at most polynomially in the system size, then one can compute the corresponding approximation to the expectation value $\text{Tr}(\rho A)$ with a computational cost scaling polynomially in the system size. Thus, for instance, global product observables can be approximated efficiently, which is not guaranteed by our Implication 3. The problem with the MPO approximation, however, is that Theorem 3 only guarantees the existence of the MPO tensors but it is not obvious how they can be computed (efficiently).

Proof of Corollary 3.—The condition $\beta < 2\beta^*$ is equivalent to $b(\beta J) < 1$. Let us denote the bound to the approximation error in Eq. (17) by ϵ. Note that the upper bound in Eq. (17) satisfies

$$\epsilon := \exp\left(|E|\frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1 \leq Cn\, b(\beta J)^L \qquad (19)$$

for distances L being at least logarithmically large in $n = |E|$ and some β-dependent constant C. Then, the distance L necessary to reach ϵ must asymptotically be at least as large as

$$L \geq \frac{\ln(Cn/\epsilon)}{|\ln[b(\beta J)]|}. \qquad (20)$$

Bounding $N(L)$ in terms of the spatial dimension D as $N(L) \leq ML^D$ with some constant M yields a tensor size bounded as

$$\log_d(\text{tensor size}) \leq 2M\left(\frac{\ln(Cn/\epsilon)}{\ln[1/b(\beta J)]}\right)^D. \qquad (21)$$

∎

D. Local topological quantum order

It is worth mentioning that Corollary 2 and Implication 2 are very reminiscent of the local topological quantum order condition for open quantum systems introduced in Ref. [41] and the results on the local stability of stationary states of local Liouvillians in Ref. [42]. A slightly different family of local topological quantum order conditions for closed quantum systems [41–44] has played a very important role in the theory of locally stable (topological) lattice systems and for rigorous proofs of entropic area laws. Corollary 2 similarly characterizes the regime where local perturbations cannot drive any thermal phase transition.

IV. DETAILS

In this section, we first discuss the generalized covariance and then provide details concerning the applicability of our results to Hamiltonians with k-body interactions. Finally, we justify Observation 1 by stating the fermionic versions of our results.

A. The generalized covariance

The generalized covariance defined in Eq. (2), which depends on a parameter $\tau \in [0, 1]$, provides more information about the correlations between two observables than the standard covariance in a similar way as the class of Rényi entropies characterizes more completely the entanglement properties of a state than simply the von Neumann entropy [45]. While it occurs quite naturally in the perturbation formula of Theorem 1, other possible applications are to be explored. Here, we discuss possible generalizations of the generalized covariance to operators of arbitrary rank, show that for operators A and A' they are always bounded by $\|A\|_\infty\|A'\|_\infty$, and comment on convexity and a symmetrized version of the generalized covariance.

A definition of the generalized covariance for states of arbitrary rank is not relevant for this work because for nonzero temperature, thermal states are full-rank operators. However, the discussion of possible generalizations also hints at the behavior of cov^τ at the end points of the unit interval. On the open interval $\tau \in]0, 1[$, it is natural to simply keep the definition from Eq. (2). There are two natural ways to define ρ^0: Either as $\rho^0 := \mathbb{1}$ or as $\rho^{0+} := \lim_{\tau \to 0}\rho^\tau$, where ρ^{0+} turns out to be the projector onto the image of the operator ρ. For each end point $\tau = 0$ and $\tau = 1$, there are hence two natural ways to define cov^τ, either such that the generalized covariance is continuous or such that $\text{cov}_\rho^0(A, A') = \text{cov}_\rho(A', A)$ and $\text{cov}_\rho^1(A, A') = \text{cov}_\rho(A, A')$, where

$$\text{cov}_\rho(A, A') := \text{Tr}(\rho A A') - \text{Tr}(\rho A)\text{Tr}(\rho A') \qquad (22)$$

defines the standard covariance.

Note that for product states and operators with disjoint support, all versions of the generalized covariance vanish. Moreover, for pure states, the continuous version of the generalized covariance vanishes also, meaning that classical correlations are needed to yield a nonzero value.

Next, we show that the generalized covariance is always bounded as

$$| \operatorname{cov}_\rho^\tau(A, A')| \le \|A\|_\infty \|A'\|_\infty, \tag{23}$$

irrespective of which definitions are chosen for cov^0 and cov^1. We consider a state ρ and define $\bar{A} := A - \operatorname{Tr}(\rho A)$. Then,

$$\operatorname{cov}_\rho^\tau(A, A') = \operatorname{Tr}(\rho^\tau \bar{A} \rho^{1-\tau} A'). \tag{24}$$

Hölder's inequality generalized to several operators and the fact that $\|X\|_p = \||X|^p\|_1^{1/p}$ then imply that

$$| \operatorname{cov}_\rho^\tau(A, A')| \le \|\rho^\tau\|_{1/\tau} \|\bar{A}\|_\infty \|\rho^\tau\|_{1/(1-\tau)} \|A'\|_\infty \tag{25}$$

$$= \|\bar{A}\|_\infty \|A'\|_\infty, \tag{26}$$

and, by noting that $\|\bar{A}\|_\infty = \|A\|_\infty$, the bound (23) is proven for the continuous version of the generalized covariance. For the noncontinuous versions, the bound follows similarly.

The variance $\operatorname{cov}_\rho^\tau(A, A)$ induced by the continuous version of the covariance is convex in τ, as can be seen by writing out ρ in its eigenbasis. As one can change the sign of $\operatorname{cov}_\rho^\tau(A, A')$ by just changing the sign of A', the generalized covariance is not convex in τ. But, it might be that its magnitude $| \operatorname{cov}_\rho^\tau(A, A')|$ is convex, which is unclear. If this were the case, it would be enough to prove the clustering Theorem 2 only for the end points $\tau \in \{0, 1\}$, and hence the proof could be significantly simplified.

Similarly, as there is a symmetrized version of the standard covariance, one can also symmetrize the generalized covariance with respect to the two operators. Because of the cyclicity of the trace, the generalized covariance satisfies the symmetry property

$$\operatorname{cov}_\rho^\tau(A, A') = \operatorname{cov}_\rho^{1-\tau}(A', A). \tag{27}$$

Hence, one can define the symmetrized version of the generalized covariance as follows:

$$\overline{\operatorname{cov}}_\rho^\tau(A, A') := \frac{1}{2}[\operatorname{cov}_\rho^\tau(A, A') + \operatorname{cov}_\rho^\tau(A', A)]. \tag{28}$$

Our results can also be phrased in terms of this symmetrized version, since the averaged generalized covariance in the perturbation formula of Theorem 1 can easily be rewritten in terms of $\overline{\operatorname{cov}}$, and a bound analogous to the clustering of Theorem 2 holds also for the symmetrized quantity.

B. Bound on the growth constant for local *k*-body interactions

In this section, we show that regular hyperlattices also have a finite growth constant, which renders our results applicable to Hamiltonians with local *k*-body interactions.

In the case of *k*-body interactions, the Hamiltonian is again a sum of local terms h_λ whose supports are hyperedges $\lambda = \operatorname{supp}(h_\lambda) \subset V$ with $|\lambda| \le k$. As before, V denotes the vertex set and E the set of hyperedges.

We assume that the interaction hypergraph (V, E) is a regular hyperlattice, i.e., that it can be embedded into a regular hypercubic lattice of a certain dimension D with hyperedges of hypercubic form. Let us denote by R the edge length of the resulting hypercubes. Note that such an embedding is, in general, not unique and changes both the number of terms in the Hamiltonian and the local interaction strength of H. Moreover, the grouping changes the values of the metric d in our results.

In order to find an exponential upper bound to the number a_m of hyperanimals composed of m hypercubes, let us define a spread-out graph of range R as the graph with the edge set consisting of all pairs $\{x, y\}$ with $0 < \|x - y\|_\infty \le R$ and $x, y \in \mathbb{Z}^D$ (see Ref. [32]). Notice that as any hypercube is uniquely specified by the coordinates of its "lower left corner," any hyperanimal of size m corresponds to a lattice animal of size $m - 1$ and range R in the spread-out graph. It follows from Lemma 2 in Ref. [32] that $a_m \le (Ke)^m$ with $K = (2R + 1)^D - 1$ being the coordination number. Hence, the hyperlattice has a growth constant bounded by $\alpha \le [(2R + 1)^D - 1]e$.

The bound obtained is, for most models, far from optimal, in particular, in situations where the supports of the local Hamiltonian terms are very different from hypercubes. For such cases, one can derive tighter but more specific bounds from known results about lattice animals in a similar way.

C. Fermionic versions of the main results

To make Observation 1 about fermions precise, we introduce the setting of interacting fermions on lattices. For each site $x \in V$, the corresponding fermionic operators, i.e., the creation and annihilation operators $f_x^\dagger$ and f_x, act on the fermionic Fock space and satisfy

$$\{f_x, f_y^\dagger\} = \delta_{x,y} \mathbb{1}, \tag{29}$$

$$\{f_x, f_y\} = 0, \tag{30}$$

where $\{A, B\} := AB + BA$ is the anticommutator. For such systems, all operators can be given in terms of polynomials in the fermionic operators. A monomial of fermionic operators is called even (odd) if it can be written as a product of an even (odd) number of fermionic operators f_x and $f_y^\dagger$. A polynomial of fermionic operators is called even (odd) if it can be written as a linear combination of only

even (odd) monomials, and an operator is called even (odd) if it can be written as an even (odd) polynomial of fermionic operators. According to the fermion-number-parity super-selection rule, only operators that are even polynomials in the fermionic operators are physical observables and Hamiltonians.

As with spin-lattice systems, we have again a finite interaction graph (V, E); however, the support of an operator is now to be understood in the picture of the second quantization as follows: The support of any operator A being a polynomial in the fermionic operators is the set of vertices of the fermionic operators that occur in the polynomial. Correspondingly, we denote the algebra of the even operators supported on a region $X \subset V$ by $\mathcal{G}_X$ and denote $\mathcal{G} := \mathcal{G}_V$ for short. The Hamiltonian of a fermionic-lattice system is of the form

$$H = \sum_{\lambda \in E} h_\lambda \qquad (31)$$

with $h_\lambda \in \mathcal{G}_\lambda$. For $B \subset V$, the truncated Hamiltonian $H_{\restriction B}$ is similarly the sum only over the edges contained in B. As for spin systems, $H_{\partial B}$ is the sum over the boundary edges of B.

Theorem 1 also holds for such fermionic-lattice systems, and we can prove statements analogous to Corollary 1, Theorem 2, and Corollary 2. Hence, all implications stated in Sec. II also hold. All proofs are presented in Sec. V C.

Corollary 4 (Fermionic truncation formula).—Let $H = \sum_{\lambda \in E} h_\lambda$ be a fermionic local Hamiltonian with local terms $h_\lambda \in \mathcal{G}$, let $B \subset V$ be a subsystem, and let the interpolating Hamiltonian by $H(s) := H - (1 - s)H_{\partial B}$ with its thermal state $g_s := g[H(s)]$. Then, for any operator A with support $\mathrm{supp}(A) \subset B$,

$$\mathrm{Tr}\{Ag[H_{\restriction B}](\beta)\} - \mathrm{Tr}[Ag(\beta)]$$
$$= \beta \int_0^1 d\tau \int_0^1 ds \, \mathrm{cov}_{g_s(\beta)}^\tau (H_{\partial B}, A). \qquad (32)$$

Theorem 4 (Clustering of correlations in fermionic systems).—Let $g(\beta)$ be the thermal state at inverse temperature β of a local fermionic Hamiltonian $H = \sum_{\lambda \in E} h_\lambda$ with finite interaction graph (V, E) having growth constant α, local terms $h_\lambda \in \mathcal{G}$, and local interaction strength J. Define the functions β^*, ξ, and L_0 as in Eqs. (9), (10), and (50). Then, for every $|\beta| < \beta^*$, $\tau \in [0, 1]$, and every two operators A and B with $d(A, B) \geq L_0(\beta, a)$, where $a := \min\{|\partial A|, |\partial B|\}$,

$$|\mathrm{cov}_{g(\beta)}^\tau (A, B)| \leq \frac{4a\|A\|_\infty \|B\|_\infty}{\ln(3)(1 - e^{-1/\xi(\beta)})} e^{-d(A,B)/\xi(\beta)}. \qquad (33)$$

Corollary 5 (Locality of fermionic thermal states).—Let H be a Hamiltonian satisfying the conditions of Theorem 4, let $|\beta| < \beta^*$, and let $S \subset B \subset V$ be subsystems with $d(S, \partial B) \geq L_0(\beta, |\partial S|)$. Then,

$$\|g^S(\beta) - g^S[H_{\restriction B}](\beta)\|_1 \leq \frac{v|\beta|J}{1 - e^{-1/\xi(\beta)}} e^{-d(S, \partial B)/\xi(\beta)}, \qquad (34)$$

where $v = 4|\partial S||\partial B|/\ln(3)$.

V. PROOFS

We start this section with the proofs of Theorems 1 and 2. One important stepping stone for the proof of the latter is a tailored version of a bound on a truncated cluster expansion (Lemma 1) from Ref. [40]. Both versions are proven in the Appendix. In the last part of the section, we prove the fermionic versions of our main results, Theorem 4 and Corollaries 4 and 5.

A. Proof of the perturbation formula (Theorem 1)

The two main ingredients in the proof of Theorem 1 are the fundamental theorem of calculus and Duhamel's formula. The generalized covariance appears as a natural measure of correlations.

Proof of Theorem 1.—Using the fundamental theorem of calculus, we obtain

$$\mathrm{Tr}[Ag_0(\beta)] - \mathrm{Tr}[Ag_1(\beta)] = -\mathrm{Tr}\left(A \int_0^1 \frac{d}{ds} \frac{e^{-\beta H(s)}}{Z_s(\beta)} ds\right)$$

with $Z_s := Z[H(s)]$. The derivative can be written as

$$\frac{d}{ds} \frac{e^{-\beta H(s)}}{Z_s(\beta)} = \frac{1}{Z_s(\beta)} \frac{d}{ds} e^{-\beta H(s)} - \frac{g_s(\beta)}{Z_s(\beta)} \mathrm{Tr}\left(\frac{d}{ds} e^{-\beta H(s)}\right).$$

After applying Duhamel's formula to both derivatives, i.e., using that

$$\frac{d}{ds} e^{-\beta H(s)} = -\beta \int_0^1 (e^{-\beta H(s)})^\tau \left(\frac{d}{ds} H(s)\right)(e^{-\beta H(s)})^{1-\tau} d\tau,$$

we obtain

$$\mathrm{Tr}(Ag_0) - \mathrm{Tr}(Ag) = -\beta \, \mathrm{Tr}\left(A \int_0^1 \int_0^1 \{-g_s^\tau (H - H_0)g_s^{1-\tau}\right.$$
$$\left. + g_s \mathrm{Tr}[g_s^\tau (H - H_0)g_s^{1-\tau}]\}d\tau ds\right). \qquad (35)$$

Together with the cyclicity of the trace and the definition of the generalized covariance in Eq. (2), the last equation finishes the proof. ∎

B. Proof of Theorem 2 on clustering of correlations

The proof of Theorem 2 builds on and develops further a cluster expansion of the power series of $e^{-\beta H}$ in terms of summands of the form

$$h(w) := h_{w_1} h_{w_2} \ldots h_{w_{|w|}}, \qquad (36)$$

where $w_j \in E$. For the sake of a compact presentation, we refer to edges as letters, to the edge set E as an alphabet, and call sequences of edges words. For any subalphabet $F \subset E$, we denote by $F^* := \bigcup_{l=0}^{\infty} F^l$ the set of words with letters in F and arbitrary length l, where the length $|w|$ of a word $w \in E^*$ is the total number of letters it contains. For two words $w, v \in E^*$, their concatenation is denoted by $w \circ v := (w_1, w_2, \ldots, w_{|w|}, v_1, v_2, \ldots, v_{|v|})$. We call a word $c \in E^*$ connected or a cluster if the set of letters in c is an animal, i.e., connected. So, clusters are connected sequences of edges where the edges can also occur multiple times, while animals are sets of edges without any order or repetition. A word v is called a subsequence of $w \in E^*$ if v can be obtained from w by omitting letters, i.e., if there is an increasing sequence $j_1 < j_2 < \cdots < j_{|v|}$ such that $v_i = w_{j_i}$. This will be denoted by $v \subset w$. A connected subsequence $c \subset w$ is called a maximal cluster of w if c is not a subsequence of any other connected subsequence of w. Importantly, for any word $w \in E^*$, one can permute its letters to a new word w' such that $h(w') = h(w)$, irrespective of the choice of the local terms h_λ and such that $w' = c_1 \circ c_2 \circ \cdots \circ c_k$ is a concatenation of maximal clusters $c_j \subset w$ of w. Note that this decomposition is unique up to the order of the c_j.

In the following, we will consider systems that are either $n = 2$ or $n = 4$ copies of the original system with Hilbert space $\mathcal{H}$. For any operator A on $\mathcal{H}$, we denote by $A^{(j)}$ the operator on $\mathcal{H}^{\otimes n}$ that acts as A on the jth copy, e.g., $A^{(2)} := \mathbb{1} \otimes A$ for $n = 2$. By $\mathcal{S}^{(i,j)}$, we denote the swap operator on $\mathcal{H}^{\otimes n}$ that swaps the ith and jth tensor factors, e.g., $\mathcal{S}^{1,2}|k_1, k_2, k_3, k_4\rangle = |k_2, k_1, k_3, k_4\rangle$ for $n = 4$. For $n = 2$, we write $\mathcal{S}$ instead of $\mathcal{S}^{1,2}$.

We can now state the subsequent lemma, which is a bound on a truncated cluster expansion that is based on a more general, but for our purposes not tight enough bound, used previously in Ref. [40]. (See Lemma 2 in the Appendix.) The lemma will play an important role in the subsequent proof of Theorem 2.

Lemma 1 (Truncated cluster expansion).—Let $\tau \in [0, 1]$ and $H = \sum_{\lambda \in E} h_\lambda$ be a local Hamiltonian on $\mathcal{H}$ with finite interaction graph (V, E) having growth constant α and local interaction strength $J = \max_{\lambda \in E} \|h_\lambda\|_\infty$. We denote by $\tilde{H}$ the Hamiltonian of two weighted copies with local terms $\tilde{h}_\lambda := \tau h_\lambda^{(1)} + (1 - \tau) h_\lambda^{(2)}$. Consider two operators A and B on $\mathcal{H}$, define $b(x) := \alpha e^{|x|}(e^{|x|} - 1)$, and let $|\beta|$ be small enough such that $b(\beta J) < 1$. For some set of edges $F \subset E$, let $\mathcal{C}_{\geq L}(F) \subset E^*$ be the set of words containing at least one cluster c that contains at least one letter of F and has size $|c| \geq L$ (see Fig. 5) and let us denote the corresponding truncated cluster expansion of $e^{-\beta \tilde{H}}$ by

$$\Omega[\tilde{H}](\beta) := \sum_{w \in \mathcal{C}_{\geq L}(F)} \frac{(-\beta)^{|w|}}{|w|!} \tilde{h}(w), \quad (37)$$
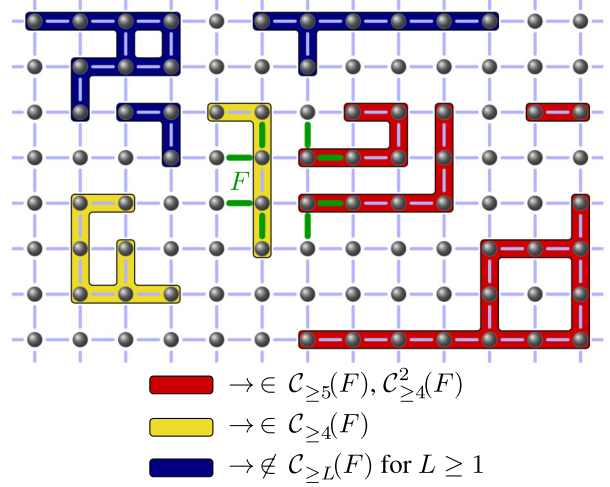

FIG. 5.   A 2D square lattice. Three different subalphabets are indicated: Words that contain all letters in those alphabets are members of different sets $\mathcal{C}_{\geq L}(F)$.

with $\tilde{h}(\omega) := \tilde{h}_{\omega_1} \tilde{h}_{\omega_2} \ldots \tilde{h}_{\omega_{|\omega|}}$. Then, for all $\tau \in [0, 1]$,

$$\frac{|\text{Tr}[\mathcal{S} A^{(1)} B^{(2)} \Omega[\tilde{H}](\beta)]|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq \exp\left(|F| \frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1. \quad (38)$$

We provide a detailed proof of this lemma in the Appendix. The terms resulting from the expansion of the exponential series of $e^{-\beta H}$ are classified according to whether they contain a cluster of size at least L that contains a letter from F. One can then show that there is a percolation transition at $\beta^* = b^{-1}(1)/(2J)$ such that for $|\beta| < \beta^*$, the contribution of long clusters is exponentially suppressed.

In the following proof of the exponential clustering, we will use the so-called swap trick: For any two operators A and B, it holds that

$$\text{Tr}(AB) = \text{Tr}[\mathcal{S}(A \otimes B)], \quad (39)$$

which can be checked by a straightforward calculation.

Proof of Theorem 2.—Fix some $\tau \in [0, 1]$. For any operator $A \colon \mathcal{H} \to \mathcal{H}$, we define $A^{(\pm)} := A \otimes \mathbb{1} \pm \mathbb{1} \otimes A$ and $\tilde{A}^{(+)} := \tau(A^{(1)} + A^{(2)}) + (1 - \tau)(A^{(3)} + A^{(4)})$.

As the first step, we write the covariance as

$$\text{cov}_\rho^\tau(A, B) = \frac{1}{2} 2 \text{Tr}[A^{(-)}(\rho^\tau \otimes \rho^\tau) B^{(-)}(\rho^{1-\tau} \otimes \rho^{1-\tau})].$$

Using the swap trick (39) yields (see Fig. 6)

$$\text{cov}_\rho^\tau(A, B) = \frac{1}{2} 2 \text{Tr}[\mathcal{S}^{1,3} \mathcal{S}^{2,4}(A^{(-)} \otimes B^{(-)}) \rho_4], \quad (40)$$

where $\rho_4 := \rho^\tau \otimes \rho^\tau \otimes \rho^{1-\tau} \otimes \rho^{1-\tau}$. For the case $\rho = g(\beta)$, the operator ρ_4 turns out to be

$$\rho_4 = \frac{e^{-\beta \tilde{H}^{(+)}}}{Z(\beta)^2}. \tag{41}$$

Writing out ρ_4 as a power series yields

$$\mathrm{cov}^\tau_{g(\beta)}(A, B) = \frac{1}{2Z(\beta)^2} \sum_{w \in E^*} \frac{(-\beta)^{|w|}}{|w|!} t(w) \tag{42}$$

with

$$t(w) := \mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}(A^{(-)} \otimes B^{(-)})\tilde{h}^{(+)}(w)] \tag{43}$$

and $\tilde{h}^{(+)}(w) := \tilde{h}^{(+)}_{w_1} \tilde{h}^{(+)}_{w_2} \ldots \tilde{h}^{(+)}_{w_{|w|}}$. Next, we argue that $t(w)$ vanishes whenever w does not contain a cluster connecting the supports of A and B. Without loss of generality, we assume that $|\partial A| \leq |\partial B|$ and consider $\mathcal{C}_{\geq L}(\partial A)^c = E^* \backslash \mathcal{C}_{\geq L}(\partial A)$, the set of words that do not contain a cluster containing an edge in ∂A of size $L := d(A, B)$ or larger. The set $\mathcal{C}_{\geq L}(\partial A)^c$ hence contains no words with clusters that connect $\mathrm{supp}(A)$ and $\mathrm{supp}(B)$. Any word $w \in \mathcal{C}_{\geq L}(\partial A)^c$ can be replaced by a concatenation of two words w_A and w_B such that $\tilde{h}^{(+)}(w) = \tilde{h}^{(+)}(w_A)\tilde{h}^{(+)}(w_B)$, where w_A contains all maximal clusters of w that overlap with $\mathrm{supp}(A)$ and w_B all other maximal clusters of w. The operators $\tilde{h}^{(+)}(w_A)$ and $\mathbb{1} \otimes \mathbb{1} \otimes B^{(-)} =: \hat{B}$, and $\tilde{h}^{(+)}(w_B)$ and $A^{(-)} \otimes \mathbb{1} \otimes \mathbb{1} =: \hat{A}$, then have disjoint supports, respectively, and the trace in Eq. (43) factorizes into a product of two traces, one over the subsystem $X := \mathrm{supp}(\hat{A}) \cup \mathrm{supp}[\tilde{h}^{(+)}(w_A)]$ and the other over the rest of the system. It turns out that both vanish: By using the symmetries $\hat{A} = -\mathcal{S}^{1,2}\hat{A}\mathcal{S}^{1,2}$, $\tilde{h}^{(+)}(w_A) = \mathcal{S}^{1,2}\mathcal{S}^{3,4}\tilde{h}^{(+)}(w_A)\mathcal{S}^{3,4}\mathcal{S}^{1,2}$, $\hat{A}\mathcal{S}^{3,4} = \mathcal{S}^{3,4}\hat{A}$, and that $(\mathcal{S}^{i,j})^2 = \mathbb{1}$, one can show, e.g., that

$$\mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\tilde{h}^{(+)}(w_A)] = -\mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\tilde{h}^{(+)}(w_A)]. \tag{44}$$

This equation implies that for every $w \in \mathcal{C}_{\geq L}(\partial A)^c$,

$$t(w) \propto \mathrm{Tr}[\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\tilde{h}^{(+)}(w_A)] = 0. \tag{45}$$

Together with Eq. (42), realizing that $Z(\beta)^2 = Z[H^{(+)}](\beta)$, and using the notation from Eq. (37) with $F = \partial A$ and $L = d(A, B)$, it follows that

$$\mathrm{cov}^\tau_{g(\beta)}(A, B) = \mathrm{Tr}\left(\frac{\mathcal{S}^{1,3}\mathcal{S}^{2,4}\hat{A}\,\hat{B}}{2Z(\beta)^2}\Omega[\tilde{H}^{(+)}](\beta)\right). \tag{46}$$

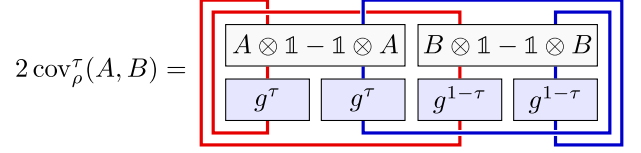After applying Lemma 1 and using that $\|\hat{A}\|_\infty \leq 2\|A\|_\infty$, and similarly for B, we obtain

FIG. 6. The "multiple swap trick": Eq. (40) as a tensor network.

$$\frac{|\mathrm{cov}^\tau_{g(\beta)}(A, B)|}{\|A\|_\infty \|B\|_\infty} \leq 2(e^{|\partial A|b(2\beta J)^L/[1-b(2\beta J)]} - 1). \tag{47}$$

The fact that the condition $\beta < \beta^*$ is equivalent to $b(2\beta J) < 1$ implies that $b(2\beta J)^L$ decays exponentially with L. In order to obtain the desired exponential bound (11), we apply the bound $\forall x \in [0, x_0]$: $\exp(x) - 1 \leq x(e^{x_0} - 1)/x_0$ with the choice $x_0 = \ln(3)$. In order to have $|\partial A|\frac{b(2\beta J)^L}{1-b(2\beta J)} \leq \ln(3)$, we impose

$$L \geq \left| \ln\left(\frac{|\partial A|}{\ln(3)[1 - b(2\beta J)]}\right) / \ln(2\beta J) \right| \tag{48}$$

$$= \xi(\beta)|\ln[\ln(3)(1 - e^{-1/\xi(\beta)})/|\partial A|]| \tag{49}$$

$$=: L_0(\beta, |\partial A|). \tag{50}$$

This inequality guarantees the exponential bound (11) and finishes the proof. ∎

C. Proofs of the fermionic versions of the main results

In order to also establish our main results for fermionic systems, we go through the proofs for spin systems and discuss the necessary modifications.

Proof of Corollary 4.—In Theorem 1, we choose $H_0 = H - H_{\partial B}$. As the local terms are all in $\mathcal{G}$, we have that the thermal state of H_0 factorizes, i.e., $g_0 = g[H_{\restriction B}]g[H_{\restriction B^c}]$. After tracing over B^c, the statement follows. ∎

Proof of Theorem 4.—We use the same tensor copy trick as in the proof of Theorem 2. Equation (40) still holds in the fermionic setting. Note that the Hilbert space over which the trace is performed in Eq. (40) is not the Fock space of a system of 4 times the number of modes but the tensor product of four identical fermionic Fock spaces with the canonical inner product. This Hilbert space can be interpreted as that of a system of four types of fermionic particles that are each mutually indistinguishable and subject to (up to τ-dependent prefactors) identical Hamiltonians but do not interact with each other and can be distinguished from each other. It is spanned by tensor products of Fock states. The state $g[\tilde{H}^{(+)}](\beta)$ is the thermal state of this system. Equation (42) with $t(w)$ as defined as in Eq. (43) still holds. Note that the swap operators swap tensor factors, not fermionic modes. Thus, they still satisfy the symmetry relations that are

used to prove that only terms corresponding to the words $w \in \mathcal{C}_{\geq L}(\partial A)$ can contribute to the covariance.

It remains to show that Lemma 1 still holds in the fermionic setting. Lemmas 3 and 8 are purely combinatorial. Lemmas 4–7, 9, and 10 only use the local boundedness of the Hamiltonian and that Hamiltonian terms with disjoint support commute. The same holds in the fermionic setting because the Hamiltonian terms must be physical operators, i.e., even polynomials in the fermionic operators. Hence, all lemmas used in the proof of Lemma 1 carry over to the fermionic setting. It is then straightforward to see that the proof itself also goes through without any modifications. ∎

Proof of Corollary 5.—Tracing out B^c in the second trace in Eq. (32) and bounding the integral yields

$$|\mathrm{Tr}[Ag(\beta)] - \mathrm{Tr}\{Ag[H_{\restriction B}](\beta)\}|$$
$$\leq |\beta| \sup_{s \in [0,1]} \sup_{\tau \in [0,1]} |\, \mathrm{cov}^{\tau}_{g_s(\beta)}(A, H_{\partial B})|. \quad (51)$$

Taking the supremum over all A with $\|A\|_{\infty} = 1$ and $\mathrm{supp}(A) \subseteq S$ and using Theorem 4 finish the proof. ∎

VI. CONCLUSIONS

In this work, we clarify the limitations of a universal concept of scale-independent temperature by showing that temperature is intensive on a given length scale if and only if correlations are negligible. The corresponding correlation measure turns out to also quantitatively capture the stability of thermal states against perturbations of the Hamiltonian. Moreover, we find a universal critical temperature above which correlations always decay exponentially with the distance. We compare our results to known results on phase transitions, comment on recent advances concerning thermalization in closed quantum systems (e.g., concerning the eigenstate-thermalization hypothesis), and discuss known matrix product operator approximations of thermal states. More concretely, our results imply that at high enough temperatures, the error made when truncating a Hamiltonian at some distance away from the system of interest is exponentially suppressed with the distance. As a computational consequence, expectation values of local observables can be approximated efficiently.

APPENDIX: CLUSTER EXPANSIONS AND PROOF OF LEMMA 1

The following discussion of cluster expansions is expected to be interesting in its own right, as it contains

a rigorous formulation of the ideas outlined in Ref. [40]. We will provide a proof of the original statement used to establish Theorem 3 as well as of the tailored statement in Lemma 1, which is used to prove Theorem 2 on the clustering of correlations.

1. The original cluster expansion from Ref. [40]

The original cluster expansion is similar to Lemma 1 with just one copy of the system instead of two weighted ones.

Lemma 2 (Truncated cluster expansion [40]).—Let $H = \sum_{\lambda \in E} h_\lambda$ be a local Hamiltonian with finite interaction graph (V, E) having growth constant α and local interaction strength $J = \max_{\lambda \in E} \|h_\lambda\|_\infty$, and define $b(x) := \alpha e^{|x|}(e^{|x|} - 1)$. Moreover, let β be small enough such that $b(\beta J) < 1$. For some subset of edges $F \subset E$, let $\mathcal{C}_{\geq L}(F) \subset E^*$ be the set of words containing at least one cluster c that contains at least one letter of F and has size $|c| \geq L$ and denote the corresponding truncated cluster expansion by

$$\Omega[H](\beta) := \sum_{w \in \mathcal{C}_{\geq L}(F)} \frac{(-\beta)^{|w|}}{|w|!} h(w). \quad (A1)$$

Then,

$$\frac{\|\Omega[H](\beta)\|_1}{Z(\beta)} \leq \exp\left(|F| \frac{b(\beta J)^L}{1 - b(\beta J)}\right) - 1. \quad (A2)$$

If one applies this lemma to the setting of Lemma 1, one obtains a bound similar as the one in Eq. (38) but with $Z[\tilde{H}](\beta)$ instead of $Z(\beta)$, where the ratio $Z[\tilde{H}](\beta)/Z[H](\beta)$ can be exponentially large in the system size for $\tau \in]0, 1[$.

Lemma 2 was used in Ref. [40] to establish a mathematically (not algorithmically) constructive version of Theorem 3, on MPO approximations, where the MPO in Eq. (17) is given by

$$\rho(\beta, L) = \frac{1}{Z(\beta)} \sum_{w \in E^* \setminus \mathcal{C}_{\geq L}(E)} \frac{(-\beta)^{|w|}}{|w|!} h(w). \quad (A3)$$

2. Proofs of Lemmas 1 and 2

The purpose of this section is to prove Lemma 1. But, along the way, we also prove Lemma 2. In order to do so, we start with the introduction of some more notation, mainly concerning clusters and lattice animals. For $w \in E^*$ and any subalphabet $G \subset E$, we write $G \subset w$ if every letter in G also occurs in w. By $G^c := E \setminus G$, we denote the complement of $G \subset E$. The extension of G is defined to be $\bar{G} := \{\lambda \in E | \exists \lambda' \in G : \lambda' \cap \lambda \neq \emptyset\}$ and, similarly as for subsystems, its boundary is $\partial G := \bar{G} \setminus G$. Throughout the proof, we fix some subset of edges $F \subset E$. We denote by $\mathcal{C}_{\geq L}(F) \subset E^*$ the set of words that contain at least one

cluster c with $c \cap F \neq \emptyset$ and $|c| \geq L$, and we denote by $\mathcal{C}^k_{\geq L}(F)$ the set of words that contain exactly k such clusters. Note that for an animal $G \subset E$, there exists a cluster $c \in E^*$ such that $G = \{\lambda \in c\}$, and if one imposes some order on G, one obtains a cluster. We denote by $\mathcal{A}_{=l}(F)$ and $\mathcal{A}_{\geq L}(F)$ the sets of animals that contain at least one edge of F and are of size exactly l or at least L, respectively. Moreover, we denote by $\mathcal{A}^k_{\geq L}(F)$ the corresponding sets of k-fold animals, i.e.,

$$\mathcal{A}^k_{\geq L} := \{\biguplus_{j=1}^k G_j : G_j \in \mathcal{A}_{\geq L}(F) \text{ nonoverlapping}\}.$$

For a more compact notation, we write the terms in the exponential series as

$$f(w) := \frac{(-\beta)^{|w|}}{|w|!} h(w). \tag{A4}$$

We will frequently use the following fact: For any Hamiltonian with a finite interaction graph (V, E), the partial series over any set of words $\mathcal{W} \subseteq E^*$ converges absolutely, i.e.,

$$\left\| \sum_{w \in \mathcal{W}} f(w) \right\|_\infty \leq \sum_{w \in \mathcal{W}} \frac{(|\beta| J)^{|w|}}{|w|!} \tag{A5}$$

$$\leq \sum_{w \in E^*} \frac{(|\beta| J)^{|w|}}{|w|!} \tag{A6}$$

$$= \exp(|\beta| J |E|). \tag{A7}$$

In particular, this bound implies that the order of the terms in the series over any subset of words $\mathcal{W}$ does not matter.

In the following proofs of Lemmas 1 and 2, we use several technical auxiliary lemmas, which we will only state and prove subsequently.

Proof of Lemma 1.—During this proof, we indicate quantities corresponding to $\tilde{H}$ by a tilde accent, e.g., $\tilde{f}(w)$ is defined as in Eq. (A4) but with respect to the local terms $\tilde{h}_\lambda$ of $\tilde{H}$ while $f(w)$ is defined with respect to the local terms h_λ of H.

We start the proof by rearranging the terms in the series over $\mathcal{C}_{\geq L}(F)$ in Eq. (37) according to the number of relevant clusters they contain and use Lemma 3 with b_k being the series over $\mathcal{C}^k_{\geq L}(F)$ to obtain

$$\Omega[\tilde{H}](\beta) = \sum_{k=1}^\infty \sum_{w \in \mathcal{C}^k_{\geq L}(F)} \tilde{f}(w)$$

$$= -\sum_{m=1}^\infty (-1)^m \sum_{k=m}^\infty \binom{k}{m} \sum_{w \in \mathcal{C}^k_{\geq L}(F)} \tilde{f}(w). \tag{A8}$$

Lemmas 5, 6, and 9 are the core of the proof. They define a series of operators $(\tilde{\rho}_m)_{m=1}^\infty$ that have a particularly useful

form given in Lemma 10. This form exactly matches the series over k in Eq. (A8), which leads to the following identity:

$$\Omega[\tilde{H}](\beta) = -\sum_{m=1}^\infty (-1)^m \tilde{\rho}_m. \tag{A9}$$

The operators $\tilde{\rho}_m$ are defined in Eq. (A57) as series over m-fold lattice animals G of operators $\rho(G)$ [defined in Eq. (A31)]. This definition implies

$$\mathrm{Tr}(\mathcal{S}AB\tilde{\rho}_m) = \sum_{G \in \mathcal{A}^m_{\geq L}(F)} \mathrm{Tr}[\mathcal{S}AB\tilde{\rho}(G)]. \tag{A10}$$

In the previous steps, the series over words has been rewritten as a series over m-fold animals. Lemma 7 provides a bound on $\tilde{\rho}(G)$ that, together with Eqs. (A9) and (A10), yields

$$\frac{|\mathrm{Tr}\{\mathcal{S}AB\Omega[\tilde{H}](\beta)\}|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq \sum_{m=1}^\infty \sum_{G \in \mathcal{A}^m_{\geq L}(F)} y(\beta J)^{|G|}. \tag{A11}$$

Now, a counting argument for lattice animals from Lemma 8 allows us to bound the series over m-fold animals G in terms of a series of animals

$$\frac{|\mathrm{Tr}\{\mathcal{S}AB\Omega[\tilde{H}](\beta)\}|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq \sum_{m=1}^\infty \frac{1}{m!} \left(\sum_{G \in \mathcal{A}_{\geq L}(F)} y(\beta J)^{|G|} \right)^m.$$

Using that the number a_l [see Eq. (7)] of lattice animals G with $G \cap F \neq \emptyset$ and of size $|G| = l$ is bounded by $|F| a_l$ and that $a_l \leq \alpha^l$ [see Eq. (8)], we obtain

$$|\mathrm{Tr}\{\mathcal{S}AB\Omega[\tilde{H}](\beta)\}| \leq Z(\beta) \sum_{m=1}^\infty \frac{1}{m!} \left(|F| \sum_{l=L}^\infty b(\beta J)^l \right)^m$$

with $b(x) := \alpha y(x)$. Performing the partial geometric series over l with argument $b(\beta J) < 1$ and the exponential series over m yields Eq. (A2) and completes the proof. ∎

Similarly, we prove Lemma 2.

Proof of Lemma 2.—By the same argument that led us to Eq. (A9) in the proof of Lemma 1, we obtain

$$\Omega[H](\beta) = -\sum_{m=1}^\infty (-1)^m \rho_m. \tag{A12}$$

Applying the triangle inequality and using the bound on ρ_m from Lemma 9 yields

$$\|\Omega[H](\beta)\|_1 \leq Z(\beta) \sum_{m=1}^\infty \frac{1}{m!} \left(|F| \sum_{l=L}^\infty b(\beta J)^l \right)^m. \tag{A13}$$

Performing the partial geometric series over l with argument $b(\beta J) < 1$ and the exponential series over m yields Eq. (A2) and completes the proof. ∎

We now prove various lemmas that are used in the previous proofs of Lemmas 1 and 2.

Lemma 3.—Let $(b_k)_{k=1}^{\infty}$ be a sequence of complex matrices

$$A_K := \sum_{k=1}^{K} b_k \tag{A14}$$

and

$$B_K := -\sum_{m=1}^{K} (-1)^m \sum_{k=m}^{K} \binom{k}{m} b_k. \tag{A15}$$

Then, $A_K = B_K$ for all $K \in \mathbb{N}$. In particular, if both sequences converge, then their limits are the same, i.e., $\lim_{K\to\infty} A_K = \lim_{K\to\infty} B_K$.

Proof.—Applying the binomial theorem to $(1-1)^k = 0$ yields

$$\sum_{l=0}^{k} (-1)^l \binom{k}{l} = 0, \tag{A16}$$

which we will use. We prove the identity by induction. $A_1 = B_1$ is easy to see. Under the assumption that $A_K = B_K$ for some $K \in \mathbb{N}$, we obtain

$$B_{K+1} = B_K - (-1)^{K+1} \binom{K+1}{K+1} b_{K+1} \tag{A17}$$

$$-\sum_{m=1}^{K} (-1)^m \binom{K+1}{m} b_{K+1} \tag{A18}$$

$$= A_K + \left[-(-1)^{K+1} - \sum_{m=1}^{K} (-1)^m \binom{K+1}{m} \right] b_{K+1}$$

$$= A_{K+1}, \tag{A19}$$

where we have used Eq. (A16) in the last step. ∎

The goal of the following lemmas is to show that ρ_m is well-defined and to upper bound it in trace norm. The order of the lemmas is chosen in a way that makes clear that the two quantities ρ_m and $\rho(G)$, which will be defined shortly, are actually well-defined.

We start with a trace norm bound on the perturbed exponential series.

Lemma 4 [Eq. (21) from Ref. [40]].—Let H be a Hamiltonian with finite interaction graph (V, E). For any sequence $(G_j)_{j=1}^{k}$ of subalphabets $G_j \subset E$,

$$\left\| e^{-\beta(H-\sum_{j=1}^{k} H_{G_j})} \right\|_1 \le Z(\beta) \prod_{j=1}^{k} \left\| e^{|\beta| H_{G_j}} \right\|_\infty. \tag{A20}$$

Proof.—The lemma is essentially a consequence of the Golden-Thompson inequality and the fact that the trace norm of a positive operator coincides with its trace. Using first the Golden-Thompson and then Hölder's inequality, we obtain

$$\left\| e^{-\beta(H-\sum_{j=1}^{k} H_{G_j})} \right\|_1 \le \mathrm{Tr}[e^{-\beta(H-\sum_{j=1}^{k-1} H_{G_j})} e^{\beta H_{G_k}}]$$

$$\le \mathrm{Tr}[e^{-\beta(H-\sum_{j=1}^{k-1} H_{G_j})}] \left\| e^{|\beta| H_{G_k}} \right\|_\infty. \tag{A21}$$

Now, iteration completes the proof. ∎

We will use the following lemma to bound the operator norm of certain subseries of $f(w)$.

Lemma 5.—Let (V, E) be a finite graph and $J \ge 0$. For any $G \subset E$,

$$\sum_{w \in G^*: G \subset w} \frac{|\beta J|^{|w|}}{|w|!} = (e^{|\beta J|} - 1)^{|G|}. \tag{A22}$$

Proof.—Ordering the words in the sum in Eq. (A22) with respect to their length yields

$$\sum_{w \in G^*: G \subset w} \frac{|\beta J|^{|w|}}{|w|!} = \sum_{l=|G|}^{\infty} \sum_{w \in G^l: G \subset w} \frac{|\beta J|^{|w|}}{|w|!} \tag{A23}$$

$$= \sum_{l=|G|}^{\infty} \frac{|\beta J|^l}{l!} |\{w \in G^l: G \subset w\}|. \tag{A24}$$

From basic combinatorial considerations, we obtain

$$|\{w \in G^l: G \subset w\}| = \sum_{\substack{j_1, j_2, \dots, j_n \ge 1, \\ j_1 + j_2 + \dots + j_n = l}} \binom{l}{j}, \tag{A25}$$

where $\binom{l}{j}$ is a multinomial coefficient. Therefore, the right-hand side of Eq. (A24) only depends on $n := |G|$ and we denote it by

$$\gamma(n) := \sum_{l=n}^{\infty} \gamma(n, l) \tag{A26}$$

with

$$\gamma(n,l) := \frac{|\beta J|^l}{l!} \sum_{\substack{j_1,j_2,\ldots,j_n \geq 1, \\ j_1+j_2+\ldots+j_n=l}} \binom{l}{j}$$

$$= \sum_{\substack{j_1,j_2,\ldots,j_n \geq 1, \\ j_1+j_2+\ldots+j_n=l}} \frac{|\beta J|^{j_1}}{j_1!} \frac{|\beta J|^{j_2}}{j_2!} \cdots \frac{|\beta J|^{j_n}}{j_n!}. \quad (A27)$$

Then,

$$\gamma(n) = \sum_{l=n}^{\infty} \sum_{\substack{j_1,j_2,\ldots,j_n \geq 1, \\ j_1+j_2+\ldots+j_n=l}} \frac{|\beta J|^{j_1}}{j_1!} \frac{|\beta J|^{j_2}}{j_2!} \cdots \frac{|\beta J|^{j_n}}{j_n!}$$

$$= \sum_{l=n}^{\infty} \sum_{j_1=1}^{l-(n-1)} \frac{|\beta J|^{j_1}}{j_1!} \sum_{\substack{j_2,\ldots,j_n \geq 1, \\ j_2+\ldots+j_n=l-j_1}} \frac{|\beta J|^{j_2}}{j_2!} \cdots \frac{|\beta J|^{j_n}}{j_n!}$$

$$= \sum_{l=1}^{\infty} \sum_{j_1=1}^{l} \frac{|\beta J|^{j_1}}{j_1!} \gamma(n-1, l+n-1-j_1) \quad (A28)$$

and, after realizing that the last series is a Cauchy product,

$$\gamma(n) = \sum_{j_1=1}^{\infty} \frac{|\beta J|^{j_1}}{j_1!} \sum_{l=n-1}^{\infty} \gamma(n-1, l) \quad (A29)$$

$$= (e^{|\beta J|} - 1)\gamma(n-1). \quad (A30)$$

We note that $\gamma(1) = e^{|\beta J|} - 1$, and iteration finishes the proof. ∎

The following lemma provides a factorization of the series $\rho(G)$ in Eq. (A33) over words that have no letters on the boundary of an m-fold animal $G \in \mathcal{A}^m_{=l}(F)$ and contain all letters in G, into $\exp(-\beta H_{(\bar{G})^c})$, whose norm we have bounded in Lemma 4, times a product of operators $\eta(G_j)$. The $\eta(G_j)$ are supported on the single animals G_j composing the m-fold animal G. As we will see, a norm bound for $\eta(G_j)$ follows immediately from the previous lemma, which, in turn, also yields an upper bound on $\rho(G)$. The form of $\rho(G)$ given in Eq. (A33) together with this upper bound plays an important role in the main cluster expansion.

Lemma 6.—Let H be a Hamiltonian with finite interaction graph (V, E). For $G \subset E$, let $G = \biguplus_{j=1}^m G_j$ be the decomposition of G into nonoverlapping animals $G_j \subset E$ and define

$$\rho(G) := e^{-\beta H_{(\bar{G})^c}} \prod_{j=1}^m \eta(G_j) \quad (A31)$$

with

$$\eta(G) := \sum_{w \in G^*: G \subset w} f(w). \quad (A32)$$

Then,

$$\rho(G) = \sum_{w \in [(\partial G)^c]^*: G \subset w} f(w). \quad (A33)$$

Proof.—To simplify the notation, we denote the relevant set of words that contain no letters in ∂G and each letter in G at least once by

$$\mathcal{W}^{\supset G} := \{w \in [(\partial G)^c]^* : G \subset w\}. \quad (A34)$$

The idea is to group these words into subsets $[w] \subset \mathcal{W}^{\supset G}$ that coincide on the connected components of G and on $(\bar{G})^c$ and correspondingly split up the series (A33). We formalize this idea by introducing an equivalence relation on $\mathcal{W}^{\supset G}$. For $v, w \in \mathcal{W}$, we define

$$v \sim w : \Leftrightarrow \begin{cases} v \restriction G^c = w \restriction G^c \\ v \restriction G_j = w \restriction G_j & \forall j = 1, 2, \ldots, k, \end{cases}$$

where, for any subalphabet $G' \subset E$, the restriction $w \restriction G'$ of a word $w \in E^*$ is obtained from w by omitting all letters that are not in G'. Then, the size of each equivalence class $[w] \in \mathcal{W}^{\supset G}/\sim$ is given by the multinomial coefficient

$$|[w]| = \binom{|w|}{(|w \restriction G^c|, |w \restriction G_1|, \ldots, |w \restriction G_k|)}. \quad (A35)$$

Note also that $h([w]) := h(w \restriction \overline{G}^c) \prod_{j=1}^k h(w \restriction G_j) = h(w)$ is well-defined as a function on the classes. Let us denote the set of words over the alphabet G_j that contain all letters at least once by

$$\mathcal{W}^{=G_j} := \{w \in (G_j)^* : G_j \subset w\}. \quad (A36)$$

Then, the quotient set can be identified with a Cartesian product of these sets

$$\mathcal{W}^{\supset G}/\sim \; \cong \; [(\bar{G})^c]^* \times \bigtimes_{j=1}^{k} \mathcal{W}^{=G_j}. \quad (A37)$$

For each equivalence class $K \in \mathcal{W}^{\supset G}/\sim$, we pick an arbitrary representative $w_K \in \mathcal{W}^{\supset G}$, use the definition of f in Eq. (A4), and determine that k is the number of connected components of G to obtain

$$\sum_{w \in [(\partial G)^c]^* : G \subset w} f(w) = \sum_{K \in \mathcal{W}^{\supset G}/\sim} |K| \frac{(-\beta)^{|w_K|}}{|w_K|!} h(w_K) \tag{A38}$$

$$= \sum_{v \in [(\tilde{G})^c]^*} \sum_{w_1 \in \mathcal{W}^{=G_1}} \sum_{w_2 \in \mathcal{W}^{=G_2}} \cdots \sum_{w_k \in \mathcal{W}^{=G_k}} \binom{|v| + \sum_{j=1}^k |w_j|}{(|v|, |w_1|, \ldots, |w_k|)} \frac{(-\beta)^{|v| + \sum_{j=1}^k |w_j|}}{(|v| + \sum_{j=1}^k |w_j|)!} h(v) \prod_{j=1}^k h(w_j)$$

$$= \sum_{v \in [(\tilde{G})^c]^*} f(v) \left(\sum_{w_1 \in \mathcal{W}^{=G_1}} f(w_1) \right) \left(\sum_{w_2 \in \mathcal{W}^{=G_2}} f(w_2) \right) \cdots \left(\sum_{w_k \in \mathcal{W}^{=G_k}} f(w_k) \right). \tag{A39}$$

Using the definition of η from Eq. (A32) on the last factors yields

$$\sum_{w \in [(\partial G)^c]^* : G \subset w} f(w) = e^{-\beta H_{(\tilde{G})^c}} \prod_{j=1}^k \eta(G_j) = \rho(G). \tag{A40}$$

∎

The following lemma is a tighter variant of some of the original arguments leading to Lemma 2 for Hamiltonians consisting of two weighted copies of a local Hamiltonian. Its purpose is to provide a specialized tighter bound on $\rho(G)$, which turns out to be sufficient for our purposes. The central idea of the lemma is to expand $\rho(G)$ in the left-hand side of Eq. (A41) in order to be able to bound the trace using the generalized Hölder's inequality.

Lemma 7.—Let τ, H, $\tilde{H}$, A, and B be as in Lemma 1 and let $G \in \mathcal{A}_{\geq L}^{=m}(F)$ be an m-fold lattice animal with $G = \bigcup_{j=1}^m G_j$ and $G_j \in \mathcal{A}_{\geq L}(F)$. Moreover, let $\tilde{\rho}(G)$ be defined as $\rho(G)$ in Eq. (A31) but with respect to $\tilde{H}$. Then,

$$\frac{|\mathrm{Tr}[SA^{(1)}B^{(2)} e^{-\beta \tilde{H}_{\tilde{G}^c}} \tilde{\rho}(G)]|}{\|A\|_\infty \|B\|_\infty Z(\beta)} \leq y(\beta J)^{|G|}, \tag{A41}$$

where $y(x) := e^{|x|}(e^{|x|} - 1)$.

Proof.—Let us denote $k_\lambda^{(1)} := \tau h_\lambda^{(1)}$ and $k_\lambda^{(2)} := (1-\tau)h_\lambda^{(2)}$. For $w \in E^*$ and $v \in \{1,2\}^{|w|}$, we define $\tilde{h}(w, v) := k_{w_1}^{v_1} k_{w_2}^{v_2} \cdots k_{w_{|w|}}^{v_{|w|}}$. Then, by expanding the product $\tilde{h}(w)$, it can be written as

$$\tilde{h}(w) = \sum_{v \in \{1,2\}^{|w|}} \tilde{h}(w, v). \tag{A42}$$

Importantly, we can reorder the terms in $\tilde{h}(w, v)$ so that

$$\tilde{h}(w, v) = \tilde{h}^{(1)}(w, v) \tilde{h}^{(2)}(w, v), \tag{A43}$$

where $\tilde{h}^{(1)}(w, v) = h^{(i)}(w, v) \otimes \mathbb{1}$ and $\tilde{h}^{(2)}(w, v) = \mathbb{1} \otimes h^{(ii)}(w, v)$. Factorizing the operators and using the swap trick (39), we obtain

$$\mathrm{Tr}[SA^{(1)}B^{(2)} e^{-\beta \tilde{H}_{\tilde{G}^c}} \tilde{h}(w, v)] = \mathrm{Tr}\{S[Ae^{-\beta \tau H_{\tilde{G}^c}} h^{(i)}(w, v)] \otimes [Be^{-\beta(1-\tau)H_{\tilde{G}^c}} h^{(ii)}(w, v)]\} \tag{A44}$$

$$= \mathrm{Tr}\{[Ae^{-\beta \tau H_{\tilde{G}^c}} h^{(i)}(w, v)][Be^{-\beta(1-\tau)H_{\tilde{G}^c}} h^{(ii)}(w, v)]\}. \tag{A45}$$

Bounding the trace by the trace norm and applying Hölder's inequality generalized to several operators yields

$$|\mathrm{Tr}[SA^{(1)}B^{(2)} e^{-\beta \tilde{H}_{\tilde{G}^c}} \tilde{h}(w, v)]| \leq \|A\|_\infty \|B\|_\infty \|e^{-\beta \tau H_{\tilde{G}^c}}\|_{1/\tau} \|e^{-\beta(1-\tau)H_{\tilde{G}^c}}\|_{1/(1-\tau)} \|h^{(i)}(w, v)\|_\infty \|h^{(ii)}(w, v)\|_\infty$$

$$\leq \|A\|_\infty \|B\|_\infty \|e^{-\beta H_{\tilde{G}^c}}\|_1 J^{|w|} \tau^{n^{(1)}(v)} (1 - \tau)^{n^{(2)}(v)}, \tag{A46}$$

where in the second step, we have used that $\|X\|_p = \||X|^p\|_1^{1/p}$ and that with $n^{(j)}(v) := |\{v_k : v_k = j\}|$ for $j \in \{1, 2\}$, the bounds $\|h^{(i)}(w, v)\|_\infty \leq (\tau J)^{n^{(1)}(v)}$ and $\|h^{(ii)}(w, v)\|_\infty \leq [(1 - \tau)J]^{n^{(2)}(v)}$ hold. Now, we apply Lemma 4 and use that $\|e^{|\beta|H_{G_j}}\|_\infty \leq e^{|\beta|J|G_j|}$ to arrive at

$$|\mathrm{Tr}[SA^{(1)}B^{(2)} e^{-\beta \tilde{H}_{\tilde{G}^c}} \tilde{h}(w, v)]| \leq \|A\|_\infty \|B\|_\infty Z(\beta) e^{|\beta|J|G|} J^{|w|} \tau^{n^{(1)}(v)} (1 - \tau)^{n^{(2)}(v)}. \tag{A47}$$

From the definition of η in Eq. (A32), it follows that

$$\prod_{j=1}^{m} \tilde{\eta}(G_j) = \sum_{\left\{ \substack{w^{(j)} \in G_j^* : \\ G_j \subset w^{(j)}} \right\}_{j=1}^{m}} \prod_{i=1}^{m} \frac{(-\beta)^{|w^{(i)}|}}{|w^{(i)}|!} \tilde{h}(w^{(i)}) \quad \text{(A48)}$$

$$= \sum_{\left\{ \substack{w^{(j)} \in G_j^* : \\ G_j \subset w^{(j)}} \right\}_{j=1}^{m}} \frac{(-\beta)^{|w|}}{\prod_{i=1}^{m} |w^{(i)}|!} \sum_{v \in \{1,2\}^{|w|}} \tilde{h}(w, v),$$

$$\text{(A49)}$$

where $w := w^{(1)} \circ w^{(2)} \circ \ldots \circ w^{(m)}$ and hence $\tilde{h}(w) = \prod_{i=1}^{m} h(w^{(i)})$. Together with the bound (A47) we obtain

$$\frac{|\mathrm{Tr}[\mathcal{S} A^{(1)} B^{(2)} e^{-\beta \tilde{H}_{\bar{G}^c}} \prod_{j=1}^{m} \tilde{\eta}(G_j)]|}{\|A\|_{\infty} \|B\|_{\infty} Z(\beta)}$$

$$\leq e^{|\beta| J |G|} \sum_{\left\{ \substack{w^{(j)} \in G_j^* : \\ G_j \subset w^{(j)}} \right\}_{j=1}^{m}} \frac{|\beta|^{|w|}}{\prod_{i=1}^{m} |w^{(i)}|!} J^{|w|}$$

$$\times \sum_{v \in \{1,2\}^{|w|}} \tau^{n^{(1)}(v)} (1 - \tau)^{n^{(2)}(v)}. \quad \text{(A50)}$$

Using the definition (A31) of $\rho(G)$ and the multinomial formula yields

$$\frac{|\mathrm{Tr}[\mathcal{S} A^{(1)} B^{(2)} e^{-\beta \tilde{H}_{\bar{G}^c}} \tilde{\rho}(G)]|}{\|A\|_{\infty} \|B\|_{\infty} Z(\beta)} \quad \text{(A51)}$$

$$= e^{|\beta| J |G|} \sum_{\left\{ \substack{w^{(j)} \in G_j^* : \\ G_j \subset w^{(j)}} \right\}_{j=1}^{m}} \prod_{i=1}^{m} \frac{(|\beta| J)^{|w^{(i)}|}}{|w^{(i)}|!} \quad \text{(A52)}$$

$$= e^{|\beta| J |G|} \prod_{i=1}^{m} \left(\sum_{\substack{w^{(i)} \in G_i^* : \\ G_i \subset w^{(i)}}} \frac{(|\beta| J)^{|w^{(i)}|}}{|w^{(i)}|!} \right) \quad \text{(A53)}$$

$$\leq e^{|\beta| J |G|} \prod_{i=1}^{m} (e^{|\beta| J} - 1)^{|G_i|}, \quad \text{(A54)}$$

where in the second-to-last step, we have factorized the series and in the last step, we have used Lemma 5. ∎

We will need the following combinatorial lemma.

Lemma 8.—Let (V, E) be a finite (hyper)graph and $y \in [0, 1[$. Then, for any $F \subset E$,

$$\sum_{G \in \mathcal{A}_{\geq L}^m(F)} y^{|G|} \leq \frac{1}{m!} \left(\sum_{G \in \mathcal{A}_{\geq L}(F)} y^{|G|} \right)^m. \quad \text{(A55)}$$

Proof.—Remember that $\mathcal{A}_{\geq L}^m(F)$ is the set of m-fold (edge) animals of size at least L that contain a letter from F.

For every $G \in \mathcal{A}_{\geq L}^m(F)$, one finds m pairs (G_1, G_2) with $G_1 \in \mathcal{A}_{\geq L}^{m-1}(F)$ and $G_2 \in \mathcal{A}_{\geq L}^m(F)$ such that $G = G_1 \uplus G_2$; hence,

$$m \sum_{G \in \mathcal{A}_{\geq L}^m(F)} y^{|G|} \leq \sum_{G_1 \in \mathcal{A}_{\geq L}^{m-1}(F)} \sum_{G_2 \in \mathcal{A}_{\geq L}(F)} y^{|G_1| + |G_2|}$$

$$= \left(\sum_{G \in \mathcal{A}_{\geq L}^{m-1}(F)} y^{|G|} \right) \left(\sum_{G \in \mathcal{A}_{\geq L}(F)} y^{|G|} \right).$$

By iterating this inequality, we obtain

$$\sum_{G \in \mathcal{A}_{\geq L}^m(F)} y^{|G|} \leq \frac{1}{m!} \left(\sum_{G \in \mathcal{A}_{\geq L}(F)} y^{|G|} \right)^m. \quad \text{(A56)}$$

∎

In the following lemma, we define a family of operators ρ_m and bound their trace norms. The bounds, in particular, guarantee that the ρ_m are well-defined. In addition, they are useful for the proof of Lemma 2, albeit they are not explicitly needed for the proof of Lemma 1.

Lemma 9.—Let $\rho(G)$ be defined as in Lemma 6 with respect to a Hamiltonian H having a finite interaction (hyper)graph (V, E) with growth constant α and let

$$\rho_m := \sum_{G \in \mathcal{A}_{\geq L}^m(F)} \rho(G) \quad \text{(A57)}$$

for some $F \subset E$. Then,

$$\|\rho_m\|_1 \leq \frac{Z(\beta)}{m!} \left(|F| \sum_{l=L}^{\infty} b(\beta J)^l \right)^m, \quad \text{(A58)}$$

where $b(x) := \alpha e^{|x|}(e^{|x|} - 1)$.

Proof.—Consider a k-fold animal $G \in \mathcal{A}_{\geq L}^m(F)$ and decompose it into its k nonoverlapping animals $G_j \in \mathcal{A}_{\geq L}(F)$ as $G = \uplus_{j=1}^k G_j \subset E$. Then, Eq. (A31) and Hölder's inequality imply

$$\|\rho(G)\|_1 \leq \|e^{-\beta H_{(\bar{G})^c}}\|_1 \prod_{j=1}^{k} \|\eta(G_j)\|_{\infty}, \quad \text{(A59)}$$

and it follows from Lemmas 4 and 5 in conjunction with the definition of η in Eq. (A32) that

$$\|\rho(G)\|_1 \leq Z(\beta) y(\beta J)^{|G|}. \quad \text{(A60)}$$

Hence, by the definition from Eq. (A57) and Lemma 8, we obtain

$$\|\rho_m\|_1 \le Z(\beta) \sum_{G \in \mathcal{A}^m_{\ge L}(F)} y(\beta J)^{|G|} \tag{A61}$$

$$\le \frac{Z(\beta)}{m!} \left(\sum_{G \in \mathcal{A}_{\ge L}(F)} y(\beta J)^{|G|} \right)^m. \tag{A62}$$

By decomposing the set of animals of size at least L into a union of sets of animals of fixed size l, i.e., $\mathcal{A}_{\ge L}(F) = \biguplus_{l=L}^{\infty} \mathcal{A}_{=l}(F)$, we can write

$$\|\rho_m\|_1 \le \frac{Z(\beta)}{m!} \left(\sum_{l=L}^{\infty} |\mathcal{A}_{=l}(F)| y(\beta J)^l \right)^m. \tag{A63}$$

The bound (8) on the number of lattice animals, the fact that the number $|F|$ of edges in F upper bounds the number of possibilities of translating an animal G such that $G \subset F$, and $b = \alpha y$ finish the proof. ∎

While the last lemma provides a bound on ρ_m and, in particular, implies that ρ_m is well-defined, the next lemma provides a useful form of ρ_m.

Lemma 10.—Let ρ_m be defined as in Eq. (A57). Then,

$$\rho_m = \sum_{k=m}^{\infty} \binom{k}{m} \sum_{w \in \mathcal{C}^k_{\ge L}(F)} f(w). \tag{A64}$$

Proof.—For $G \in \mathcal{A}^m_{\ge L}(F)$, let

$$\mathcal{W}(G) := \{w \in [(\partial G)^c]^* : G \subset w\}. \tag{A65}$$

According to Eqs. (A33) and (A57),

$$\rho_m = \sum_{G \in \mathcal{A}^m_{\ge L}(F)} \sum_{w \in [(\partial G)^c]^* : G \subset w} f(w). \tag{A66}$$

As

$$\bigcup_{G \in \mathcal{A}^m_{\ge L}(F)} \mathcal{W}(G) = \biguplus_{k=m}^{\infty} \mathcal{C}^k_{\ge L}(F), \tag{A67}$$

the sums in Eqs. (A57) and (A64) contain the same terms. It remains to show that the multiplicities are correct, i.e., are given by the binomial factor. Every word in $\mathcal{W}(G)$ contains at least m maximal clusters of size at least L, each of which contains a letter in F. The key is to decompose this set as

$$\mathcal{W}(G) = \biguplus_{k=m}^{\infty} \mathcal{W}^k(G) \tag{A68}$$

with

$$\mathcal{W}^k(G) := \{w \in \mathcal{W}(G) : \exists \text{ exactly } k \text{ maximal clusters}$$
$$c \subset w : c \in \mathcal{C}_{\ge L}(F)\},$$

i.e., into sets of words having exactly $k \ge m$ such clusters. Then, the observation that for every $w \in \mathcal{W}^k(G)$ there are exactly $\binom{k}{m}$ many m-fold animals $G' \in \mathcal{A}^m_{\ge L}(F)$ with $w \subset G'$ completes the proof. ∎

[1] H. Pothier, S. Guéron, N. O. Birge, D. Esteve, and M. H. Devoret, *Energy Distribution Function of Quasiparticles in Mesoscopic Wires*, Phys. Rev. Lett. **79**, 3490 (1997).

[2] P. Peng, Z. Su, Z. Liu, Q. Yu, Z. Cheng, and J. Bao, *Nanowire Thermometers*, Nanoscale **5**, 9532 (2013).

[3] Y. Gao and Y. Bando, *Nanotechnology: Carbon Nanothermometer Containing Gallium*, Nature (London) **415**, 599 (2002).

[4] N. Linden, S. Popescu, and P. Skrzypczyk, *How Small Can Thermal Machines Be? The Smallest Possible Refrigerator*, Phys. Rev. Lett. **105**, 130401 (2010).

[5] A. Mari and J. Eisert, *Cooling by Heating*, Phys. Rev. Lett. **108**, 120602 (2012).

[6] M. Hartmann, G. Mahler, and O. Hess, *Existence of Temperature on the Nanoscale*, Phys. Rev. Lett. **93**, 080402 (2004).

[7] M. Hartmann and G. Mahler, *Measurable Consequences of the Local Breakdown of the Concept of Temperature*, Europhys. Lett. **70**, 579 (2005).

[8] M. Hartmann, *Minimal Length Scales for the Existence of Local Temperature*, Contemp. Phys. **47**, 89 (2006).

[9] A. Ferraro, A. Garcia-Saez, and A. Acin, *Intensive Temperature and Quantum Correlations for Refined Quantum Measurements*, Europhys. Lett. **98**, 10009 (2012).

[10] M. Vojta, *Quantum Phase Transitions*, Rep. Prog. Phys. **66**, 2069 (2003).

[11] D. Ruelle, *Statistical Mechanics: Rigorous Results* (Benjamin, New York, 1969).

[12] J. Ginibre, *Reduced Density Matrices of Quantum Gases. II. Cluster Property*, J. Math. Phys. (N.Y.) **6**, 252 (1965).

[13] W. Greenberg, *Critical Temperature Bounds of Quantum Lattice Gases*, Commun. Math. Phys. **13**, 335 (1969).

[14] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (Springer, New York, 1981).

[15] Y. M. Park and H. J. Yoo, *Uniqueness and Clustering Properties of Gibbs States for Classical and Quantum Unbounded Spin Lattices*, J. Stat. Phys. **80**, 223 (1995).

[16] S. M. Bhattacharjee and A. Khare, *Fifty Years of the Exact Solution of the Two-Dimensional Ising Model by Onsager*, Curr. Sci. **69**, 816 (1995).

[17] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, *Canonical Typicality*, Phys. Rev. Lett. **96**, 050403 (2006).

[18] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, *Exact Relaxation in a Class of Nonequilibrium Quantum Lattice Systems*, Phys. Rev. Lett. **100**, 030602 (2008).

[19] N. Linden, S. Popescu, A. J. Short, and A. Winter, *Quantum Mechanical Evolution towards Thermal Equilibrium*, Phys. Rev. E **79**, 061103 (2009).

[20] M. Rigol, V. Dunjko, and M. Olshanii, *Thermalization and Its Mechanism for Generic Isolated Quantum Systems,* Nature (London) **452,** 854 (2008).

[21] V. I. Yukalov, *Equilibration and Thermalization in Finite Quantum Systems,* Laser Phys. Lett. **8,** 485 (2011).

[22] P. Reimann and M. Kastner, *Equilibration of Isolated Macroscopic Quantum Systems,* New J. Phys. **14,** 043020 (2012).

[23] A. Riera, C. Gogolin, and J. Eisert, *Thermalization in Nature and on a Quantum Computer,* Phys. Rev. Lett. **108,** 080402 (2012).

[24] M. P. Mueller, E. Adlam, L. Masanes, and N. Wiebe, *Thermalization and Canonical Typicality in Translation-Invariant Quantum Lattice Systems,* arXiv:1312.7420.

[25] L. Masanes, A. J. Roncaglia, and A. Acín, *The Complexity of Energy Eigenstates as a Mechanism for Equilibration,* Phys. Rev. E **87,** 032137 (2013).

[26] M. Troyer, F. Alet, S. Trebst, and S. Wessel, *Non-local Updates for Quantum Monte Carlo Simulations,* AIP Conf. Proc. **690,** 156 (2003).

[27] M. B. Hastings and T. Koma, *Spectral Gap and Exponential Decay of Correlations,* Commun. Math. Phys. **265,** 781 (2006).

[28] B. Nachtergaele and R. Sims, *Lieb-Robinson Bounds and the Exponential Clustering Theorem,* Commun. Math. Phys. **265,** 119 (2006).

[29] B. Nachtergaele, Y. Ogata, and R. Sims, *Propagation of Correlations in Quantum Lattice Systems,* J. Stat. Phys. **124,** 1 (2006).

[30] Z. Landau, U. Vazirani, and T. Vidick, *A Polynomial-Time Algorithm for the Ground State of 1D Gapped Local Hamiltonians,* arXiv:1307.5143.

[31] Everything also works with more technical assumptions for certain cases where the local Hilbert spaces are not finite dimensional. But, in this case, the local terms of the Hamiltonian need to be bounded in operator norm by a constant. Hence, our results are not applicable to bosonic systems.

[32] Y. M. Miranda and G. Slade, *The Growth Constants of Lattice Trees and Lattice Animals in High Dimensions,* Electron. Commun. Probab. **16,** 129 (2011).

[33] M. Penrose, *Self-Avoiding Walks and Trees in Spread-Out Lattices,* J. Stat. Phys. **77,** 3 (1994).

[34] T. Barthel and M. Kliesch, *Quasi-locality and Efficient Simulation of Markovian Quantum Dynamics,* Phys. Rev. Lett. **108,** 230504 (2012).

[35] D. Poulin, *Lieb-Robinson Bound and Locality for General Markovian Quantum Dynamics,* Phys. Rev. Lett. **104,** 190401 (2010).

[36] B. Nachtergaele, A. Vershynina, and V. A. Zagrebnov, *Lieb-Robinson Bounds and Existence of the Thermodynamic Limit for a Class of Irreversible Quantum Dynamics,* AMS Contemporary Mathematics **552,** 161 (2011).

[37] M. B. Hastings, *Decay of Correlations in Fermi Systems at Non-zero Temperature,* Phys. Rev. Lett. **93,** 126402 (2004).

[38] N. D. Mermin and H. Wagner, *Absence of Ferromagnetism or Antiferromagnetism in One-or Two-Dimensional Isotropic Heisenberg Models,* Phys. Rev. Lett. **17,** 1133 (1966).

[39] J. Fröhlich and T. Spencer, *The Kosterlitz-Thouless Transition in Two-Dimensional Abelian Spin Systems and the Coulomb Gas,* Commun. Math. Phys. **81,** 527 (1981).

[40] M. B. Hastings, *Solving Gapped Hamiltonians Locally,* Phys. Rev. B **73,** 085115 (2006).

[41] T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia, *Stability of Local Quantum Dissipative Systems,* arXiv:1303.4744.

[42] M. J. Kastoryano and J. Eisert, *Rapid Mixing Implies Exponential Decay of Correlations,* J. Math. Phys. (N.Y.) **54,** 102201 (2013).

[43] S. Michalakis and J. Pytel, *Stability of Frustration-Free Hamiltonians,* Commun. Math. Phys. **322,** 277 (2013).

[44] S. Bravyi, M. Hastings, and S. Michalakis, *Topological Quantum Order: Stability under Local Perturbations,* J. Math. Phys. (N.Y.) **51,** 093512 (2010).

[45] H. Li and F. D. M. Haldane, *Entanglement Spectrum as a Generalization of Entanglement Entropy: Identification of Topological Order in Non-Abelian Fractional Quantum Hall Effect States,* Phys. Rev. Lett. **101,** 010504 (2008).