Commun. Math. Phys. 276, 437–472 (2007)
Digital Object Identifier (DOI) 10.1007/s00220-007-0342-z

Communications in
Mathematical Physics

A Multi-Dimensional Lieb-Schultz-Mattis Theorem

Bruno Nachtergaele, Robert Sims

Department of Mathematics, University of California at Davis, Davis, CA 95616, USA.
E-mail: bxn@math.ucdavis.edu; rjsims@math.ucdavis.edu

Received: 21 August 2006 / Accepted: 26 April 2007
Published online: 9 September 2007 –

Abstract: For a large class of finite-range quantum spin models with half-integer spins, we prove that uniqueness of the ground state implies the existence of a low-lying excited state. For systems of linear size L, with arbitrary finite dimension, we obtain an upper bound on the excitation energy (i.e., the gap above the ground state) of the form $(C \log L)/L$. This result can be regarded as a multi-dimensional Lieb-Schultz-Mattis theorem [14] and provides a rigorous proof of the main result in [8].

1. Introduction and Main Result

1.1. Introduction. Ground state properties of Heisenberg-type antiferromagnets on a variety of lattices are of great interest in condensed matter physics and material science. Antiferromagnetic Heisenberg models are directly relevant for the low-temperature behavior of many materials, most notably the cuprates that exhibit high-T_c superconductivity [16].

There are several general types of ground states that are known, or expected, to occur in specific models: a disordered ground state or spin liquid, critical correlations (power law decay), dimerization (spin-Peierls states), columnar phases, incommensurate phases, and Néel order. More exotic phenomena such as chiral symmetry breaking have also been considered [21, 22].

Which behavior occurs in a given model depends on the lattice, in particular the dimension and whether or not the lattice is bipartite, on the type of spin (integer versus half-integer) and, of course, also on the interactions. In this paper we are considering a class of half-integral spin models (or models where the magnitude of at least some of the spins is half-integral). Our aim is to prove a generalization of the Lieb-Schultz-Mattis Theorem [14]. Such a generalization was presented by Hastings in [8] and a substantial part of our proof is based on his work. Our main contribution is to provide what we hope is a more transparent argument which in addition is mathematically rigorous.

The well known theorem by Lieb and Mattis [13] implies, among other things, that the ground state of the Heisenberg antiferromagnet on a bipartite lattice with isomorphic sublattices, is non-degenerate. For one-dimensional and quasi-one-dimensional systems of even length and with half-integral spin Affleck and Lieb [1], generalizing the original result by Lieb, Schultz, and Mattis [14], proved that the gap in the spectrum above the ground state is bounded above by $constant/L$. A vanishing gap can be expected to lead to a gapless continuous spectrum above the ground state in the thermodynamic limit. Such an excitation spectrum is generically associated with power-law (as opposed to exponential) decay of correlations. Aizenman and Nachtergaele proved for the spin-1/2 antiferromagnetic chain that if translation invariance is not broken (in particular, when the ground state is unique), the spin-spin correlation function can decay no faster than $1/r^3$ [3]. In other words, uniqueness of the ground state implies slow (power-law) decay of correlations. Recently, it was proved rigorously that a non-vanishing spectral gap implies exponential decay of correlations [9, 20]. Therefore, non-exponential decay of correlations implies the absence of a gap. In particular, the result by Aizenman and Nachtergaele implies the absence of a gap in the infinite spin-1/2 antiferromagnetic chain if the translation invariance is not broken, e.g., if the ground state is unique. This result can be generalized to an interesting class of antiferromagnetic chains of half-integer spins [17]. The Lieb-Schultz-Mattis Theorem has also been extended to fermion systems on the lattice [24, 25]. All these results are for one-dimensional systems. The bulk of the applications of the spin-1/2 Heisenberg antiferromagnet is in two-dimensional physics and therefore, the rigorous proof we provide here, based in part on ideas of Hastings [8], should be of considerable interest as it is applicable to higher-dimensional models.

The most common argument employed to bound a spectral gap from above uses the variational principle. Often, the variational state is a perturbation of the ground state. The proofs in [14] and [1] are of this kind. However, since the ground state is not known, and no assumptions are made about it except for its uniqueness, these proofs are not a variational calculation in the usual sense. The variational states are defined by acting with suitable local operators A on the (unknown) ground state.

For a finite volume Hamiltonian H_L generated by a potential Φ of the type we consider (see the paragraph including (1.7) and (1.8) in Sect. 1.2 for the relevant definitions), and with a unique ground state, it is straightforward to show that the gap above the ground state, γ_L, is bounded uniformly in L. To see this, note that for any ground state vector Ω and for any site x, there exists a unitary on the state space of x with vanishing expectation in the state Ω, i.e., $U\Omega \perp \Omega$. Since Ω is the unique ground state by assumption, $U\Omega$ is a variational state for the gap. Therefore, we have the bound

$$\gamma_L \leq \langle \Omega, [H_L, U]\Omega \rangle \leq 2 \inf_x \sum_{X \ni x} \|\Phi(X)\| \leq 2|||\Phi|||_1 , \tag{1.1}$$

which is uniform in the system size L. Here, $|||\Phi|||_1$ is as defined in (1.13). See Sect. 5.5 for the proof that such a unitary exists.

In order to obtain a better bound on the energy of the first excited state one has to exploit the few properties assumed of the ground state, such as its uniqueness and symmetries. Furthermore, one must show that any proposed variational state has a sufficiently large component in the orthogonal complement of the ground state. In Sect. 2.2, we propose a variational state for finite systems of size L and then demonstrate the relevant estimates, as mentioned above, in Sects. 3 and 4. It is interesting to note that the energy estimate we obtain will itself contain the spectral gap of the finite system in such a way that assuming a large gap leads to an upper bound less than the assumed gap. From this contradiction one can conclude an upper bound on the finite-volume gaps.

Our results apply to a rather general class of models, which we will define precisely in the next section. The application of our general result to spin-1/2 Hamiltonians with translation invariant (or periodic) isotropic finite-range spin-spin interactions on a d-dimensional lattice is easy to state. First, let $\Lambda_L = [1, L] \times V_L$ with L even and with periodic boundary conditions in the 1-direction, i.e., in the direction that is of even size. It will be important that the number of spins in V_L, $|V_L|$, is odd, and satisfies $|V_L| \leq cL^{d-1}$, for some $d \geq 1$ and a suitable constant c. Assuming that the model defined on Λ_L has a unique ground state, we prove that the spectral gap γ_L satisfies the bound

$$\gamma_L \leq C\frac{\log L}{L}, \tag{1.2}$$

where C depends on d and the specifics of the interaction, but not on L.

Because of the presence of the factor $\log L$, the bound (1.2) applied to one-dimensional models does not fully recover the original Lieb-Schultz-Mattis Theorem in [14] or the bound proved by Affleck and Lieb in [1]. This indicates that in general our bound is not optimal. Our proof uses in an essential way Lieb-Robinson bounds [9, 15, 20], as does Hastings' argument in [8], and the appearance of the factor $\log L$ seems to be an inevitable consequence of this. In fact, it is known that the standard Heisenberg antiferromagnets with spin ≥ 1 on the two-dimensional square lattice or with spin $\geq 1/2$ on $\mathbb{Z}^d$ with $d \geq 3$, have Néel ordered ground states [5, 10] and in that case one can show that the gap is bounded by C/L (see, e.g., [11, 12]).

1.2. Setup and main result. The arguments we develop below can be applied to a rather general class of quantum spin Hamiltonians defined on a large variety of lattices. We believe it is useful to present them in a suitably general framework which applies to many interesting models. Attempting to be as general as possible, however, would lead us into a morass of impenetrable notation. Therefore, we have limited the discussion of further generalizations to some brief comments in Sect. 1.5.

We assume that the Hamiltonians describe interactions between spins that are situated at the points of some underlying set Λ. For simplicity, one may think of $\Lambda = \mathbb{Z}^d$, but we need only assume that the set Λ has one direction of translational invariance, which we will refer to as the 1-direction. We assume that there is an increasing sequence of sets $\{\Lambda_L\}_{L=1}^{\infty}$ which exhaust Λ of the form $\Lambda_L = [1, L] \times V_L$, where $|V_L| \leq cL^{d-1}$ for some $d \geq 1$. Here each $x \in \Lambda_L$ can be written as $x = (n, v)$, where $n \in \{1, 2, \ldots, L\}$ and $v \in V_L$, and we will denote by (n, V_L) the set of all $x \in \Lambda_L$ of the form $x = (n, v)$ for some $v \in V_L$.

Estimates on the decay of correlations in the ground state and Lieb-Robinson bounds on the dynamics will play an important role in the proof of the main result. Both are expressed in terms of a distance function on Λ, which we will denote by d. Often, Λ has the structure of a connected graph and $d(x, y)$ is the minimum number of edges in a path from x to y. In any case, we will assume that d is a metric and furthermore that there is a function $F : [0, \infty) \to (0, \infty)$ satisfying the following two conditions.
Condition F1: F is uniformly integrable over Λ in the sense that

$$\| F \| := \sup_{x\in\Lambda} \sum_{y\in\Lambda} F(d(x, y)) < \infty. \tag{1.3}$$

Condition F2: F satisfies

$$C(F) := \sup_{x,y\in\Lambda} \sum_{z\in\Lambda} \frac{F\,(d(x,z))\ F\,(d(z,y))}{F\,(d(x,y))} < \infty, \tag{1.4}$$

which means that the "convolution" of F with itself is bounded by a multiple of itself.

F1 and F2 are restricitve conditions only when Λ is infinite, however, for finite Λ, the constants $\|F\|$ and $C(F)$ will be useful in our estimates. It is also important to note that for any given set Λ and function F that satisfies F1 and F2 above, we can define a one-parameter family of functions, F_λ, $\lambda \geq 0$, by

$$F_\lambda(x) := e^{-\lambda x}\, F(x), \tag{1.5}$$

and easily verify that $F1$ and $F2$ hold for F_λ, with $\|F_\lambda\| \leq \|F\|$ and $C_\lambda(F) \leq C(F)$.

As a concrete example, take $\Lambda = \mathbb{Z}^d$ and $d(x,y) = |x-y|$. In this case, one may take the function $F(x) = (1+x)^{-d-\varepsilon}$ for any $\varepsilon > 0$. Clearly, (1.3) is satisfied, and a short calculation demonstrates that (1.4) holds with

$$C(F) \leq 2^{d+\varepsilon+1} \sum_{n\in\mathbb{Z}^d} \frac{1}{(1+|n|)^{d+\varepsilon}}. \tag{1.6}$$

Each $x \in \Lambda$ is assigned a finite-dimensional Hilbert space $\mathcal{H}_x$. For any finite subset $X \subset \Lambda$, the Hilbert space associated with X is the tensor product $\mathcal{H}_X = \bigotimes_{x\in X} \mathcal{H}_x$, and the set of corresponding observables supported in X is denoted by $\mathcal{A}_X = \mathcal{B}(\mathcal{H}_X)$, the bounded linear operators over $\mathcal{H}_X$. These local observables form an algebra, and with the natural embedding of $\mathcal{A}_{X_1}$ in $\mathcal{A}_{X_2}$ for any $X_1 \subset X_2$, one can define the C^*-algebra of all observables, $\mathcal{A}$, as the norm completion of the union of all local observable algebras $\mathcal{A}_X$ for finite $X \subset \Lambda$. Since we have assumed that Λ_L is of the form $[l,r] \times V_L$ with $r - l = L - 1$, we can define translation automorphisms τ_n, for $n \in \mathbb{Z}$, which map $\mathcal{A}_{(m,V_L)}$ into $\mathcal{A}_{(n+m,V_L)}$ for all $m \in \mathbb{Z}$.

An interaction for the system is a map Φ from the finite subsets of Λ to $\mathcal{A}$ such that for each finite $X \subset \Lambda$, $\Phi(X)^* = \Phi(X) \in \mathcal{A}_X$. For given Λ and F, and any $\lambda \geq 0$, let $\mathcal{B}_\lambda(\Lambda)$ be the set of interactions that satisfy

$$\|\Phi\|_\lambda := \sup_{x,y\in\Lambda} \sum_{X\ni x,y} \frac{\|\Phi(X)\|}{F_\lambda\,(d(x,y))} < \infty. \tag{1.7}$$

All interactions considered in this paper are assumed to belong to $\mathcal{B}_\lambda(\Lambda)$ for some choice of F and $\lambda > 0$. The constant $\|\Phi\|_\lambda$ will show up in many estimates. The finite volume Hamiltonians are defined in terms of the interaction Φ in the usual way by

$$H_L = \sum_{X\subset\Lambda_L} \Phi(X) + \text{ boundary terms}. \tag{1.8}$$

We will always assume periodic boundary conditions in the 1-direction and arbitrary boundary conditions in the other directions (i.e., any boundary terms in the other directions are included in the definition of Φ).

The condition that $\|\Phi\|_\lambda$ is finite is sufficient to guarantee the existence of the dynamics in the thermodynamic limit as a one-parameter group of automorphisms on $\mathcal{A}$. In particular this means that the limits

$$\alpha_t^{\Phi}(A) := \lim_{L\to\infty} \alpha_t^{\Phi,L}(A) := \lim_{L\to\infty} e^{itH_L} A e^{-itH_L} \tag{1.9}$$

exist in norm for all $t \in \mathbb{R}$, and all observables $A \in \mathcal{A}_X$, for any finite $X \subset \Lambda$. We will often suppress the L or Φ dependence in the notation $\alpha_t^{\Phi,L}$. See [4, 18, 23] for more details.

Next, we turn to a set of conditions that more specifically describe the class of models to which the Lieb-Schultz-Mattis Theorem may be applied.

Condition LSM1: We assume that the interaction is translation invariant in at least one direction, which we will take to be the 1-direction. This means

$$\Phi\left(X+e_1\right)=\tau_1\left(\Phi(X)\right), \tag{1.10}$$

where, for any $X \subset \Lambda$, $X + e_1$ is translation of all points in X by one unit in the 1-direction. We will consider finite systems with Hamiltonians H_L defined with periodic boundary conditions in the 1-direction. For convenience of the presentation we will assume free boundary conditions in the other directions but this is not crucial. Since we have assumed periodicity in the 1-direction, we can implement the translation invariance for finite systems by a unitary $T \in \mathcal{A}_{\Lambda_L}$ such that $\Phi(X+e_1) = T^*\Phi(X)T$, for all $X \subset \Lambda_L$. Here T depends on L, but we suppress this dependence in the notation.

Condition LSM2: The interactions are assumed to be of finite range in the 1-direction, i.e., there exists $R > 0$ (the range), such that if $X \subset \Lambda$ and $X \ni x_i = (n_i, v_i)$ for $i = 1, 2$ with $|n_1 - n_2| \geq R$, then $\Phi(X) = 0$.

Condition LSM3: We assume rotation invariance about one axis. More precisely, we assume that there is a hermitian matrix in every $\mathcal{A}_{\{x\}}$, $x \in \Lambda$, which we will denote by S_x^3, with eigenvalues that are either all integer or all half-integer (i.e. belonging to $\mathbb{Z}+1/2$). We also require that $\tau_m(S_x^3) = S_{x+me_1}^3$. Define, for $\theta \in \mathbb{R}$, the unitary $U(\theta) \in \mathcal{A}_{\Lambda_L}$ by

$$U(\theta)=\bigotimes_{x\in\Lambda_L} e^{i\theta S_x^3}. \tag{1.11}$$

The interaction is taken to be rotation invariant in the sense that for each finite $X \subset \Lambda$,

$$U^*(\theta)\Phi(X)U(\theta)=\Phi(X) \ \text{ for all } \theta\in\mathbb{R}. \tag{1.12}$$

Condition LSM4: We assume that the S_x^3 are uniformly bounded: there exists S such that $\|S_x^3\| \leq S$, for all $x \in \Lambda$. The following condition, which we call *odd parity*, is crucial: define the parity of x, p_x to be 0 if the eigenvalues of S_x^3 are integers, and $p_x = 1/2$ if they are half-integers. We assume that $\sum_{v\in V_L} p_{(n,v)} \in \mathbb{Z} + 1/2$, for all $n \in \mathbb{Z}$. The simplest and most important case where this is satisfied is when we have a spin 1/2 at each site, and $|V_L|$ is odd.

Condition LSM5: The ground state of H_L is assumed non-degenerate. This implies it is an eigenvector of the translation T and rotations $U(\theta)$. Without loss of generality we can assume that 1 is the corresponding eigenvalue of T (if the eigenvalue is $e^{i\phi}$, replace T by $e^{-i\phi}T$). We also assume that the ground state has eigenvalue 1 for the rotations $U(\theta)$.

Condition LSM6: We assume that there are orthonormal bases of the Hilbert spaces $\mathcal{H}_{\Lambda_L}$ with respect to which S_x^3 and $\Phi(X)$ are *real*, for all $x \in \Lambda_L$, $X \subset \Lambda_L$. This condition is only used in the proof of Lemma 1. Therefore, this condition may be replaced by the property proved in that lemma.

We will also use the following quantities:

$$|||\Phi|||_1 := \sup_{x\in\Lambda}\sum_{X\ni x}\|\Phi(X)\| < \infty, \tag{1.13}$$

and

$$|||\Phi|||_2 := \sup_{x\in\Lambda} \sum_{X\ni x} |X| \sum_{x'\in X} \|[S^3_{x'}, \Phi(X)]\| < \infty. \tag{1.14}$$

It is not hard to show that the conditions F1 and F2 are sufficient to imply that $|||\Phi|||_1$ and $|||\Phi|||_2$ are finite.

We can now state our main result.

Theorem 1. *Let γ_L be the spectral gap, i.e., the difference between the lowest and next-lowest eigenvalue of the Hamiltonian H_L of a model satisfying conditions F1, F2, and LSM1-6. Then, there exists a constant C, depending only on properties of Λ (such as the dimension), the constants $\|F\|$ and $C(F)$, and the interaction ($\|\Phi\|_\lambda$, for some $\lambda > 0$, $|||\Phi|||_1$, and $|||\Phi|||_2$), such that*

$$\gamma_L \le C\frac{\log L}{L}. \tag{1.15}$$

1.3. Structure of the proof. The simplest way to present the proof is as a proof by contradiction. Under the assumption that there exists a sufficiently large constant $C > 0$, such that γ_L *exceeds* $(C \log L)/L$ for large L, we will construct a state orthogonal to the ground state with an energy difference that is boundable by a quantity that is *strictly less than* the assumed gap for sufficiently large L. Thus, the proof is in essence a variational argument. The variational state is constructed as a perturbation of the ground state, as the solution of the differential equation proposed by Hastings [8] with the ground state as initial condition (see Sect. 2, in particular (2.25), for this equation). The important idea is that this equation will lead to a state which resembles the ground state of the Hamiltonian with twisted rather than periodic boundary conditions (see Sect. 2.1 for the definition of the twists), at least in part of the system, say the left half. In the right half the ground state will be left essentially unperturbed. This state is defined in Sect. 2.

After the variational state has been defined, there are two main steps in the proof: estimating its excitation energy and verifying that it is "sufficiently orthogonal" to the ground state. In general, one may also have to consider the normalization of the variational state, but in our case the differential equation defining it will be manifestly norm preserving. Hence, this is not an issue for our proof.

The main difficulty is that under the general assumptions we have made, no explicit information about the ground state is available. Its uniqueness, translation, and rotation invariance are the only properties we can use. In combination with the general assumptions on the interactions and the assumption on the magnitude of the spectral gap above the ground state, however, one can obtain an upper bound on the decay of correlations of the ground state in the 1-direction. The recently proved Lieb-Robinson bounds [9, 18, 20] will be essential to show that the effects of the perturbations we define in the left half of the system remain essentially localized there. This allows us to compare the energy of the variational state with the ground state energy of a Hamiltonian, $H_{\theta,-\theta}$ introduced in (2.5), which, instead of twisted boundary conditions, has two twists that cancel each other. The twisted Hamiltonian is unitarily equivalent to the original one and therefore has the same ground state energy. We work out this argument in Sect. 3. The result is

$$|\langle\psi_1, H_L\psi_1\rangle - E_0| \le CL^\nu e^{-c\gamma_L L}\ (1 + \textit{corrections}), \tag{1.16}$$

where ψ_1 is the normalized variational state we construct, and E_0 is the ground state energy. The dependence of both quantities on L is surpressed in the notation. ν, C and c are positive constants that only depend on properties of the lattice and the interactions. The *correction terms* appearing above, and also in (1.17) below, can be made explicit by the estimates provided in Sect. 5. They depend on the quantity $\gamma_L L$ in such a way that assuming there exists a constant $C > 0$ for which $\gamma_L L \geq C$ for sufficiently large L, they are uniformly bounded in L. Due to the nature of our proof of Theorem 1, see below, we do not write these additional terms out explicitly.

For the orthogonality, our strategy is to show that ψ_1 is almost an eigenvector of the translation T with eigenvalue -1. Since the ground state ψ_0 is an eigenvector of T with eigenvalue 1, by assumption, this shows that ψ_1 is nearly orthogonal to ψ_0. In Sect. 4 we obtain a bound on their inner product of the form:

$$|\langle \psi_1, \psi_0 \rangle| \leq C' L^{\nu'} e^{-c' \gamma_L L} \ (1 + corrections) , \tag{1.17}$$

where ν', C' and c' are positive constants similar to ν, C and c.

The proof of Theorem 1 then easily follows.

Proof of Theorem 1. Suppose that $\gamma_L L \geq C \log L$ with a sufficiently large constant C. In this case, the *correction terms* which appear in the bounds (1.16) and (1.17) above are negligible. It is easy to see then that one obtains a contradiction for L large enough. □

To help the reader see the forest through the trees we have tried to streamline the estimates in Sects. 3 and 4 by collecting some results of a more technical nature in Sect. 5.

1.4. Examples. The conditions LSM1-6 we have imposed on the models are not unreasonable. We will illustrate this by considering various antiferromagnetic Heisenberg models defined on $\Lambda_L = [1, L] \times V_L$, where for each L, V_L is a finite set. As before, at each $x \in \Lambda_L$, we have a finite-dimensional spin with spin-matrices S_x^i, $i = 1, 2, 3$, and we consider Hamiltonians of the form

$$H = \sum_{x,y \in \Lambda_L, x \neq y} J(x, y) \boldsymbol{S}_x \cdot \boldsymbol{S}_y, \tag{1.18}$$

where $J(x, y) \in \mathbb{R}$ are the coupling constants.

If $V_L \subset \mathbb{Z}^{d-1}$, with $d \geq 1$, and such $|V_L| \leq cL^{d-1}$, for a suitable constant c, which describes the case for d-dimensional systems defined on subsets of $\mathbb{Z}^d$, there exists a function F satisfying Conditions F1 and F2 as we have indicated in the paragraph containing (1.6). It is also easy to see that if V_L is a fixed finite set independent of L, in which case the system is (quasi) one-dimensional, any function F that works for the one-dimensional lattice will suffice. All the examples we discuss below will be of this type.

Certainly, there are still many Hamiltonians of the form (1.18) that fail to satisfy all six conditions, but this is generally for a good reason. For example, without translation invariance in at least one direction one can easily have a non-vanishing gap above the ground state.

Condition LSM2, finite-range, does not need to be satisfied in the strict sense. Sufficiently rapidly decaying interactions could also be treated. For the present discussion, let's assume that the model is translation invariant in the 1-direction, and that the interactions are nearest neighbor in the 1-direction in the sense that for any $x = (n_1, v)$ and $y = (n_2, u)$, with $|n_1 - n_2| > 1$, we have $J(x, y) = 0$.

The rotation invariance about at least one axis imposed in Condition LSM3 is essential for the type of result we prove. The models (1.18) have full rotation invariance, so they clearly satisfy this condition. Anisotropic models of the XXZ type would still satisfy LSM3.

In order to satisfy LSM4, we have to assume a uniform bound on the size of the spin. Clearly, all models with only one kind of spins or a periodic arrangement of spin magnitudes satisfy this condition. Since we already assumed translation invariance in the 1-direction, we can verify the *odd parity* condition by adding the magnitudes of all spins in the "slice" $(1, V_L)$. If we have only spin 1/2's, e.g., we simply need that $|V_L|$ is odd. For the one-dimensional chains of identical spins of magnitude S, the condition requires that S is half-integral. Haldane's Conjecture [6, 7] predicts that for integer values of S there exists a non-vanishing gap. There are examples of isotropic integer-spin chains which satisfy all the other conditions and for which the existence of a non-vanishing gap has been rigorously established, such as the AKLT chain [2]. For p-periodic spin chains with a repeating pattern of spin magnitudes $S_1, \ldots, S_p$, LSM4 is satisfied if $S_1 + \cdots + S_p$ is half-integral. Similarly, for spin ladders LSM4 is satisfied if the total spin in each rung is half-integral.

There is a large class of models for which the uniqueness of the ground state demanded by LSM5 follows from the Lieb-Mattis Theorem [13]. For Hamiltonians of the form (1.18), a simple case where the Lieb-Mattis Theorem applies is the following: if Λ_L is the union of two disjoint subsets $\Lambda_{L,A}$ and $\Lambda_{L,B}$ of equal size, with $J(x, y) \leq 0$ whenever x and y do not belong to the same subset, and sufficiently many $J(x, y)$ are non-vanishing such that the graph formed by the edges with non-zero coupling constants is connected. This is satisfied if $V_L \subset \mathbb{Z}^{d-1}$ is connected and the Hamiltonian is the usual nearest neighbor antiferromagnetic Heisenberg model.

All models of the form (1.18) satisfy LSM6.

The above discussion demonstrates that there is a large variety of models that satisfy all conditions of our main theorem. In particular, all nearest-neighbor half-integer spin Heisenberg antiferromagnets defined on subsets $\Lambda_L = [1, 2L] \times V_L$ of d-dimensional hypercubic lattice with $|V_L|$ odd and such that $|V_L| \leq cL^\alpha$, for some $\alpha \geq 0$ (it is natural but not necessary to assume $\alpha = d - 1$), have a unique ground state with a gap γ_L above it satisfying $\gamma_L \leq C(\log L)/L$, for some constant C.

1.5. Generalizations. One can envision several generalizations of Theorem 1. An obvious one is to remove the condition that the interaction is strictly finite range in the 1-direction. It is not hard to see that the arguments given in the following sections can be extended to long-range interactions with sufficiently fast decay.

One may wonder whether the assumption that L is *even* is essential. It is used in the proof of near orthogonality of the variational state, which is based on investigating the behavior under translations of the state: the variational state is close to an eigenvector with eigenvalue -1 of the translation operator T, whereas the ground state has eigenvalue 1. Our proof of this fact assumes that the ground state is an eigenvector of the rotations with eigenvalue 1. For L odd, our assumptions preclude the existence of such an eigenvector. However, it seems plausible that for odd L a slight modification of our proof will work to show that the ground state and the variational excited state have opposite eigenvalues for translations.

The main applications we think of are to SU(2)-invariant Hamiltonians with antiferromagnetic interactions. Affleck and Lieb [1] pointed out that their proof easily extends to a class of models with SU(N) symmetry. There are no obstructions to generalizing our arguments to such models with SU(N) symmetry given by suitable representations.

It may also be of interest to consider different topologies of the underlying lattice and/or the twistings. Instead of cylindrical systems with periodic boundary conditions which can be deformed by a twist, one could apply a similar strategy to systems defined on a ball or a sphere. We do not explore such possibilities here.

Another question we do not address in this paper is under what circumstances the trial state we construct is actually a good approximation of a low-lying eigenstate with energy close to the first excited state, or even whether it is a state orthogonal to the ground state and with energy bounded by $C(\log L)/L$. We do not believe that statements of this kind hold under the general conditions we impose. It is expected that in some cases the true gap of the system is much smaller than the bound we prove. This is of course not in contradiction with our result, but under such circumstances our method does not provably construct a good variational state. There is no reason to assume that it always would.

2. Construction of the Variational State

2.1. Twisted Hamiltonians. The main motivation behind the construction of the variational excited state is that it should resemble the ground state of the model with twisted (as opposed to periodic) boundary conditions. Therefore, we first describe some elementary properties of a family of perturbations of the Hamiltonian, which we will call twisted Hamiltonians for reasons that will become obvious.

Given an interaction Φ which satisfies the general assumptions outlined in Sect. 1.2, we will now define a two parameter family of "twisted Hamiltonians" to analyze. These Hamiltonians will be defined on a finite volume $\Lambda_L = [1, L] \times V_L$, where $[1, L]$ is considered with periodic boundary conditions for some even $L > 4R$, where $R > 0$ is the range of Φ in the 1-direction. Let Φ_L be the periodic extension of Φ restricted to Λ_L. Recall that each point $x \in \Lambda_L$ can be written as $x = (n, v)$, where $n \in \{1, 2, \ldots, L\}$ and $v \in V_L$, and we will denote by $(n, V_L) = \{x \in \Lambda_L \,:\, x = (n, v) \text{ for some } v \in V_L\}$. For any $\theta \in \mathbb{R}$ and $n \in \{1, 2, \ldots, L\}$, define the "column" rotations $U_n(\theta)$ by

$$U_n(\theta) = \bigotimes_{x \in (n, V_L)} e^{i\theta S_x^3}. \tag{2.1}$$

For $m \in \{1, 2, \ldots, L-1\}$, we will denote by $V_m(\theta)$ the unitary given by

$$V_m(\theta) = \bigotimes_{m < n \leq L} U_n(\theta). \tag{2.2}$$

The "twisted Hamiltonians" are defined to be perturbations of the initial Hamiltonian with periodic boundary conditions defined by

$$H = \sum_{X \subset \Lambda_L} \Phi_L(X). \tag{2.3}$$

The perturbations have the following form:

$$H_\theta(m) := \sum_{X \subset \Lambda_L} V_m(\theta)^* \Phi(X) V_m(\theta) - \Phi(X), \tag{2.4}$$

for $m \in [R, L-R]$ to avoid interactions across the seam created by identifying $L+1$ with 1. Note that here we use the original potential Φ, and not its periodic extension Φ_L. Clearly, if $X \subset \bigcup_{m<n\leq L}(n, V_L)$ or $X \subset \bigcup_{1\leq n<m}(n, V_L)$, then $V_m(\theta)^*\Phi(X)V_m(\theta) - \Phi(X)$ will vanish by rotation invariance of the interaction, and therefore only those interactions across the column (m, V_L) contribute in (2.4). For $\theta, \theta' \in \mathbb{R}$ and $m \in \{R, R+1, \ldots, L/2 - R\}$ fixed, we define

$$H_{\theta,\theta'} := H + H_\theta(m) + H_{\theta'}(m + L/2), \tag{2.5}$$

to be a doubly twisted Hamiltonian. With m fixed, we regard Λ_L as the disjoint union of two sets

$$\Lambda_L = \Lambda_L^{(W)} \cup \Lambda_L^{(S)}, \tag{2.6}$$

where $\Lambda_L^{(W)}$ consists of two windows, one about each column at which a twist has been applied; namely

$$\Lambda_L^{(W)} := \Lambda_L^{(W)}(m) \cup \Lambda_L^{(W)}(m + L/2) \quad \text{and} \quad \Lambda_L^{(W)}(y) := \bigcup_{|n-y|\leq \frac{L}{4}-R} (n, V_L), \tag{2.7}$$

for $y \in \{m, m + L/2\}$. Moreover, $\Lambda_L^{(S)}$ comprises the remaining strips in Λ_L. Given this decomposition of the underlying space, the twisted Hamiltonian can be written as

$$H_{\theta,\theta'} = H_{\theta,\theta'}^{(W)} + H^{(S)}, \tag{2.8}$$

where

$$H^{(S)} = \sum_{\substack{X\subset\Lambda_L: \\ X\cap\Lambda_L^{(S)}\neq\emptyset}} \Phi_L(X), \tag{2.9}$$

and $H_{\theta,\theta'}^{(W)}$ denotes the remaining terms in $H_{\theta,\theta'}$ which, due to (2.9), are supported strictly within the windows.

There are a variety of useful symmetries the Hamiltonians $H_{\theta,\theta'}$, introduced in (2.5), possess. With $m \in \{R, R+1, \ldots, L/2 - R\}$ fixed as above, one may define

$$W(\phi) := \bigotimes_{m<n\leq m+L/2} U_n(-\phi), \tag{2.10}$$

for any real ϕ. See (2.1) for the definition of the column rotations U_n. It is easy to check that for any angles $\theta, \theta', \phi \in \mathbb{R}$, one has that

$$W^*(\phi)\, H_{\theta,\theta'}\, W(\phi) = H_{\theta-\phi,\theta'+\phi}, \tag{2.11}$$

due to the (term by term) rotation invariance of the interactions. Given this relation, it is clear that along the path $\theta' = -\theta$ the twisted Hamiltonian is unitarily equivalent to the untwisted Hamiltonian, i.e.,

$$W(\theta)^*\, H_{\theta,-\theta}\, W(\theta) = H_{0,0} = H, \tag{2.12}$$

which, due to the periodic boundary conditions, is not true for general pairs θ, θ'.

The untwisted Hamiltonian is translation invariant (in the 1-direction), and it is important that the twisted Hamiltonians inherit a "twisted" translation invariance. Define

$$T_{\theta,\theta'} = T\, U_m(\theta)\, U_{m+L/2}(\theta'), \tag{2.13}$$

where T is the unitary implementing the translation by 1 in the 1-direction. It is then straightforward to verify that

$$H_{\theta,\theta'} = T^*_{\theta,\theta'}\, H_{\theta,\theta'}\, T_{\theta,\theta'}. \tag{2.14}$$

Note that under the odd parity condition LSM4 we have

$$T_{2\pi,0} = -T, \tag{2.15}$$

which will be important in the proof of the almost orthogonality of the trial state in Sect. 4.

If we denote by ψ_0 the (unique) ground state of H, i.e., $H\psi_0 = E_0\psi_0$, then by translation invariance, and specifically LSM5, we have that $T\psi_0 = \psi_0$. Moreover, using the unitary equivalence (2.12), we see that the ground state of the twisted Hamiltonian $H_{\theta,-\theta}$ satisfies $H_{\theta,-\theta}\psi_0(\theta,-\theta) = E_0(\theta,-\theta)\psi_0(\theta,-\theta)$ with $E_0(\theta,-\theta) = E_0$ and $\psi_0(\theta,-\theta) = W(\theta)\psi_0$. Although the twisted ground state $\psi_0(\theta,-\theta)$ is not translation invariant, it does satisfy invariance with respect to the twisted translations, i.e., $T_{\theta,-\theta}\psi_0(\theta,-\theta) = \psi_0(\theta,-\theta)$. As a consequence, we have the following simple but important property of E_0.

Lemma 1. *Let $E_0(\theta,\theta')$ denote the ground state energy of $H_{\theta,\theta'}$.Then, the partial derivatives of E_0 vanish on the line $\theta' = -\theta$:*

$$\partial_1 E_0(\theta,-\theta) = \partial_2 E_0(\theta,-\theta) = 0. \tag{2.16}$$

Proof. First, we note that E_0 is differentiable in its two variables in a neighborhood of $(0,0)$ by the non-degeneracy condition LSM5. By unitary equivalence E_0 is then differentiable in a neighborhood of the line $(\theta,-\theta)$. For $\psi,\phi \in \mathbb{R}$, let $\mathcal{E}(\psi,\phi) = E_0(\psi-\phi,\psi+\phi)$ denote the ground state energy of $H_{\psi-\phi,\psi+\phi}$. Due to the unitary equivalence Eq. (2.11), $\mathcal{E}$ depends only on ψ. Hence, $\partial_\phi\mathcal{E}(\psi,\phi) = 0$, for all ψ,ϕ. Under the additional assumption that the interactions $\Phi(X)$ are *real* (LSM6), we have that $\overline{H_{\theta,\theta'}} = H_{-\theta,-\theta'}$, and therefore $\mathcal{E}(\psi,0) = \mathcal{E}(-\psi,0)$. Hence, $\mathcal{E}$ is an even function of ψ and $\partial_\psi\mathcal{E}(\psi,\phi)|_{\psi=0} = 0$. Using these properties and the fact that the partial derivatives of E_0 are linear combinations of the partial derivatives of $\mathcal{E}$, we find that both partial derivatives of E_0 vanish on the line $\theta' = -\theta$. □

2.2. The variational state. Our aim is to construct a state that resembles the ground state of $H_{\theta,-\theta}$ in a region surrounding those spins that were twisted by an angle of θ, while it otherwise resembles the ground state of $H = H_{0,0}$.

From the unitary equivalence (2.12) we have that $E_0(\theta,-\theta)$ is independent of θ, i.e., $\partial_\theta E_0(\theta,-\theta) = 0$. Moreover, the partial derivatives of E_0 vanish on the line $(\theta,-\theta)$, as was proven in Lemma 1. This property, in general, allows one to derive an equation for the ground state.

Consider a differentiable one-parameter family of self-adjoint operators $H(x)$, $x \in [a,b] \subset \mathbb{R}$, and let $E_0(x)$ denote the ground state energy of $H(x)$ with a differentiable family of ground state eigenvectors $\psi_0(x)$. Suppose $\partial_x E_0(x) = 0$ for $x \in [a, b]$. Then, it is easy to see that $\psi_0(x) \perp (\partial_x H(x))\psi_0(x)$, from which we obtain:

$$\partial_x \psi_0(x) \;=\; -\; [H(x) - E_0(x)]^{-1}\; \partial_x H(x)\, \psi_0(x). \tag{2.17}$$

For any vector ψ, this leads to

$$\begin{aligned}\langle \psi, \partial_x \psi_0(x)\rangle &= -\int_{E_0(x)}^{\infty} \frac{1}{E - E_0(x)}\, d\,\left\langle \psi, P_E^x \partial_x H(x)\psi_0(x)\right\rangle \\ &= -\int_{E_0(x)}^{\infty}\int_0^{\infty} e^{-(E-E_0(x))t}\, dt\, d\,\left\langle \psi, P_E^x \partial_x H(x)\psi_0(x)\right\rangle \\ &= -\int_0^{\infty} \left\langle \psi, \alpha_{it}^x\,(\partial_x H(x))\,\psi_0(x)\right\rangle dt,\end{aligned} \tag{2.18}$$

where P_E^x is the spectral resolution for $H(x)$ and α_{it}^x is the imaginary-time evolution corresponding to the Hamiltonian.

Motivated by this calculation, we introduce the family of operators $B(A, H)$, where H is a Hamiltonian for which the dynamics $\{\alpha_t \mid t \in \mathbb{R}\}$ exists as a strongly continuous group of $*$-automorphisms and A is any local observable, defined by

$$B(A, H) = -\int_0^{\infty} \alpha_{it}(A)dt, \tag{2.19}$$

where α_{it} is the imaginary time evolution generated by H. For unbounded Hamiltonians H, it may not be obvious that $B(A, H)$ can be defined on a dense domain. However, if ψ is a ground state corresponding to the Hamiltonian H, then $B(A, H)\psi$ exists. Moreover, from (2.18), we conclude that

$$\partial_x \psi_0(x) = B(x)\psi_0(x), \tag{2.20}$$

where $B(x) = B\,(\partial_x H(x), H(x))$. Similarly, in the density matrix formalism, for

$$\rho_0(x) \;:=\; |\psi_0(x)\rangle\,\langle\psi_0(x)|, \tag{2.21}$$

Eq. (2.20) implies that

$$\partial_x \rho_0(x) \;=\; B(x)\rho_0(x) + \rho_0(x)B(x)^*. \tag{2.22}$$

We will define the proposed excited state ψ as the solution of a differential equation analogous to (2.20). First, we need to introduce some further notation. Let H be a Hamiltonian for which the dynamics $\{\alpha_t\}$ exists; finite volume is sufficient. For any $a > 0$, $t \in \mathbb{R} \setminus \{0\}$, and local observable $A \in \mathcal{A}$, we may define

$$A_a(it, H) \;:=\; \frac{1}{2\pi i} e^{-at^2} \int_{-\infty}^{\infty} \alpha_s(A)\, \frac{e^{-as^2}}{s - it}\, ds. \tag{2.23}$$

In addition, for $T > 0$ the quantity

$$B_{a,T}(A, H) \;:=\; -\int_0^T A_a(it, H) - A_a(it, H)^* dt, \tag{2.24}$$

will play a crucial role. In Lemma 7 of Sect. 5, we will show that when projected onto the ground state of a gapped Hamiltonian H, the quantity $B_{a,T}(A, H)$ well approximates $B(A, H)$ for a judicious choice of parameters, e.g., $a = \gamma_L/L$ and $T = L/2$; we note that the observable A must also satisfy the constraint that its range is orthogonal to the ground state. With this in mind, consider the solution of the differential equation introduced by Hastings in [8]:

$$\partial_\theta \psi_{a,T}(\theta) = B_{a,T}(\theta)\, \psi_{a,T}(\theta), \tag{2.25}$$

where $B_{a,T}(\theta) = B_{a,T}(\partial_\theta H_{\theta,0}, H_{\theta,-\theta})$, subject to the boundary condition $\psi_{a,T}(0) = \psi_0$. Note that $B_{a,T}(\theta)$ is anti-hermitian, and hence any $\psi_{a,T}(\theta)$ solving (2.25) will have constant norm.

To be explicit, the proposed state $\psi_{a,T}(\theta)$ differs from the actual ground state of the doubly twisted Hamiltonian $H_{\theta,-\theta}$, in three essential ways. Compare (2.19) in the case that $A = \partial_\theta H_{\theta,-\theta}$ and $H = H_{\theta,-\theta}$ with (2.24) given that $A = \partial_\theta H_{\theta,0}$ and $H = H_{\theta,-\theta}$.
i) We have introduced a cut-off at $T < \infty$.
ii) We have approximated the imaginary-time evolution of an observable A, $\alpha_{it}(A)$, by $A_a(it, H) - A_a(it, H)^*$.
iii) We have replaced the observable $\partial_\theta H_{\theta,-\theta}$ with $\partial_\theta H_{\theta,0}$.

The modifications i) and ii) are of a technical nature, i.e., to make the relevant quantities well-defined and amenable to estimation (see Sect. 5). The motivation behind the third replacement is an attempt to approximate the ground state of the singly twisted Hamiltonian $H_{\theta,0}$.

3. Energy Estimates

As is discussed in the introduction, the goal of this section is to prove an estimate of the form

$$|\langle \psi_1, H_L \psi_1 \rangle - E_0| \le C L^\nu e^{-c\gamma_L L}, \tag{3.1}$$

see (1.16) and Theorem 3 below, for the proposed variational state. Explicitly, we will take $\psi_1 = \psi_{a,T}(2\pi)$, i.e., the solution of (2.25) evaluated at $\theta = 2\pi$, with the specific choice of parameters $a = \gamma_L/L$ and $T = L/2$. Since the operator $B_{a,T}(\theta)$, defined in (2.25), is anti-hermitian, it is clear that ψ_1 remains normalized, and the bound stated above demonstrates that if the gap is sufficiently large, $\gamma_L \ge C \log(L)/L$, then ψ_1 corresponds to a state with small (depending on C) excitation energy. An estimate of the form (3.1), *with correction terms*, can be proven based on the general results in Sect. 5. In the proof of Theorem 1, which is a proof by contradiction, we will be assuming $\gamma_L \ge C \log(L)/L$. Therefore, we can assume here, without loss of generality, that there exists a constant $c > 0$ such that $\gamma_L L \ge c$ for sufficiently large L. This assumption, which is not necessary, will simplify the presentation in Sects. 3 and 4.

3.1. Local estimates on the states. In this subsection we prove a technical result which estimates, *uniformly in* θ, the norm difference between the ground state of $H_{\theta,-\theta}$ and the proposed state in the left half of the system, more precisely in the window centered around the location, (m, V_L) where the θ-twist occurs. Since the restrictions of the states to the half-systems are described by density matrices, it is natural to use the trace norm for this estimate. Recall that for any bounded operator A on a Hilbert space $\mathcal{H}$, the trace norm is defined by

$$\|A\|_1 = \mathrm{Tr}\sqrt{A^*A}, \tag{3.2}$$

assuming this quantity is finite. Using the polar decomposition for bounded linear operators, it is easy to see that, alternatively,

$$\|A\|_1 = \sup_{\substack{B\in\mathcal{B}(\mathcal{H}):\\ \|B\|=1}} |\mathrm{Tr}AB|. \tag{3.3}$$

Recall that the density matrix corresponding to the ground state of the $H_{\theta,-\theta}$ Hamiltonian satisfies the equation

$$\partial_\theta \rho_0(\theta,-\theta) \;=\; B(\theta)\,\rho_0(\theta,-\theta) \;+\; \rho_0(\theta,-\theta)\,B(\theta)^*, \tag{3.4}$$

compare with (2.22), where we have used the notation $B(\theta) = B(A(\theta), H_{\theta,-\theta})$ for the operator $B(A, H)$ as defined in (2.19) and the observable $A(\theta) = \partial_\theta H_{\theta,-\theta}$. Note that by construction $\rho_0(\theta,-\theta)$ remains normalized. We will often write $A(\theta) = A_1(\theta) - A_2(\theta)$ where, the observables $A_1(\theta) = \partial_\theta H_{\theta,0}$ and $A_2(\theta) = \partial_\theta H_{0,\theta}$ are supported in the window about the twists of angle θ and $-\theta$, respectively. Regarding $H_{\theta,\theta'}$ as a function of two variables, we may write $A_i(\theta) = \partial_i H_{\theta,-\theta}$ for convenience. The notation $B_i(\theta) = B(A_i(\theta), H_{\theta,-\theta})$ will also be useful.

The proposed state is the solution of

$$\partial_\theta \rho_{a,T}(\theta) = \left[\, B_{a,T}(\theta), \rho_{a,T}(\theta) \,\right], \tag{3.5}$$

where the operator

$$B_{a,T}(\theta) = B_{a,T}(A_1(\theta), H_{\theta,-\theta})$$

is defined in (2.24) with observable $A_1(\theta) = \partial_1 H_{\theta,-\theta}$. The parametrization we choose is $a = \gamma_L/L$ and $T = L/2$. Since the operator $B_{a,T}(\theta)$ is anti-hermitian, the solution $\rho_{a,T}(\theta)$ is a density matrix. We will denote by $\mathrm{Tr}_{m^c}[\cdot]$ the partial trace over the Hilbert space corresponding to $\Lambda_L^{(S)} \cup \Lambda_L^{(W)}(m + L/2)$. Note that the terms in the Hamiltonian that have been twisted by an angle θ are supported in the complementary region $\Lambda_L^{(W)}(m)$. Given a gap $\gamma_L > 0$ above the ground state of the $H = H_{0,0}$ Hamiltonian, we will be able to estimate the trace norm of the difference in the two states restricted to $\Lambda_L^{(W)}(m)$. We will show this by estimating $\partial_\theta\, \mathrm{Tr}_{m^c}\left[\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right]$.

Theorem 2. *As described in the introduction, we assume F1, F2, and LSM1-6. If there exists a constant $c > 0$ such that $\gamma_L L \geq c$ for sufficiently large L and we choose the parameters $a = \gamma_L/L$ and $T = L/2$, then there exist constants $C > 0$ and $k > 0$ so that*

$$\sup_{\theta\in[0,2\pi]} \left\|\mathrm{Tr}_{m^c}\left[\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right]\right\|_1 \;\leq\; C\, L^{2d}\, e^{-k\gamma_L L}\,, \tag{3.6}$$

for L large enough. Here C and k depend only on the interaction Φ and the underlying set Λ.

We note that the assumption concerning the existence of a constant $c > 0$ such that $\gamma_L L \geq c$ for sufficiently large L is not necessary. We impose it here for simplicity of presentation. Without this additional assumption, one may prove an analogue of (3.6), which contains correction terms, by inserting the bounds proven in Sect. 5 directly into the proof given below. Since we make this assumption, it is convenient to state a lemma which compiles many of the technical results found in Sect. 5 and applies them to the present set-up. For this, we need two more definitions. Denote by $B_{a,T}^{(W)}(\theta)$ the operator defined by $B_{a,T}\left(A_1(\theta),\, H_{\theta,-\theta}^{(W)}\right)$, where the Hamiltonian $H_{\theta,-\theta}^{(W)}$ is the full Hamiltonian $H_{\theta,-\theta}$ restricted to the windows about the twists, see (2.8). Lastly, set P_0^{θ} to be the projection onto the ground state $\psi_0(\theta,-\theta)$ of the twisted Hamiltonian $H_{\theta,-\theta}$.

Lemma 2. *Assume F1, F2, and LSM1-6. If there exists a constant $c > 0$ such that $\gamma_L L \geq c$ for sufficiently large L and we choose the parameters $a = \gamma_L/L$ and $T = L/2$, then there exists constants $C > 0$ and $k > 0$ for which both*

$$\sup_{\theta\in[0,2\pi]} \|B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta)\| \;\leq\; C\, L^{2d}\, e^{-k\gamma_L L} \tag{3.7}$$

and

$$\sup_{\theta\in[0,2\pi]} \|\left(B_{a,T}(\theta) - B_1(\theta)\right) P_0^{\theta}\| \;\leq\; C\, L^{d}\, e^{-k\gamma_L L} \tag{3.8}$$

when L is large enough.

Proof of Lemma 2. Equation (3.7) follows by combining Lemma 4 and Remark 1. Using Lemma 7 and Remark 3, one obtains (3.8). □

Proof of Theorem 2. The proof of Theorem 2 follows by deriving a uniform bound on the θ-derivative of the differences in these density matrices. Specifically, the bound is in trace norm, and the uniformity is with respect to $\theta \in [0, 2\pi]$.

Using (3.4), (3.5), and inserting the local operator $B_{a,T}^{(W)}(\theta)$ for comparison, one may easily verify that

$$\partial_\theta \mathrm{Tr}_{m^c}\left[\,\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\,\right] \;=\; \mathrm{Tr}_{m^c}\left(\left[\,B_{a,T}^{(W)}(\theta),\; \rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\,\right]\right) + \sum_{i=1}^{3} r_i(\theta), \tag{3.9}$$

where the three remainder terms are given by

$$r_1(\theta) \;:=\; \mathrm{Tr}_{m^c}\left(\left[\,B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta),\; \rho_{a,T}(\theta)\,\right]\right), \tag{3.10}$$

$$r_2(\theta) \;:=\; \mathrm{Tr}_{m^c}\left(\left[\,B_{a,T}^{(W)}(\theta),\; \rho_0(\theta,-\theta)\,\right] - \partial_1\rho_0(\theta,-\theta)\right), \tag{3.11}$$

and

$$r_3(\theta) \;:=\; \mathrm{Tr}_{m^c}\left[\partial_1\rho_0(\theta,-\theta) \;-\; \partial_\theta\rho_0(\theta,-\theta)\right]. \tag{3.12}$$

As $A_1(\theta)$ is supported near (m, V_L) and $H^{(W)}_{\theta,-\theta}$ contains only those interaction terms over sets $X \subset \Lambda_L^{(W)}$, it is clear that $B^{(W)}_{a,T}(\theta)$ is contained in the algebra of local observables with support in $\Lambda_L^{(W)}(m)$; we will denote this algebra by $\mathcal{A}(m)$. Therefore,

$$\mathrm{Tr}_{m^c}\left(\left[B^{(W)}_{a,T}(\theta), \rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right]\right) = \left[B^{(W)}_{a,T}(\theta), \mathrm{Tr}_{m^c}\left(\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right)\right]. \tag{3.13}$$

Since $B^{(W)}_{a,T}(\theta)$ is anti-hermitian, we may apply norm preservation, i.e. Theorem 7, to (3.9) and conclude that

$$\left\| \mathrm{Tr}_{m^c}\left[\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right] \right\|_1 \le \sum_{i=1}^{3} \int_0^{\theta} \|r_i(\theta')\|_1 \, d\theta'. \tag{3.14}$$

We need only bound the trace norms of the remainder terms $r_i(\theta)$.

As $\rho_{a,T}(\theta)$ is a density matrix, in particular non-negative with a normalized trace, one has that

$$\|r_1(\theta)\|_1 \le 2 \left\| B_{a,T}(\theta) - B^{(W)}_{a,T}(\theta) \right\| \le C\, L^{2d}\, e^{-k\gamma_L L}, \tag{3.15}$$

using Lemma 2 above.

To estimate $r_2(\theta)$, we note that as in (3.4),

$$\partial_1 \rho_0(\theta,-\theta) = B_1(\theta)\,\rho_0(\theta,-\theta) + \rho_0(\theta,-\theta)\, B_1(\theta)^*, \tag{3.16}$$

where ∂_1 denotes differentiation with respect to only the first twist angle, namely θ, which is situated near the sites (m, V_L). Here we have also used that $\partial_1 E_0(\theta,-\theta) = 0$, see Lemma 1. A simple norm estimate yields that

$$\begin{aligned} \|r_2(\theta)\|_1 &\le 2 \left\| \left(B^{(W)}_{a,T}(\theta) - B_1(\theta)\right) P_0^{\theta} \right\| \\ &\le 2 \left\| B_{a,T}(\theta) - B^{(W)}_{a,T}(\theta) \right\| + 2 \left\| \left(B_{a,T}(\theta) - B_1(\theta)\right) P_0^{\theta} \right\|. \end{aligned} \tag{3.17}$$

Appealing again to Lemma 2, we see that $r_2(\theta)$ satisfies the desired bound.

Lastly, $r_3(\theta) = \mathrm{Tr}_{m^c}\left[\partial_2 \rho_0(\theta,-\theta)\right]$. Since we have shown in Lemma 1 that $\partial_2 E_0(\theta,-\theta) = 0$ as well, the analogue of (3.16) holds for $\partial_2 \rho_0(\theta,-\theta)$. Thus,

$$\begin{aligned} \|r_3(\theta)\|_1 &= \sup_{\substack{O \in \mathcal{A}(m): \\ \|O\|=1}} \left| \mathrm{Tr}\left[O \left(B_2(\theta)\,\rho_0(\theta,-\theta) + \rho_0(\theta,-\theta)\, B_2(\theta)^* \right) \right] \right| \\ &\le 2 \sup_{\substack{O \in \mathcal{A}(m): \\ \|O\|=1}} \int_0^{\infty} \left| \langle\, \psi_0(\theta,-\theta),\ O\, \alpha_{it}\left(A_2(\theta)\right) \psi_0(\theta,-\theta) \,\rangle \right| dt, \end{aligned} \tag{3.18}$$

where the observables O are arbitrary elements of $\mathcal{A}(m)$, again, the algebra of local observables with support in $\Lambda_L^{(W)}(m)$. Integrals of this type are bounded using Lemma 6; see also Remark 2. Since the observables we are considering have a separation distance proportional to L, we may estimate

$$\|r_3(\theta)\|_1 \le C\, L^{2d}\, e^{-k\gamma_L L}. \tag{3.19}$$

Combining the results found on each of the remainders, we arrive at the estimate claimed in (3.6). □

3.2. Bound on the energy. Equipped with Theorem 2 and Lemma 2, we may now bound the excitation energy corresponding to the proposed state.

Theorem 3. *Assume F1, F2, and LSM1-6. If there exists a constant $c > 0$ such that $\gamma_L L \geq c$ for sufficiently large L and we choose the parameters $a = \gamma_L/L$ and $T = L/2$, then there exists constants $C > 0$ and $k > 0$ so that*

$$|\langle \psi_1, H_L \psi_1 \rangle - E_0| \leq C L^{3d-1} e^{-k\gamma_L L} \tag{3.20}$$

for large enough L. Here, we take $\psi_1 = \psi_{a,T}(2\pi)$.

The proof of this theorem may be understood as follows. Recall that the ground state energy of the doubly twisted Hamiltonian is independent of θ, i.e.,

$$E_0 = \langle \psi_0, H_L \psi_0 \rangle = \langle \psi_0(\theta, -\theta), H_{\theta,-\theta} \psi_0(\theta, -\theta) \rangle. \tag{3.21}$$

Moreover, the separation between the twists of angle θ and $-\theta$ grows with the volume. Locality should enable one to estimate the energy difference between performing two twists, the ground state, and performing only one twist, the excited state. A rigorous version of this idea is described below.

First, we recall some of the notation introduced in Sect. 2.1. We have written the twisted Hamiltonian as the sum of two terms

$$H_{\theta,-\theta} = H^{(W)}_{\theta,-\theta} + H^{(S)}. \tag{3.22}$$

It is useful to further subdivide the twisted terms as

$$H^{(W)}_{\theta,-\theta} = H^{(W)}_{\theta}(m) + H^{(W)}_{-\theta}(m + L/2), \tag{3.23}$$

where $H^{(W)}_{\theta}(m)$ contains all those interaction terms in $H^{(W)}_{\theta,-\theta}$ with support in a window about the twist of angle θ, i.e. $\Lambda^{(W)}_L(m)$, and similarly, $H^{(W)}_{-\theta}(m + L/2)$ contains all those interaction terms in $H^{(W)}_{\theta,-\theta}$ with support in $\Lambda^{(W)}_L(m + L/2)$. The untwisted terms in (3.22) are supported in the remaining strips. We refer to Eqs. (2.5)–(2.9) for more details. It was also noted in Sect. 2.1 that

$$W(\theta)^* H_{\theta,-\theta} W(\theta) = H, \tag{3.24}$$

see (2.12).

Now, for any state ψ, one may calculate the expected energy due to a single twist:

$$\begin{aligned} \langle \psi, H_{\theta,0} \psi \rangle &= \langle \psi, H^{(W)}_{\theta}(m) \psi \rangle + \langle \psi, \left(H^{(W)}_0(m + L/2) + H^{(S)} \right) \psi \rangle \\ &= E_0 + R_1(\theta) + R_2(\theta). \end{aligned} \tag{3.25}$$

Here, we inserted appropriate terms so that we may compare $\langle \psi, H_{\theta,0} \psi \rangle$ to the ground state energy; the remainder terms are given by

$$R_1(\theta) := \langle \psi, H^{(W)}_{\theta}(m) \psi \rangle - \langle \psi_0(\theta, -\theta), H^{(W)}_{\theta}(m) \psi_0(\theta, -\theta) \rangle \tag{3.26}$$

and

$$R_2(\theta) := \langle \psi, \left(H^{(W)}_0(m + L/2) + H^{(S)} \right) \psi \rangle - \langle \psi_0, \left(H^{(W)}_0(m + L/2) + H^{(S)} \right) \psi_0 \rangle, \tag{3.27}$$

where ψ_0 is the ground state of $H = H_{0,0}$. The bound

$$\left| \langle \psi, H_{\theta,0} \psi \rangle - E_0 \right| \leq |R_1(\theta)| + |R_2(\theta)|, \tag{3.28}$$

readily follows for any state ψ.

Proof of Theorem 3. For each fixed θ, the bound (3.28) is valid for our proposed state $\psi_{a,T}(\theta)$. We will estimate the resulting remainders uniformly for $\theta \in [0, 2\pi]$ and thereby prove the claimed result.

To see this, we first rewrite the remainders in terms of the density matrices of the states restricted to the region containing the first twist. It is clear that

$$\begin{aligned} R_1(\theta) &= \mathrm{Tr}\left[\left(\rho_{a,T}(\theta) - \rho_0(\theta, -\theta) \right) \, H_\theta^{(W)}(m) \right] \\ &= \mathrm{Tr}_m \left[\mathrm{Tr}_{m^c} \left[\rho_{a,T}(\theta) - \rho_0(\theta, -\theta) \right] \, H_\theta^{(W)}(m) \right], \end{aligned} \tag{3.29}$$

where the partial traces are as defined just prior to Theorem 2. Thus

$$\begin{aligned} |R_1(\theta)| &\leq \| H_\theta^{(W)}(m) \| \, \| \mathrm{Tr}_{m^c} \left[\rho_{a,T}(\theta) - \rho_0(\theta, -\theta) \right] \|_1 \\ &\leq C \, L^{3d-1} \, e^{-k \gamma_L L}, \end{aligned} \tag{3.30}$$

where we have used Theorem 2. In fact, from the assumptions we have made, one verifies that

$$\| H_\theta^{(W)}(m) \| \leq 2 \sum_{x \in \Lambda_L^{(W)}(m)} \sum_{X \ni x} \| \Phi(X) \| \leq 2 |||\Phi|||_1 \, |\Lambda_L^{(W)}(m)| \leq C L^{d-1}. \tag{3.31}$$

For the second remainder,

$$R_2(\theta) = \mathrm{Tr}\left[\left(\rho_{a,T}(\theta) - \rho_0(0,0) \right) \left(H_0^{(W)}(m + L/2) + H^{(S)} \right) \right], \tag{3.32}$$

we observe that the only θ dependence is in the density matrix corresponding to the proposed state. Using the differential equation (3.5), we find that

$$\begin{aligned} R_2'(\theta) &= \mathrm{Tr}\left(\left[B_{a,T}(\theta), \rho_{a,T}(\theta) \right] \left(H_0^{(W)}(m + L/2) + H^{(S)} \right) \right) \\ &= - \mathrm{Tr}\left(\left[B_{a,T}(\theta), \left(H_0^{(W)}(m + L/2) + H^{(S)} \right) \right] \rho_{a,T}(\theta) \right). \end{aligned} \tag{3.33}$$

The first term above is easy to estimate. Recall that the quantity $B_{a,T}^{(W)}(\theta)$ is supported in $\Lambda_L^{(W)}(m)$, whereas $H_0^{(W)}(m + L/2)$ has support in $\Lambda_L^{(W)}(m + L/2)$. Thus

$$\left[B_{a,T}(\theta), H_0^{(W)}(m + L/2) \right] = \left[B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta), H_0^{(W)}(m + L/2) \right], \tag{3.34}$$

and moreover,

$$\begin{aligned} &\left| \mathrm{Tr}\left(\left[B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta), H_0^{(W)}(m + L/2) \right] \rho_{a,T}(\theta) \right) \right| \\ &\quad \leq 2 \left\| H_0^{(W)}(m + L/2) \right\| \left\| B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta) \right\| \leq C \, L^{3d-1} \, e^{-k \gamma L}. \end{aligned} \tag{3.35}$$

The second term may be similarly estimated. Let $\tilde{H}_{\theta,-\theta}^{(W)}$ be defined as in (2.8), excepting that the windows are slightly smaller: of size $\frac{L}{4} - 2R$. Then $[\tilde{B}_{a,T}^{(W)}(\theta), H^{(S)}] = 0$, and the argument above applies. We have bounded $R_2(\theta)$. □

4. Orthogonality

We will now prove that, under the assumptions given in the Introduction, the proposed state is nearly orthogonal to the ground state. As in Sect. 3, we again make the assumption that $\gamma_L L \geq c > 0$, for sufficiently large L.

The reasoning behind orthogonality is simple. From LSM5, we know that the ground state is an eigenvector of the translation operator with eigenvalue 1, i.e., $T\psi_0 = \psi_0$. On the other hand, the proposed state will very nearly be an eigenvector of $T_{2\pi,0}$, as defined in Sect. 2.1, with eigenvalue 1. Due to the odd parity condition $T_{2\pi,0} = -T$ and, hence, we find that the ground state and the proposed state are eigenvectors corresponding to distinct eigenvalues.

More concretely, it is easy to check that

$$\langle \psi_{a,T}(2\pi), \psi_0 \rangle = \langle T_{2\pi,0}\psi_{a,T}(2\pi), T\psi_0 \rangle + \langle \left(I - T_{2\pi,0}\right) \psi_{a,T}(2\pi), \psi_0 \rangle, \quad (4.1)$$

from which the estimate

$$\left| \langle \psi_{a,T}(2\pi), \psi_0 \rangle \right| \leq \frac{1}{2} \left\| \left(T_{2\pi,0} - I\right) \psi_{a,T}(2\pi) \right\| \quad (4.2)$$

immediately follows. The remainder of this section will be used to prove a bound on

$$\| \, T_{\theta,0}\psi_{a,T}(\theta) - \psi_{a,T}(\theta) \, \| \quad (4.3)$$

uniformly for $\theta \in [0, 2\pi]$. This is the content of Theorem 4.

4.1. Observations concerning the twisted ground state. We begin with a warm-up exercise involving the twisted ground state. In Sect. 2.1, we saw that the twisted ground state is invariant with respect to the twisted translations; i.e.,

$$T_{\theta,-\theta} \, , \psi_0(\theta, -\theta) = \psi_0(\theta, -\theta) \, ,$$

and therefore

$$\partial_\theta \left[\, T_{\theta,-\theta}\psi_0(\theta, -\theta) \; - \; \psi_0(\theta, -\theta) \, \right] = 0. \quad (4.4)$$

One may rewrite this derivative in the form of an operator acting on $\psi_0(\theta, -\theta)$, i.e., (4.4) is equivalent to

$$D(\theta)\psi_0(\theta, -\theta) = 0, \quad (4.5)$$

where $D(\theta)$ is given by

$$D(\theta) = \partial_\theta T_{\theta,-\theta} \, T^*_{\theta,-\theta} + T_{\theta,-\theta} \, B(\theta) \, T^*_{\theta,-\theta} - B(\theta). \quad (4.6)$$

Here we have used the differential equation for $\psi_0(\theta, -\theta)$, i.e. (2.20), and the notation from the beginning of Sect. 3.1, which will be used throughout this section.

It will be easy to see that the operator $D(\theta)$ can be written as the sum of two terms, $D_1(\theta)$ and $D_2(\theta)$, corresponding to the twists at m and $m + L/2$, respectively. The goal of this subsection is to estimate $\|D_1(\theta)\psi_0(\theta, -\theta)\|$, see Lemma 3 below.

Using (2.13), one finds that

$$\partial_\theta T_{\theta,-\theta} \, T^*_{\theta,-\theta} = i \sum_{v \in V_L} S^3_{(m+1,v)} - i \sum_{v \in V_L} S^3_{(m+L/2+1,v)}. \quad (4.7)$$

One has that $D(\theta) = D_1(\theta) - D_2(\theta)$, where

$$D_1(\theta) \; = \; i \sum_{v \in V_L} S^3_{(m+1,v)} \, + \, T_{\theta,-\theta} \, B_1(\theta) \, T^*_{\theta,-\theta} \; - \; B_1(\theta), \tag{4.8}$$

and

$$D_2(\theta) \; = \; i \sum_{v \in V_L} S^3_{(m+L/2+1,v)} \, + \, T_{\theta,-\theta} \, B_2(\theta) \, T^*_{\theta,-\theta} \; - \; B_2(\theta). \tag{4.9}$$

For what follows, we will denote by $\langle A \rangle_\theta \; = \; \langle \psi_0(\theta, -\theta), A\psi_0(\theta, -\theta) \rangle$ the twisted ground state expectation of a local observable A. We have demonstrated in Lemma 1 that

$$0 \; = \; \partial_i E_0(\theta, -\theta) \; = \; \langle \partial_i H_{\theta,-\theta} \rangle_\theta \; = \; \langle A_i(\theta) \rangle_\theta, \tag{4.10}$$

for $i = 1, 2$. From this, we conclude that

$$\left\langle T_{\theta,-\theta} \, B_i(\theta) \, T^*_{\theta,-\theta} \right\rangle_\theta \; = \; \langle \, B_i(\theta) \, \rangle_\theta \; = \; \langle A_i(\theta) \rangle_\theta \; = \; 0, \tag{4.11}$$

as well. Moreover, we similarly have that

$$\langle D_i(\theta) \rangle_\theta \; = \; 0 \qquad \text{as} \qquad \left\langle \sum_{v \in V_L} S^3_{(x,v)} \right\rangle_\theta \; = \; \left\langle \sum_{v \in V_L} S^3_{(x,v)} \right\rangle_0 \; = \; 0, \tag{4.12}$$

for any $x \in [1, L]$. For the last equality above, we used that $\psi_0(\theta, -\theta) = W(\theta)\psi_0$, $W(\theta)$ commutes with the third component of the spins, rotation invariance implies that the total spin is zero, and translation invariance in the 1-direction.

Since $D(\theta)\psi_0(\theta, -\theta) = 0$, we have that $D_1(\theta)\psi_0(\theta, -\theta) = D_2(\theta)\psi_0(\theta, -\theta)$ from which it is clear that

$$0 \; = \; \langle D(\theta)^* D(\theta) \rangle_\theta \; = \; 2\langle D_1(\theta)^* D_1(\theta) \rangle_\theta \; - \; 2\langle D_1(\theta)^* D_2(\theta) \rangle_\theta. \tag{4.13}$$

As indicated above, we wish to estimate the first term on the right-hand side above. We do so by estimating the second term. Observe that

$$\begin{aligned} \langle D_1(\theta)^* D_2(\theta) \rangle_\theta \; &= \; - \sum_{v, v' \in V_L} \langle S^3_{(m+1,v)} \, S^3_{(m+L/2+1,v')} \rangle_0 \\ &+ i \sum_{v \in V_L} \int_0^\infty \left\langle \left(S^3_{(m,v)} \; - \; S^3_{(m+1,v)} \right) \alpha_{it} \left(A_2(\theta) \right) \right\rangle_\theta \, dt \\ &+ \int_0^\infty \int_0^\infty \left\langle \alpha_{it}(A_1(\theta))^* \left(\alpha_{is}(A_2(\theta)) - \alpha_{is} \left(T^*_{\theta,-\theta} A_2(\theta) T_{\theta,-\theta} \right) \right) \right\rangle_\theta \, ds \, dt \\ &+ i \sum_{v' \in V_L} \int_0^\infty \left\langle \alpha_{it}(A_1(\theta))^* \left(S^3_{(m+L/2+1,v')} - S^3_{(m+L/2,v')} \right) \right\rangle_\theta \, dt \\ &+ \int_0^\infty \int_0^\infty \left\langle \alpha_{it}(A_1(\theta))^* \left(\alpha_{is}(A_2(\theta)) - \alpha_{is} \left(T_{\theta,-\theta} A_2(\theta) T^*_{\theta,-\theta} \right) \right) \right\rangle_\theta \, ds \, dt. \end{aligned} \tag{4.14}$$

That each of these terms is bounded follows from our decay of correlation results found in Sect. 5.3. In fact, we have proven the following lemma.

Lemma 3. *Assume F1, F2, and LSM1-6. If there exist a constant $c > 0$ such that $\gamma_L L \geq c$ for sufficiently large L, then there exist constants $C > 0$ and $k > 0$ so that*

$$\|D_1(\theta)\psi_0(\theta,-\theta)\|^2 \;\leq\; C L^{3d-1} e^{-k\gamma_L L}, \tag{4.15}$$

for L large enough.

Proof. Clearly, one has that

$$\|D_1(\theta)\psi_0(\theta,-\theta)\|^2 = \left\langle D_1(\theta)^* D_1(\theta) \right\rangle_\theta \;=\; \left\langle D_1(\theta)^* D_2(\theta) \right\rangle_\theta, \tag{4.16}$$

from (4.13) above. Applying Theorem 6, Lemma 6, and Remark 2, as appropriate, to the terms found in (4.14), one arrives at (4.15). □

4.2. Orthogonality of the excited state. We are now ready to provide the orthogonality estimate.

Theorem 4. *Assume F1, F2, and LSM1-6. If there exist a constant $c > 0$ such that $\gamma_L L \geq c$ for sufficiently large L and we choose the parameters $a = \gamma_L/L$ and $T = L/2$, then there exist constants $C > 0$ and $k > 0$ so that*

$$\left|\langle \psi_{a,T}(2\pi), \psi_0\rangle\right| \;\leq\; C L^{2d} e^{-k\gamma_L L} \tag{4.17}$$

when L is large enough.

Proof. As is clear from (4.2), Theorem 4 follows from bounding the quantity appearing in (4.3) uniformly for $\theta \in [0, 2\pi]$. A short calculation, using (2.25), shows that

$$\begin{aligned}\partial_\theta \left[T_{\theta,0}\psi_{a,T}(\theta) - \psi_{a,T}(\theta)\right] = C_{a,T}(\theta)\, &\left[T_{\theta,0}\psi_{a,T}(\theta) - \psi_{a,T}(\theta)\right] \\ &+ D_{a,T}(\theta)\, \psi_{a,T}(\theta),\end{aligned} \tag{4.18}$$

where

$$C_{a,T}(\theta) \;=\; \partial_\theta T_{\theta,0}\, T_{\theta,0}^* \;+\; T_{\theta,0}\, B_{a,T}(\theta)\, T_{\theta,0}^*, \tag{4.19}$$

and

$$D_{a,T}(\theta) \;=\; \partial_\theta T_{\theta,0}\, T_{\theta,0}^* \;+\; T_{\theta,0}\, B_{a,T}(\theta)\, T_{\theta,0}^*, \;-\; B_{a,T}(\theta), \tag{4.20}$$

are both anti-Hermitian operators. The first term on the right-hand side of (4.18) is norm-preserving, and therefore, we need only bound the norm of the second by Theorem 7.

The norm of $D_{a,T}(\theta)\psi_{a,T}(\theta)$ will now be estimated by rewriting it in terms of quantities for which we have already proven bounds. Each term will be shown to satisfy a bound of the form (4.17).

We begin by writing

$$\begin{aligned}\|\, D_{a,T}(\theta)\, \psi_{a,T}(\theta)\, \|^2 &= \mathrm{Tr}\left[D_{a,T}(\theta)^* D_{a,T}(\theta)\, \rho_{a,T}(\theta)\right] \\ &= \mathrm{Tr}\left[D_{a,T}(\theta)^* D_{a,T}(\theta)\, \rho_0(\theta,-\theta)\right] \\ &\quad + \mathrm{Tr}\left[D_{a,T}(\theta)^* D_{a,T}(\theta)\, \left(\rho_{a,T}(\theta) \;-\; \rho_0(\theta,-\theta)\right)\right].\end{aligned} \tag{4.21}$$

The first term on the right-hand side above, which is equal to $\|D_{a,T}(\theta)\psi_0(\theta,-\theta)\|^2$, may be estimated by comparing it with the vector $D_1(\theta)\psi_0(\theta,-\theta)$ introduced in the previous subsection. In fact,

$$\|D_{a,T}(\theta)\psi_0(\theta,-\theta)\| \leq \|D_1(\theta)\psi_0(\theta,-\theta)\|+\| \left(D_{a,T}(\theta)-D_1(\theta)\right)\psi_0(\theta,-\theta)\|. \quad (4.22)$$

We bounded the first term above in Lemma 3. For the second, observe that

$$\begin{aligned} D_{a,T}(\theta) - D_1(\theta) = {} & T_{\theta,-\theta}\left(B_{a,T}(\theta) - B_1(\theta)\right) T^*_{\theta,-\theta} + T_{\theta,0}\left(B_{a,T}(\theta) - B^{(W)}_{a,T}(\theta)\right) T^*_{\theta,0} \\ & - T_{\theta,-\theta}\left(B_{a,T}(\theta) - B^{(W)}_{a,T}(\theta)\right) T^*_{\theta,-\theta} + B_1(\theta) - B_{a,T}(\theta), \end{aligned} \quad (4.23)$$

from which it is clear that

$$\begin{aligned} \| \left(D_{a,T}(\theta) - D_1(\theta)\right)\psi_0(\theta,-\theta)\| \leq {} & 2\,\|B_{a,T}(\theta) - B^{(W)}_{a,T}(\theta)\| \\ & + 2\,\| \left(B_{a,T}(\theta) - B_1(\theta)\right) P^{\theta}_0\,\|. \end{aligned} \quad (4.24)$$

That each of these terms satisfies the desired bound follows from Lemma 2.

For the final term on the right-hand side of (4.21), we insert and remove

$$D^{(W)}_{a,T}(\theta) = \partial_\theta T_{\theta,0}\,T^*_{\theta,0} + T_{\theta,0}\,B^{(W)}_{a,T}(\theta)\,T^*_{\theta,0}, - B^{(W)}_{a,T}(\theta), \quad (4.25)$$

a local observable supported in $\Lambda^{(W)}_L(m)$. Observe that

$$\|D^{(W)}_{a,T}(\theta)\| \leq \|\partial_\theta T_{\theta,0}T^*_{\theta,0}\| + 2\|B^{(W)}_{a,T}(\theta)\| \leq CL^d, \quad (4.26)$$

where we have used Proposition 2. We may write

$$\begin{aligned} & \mathrm{Tr}\left[D_{a,T}(\theta)^* D_{a,T}(\theta)\left(\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right)\right] \\ & = \mathrm{Tr}\left[D^{(W)}_{a,T}(\theta)^* D^{(W)}_{a,T}(\theta)\left(\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right)\right] \\ & \quad + \mathrm{Tr}\left[\left(D_{a,T}(\theta)^* D_{a,T}(\theta) - D^{(W)}_{a,T}(\theta)^* D^{(W)}_{a,T}(\theta)\right)\left(\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right)\right]. \end{aligned} \quad (4.27)$$

The first term above may be estimated by

$$\begin{aligned} & \left|\mathrm{Tr}_m\left[D^{(W)}_{a,T}(\theta)^*\,D^{(W)}_{a,T}(\theta)\,\mathrm{Tr}_{m^c}\left[\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right]\right]\right| \\ & \leq \left\|D^{(W)}_{a,T}(\theta)^*\,D^{(W)}_{a,T}(\theta)\right\|\,\left\|\mathrm{Tr}_{m^c}\left[\rho_{a,T}(\theta) - \rho_0(\theta,-\theta)\right]\right\|_1 \\ & \leq CL^{4d}e^{-k\gamma L}, \end{aligned} \quad (4.28)$$

where for the final inequality above we used Theorem 2 again.

For the second term, we rewrite the difference as

$$\begin{aligned} D_{a,T}(\theta)^* D_{a,T}(\theta) - D^{(W)}_{a,T}(\theta)^* D^{(W)}_{a,T}(\theta) = {} & \left(D_{a,T}(\theta) - D^{(W)}_{a,T}(\theta)\right)^* D_{a,T}(\theta) \\ & + D^{(W)}_{a,T}(\theta)^*\left(D_{a,T}(\theta) - D^{(W)}_{a,T}(\theta)\right), \end{aligned} \quad (4.29)$$

and apply the norm estimate

$$\left\| D_{a,T}(\theta) - D_{a,T}^{(W)}(\theta) \right\| \leq 2 \left\| B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta) \right\|. \tag{4.30}$$

We find that

$$\begin{aligned} \operatorname{Tr}&\left[\left(D_{a,T}(\theta)^* D_{a,T}(\theta) - D_{a,T}^{(W)}(\theta)^* D_{a,T}^{(W)}(\theta) \right) \left(\rho_{a,T}(\theta) - \rho_0(\theta, -\theta) \right) \right] \\ &\leq 4 \left\| B_{a,T}(\theta) - B_{a,T}^{(W)}(\theta) \right\| \left(\| D_{a,T}(\theta) \| + \| D_{a,T}^{(W)}(\theta) \| \right), \end{aligned} \tag{4.31}$$

which satisfies the required bound by Lemma 2 and an estimate analogous to (4.26). This completes the proof of Theorem 4. □

5. Auxiliary Results

In this section, we collect a number of auxiliary results, technical estimates as well as a few lemmas of a more general nature, which are needed for the proofs in Sects. 3 and 4.

We first recall the Lieb-Robinson bounds which are used to demonstrate quasi-locality of the dynamics associated to general quantum spin systems, see Theorem 5. Then, we observe in Proposition 1 that these Lieb-Robinson bounds may be used to compare the dynamics of a Hamiltonian defined on a given system with the dynamics of the same Hamiltonian restricted to a subsystem. Next, we provide in Lemma 4 an explicit bound which applies to the specific type of interactions we consider in this work.

In Sect. 5.2, we introduce the operators $B_{a,T}(A, H)$ which play a prominent role in our argument. We first discuss a few of their basic properties, and then use Proposition 1 to estimate the difference that arises in defining the operator with the full Hamiltonian as opposed to the Hamiltonian restricted to a subsystem; this is the content of Lemma 5. Lastly, we remark on exactly how this estimate will be used in the main text.

We review the Exponential Clustering Theorem in Sect. 5.3, and use it to prove a technical estimate, see Lemma 6. Moreover, in this section we also prove Lemma 7. This result provides an estimate on the quantity $\| \left(B_{a,T}(A, H) - B(A, H) \right) P_0 \|$ in terms of the parameters a, T, and the spectral gap of H, see (5.64). Here P_0 denotes the spectral projection onto the ground state of H, and the bound is valid for local observables A satisfying $P_0 A P_0 = 0$.

Lastly, we formulate a statement concerning solutions of certain simple differential equations in Sect. 5.4.

5.1. Lieb-Robinson bounds. For what follows, we adopt the same general framework for quantum spin models that was described in Sect. 1.2, including Conditions F1, F2, and the assumption that $\|\Phi\|_\lambda < +\infty$ for some $\lambda > 0$ (see (1.7) for the definition of the norm $\| \cdot \|_\lambda$).

We will use the following version of the Lieb-Robinson bound [18], which is a variant of the results proven in [20, 9].

Theorem 5 (Lieb-Robinson Bound). *Let $\lambda \geq 0$ and take $\Phi \in \mathcal{B}_\lambda(\Lambda)$. For any pair of local observables $A \in \mathcal{A}_X$ and $B \in \mathcal{A}_Y$ with $X, Y \subset \Lambda$, one may estimate*

$$\| [\alpha_t(A), B] \| \leq \frac{2 \, \|A\| \, \|B\|}{C_\lambda(F)} \, g_\lambda(t) \sum_{x \in X} \sum_{y \in Y} F_\lambda \left(d(x, y) \right), \tag{5.1}$$

for any $t \in \mathbb{R}$. *Here* $\{\alpha_t\}$ *is the dynamics generated by* Φ, *and one may take*

$$g_\lambda(t) = \begin{cases} \left(e^{2\,\|\Phi\|_\lambda\, C_\lambda(F)\,|t|} - 1\right) & \text{if } d(X,Y) > 0, \\ e^{2\,\|\Phi\|_\lambda\, C_\lambda(F)\,|t|} & \text{otherwise.} \end{cases} \tag{5.2}$$

Our proof of the Lieb-Schultz-Mattis theorem relies heavily on comparing the time evolution corresponding to a given Hamiltonian to that of the Hamiltonian restricted to a subsystem. The errors that result from such a comparison can be estimated in terms of a specific commutator to which the Lieb-Robinson bounds readily apply.

We begin with some notation. Let $\lambda \geq 0$ and consider $\Phi \in \mathcal{B}(\Lambda)$. For finite $\Lambda_0 \subset \Lambda$, the Hamiltonian corresponding to Φ restricted to Λ_0 is given by the self-adjoint operator

$$H_0 = \sum_{X \subset \Lambda_0} \Phi(X). \tag{5.3}$$

We will denote by $\alpha_t^{(0)}$ the time evolution corresponding to H_0, i.e., for any local observable A, $\alpha_t^{(0)}(A) = e^{itH_0} A e^{-itH_0}$ for all $t \in \mathbb{R}$.

Proposition 1. *Let* $\lambda \geq 0$ *and* $\Phi \in \mathcal{B}_\lambda(\Lambda)$. *Suppose the Hamiltonian corresponding to* Φ *restricted to a finite volume* $\Lambda_0 \subset \Lambda$ *is written as the sum of two self-adjoint operators, i.e.,* $H_0 = H_1 + H_2$. *Denoting by* $\alpha_t^{(i)}$ *the time evolution corresponding to* H_i, *for* $i = 0, 1, 2$, *then for any local observable* A *and* $t \in \mathbb{R}$, *one has that*

$$\| \alpha_t^{(0)}(A) - \alpha_t^{(1)}(A) \| \leq \int_0^{|t|} \left\| [H_2, \alpha_s^{(1)}(A)] \right\| ds. \tag{5.4}$$

Proof. Define the function $f : \mathbb{R} \to \mathcal{A}$ by

$$f(t) := \alpha_t^{(0)}(A) - \alpha_t^{(1)}(A). \tag{5.5}$$

A simple calculation shows that f satisfies the following differential equation:

$$f'(t) = i\left[H_0 - H_1, \alpha_t^{(1)}(A)\right] + i\,[H_0, f(t)], \tag{5.6}$$

subject to the boundary condition $f(0) = 0$. As this is a first order equation, the solution can be found explicitly:

$$f(t) = \alpha_t^{(0)}\left(\int_0^t \alpha_{-s}^{(0)}\left(i\left[H_2, \alpha_s^{(1)}(A)\right]\right) ds \right). \tag{5.7}$$

Using expression (5.7) and the automorphism property of $\alpha_t^{(0)}$, it is clear that

$$\| f(t) \| \leq \int_0^{|t|} \left\| [H_2, \alpha_s^{(1)}(A)] \right\| ds, \tag{5.8}$$

as claimed. □

To estimate the norm of the commutator appearing in Proposition 1, specifically in the bound (5.4), it is useful to specialize the general Lieb-Robinson bounds described above to the exact context we encounter in the present work. For example, we will be interested in specific finite volume Hamiltonians, those defined in Sect. 2 as $H_{\theta,\theta'}$, and particular observables, such as $A_1(\theta) = \partial_1 H_{\theta,-\theta}$ and $A_2(\theta) = \partial_2 H_{\theta,-\theta}$. Let α_t be the time evolution corresponding to the $H_{\theta,\theta'}$ Hamiltonian, and let $\alpha_t^{(W)}$ denote the dynamics associated with the Hamiltonian $H_{\theta,\theta'}^{(W)}$ which is defined in (2.8). We use the following estimate several times.

Lemma 4. *Let $\Phi \in \mathcal{B}_\lambda(\Lambda)$, then there exists constants $C > 0$ and $k > 0$ for which*

$$\max_{i=1,2} \sup_{\theta\in[0,2\pi]} \left\| \left[H^{(S)}, \alpha_t^{(W)}(A_i(\theta)) \right] \right\| \leq C\, e^{k|t|}\, L^{2(d-1)}\, e^{-\lambda L/4}. \tag{5.9}$$

Here it is important that C and k depend only on the properties of the underlying set Λ and the interaction Φ; they do not depend on the length scale L.

Proof. We will estimate the above commutator in the case that the observable is $A_1(\theta)$; an analogous result holds for $A_2(\theta)$. Recall that in (2.9) we wrote $H^{(S)}$ as a sum of interaction terms. Similarly, if one denotes by $P_m(\theta; Y) := V_m(\theta)^* \Phi(Y) V_m(\theta) - \Phi(Y)$, then $A_1(\theta)$ may be written as

$$\begin{aligned} A_1(\theta) &= \sum_{Y\subset\Lambda_L} \partial_\theta P_m(\theta; Y) \\ &= -i \sum_{\substack{Y\subset\Lambda_L:\\ P_m(\theta;Y)\neq 0}} \sum_{y\in Y_+} V_m(\theta)^* \left[S_y^3, \Phi(Y) \right] V_m(\theta), \end{aligned} \tag{5.10}$$

where Y_+ is the set of sites $y \in Y$ strictly to the right of m. Inserting both of these expressions into the right-hand side of (5.9) and applying the triangle inequality, it is clear that we must bound many terms of the form

$$\left\| \left[\Phi(X),\, \alpha_t^{(W)} \left(V_m(\theta)^* \left[S_y^3,\, \Phi(Y) \right] V_m(\theta) \right) \right] \right\|. \tag{5.11}$$

Term by term, we apply the Lieb-Robinson bound provided by Theorem 5, and use that the distance between the supports of X and Y is linear in L; concretely for any $x \in X$ and $y \in Y$, $d(x, y) \geq d(X, Y) \geq \frac{L}{4} - 3R$. We find that each term described by (5.11) satisfies an upper bound of the form

$$C(t)\, \|\Phi(X)\|\, |Y|\, \|[S_y^3, \Phi(Y)]\|\, e^{-\lambda L/4}, \tag{5.12}$$

where $C(t)$ may be taken as

$$C(t) = \frac{2\|F\|}{C_\lambda(F)} e^{2C_\lambda(F)\|\Phi\|_\lambda |t| + 3\lambda R}. \tag{5.13}$$

We need only count the number of terms. The combinatorics of the sums may be naively estimated as follows: $H^{(S)}$ corresponds to a sum of the form

$$\sum_{\substack{X\subset\Lambda_L:\\ X\cap\Lambda_L^{(S)}\neq\emptyset}} \leq \sum_{n=\frac{L}{4}-R+1}^{\frac{L}{4}+R-1} \sum_{v\in V_L} \sum_{X\ni(m+n,v)} + \sum_{n=\frac{3L}{4}-R+1}^{\frac{3L}{4}+R-1} \sum_{v\in V_L} \sum_{X\ni(m+n,v)}, \tag{5.14}$$

whereas for $A_1(\theta)$ we have that the sum

$$\sum_{\substack{Y\subset\Lambda_L:\\ P_m(\theta;Y)\neq 0}} \sum_{y\in Y} \leq \sum_{n=m-R}^{m+R} \sum_{v\in V_L} \sum_{Y\ni(n,v)} \sum_{y\in Y}. \tag{5.15}$$

Putting everything together, we have obtained that

$$\left\| \left[H^{(S)}, \alpha_t^{(W)}\left(A_1(\theta)\right) \right] \right\| \leq 2\,C(t)\,|||\Phi|||_1\,|||\Phi|||_2\,|V_L|^2\,(2R+1)(2R-1)\,e^{-\lambda\frac{L}{4}}, \tag{5.16}$$

which proves the claim. Recall,

$$|||\Phi|||_1 := \sup_{x\in\Lambda} \sum_{X\ni x} \|\Phi(X)\| \tag{5.17}$$

and

$$|||\Phi|||_2 := \sup_{x\in\Lambda} \sum_{X\ni x} |X| \sum_{x'\in X} \|\,[S^3_{x'}, \Phi(X)]\,\|. \tag{5.18}$$

□

5.2. Approximation of the imaginary time evolution. For our proof of the Lieb-Schultz-Mattis Theorem, we introduce an operator which, under certain assumptions, approximates the imaginary time evolution corresponding to a given Hamiltonian. In this section, we provide several basic estimates of this operator to which we will often refer in the main text.

Let $\lambda \geq 0$, $\Phi \in \mathcal{B}_\lambda(\Lambda)$, $\Lambda_0 \subset \Lambda$ be a finite set, and H be the Hamiltonian corresponding to Φ restricted to Λ_0. Denote by α_t, for $t \in \mathbb{R}$, the time evolution determined by H. For any local observable A, $a > 0$, $M > 0$, and $t \neq 0$, define

$$A_{a,M}(it, H) = \frac{e^{-at^2}}{2\pi i} \int_{-M}^{M} \alpha_s(A)\, \frac{e^{-as^2}}{s - it}\, ds, \tag{5.19}$$

and set $A_a(it, H) = \lim_{M\to\infty} A_{a,M}(it, H)$. We use the operator

$$B_{a,T}(A, H) = -\int_0^T A_a(it, H) - A_a(it, H)^*\, dt, \tag{5.20}$$

to define our variational state in the main text, see (2.25). We begin with some basic properties.

Proposition 2 (Shanti's Bound). *Let $\Phi \in \mathcal{B}_\lambda(\Lambda)$, A be a local observable, $a > 0$, and $T > 0$. The operator $B_{a,T}(A, H)$ is anti-hermitian and bounded. In particular,*

$$\|B_{a,T}(A, H)\| \leq \frac{\|A\|}{2}\sqrt{\frac{\pi}{a}}. \tag{5.21}$$

It is important to note that the bound above is independent of the finte volume Λ_0 on which the Hamiltonian H is defined.

Proof. That $B_{a,T}(A, H)$ is anti-hermitian follows immediately from (5.20). Combining (5.19) and (5.20), one finds that

$$B_{a,T}(A, H) = \frac{i}{\pi} \int_0^T \int_{-\infty}^{\infty} e^{-a(s^2+t^2)} \alpha_s(A) \frac{s}{s^2+t^2} \, ds \, dt, \tag{5.22}$$

from which (5.21) easily follows as

$$\|B_{a,t}(A, H)\| \leq \frac{\|A\|}{\pi} \int_{-\infty}^{\infty} e^{-as^2} |s| \int_0^T \frac{1}{s^2+t^2} \, dt \, ds \leq \frac{\|A\|}{2} \sqrt{\frac{\pi}{a}}. \tag{5.23}$$

□

In situations where the local observable A and the Hamiltonian H are fixed, we will often write $A_a(it)$ and $B_{a,T}$ to simplify notation. The following estimate is a simple consequence of (5.19).

Proposition 3. *Let $\Phi \in \mathcal{B}_\lambda(\Lambda)$ and A be a local observable. One has that*

$$\left\| \int_0^T A_a(it) - A_{a,M}(it) \, dt \right\| \leq \frac{T}{2M} \frac{\|A\|}{\sqrt{\pi a}} e^{-aM^2}. \tag{5.24}$$

Proof. For any $t \neq 0$,

$$A_a(it) - A_{a,M}(it) = \frac{e^{-at^2}}{2\pi i} \int_{|s|>M} \alpha_s(A) \frac{e^{-as^2}}{s - it} \, ds, \tag{5.25}$$

and therefore, one has the pointwise estimate

$$\| A_a(it) - A_{a,M}(it) \| \leq e^{-at^2} \frac{\|A\|}{2\pi M} e^{-aM^2} \sqrt{\frac{\pi}{a}}. \tag{5.26}$$

Upon integration, (5.24) readily follows. □

We will now prove an analogue of Proposition 1 for the quantities $B_{a,T}(A, H)$ introduced in (5.20). The estimate provided below is made explicit in terms of an a priori input, an assumed form of the Lieb-Robinson bound, see (5.27) below.

Lemma 5. *Let $\lambda \geq 0$ and $\Phi \in \mathcal{B}_\lambda(\Lambda)$. Suppose the Hamiltonian corresponding to Φ restricted to a finite volume $\Lambda_0 \subset \Lambda$ is written as the sum of two self-adjoint operators, i.e., $H_0 = H_1 + H_2$. Denote by $\alpha_t^{(i)}$ the time evolution corresponding to H_i, for $i = 0, 1, 2$. If, for a given local observable A, there exists numbers $c_i > 0$, $i = 1, 2, 3$, for which*

$$\left\| [H_2, \alpha_t^{(1)}(A)] \right\| \leq c_1 e^{c_2|t| - c_3}, \tag{5.27}$$

for all $t \in \mathbb{R}$, then the following estimate holds:

$$\| B_{a,T}(A, H_0) - B_{a,T}(A, H_1) \| \leq \frac{2T}{M} e^{-aM^2} \left(\frac{\|A\|}{\sqrt{\pi a}} + \frac{c_1 M^2}{\pi} \right), \tag{5.28}$$

where M has to be chosen as the positive solution of

$$aM^2 + c_2 M - c_3 = 0. \tag{5.29}$$

We note that in our applications the numbers c_i will depend on the observables A and H_2; in fact, they will be functions of the length scale L. We articulate this dependence explicitly in Remark 1 below.

Proof. One may write

$$B_{a,T}(A, H_0) - B_{a,T}(A, H_1) = -\int_0^T A_a(it, H_0) - A_a(it, H_1)\, dt + \int_0^T A_a(it, H_0)^* - A_a(it, H_1)^*\, dt, \quad (5.30)$$

and therefore

$$\| B_{a,T}(A, H_0) - B_{a,T}(A, H_1) \| \leq 2 \left\| \int_0^T A_a(it, H_0) - A_a(it, H_1)\, dt \right\|. \quad (5.31)$$

Moreover, the integrand may be expressed as

$$A_a(it, H_0) - A_a(it, H_1) = A_a(it, H_0) - A_{a,M}(it, H_0) + A_{a,M}(it, H_0) - A_{a,M}(it, H_1) + A_{a,M}(it, H_1) - A_a(it, H_1), \quad (5.32)$$

and for $j = 0, 1$, the bounds

$$\left\| \int_0^T A_a(it, H_j) - A_{a,M}(it, H_j)\, dt \right\| \leq \frac{T}{2M} \frac{\|A\|}{\sqrt{\pi a}} e^{-aM^2}, \quad (5.33)$$

follow immediately from Proposition 3. From this we conclude that for any $M > 0$,

$$\| B_{a,T}(A, H_0) - B_{a,T}(A, H_1) \| \leq 2 \left\| \int_0^T A_{a,M}(it, H_0) - A_{a,M}(it, H_1)\, dt \right\| + \frac{2T}{M} \frac{\|A\|}{\sqrt{\pi a}} e^{-aM^2}. \quad (5.34)$$

Clearly, the pointwise estimate

$$\|A_{a,M}(it, H_0) - A_{a,M}(it, H_1)\| \leq \frac{e^{-at^2}}{2\pi} \int_{-M}^{M} \frac{\|\alpha_s^{(0)}(A) - \alpha_s^{(1)}(A)\|}{|s|} e^{-as^2}\, ds, \quad (5.35)$$

follows directly from (5.19). By Proposition 1, we have that

$$\|\alpha_s^{(0)}(A) - \alpha_s^{(1)}(A)\| \leq \int_0^{|s|} \left\| [H_2, \alpha_x^{(1)}(A)] \right\| dx, \quad (5.36)$$

and by assumption (5.27), the integrand satisfies a uniform bound for $|s| \leq M$. The implication is that for all $|s| \leq M$,

$$\frac{\|\alpha_s^{(0)}(A) - \alpha_s^{(1)}(A)\|}{|s|} \leq c_1 e^{c_2 M - c_3}. \quad (5.37)$$

Putting everything together, we obtain that

$$\| B_{a,T}(A, H_0) - B_{a,T}(A, H_1) \| \leq \frac{2T}{M} \frac{\|A\|}{\sqrt{\pi a}} e^{-aM^2} + \frac{c_1 e^{c_2 M - c_3}}{\pi} \int_0^T \int_{-M}^{M} e^{-a(t^2+s^2)} \, ds dt. \quad (5.38)$$

As M here was arbitrary, we chose it as the (positive) solution of the following quadratic equation $aM^2 + c_2 M - c_3 = 0$. In this case,

$$\| B_{a,T}(A, H_0) - B_{a,T}(A, H_1) \| \leq \frac{2T}{M} e^{-aM^2} \left(\frac{\|A\|}{\sqrt{\pi a}} + \frac{c_1 M^2}{\pi} \right) \quad (5.39)$$

as claimed. □

Remark 1. In the main text of the paper, we will use Lemma 5 for the Hamiltonians $H_0 = H_{\theta,-\theta}$ and $H_1 = H^{(W)}_{\theta,-\theta}$ each of which depends on the length scale L; see Sect. 2 for the relevant definitions. It is assumed that H_0 has a gap $\gamma_L > 0$ above the ground state energy. The local observable A will be exactly as in Lemma 4, and therefore, the numbers c_i, $i = 1, 2, 3$, may be taken as follows: $c_1 = CL^{2(d-1)}$, $c_2 = k$, and $c_3 = \lambda L/4$, where again C and k depend only on the interaction and the underlying set Λ. In this case, we will choose the parametrization $a = \gamma_L/L$ and $T = L/2$. With this choice, the estimate (5.28) takes the form:

$$\sup_{\theta \in [0,2\pi]} \left\| B_{a,T}(\theta) - B^{(W)}_{a,T}(\theta) \right\| \leq C\, L^{2d} e^{-k\gamma_L L} \left(1 + \frac{1}{L^d \sqrt{\gamma_L L}} \right). \quad (5.40)$$

Here we have used the notation from Sect. 3 and the fact that the gap γ_L has a uniform bound from above; see (1.1) in Sect. 1.

5.3. Estimates for gapped systems. We derive two useful results in this subsection. For the first we recall the Exponential Clustering Theorem [9, 20], and use it to prove a technical estimate Lemma 6.

The second crucial estimate in this subsection, Lemma 7 below, applies specifically to gapped systems. It provides a bound on the norm of the difference in the operators $B_{a,T}(A, H)$ and $B(A, H)$ when restricted to the space of ground states corresponding to H. The bound applies to local observables A which project off the ground state, i.e. satisfy $P_0 A P_0 = 0$, where P_0 is the spectral projection of H onto the ground states, and is explicit in the parameters a, T, and the spectral gap of H, see (5.64) below.

We will consider Hamiltonians H, of the type introduced in Sect. 5.1, with an additional feature: a gap above the ground state energy. To state the gap condition precisely, we consider a representation of the system on a Hilbert space $\mathcal{H}$. This means that there is a representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$, and a self-adjoint operator H on $\mathcal{H}$ such that

$$\pi(\alpha_t(A)) = e^{itH} \pi(A) e^{-itH},$$

for all $t \in \mathbb{R}$ and $A \in \mathcal{A}$. For the results which follow, we will assume that $H \geq 0$ and that $\Omega \in \mathcal{H}$ is a normalized ground state, i.e., a vector state for which $H\Omega = 0$ and $\|\Omega\| = 1$. We say that the system has a spectral gap in this representation if there exists

$\delta > 0$ such that $\sigma(H) \cap (0, \delta) = \emptyset$, where $\sigma(H)$ is the spectrum of the operator H. In that case, the spectral gap, γ, is defined to be

$$\gamma = \sup\{\delta > 0 \mid \sigma(H) \cap (0, \delta) = \emptyset\}. \tag{5.41}$$

Let P_0 denote the orthogonal projection onto ker H. From now on, we will work in this representation and simply write A instead of $\pi(A)$.

The following result concerning exponential clustering was proven in [20].

Theorem 6 (Exponential Clustering). *Fix $\lambda > 0$. Let $\Phi \in \mathcal{B}_\lambda(\Lambda)$ be an interaction for which the corresponding self-adjoint Hamiltonian has a representation $H \geq 0$ with a normalized ground state vector Ω, i.e., $H\Omega = 0$ and $\|\Omega\| = 1$. Let P_E denote the family of spectral projections corresponding to H. If H has a spectral gap of size $\gamma > 0$ above the ground state energy, then there exist $\mu > 0$ such that for any local observables A and B with $A \in \mathcal{A}_X$, $B \in \mathcal{A}_Y$, $d :=$ dist$(X, Y) > 0$, and $P_0 B\Omega = P_0 B^*\Omega = 0$, the estimate*

$$|\langle \Omega, A\alpha_{it}(B)\Omega \rangle| \leq C(A, B)\, e^{-\mu d\left(1 + \frac{\gamma^2 t^2}{4\mu^2 d^2}\right)}, \tag{5.42}$$

holds for all t : $0 \leq t(4\|\Phi\|_\lambda C_\lambda + \gamma) \leq 2\lambda d$. Here, one may choose

$$C(A, B) = \|A\|\,\|B\| \left(1 + \frac{2}{\pi\, C_\lambda} \sum_{x\in X}\sum_{y\in Y} F\,(d(x, y)) + \frac{1}{\sqrt{\pi\,\mu d}}\right) \tag{5.43}$$

and

$$\mu = \frac{\lambda\,\gamma}{4\|\Phi\|_\lambda\, C_\lambda + \gamma}. \tag{5.44}$$

The above result easily leads to estimates on integrals of these ground state expectations. We state two such bounds in the next lemma, as they will arise in the proof of our main result.

Lemma 6. *Under the assumptions of Theorem 6, we have the estimates*

$$\int_0^\infty |\langle \Omega, A\alpha_{it}(B)\Omega\rangle|\, dt \leq \left(2\mu d\, C(A, B) + \|A\|\,\|B\|\, e^{-\mu d}\right) \frac{e^{-\mu d}}{\gamma}, \tag{5.45}$$

and

$$\int_0^\infty \int_0^\infty \left|\langle \Omega, A\alpha_{i(s+t)}(B)\Omega\rangle\right| ds\, dt$$
$$\leq \left[(\mu d)^2\, C(A, B) + \|A\|\,\|B\| \left(2\mu d + e^{-\mu d}\right)\right] \frac{e^{-\mu d}}{\gamma^2}. \tag{5.46}$$

Proof. Define T by the equation $\gamma\, T = 2\,\mu\, d$. We have that

$$\int_0^T |\langle \Omega, A\alpha_{it}(B)\Omega\rangle|\, dt \leq C(A, B)\, T\, e^{-\mu d}, \tag{5.47}$$

and also

$$\int_T^\infty |\langle \Omega, A\alpha_{it}(B)\Omega\rangle|\, dt \;\le\; \frac{\|A\|\,\|B\|}{\gamma}\, e^{-\gamma T}. \tag{5.48}$$

Combining these two bounds, we arrive at (5.45). Similarly, one may estimate

$$\int_0^{T/2}\int_0^{T/2} \left|\langle \Omega, A\alpha_{i(s+t)}(B)\Omega\rangle\right| ds\, dt \le \frac{C(A,B)\,T^2}{4}\, e^{-\mu d}, \tag{5.49}$$

$$\int_{T/2}^\infty\int_0^{T/2} \left|\langle \Omega, A\alpha_{i(s+t)}(B)\Omega\rangle\right| ds\, dt \le \frac{\|A\|\,\|B\|\,T}{2\gamma}\, e^{-\mu d}, \tag{5.50}$$

and finally,

$$\int_{T/2}^\infty\int_{T/2}^\infty \left|\langle \Omega, A\alpha_{i(s+t)}(B)\Omega\rangle\right| ds\, dt \;\le\; \frac{\|A\|\,\|B\|}{\gamma^2}\, e^{-\gamma T}. \tag{5.51}$$

□

Remark 2. In our applications, the Hamiltonian $H = H_{\theta,-\theta}$ depends on a length scale L and has a gap $\gamma_L > 0$ above the ground state energy. The support of the observables A and B will have a minimal distance $d = L/2 - 2R - 1$, and moreover, $B = B(\theta)$ will either be $A_1(\theta)$ or $A_2(\theta)$. In this case, $\langle B(\theta)\rangle_\theta = 0$ by Lemma 1, and therefore, the assumptions of Theorem 6 hold. Here we have used $\langle\cdot\rangle_\theta$ to denote the ground state expectation corresponding to $\psi_0(\theta, -\theta)$. It is easy to see that there exist positive constants C' and C'' for which $C'\gamma_L L \le \mu d \le C'' L$, and thus ultimately constants C and k for which the bounds appearing in (5.45) and (5.46) may be estimated by

$$\sup_{\theta\in[0,2\pi]}\int_0^\infty |\langle A\alpha_{it}\,(B(\theta))\rangle_\theta|\, dt \;\le\; C\,\|A\|\,|X|\,\frac{L^d}{\gamma_L} e^{-k\gamma_L L}, \tag{5.52}$$

and

$$\sup_{\theta\in[0,2\pi]}\int_0^\infty\int_0^\infty \left|\langle A\alpha_{i(s+t)}\,(B(\theta))\rangle_\theta\right| ds\, dt \;\le\; C\,\|A\|\,|X|\,\frac{L^{d+1}}{\gamma_L^2} e^{-k\gamma_L L}, \tag{5.53}$$

respectively.

For the next lemma we will need the following basic estimate involving the decay of certain Fourier transforms.

Proposition 4. *Let $a > 0$ and $T > 0$ be given. Define a function $F_{a,T} : \mathbb{R} \to \mathbb{C}$ by*

$$F_{a,T}(E) \;:=\; \frac{1}{2\pi i}\int_0^T e^{-at^2}\int_{-\infty}^\infty \frac{e^{-iEs}\, e^{-as^2}}{s - it}\, ds\, dt. \tag{5.54}$$

For all $E \in \mathbb{R}$, $F_{a,T}(E) \ge 0$ and the estimate

$$F_{a,T}(E) \;\le\; \frac{T}{2}\, e^{-\frac{E^2}{4a}}, \tag{5.55}$$

is valid for $E \ge 0$. In the parameter range, $E \ge 2aT > 0$, one may also show that

$$\int_0^T e^{-Et}\, dt \;-\; F_{a,T}(-E) \;\le\; \frac{T}{2}\, e^{-\frac{E^2}{4a}}. \tag{5.56}$$

Proof. One may easily verify that for any $t > 0$,

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{-iEs}\, e^{-as^2}}{s - it}\, ds \;=\; \frac{1}{2\sqrt{\pi a}} \int_0^{\infty} e^{-tw}\, e^{-\frac{(w+E)^2}{4a}}\, dw, \tag{5.57}$$

for all $E \in \mathbb{R}$, see e.g. Lemma 1 in [20]. This implies the first claim. Evaluating the Gaussian integral yields

$$\frac{1}{2\sqrt{\pi a}} \int_0^{\infty} e^{-tw}\, e^{-\frac{(w+E)^2}{4a}}\, dw \;\leq\; \frac{1}{2}\, e^{-\frac{E^2}{4a}}, \tag{5.58}$$

in the case that $E \geq 0$, from which (5.55) is clear.

To obtain (5.56), we first recall that the Fourier transform of a Gaussian is a Gaussian, i.e., for all $z \in \mathbb{C}$,

$$e^{-\frac{z^2}{2}} \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, e^{-ixz}\, dx, \tag{5.59}$$

and therefore, by rescaling $z \mapsto -\sqrt{2a}z$, multiplying through by e^{iEz} (for $E \in \mathbb{R}$), and changing variables $w = \sqrt{2a}x + E$, we have that

$$e^{iEz}\, e^{-az^2} \;=\; \frac{1}{2\sqrt{\pi a}} \int_{-\infty}^{\infty} e^{iwz}\, e^{-\frac{(w-E)^2}{4a}}\, dw, \tag{5.60}$$

for all $z \in \mathbb{C}$.

Now, by direct substitution into (5.57), we have that

$$F_{a,T}(-E) \;=\; \frac{1}{2\sqrt{\pi a}} \int_0^{T} e^{-at^2} \int_0^{\infty} e^{-tw}\, e^{-\frac{(w-E)^2}{4a}}\, dw. \tag{5.61}$$

Applying (5.60), with the special choice of $z = it$, one sees that

$$\int_0^{T} e^{-Et}\, dt \;-\; F_{a,T}(-E) \;=\; \frac{1}{2\sqrt{\pi a}} \int_0^{T} e^{-at^2} \int_{-\infty}^{0} e^{-tw}\, e^{-\frac{(w-E)^2}{4a}}\, dw. \tag{5.62}$$

Since $w < 0$ and $t > 0$, the integrand above

$$e^{-tw}\, e^{-\frac{(w-E)^2}{4a}} \;=\; e^{-\frac{E^2}{4a}}\, e^{\frac{(E-2at)w}{2a}}\, e^{-\frac{w^2}{4a}} \tag{5.63}$$

satisfies a trivial bound when $E \geq 2aT$. For these values of E, (5.56) holds. □

We may now prove the main estimate for gapped systems. Recall the definitions of the operators $B = B(A, H)$ and $B_{a,T} = B_{a,T}(A, H)$ in (2.19) and (2.24), respectively.

Lemma 7. *Let $H \geq 0$ be a self-adjoint operator and P_E denote the family of spectral projections corresponding to H. Suppose H has a gap $\gamma > 0$, and let A be a local observable for which $P_0 A P_0 = 0$. If $2aT \leq \gamma$, then one has that*

$$\left\| \left(B_{a,T} - B\right) P_0 \right\| \;\leq\; T e^{-\frac{\gamma^2}{4a}} \left(\frac{\|AP_0\| + \|A^* P_0\|}{2} \right) + \frac{e^{-\gamma T}}{\gamma}\, \|AP_0\|. \tag{5.64}$$

Proof. One may rewrite the difference in these operators as

$$\big(B_{a,T} - B\big)\; P_0 = \int_0^T (\alpha_{it}(A) - A_a(it))\, dt\; P_0 + \int_T^\infty \alpha_{it}(A)\, dt\; P_0 + \int_0^T A_a(it)^* dt\; P_0. \qquad (5.65)$$

Each of these terms may be bounded in norm.

For any vectors f and g, one may calculate

$$\langle f, \int_T^\infty \alpha_{it}(A) dt P_0\, g \rangle = \int_T^\infty \langle f, e^{-tH} A P_0 g \rangle\, dt = \int_T^\infty \int_\gamma^\infty e^{-tE}\, d\langle f, P_E\, A P_0 g \rangle\, dt, \qquad (5.66)$$

where we have used the spectral theorem to rewrite the time evolution and the fact that $P_0 A P_0 = 0$. Clearly then,

$$\left|\langle f, \int_T^\infty \alpha_{it}(A) dt P_0\, g \rangle\right| \le \|f\|\, \|A P_0 g\| \int_T^\infty e^{-\gamma t}\, dt, \qquad (5.67)$$

and therefore,

$$\left\| \int_T^\infty \alpha_{it}(A)\, dt\; P_0 \right\| \le \frac{e^{-\gamma T}}{\gamma}\, \|A P_0\|. \qquad (5.68)$$

Likewise, one may similarly calculate

$$\langle f, \int_0^T A_a(it)^* dt\; P_0 g \rangle = -\int_0^T \frac{e^{-at^2}}{2\pi i} \int_{-\infty}^\infty \langle \alpha_s(A) f, P_0 g \rangle \frac{e^{-as^2}}{s + it}\, ds\, dt = \int_\gamma^\infty F_{a,T}(E)\, d\langle f, P_E\, A^* P_0 g \rangle, \qquad (5.69)$$

where we have introduced $F_{a,T}(E) = \overline{F_{a,T}(E)}$ as in (5.54) of Proposition 4 above. The estimate

$$\left\| \int_0^T A_a(it)^*\, dt\; P_0 \right\| \le \frac{T}{2}\, e^{-\frac{\gamma^2}{4a}}\, \|A^* P_0\|, \qquad (5.70)$$

readily follows from (5.55) of Proposition 4 and the fact that $0 < \gamma \le E$.

Lastly, an analogous calculation shows that

$$\int_0^T \langle f, [\alpha_{it}(A) - A_a(it)]\; P_0 g \rangle\, dt = \int_\gamma^\infty \left[\int_0^T e^{-Et}\, dt - F_{a,T}(-E) \right] d\langle f, P_E\, A P_0 g \rangle. \qquad (5.71)$$

Thus, for $2aT \le \gamma$, we may apply (5.56) of Proposition 4 and establish the bound

$$\left\| \int_0^T [\alpha_{it}(A) - A_a(it)\,]\, dt\; P_0 \right\| \le \frac{T}{2}\, e^{-\frac{\gamma^2}{4a}}\, \|A P_0\|. \qquad (5.72)$$

Compiling our estimates, we have proven that: if $2aT \leq \gamma$, then

$$\| \left(B_{a,T} - B\right) P_0 \| \leq T e^{-\frac{\gamma^2}{4a}} \left(\frac{\|AP_0\| + \|A^* P_0\|}{2} \right) + \frac{e^{-\gamma T}}{\gamma} \|AP_0\|, \quad (5.73)$$

as claimed. □

Remark 3. Applying Lemma 7 to the operator $H = H_{\theta,-\theta}$, whose spectral projections we denote by P_E^θ, and the local observable $A = A_1(\theta)$, we find that there exists a constant $C > 0$ for which, along the parametrization $a = \gamma_L / L$ and $T = L/2$,

$$\sup_{\theta \in [0,2\pi]} \| \left(B_{a,T}(\theta) - B_1(\theta)\right) P_0^\theta \| \leq C L^d e^{-\frac{\gamma_L L}{4}} \left(1 + \frac{e^{-\frac{\gamma_L L}{4}}}{\gamma_L L} \right). \quad (5.74)$$

5.4. Norm preserving flows. In this section, we collect some basic facts about the solutions of first order, inhomogeneous differential equations.

Definition 1. Let $\mathcal{B}$ be a Banach space. For each $\theta \in \mathbb{R}$, let $A(\theta) : \mathcal{B} \to \mathcal{B}$ be a bounded linear operator, and denote by $X(\theta)$ the solution of the differential equation

$$\partial_\theta X(\theta) = A(\theta) X(\theta) \quad (5.75)$$

with boundary condition $X(0) = X_0 \in \mathcal{B}$. We say that the family of operators $A(\theta)$ is *norm-preserving* if the corresponding flow is isometric, i.e., for every $X_0 \in \mathcal{B}$, the mapping $\gamma_\theta : \mathcal{B} \to \mathcal{B}$ which associates $X_0 \to X(\theta)$, i.e., $\gamma_\theta(X_0) = X(\theta)$, satisfies

$$\| \gamma_\theta(X_0) \| = \| X_0 \| \quad \text{for all } \theta \in \mathbb{R}. \quad (5.76)$$

Two typical examples are the case where $\mathcal{B}$ is a Hilbert space and $A(\theta)$ is anti-hermitian and the case where $\mathcal{B}$ is a Banach space of linear operators on a Hilbert space with a spectral norm (such as a $p-$norm with $p \in [1, +\infty]$), and where $A(\theta)$ is a symmetric derivation (e.g., i times the commutator with a self-adjoint operator).

Theorem 7. *Let $A(\theta)$, for $\theta \in \mathbb{R}$, be a family of norm preserving opeartors in some Banach space $\mathcal{B}$. For any bounded measurable function $B : \mathbb{R} \to \mathcal{B}$, the solution of*

$$\partial_\theta Y(\theta) = A(\theta) Y(\theta) + B(\theta), \quad (5.77)$$

with boundary condition $Y(0) = Y_0$, satisfies the bound

$$\| Y(\theta) - \gamma_\theta(Y_0) \| \leq \int_0^\theta \| B(\theta') \| \, d\theta'. \quad (5.78)$$

Proof. For any $\theta \in \mathbb{R}$, let $X(\theta)$ be the solution of

$$\partial_\theta X(\theta) = A(\theta) X(\theta) \quad (5.79)$$

with boundary condition $X(0) = X_0$, and let γ_θ be the linear mapping which takes X_0 to $X(\theta)$. By variation of constants, the solution of the inhomogeneous equation (5.77) may be expressed as

$$Y(\theta) = \gamma_\theta \left(Y_0 + \int_0^\theta (\gamma_s)^{-1} (B(s)) \, ds \right). \quad (5.80)$$

The estimate (5.78) follows from (5.80) as $A(\theta)$ is norm preserving. □

5.5. Existence of local unitaries with vanishing expectation. Consider a finite system with a Hamiltonian of the form

$$H = \sum_X \Phi(X), \tag{5.81}$$

where Φ is an interaction as defined at the beginning of the paragraph containing Eq. (1.7). In the introduction, (1.1), we stated a simple upper bound for the spectral gap of any such Hamiltonian with a unique ground state. The argument we gave there made use of a one-site unitary $U \in \mathcal{A}_{\{x\}}$ with the property that $\langle \Omega, U\Omega \rangle = 0$. In the following lemma we show that such a unitary always exists.

Lemma 8. *Let $\mathcal{H}$ be a complex Hilbert space of dimension at least 2. Then, for any density matrix ρ on $\mathcal{H}$, there exists a unitary $U \in \mathcal{B}(\mathcal{H})$ such that* $\mathrm{Tr}\rho U = 0$.

Proof. First consider the case where $\dim \mathcal{H}$ is finite and even, or infinite. The odd-dimensional case has to be treated slightly differently. Let $\{e_0, e_1, \ldots\}$ denote an orthonormal basis of eigenvectors of ρ, with eigenvalues ρ_i ordered in non-increasing order. If $\mathcal{H}$ is not separable, it is sufficient that $\{e_0, e_1, \ldots\}$ contain a basis for the separable subspace $\mathrm{ran}\rho$. Then, a suitable unitary U can be defined as follows:

$$U = \bigoplus_{i \geq 0} |e_{2i+1}\rangle\langle e_{2i}| + |e_{2i}\rangle\langle e_{2i+1}|. \tag{5.82}$$

If $\dim \mathcal{H}$ is odd (and hence by our assumptions ≥ 3), it is sufficient to change the first summand in (5.82) as follows

$$U = a|e_0\rangle\langle e_0| - \overline{a}|e_1\rangle\langle e_1| + b|e_1\rangle\langle e_0| + \overline{b}|e_0\rangle\langle e_1| + e^{i\phi}|e_2\rangle\langle e_2|$$
$$+ \bigoplus_{i \geq 2} |e_{2i-1}\rangle\langle e_{2i}| + |e_{2i}\rangle\langle e_{2i-1}|,$$

where $a, b \in \mathbb{C}$ and $\phi \in \mathbb{R}$, are chosen such that $|a|^2 + |b|^2 = 1$ and

$$e^{i\phi}\rho_2 = a\rho_0 - \overline{a}\rho_1.$$

This is always possible since $\rho_2^2 \leq (\rho_0 + \rho_1)^2$.

It is straightforward to check that U thus defined has the desired properties. □

Acknowledgements This article is based on work supported by the U.S. National Science Foundation under Grant # DMS06-05342. Both authors thank the Erwin Schrödinger Institute for kind hospitality during a fruitful stay.

References

1. Affleck, I., Lieb, E.H.: A proof of part of Haldane's conjecture on quantum spin chains. Lett. Math. Phys. **12**, 57–69 (1986)
2. Affleck, I., Kennedy, T., Lieb, E.H., Tasaki, H.: Valence Bond Ground States in Isotropic Quantum Antiferromagnets. Commun. Math. Phys. **115**, 477–528 (1988)
3. Aizenman, M., Nachtergaele, B.: Geometric aspects of quantum spin states. Commun. Math. Phys. **164**, 17–63 (1994)
4. Bratteli, O., Robinson, D.: *Operator algebras and quantum statistical mechanics 2*. Second ed., New York: NY: Springer Verlag, 1997

5. Dyson, F., Lieb, E.H., Simon, B.: Phase transitions in quantum spin systems with isotropic and non-isotropic interactions. J. Stat. Phys. **18**, 335–383 (1978)
6. Haldane, F.D.M.: Continuum dynamics of the 1-D Heisenberg antiferromagnet: Identification with the $O(3)$ nonlinear sigma model. Phys. Lett. **93**, 464–468 (1983)
7. Haldane, F.D.M.: Nonlinear field theory of large-spin Heisenberg antiferromagnets: Semiclassically quantized solitons of the one-dimensional easy-axis Néel state. Phys. Rev. Lett. **50**, 1153–1156 (1983)
8. Hastings, M.B.: Lieb-Schultz-Mattis in higher dimensions. Phys. Rev. B **69**, 104431–14 (2004)
9. Hastings, M.B., Koma, T.: Spectral gap and exponential decay of correlations. Commun. Math. Phys. **265**, 781–804 (2006)
10. Kennedy, T., Lieb, E.H., Shastry, B.S.: Existence of Néel order in some spin 1/2 Heisenberg antiferromagnets. J. Stat. Phys. **53**, 1019–1030 (1988)
11. Koma, T., Tasaki, H.: Symmetry breaking and finite-size effects in quantum many-body systems. J. Stat. Phys. **76**, 745–803 (1994)
12. Landau, L., Fernando-Perez, J., Wreszinski, W.F.: Energy gap, clustering, and the Goldstone theorem in statistical mechanics. J. Stat. Phys. **26**, 755–766 (1981)
13. Lieb, E., Mattis, D.: Ordering energy levels in interacting spin chains. Journ. Math. Phys. **3**, 749–751 (1962)
14. Lieb, E., Schultz, T., Mattis, D.: Two soluble models of an antiferromagnetic chain. Ann. Phys. (N.Y.) **16**, 407–466 (1961)
15. Lieb, E.H., Robinson, D.W.: The finite group velocity of quantum spin systems. Commun. Math. Phys. **28**, 251–257 (1972)
16. Manousakis, E.: The spin-$1/2$ Heisenberg antiferromagnet on a square lattice and its application to the cuprous oxides. Rev. Mod. Phys. **63**, 1–62 (1991)
17. Nachtergaele, B.: *Quasi-state decompositions for quantum spin systems*. Probability Theory and Mathematical Statistics. Proceedings of the 6th Vilnius Conference (Grigelionis, B. et al., ed.), Utrecht-Tokyo-Vilnius, VSP/Tev, 1994, pp. 565–590
18. Nachtergaele, B., Ogata, Y., Sims, R.: Propagation of correlations in quantum lattice systems. J. Stat. Phys. **124**, 1–13 (2006)
19. Nachtergaele, B., Sims, R.: Recent progress in quantum spin systems. Markov Processes Rel. Fields. **13**(2), 315–329 (2007)
20. Nachtergaele, B., Sims, R.: Lieb-Robinson bounds and the exponential clustering theorem. Commun. Math. Phys. **265**, 119–130 (2006)
21. Sachdev, S.: Quantum antiferromagnets in two dimensions. In: *Low dimensional quantum field theories for condensed matter physicists.* Yu Lu, S., Lundqvist, G., Morandi, eds., Singapore: World Scientific, 1995
22. Sachdev, S., Park, K.: Ground states of quantum antiferromagnets in two dimensions. Annals of Physics (N.Y.) **298**, 58–122 (2002)
23. Simon, B.: *The statistical mechanics of lattice gases*. Princeton Series in Physics, Vol. **1**, Princeton, NJ: Princeton University Press, 1993
24. Tasaki, H.: Low-lying excitation in one-dimensional lattice electron system. http://arXiv.org/list/cond-mat/0407616, 2004
25. Yamanaka, M., Oshikawa, M., Affleck, I.: Nonperturbative approach to Luttinger's theorem in one dimension. Phys. Rev. Lett. **79**, 1110–1113 (1997)

Communicated by H.-T. Yau