Lieb-Schultz-Mattis in higher dimensions

M. B. Hastings*
Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
(Received 25 August 2003; revised manuscript received 12 December 2003; published 29 March 2004)

A generalization of the Lieb-Schultz-Mattis theorem to higher-dimensional spin systems is shown. The physical motivation for the result is that such spin systems typically either have long-range order, in which case there are gapless modes, or have only short-range correlations, in which case there are topological excitations. The result uses a set of loop operators, analogous to those used in gauge theories, defined in terms of the spin operators of the theory. We also obtain various cluster bounds on expectation values for gapped systems. These bounds are used, under the assumption of a gap, to rule out the first case of long-range order, after which we show the existence of a topological excitation. Compared to the ground state, the topologically excited state has, up to a small error, the same expectation values for all operators acting within any local region, but it has a different momentum.

DOI: 10.1103/PhysRevB.69.104431 PACS number(s): 75.10.Jm, 71.27.+a, 75.30.Ds

I. INTRODUCTION

Lieb, Schultz, and Mattis (LSM) proved in 1961 that a one-dimensional periodic chain of length L, with half-integer spin per unit cell, has an excitation gap bounded by const/L.[1] This behavior contrasts with the possibility of a Haldane gap in the integer spin case.[2]

Despite several attempts,[3,4] this theorem has not been extended to higher dimensions. The basic difficulty in obtaining a higher-dimensional version of this theorem was pointed out in two insightful papers by Misguich and co-workers[5]: if spin correlations are short ranged, the ground-state wave function should be well described by a short-range resonating valence bond (RVB) state.[6] The short-range RVB basis decomposes into different topological sectors, depending upon the number of dimers crossing a given line through the system. This allows the construction of a low-energy excited state very similar to the twisted state of LSM.[7] Instead, if spin correlations are long ranged, such a state will not be low energy, but there will exist low-energy spin wave excitations. In contrast to the one-dimensional case, there now exist two distinct means of obtaining a low-energy excitation, significantly complicating the proof of any such theorem.

In the present paper, we show a higher-dimensional version of the LSM theorem. We consider a d-dimensional system of spin-1/2 spins, with finite-range, SU(2) invariant Hamiltonian $\mathcal{H}$, and with an odd number of spins per unit cell on the lattice. Define the total number of unit cells in the lattice to be V. Let L be the number of unit cells in one particular direction, and let L be even; this direction will be referred to as the length. Therefore, V is even (if V were odd, there would be a trivial spin degeneracy). Let the system be periodic and translationally invariant in the length direction. Let V/L^d be bounded by a constant r (this constant r is arbitrary, and imposes some bound on the behavior of the aspect ratio of the system). Define V/L to be the "width" of the system, and let this number be odd. Then, we show that if the ground state is unique, the gap ΔE to the first excited state satisfies

$$\Delta E \leqslant c \ln(L)/L, \tag{1}$$

where the constant c depends on $\mathcal{H},d$, and where the result holds for all L greater than some minimum L_0, where L_0 depends on $\mathcal{H},d,r$.[8]

In this paper, we use the term gap to deal specifically with the difference between the energy of the first excited state and the energy of the ground state. This includes two completely distinct physical cases. In the case of a one-dimensional system, a spin-1/2 Heisenberg chain has a continuous spectrum of excitations above the ground state. On the other hand, a Majumdar-Ghosh[9] chain has a doubly degenerate ground state with a gap to the next excited state. Weak perturbations of the Majumdar-Ghosh Hamiltonian can break the exact degeneracy between the two lowest states, leaving a system with a gap from the ground state to the first excited state which is exponentially small in the system size, and then a gap from the first excited state to the next excited state which is nonvanishing even in the limit of large system sizes. We consider both of these cases as systems in which the gap ΔE is vanishing in the limit of large system size. Although they are one-dimensional systems, these two cases closely match the two possibilities mentioned above for higher-dimensional systems. The first case involves a system with a continuous spectrum as it has algebraically decaying spin correlations. In the second case, the first excited state is very close to the twisted state of LSM.

The physical idea behind the proof of Eq. (1) is closely related to the two possibilities considered above for the absence of a gap. In the event of long-range order, or algebraic long-range order, one expects that there is no gap. Conversely, if there is a gap, one expects that there is no long-range order. This is the first statement we prove: we assume that the system has a gap ΔE and show, in Sec. III, that connected expectation values decay exponentially in the spacing between them. Then, to prove Eq. (1) we first assume that Eq. (1) is violated, proceeding by contradiction. We assume the existence of such a gap violating Eq. (1) in Secs. IV, V, and VI. We then use the existence of the gap ΔE and the exponential decay of correlation functions to show an insensitivity of the system to boundary conditions in Sec. IV. Then, this insensitivity is used to construct a *low-energy*, twisted state in Sec. V. The construction of this twisted state

will to some extent follow the topological attempt[4] at proving the LSM theorem in higher dimensions, with some important differences outlined below. In Sec. VI we will show that the twisted state has a different momentum than the ground state. It is here that the odd width of the system becomes essential. Despite the different momentum compared to the ground state, the twisted state has, up to a small error, the same expectation values for all operators acting within any local region. Thus, we may refer to this state as a topologically excited state. Finally, in the Appendix, we briefly consider a version of the result showing exponential decay of correlation functions for systems governed by certain Markov processes, rather than quantum systems.

Since we will constantly deal with operator equations of motion, we introduce a set of "loop operators" as a basic technique. The loop operators, which will be suitably defined products of spin operators, can be naturally interpreted as a product of gauge fields around a loop.[10] However, the use of these operators avoids the uncontrolled approximations associated with the U(1) and Z_2 gauge theory techniques.[11] The introduction of these operators is not necessary to the main development, but provides a useful notation.

II. LOOP ALGEBRA

We define operators $i^{\mu\nu} \equiv \frac{1}{2}\delta^{\mu\nu} + \Sigma_a S_i^a \sigma_a^{\mu\nu}$, where S_i^a are the spin operators at site i and σ_a are the Pauli matrices, $a = x,y,z$. Thus, $i^{\mu\nu}$ is the 2×2 matrix of spin operators

$$\begin{pmatrix} \frac{1}{2} + S_i^z & S_i^x - iS_i^y \\ S_i^x + iS_i^y & \frac{1}{2} - S_i^z \end{pmatrix}. \tag{2}$$

We consider operators of the form $i^{\mu\nu} j^{\nu\rho} k^{\rho\sigma} \ldots m^{\alpha\mu}$, which we refer to as loop operators, where a summation over repeated indices is implied. Later we will often suppress the indices μ,ν, writing i rather than $i^{\mu\nu}$ to save space. Thus, we will write the loop operator mentioned above in the form $\mathrm{tr}(ijk\ldots m)$, where the trace tr refers to a trace over the greek indices $\mu,\nu,\ldots$. Below we also use a trace Tr; this trace Tr refers to a trace of quantum operators, summing over all states in the Hilbert space of the system. Using the rule $i^{\mu\nu} i^{\rho\sigma} = \delta^{\nu\rho} i^{\mu\sigma}$, it is always possible to reduce a given product of traces to a new product such that each site appears only once. Then, an operator $\mathrm{tr}(ijk\ldots m)$ permutes the spins around the sites $i,j,\ldots$.

Given an operator $O(t)$, the operator obeys the equation of motion $\partial_t O(t) = -i[O,\mathcal{H}]$. Consider, for example, a term $\mathrm{tr}(ij)$ in $\mathcal{H}$. We have $[i,\mathrm{tr}(ij)] = ij - ji$. As an illustration, let us give the full greek indices on this commutator: we have $[i^{\mu\nu},(i^{\alpha\beta} j^{\beta\alpha})] = i^{\mu\alpha} j^{\alpha\nu} - j^{\mu\alpha} i^{\alpha\nu}$, where summation over repeated indices α,β is assumed.

Introduce coordinates (x,y) to specify sites i, where x labels the unit cells along the direction of length and is defined up to integer multiples of L. The coordinate y labels the unit cells along the other lattice directions, as well as labeling the particular spin within the unit cell. Given two sites i and j on the lattice, we define the distance between them, written $|i-j|$, as the minimum number of moves by lattice vectors needed to move from the unit cell containing i to that containing j. On a square lattice, for example, this is the Manhattan distance. Then, let R denote the range of $\mathcal{H}$, the furthest distance between two sites in any term in $\mathcal{H}$. If all distances in a product of loops are less than $L/2$, we can define a winding number of the given product around the lattice in the length direction.[12] If all distances in the product remain less than $L/2-R$, then the dynamics $\partial_t O = -i[O,\mathcal{H}]$ does not connect sectors with different winding numbers. We will make use of the coordinates later by sometimes writing loop operators $\mathrm{tr}(ij\ldots)$ in the form $\mathrm{tr}((x_1,y_1)(x_2,y_2)\ldots)$, where i has coordinates (x_1,y_1), j has coordinates (x_2,y_2), and so on.

III. LOCALITY

We consider ground-state expectation values of operators $O_1, O_2, \ldots$, written $\langle O_1(t_1) O_2(t_2) \ldots \rangle$. The expectation values are not time ordered: the ordering of operators is as written. For a system with a unique ground state and an energy gap ΔE, on physical grounds one expects that connected correlation functions, defined as $\langle A(0)B(0)\rangle_c \equiv \langle A(0)B(0)\rangle - \langle A(0)\rangle\langle B(0)\rangle$, decay exponentially in distance (without loss of generality, we will assume $\langle A\rangle = \langle B\rangle = 0$ through the rest of this section). The proof of this locality bound will be done in this section. We will do this in two steps: first, we consider commutators of the form $[A(t),B(0)]$, where $A(0)$ and $B(0)$ are separated in space. We bound the operator norm[13] of the commutator for sufficiently small t, and thus bound its expectation value, in Eq. (4) below. The proof in this subsection will just be sketched; a more rigorous derivation is given by Lieb and Robinson.[14] This result provides a bound on the velocity of the system, as will be seen below. Then, in the following subsection, from this bound on the expectation value of the commutator and the existence of a gap, we use a spectral representation of the commutator to bound the connected correlation functions, thus obtaining the desired locality bound on the expectation value, Eq. (13). Finally, we close the section by giving a similar locality bound for operators separated in time.

A. Finite velocity

We define the distance between two operators O_1, O_2 to be l if the minimum distance between any pair of sites, i,j, where S_i appears in O_1 and S_j appears in O_2, is l. $A(0),B(0)$ are sums of products of spin or loop operators, which we suppose to be distance l apart.

We start with some notation. The Hamiltonian $\mathcal{H}$ can be written as a sum of terms $\mathcal{H} = \Sigma_i \mathcal{H}_i$, such that each $\mathcal{H}_i$ only contains spins operators on sites j with $|i-j| \leq R$. Let N_A denote the number of sites appearing in $A(0)$, and N_B denote the number of sites appearing in $B(0)$. Let J denote the maximum, over sites i, of $||\mathcal{H}_i||$.

We now bound the operator norm of the commutator $[A(t),B(0)]$ for short times. On short time scales, one expects that $A(t),B(0)$ are still separated in space, up to small correction terms, as we now show. Consider first $||[A(t),\mathcal{H}_i]||$, and study the change in this quantity as a function of time:

$$\left|\frac{||[A(t)-idt[A(t),\mathcal{H}],\mathcal{H}_i]||-||[A(t),\mathcal{H}_i]||}{dt}\right|$$
$$=\left|\frac{||[A(t),\mathcal{H}_i+idt[\mathcal{H}_i,\mathcal{H}]]||-||[A(t),\mathcal{H}_i]||}{dt}\right|$$
$$\leqslant\sum_{|i-j|\leqslant 2R}\left|\frac{||[A(t),\mathcal{H}_i+idt[\mathcal{H}_i,\mathcal{H}_j]]||-||[A(t),\mathcal{H}_i]||}{dt}\right|$$
$$=\sum_{|i-j|\leqslant 2R}\left|\frac{||[A(t)-idt[A(t),\mathcal{H}_j],\mathcal{H}_i]||-||[A(t),\mathcal{H}_i]||}{dt}\right|$$
$$\leqslant 2J\sum_{|i-j|\leqslant 2R}||[A(t),\mathcal{H}_j]||. \tag{3}$$

Here, we work to linear order in dt. While the operator $[A(t),\mathcal{H}_j]$ is differentiable, its operator norm need not be. Thus, all equations here are correct when we take the lim sup as $dt\to 0$.

The first equality in Eq. (3) is obtained by moving the time derivative from $A(t)$ to $\mathcal{H}_i$ as follows: for any operator P, to linear order in dt we have $||P||=||P+idt[P,\mathcal{H}]||$. Set $P=[A(t)-idt[A(t),\mathcal{H}],\mathcal{H}_i]$. Then, to linear order in dt, $||P||=||P+idt[P,\mathcal{H}]||=||[A(t)-idt[A(t),\mathcal{H}],\mathcal{H}_i]+idt[[A(t),\mathcal{H}_i],\mathcal{H}]||=||[A(t),\mathcal{H}_i+idt[\mathcal{H}_i,\mathcal{H}]]||$.

The inequality is obtained because $[\mathcal{H}_i,\mathcal{H}_j]=0$ for $|i-j|>2R$. The next equality is obtained by moving the time derivative back to $A(t)$, using now the equality $||P||=||P-idt[P,\mathcal{H}_j]||$. The final inequality results from the bound $||\mathcal{H}_i||\leqslant J$.

Now, let S denote the maximum number of sites j within distance R of any site i. Equation (3) gives a set of differential equations which bound the operator norm of various commutators; we have also the initial conditions that $||[A(0),\mathcal{H}_j]||$ vanishes for sites j which are further than distance R from any site in $A(0)$, while $||[A(0),\mathcal{H}_j]||\leqslant 2J||A||$ for all other sites. The number of sites within distance R of $A(0)$ is bounded by N_AS.

To bound $||[A(t),\mathcal{H}_j]||$, let us then consider the following set of differential equations: for $t>0$, we take $\partial_tG_i=2J\Sigma_{|i-j|\leqslant 2R}G_j$ and for $t<0$ we take $\partial_tG_i=-2J\Sigma_{|i-j|\leqslant 2R}G_j$, with initial conditions $G_j=0$ for sites j which are further than distance R from any site in $A(0)$, and $G_j=2J||A||$ for all other sites. Then, comparing these equations to Eq. (3), we see that $||[A(t),\mathcal{H}_i]||\leqslant G_i(t)$. This set of linear equations for G_i can be solved for any given lattice. However, we are simply interested in an upper bound on G_i. Let us define G^k to be the maximum of G_i over all sites i which are at a distance greater than $(2k-1)R$ from all sites in $A(0)$. Then, we have $\partial_tG^k\leqslant 2JSG^{k-1}$ for $k>0$, and $\partial_tG^k\leqslant 2JSG^k$ for $k=0$, with initial conditions $G^0=2J||A||$ and $G^k=0$ for $k\geqslant 1$. Thus, $G^0(t)\leqslant 2J||A||e^{2JSt}$, $G^1(t)\leqslant\int_0^t dt'\,2JSG^0(t')\leqslant\int_0^t dt'(2JS)(2J||A||)e^{2JSt}=2J||A||(2JSt)e^{2JSt}$, and $G^k(t)\leqslant 2J||A||(2JSt)^ke^{2JSt}/k!$. The last set of inequalities follows inductively: $G^k(t)\leqslant\int_0^t dt'(2JS)G^{k-1}(t)\leqslant\int_0^t dt'(2JS)(2J||A||)\times(2JSt')^{k-1}/(k-1)!e^{2JSt}=2J||A||(2JSt)^ke^{2JSt}/k!$. From these inequalities, we find, for a site j which is at a distance greater than $(2k-1)R$ from $A(0)$, that $||[A(t),\mathcal{H}_j]||\leqslant 2J||A||\,|2JSt|^ke^{2JS|t|}/k!$.

Finally, consider $\partial_t||[A(t),B(0)]||$. Using a similar sequence of inequalities to Eq. (3), we find that $\partial_t||[A(t),B(0)]||\leqslant 2||B||\Sigma_j||[A(t),\mathcal{H}_j]||$, where the sum over j extends over sites j which are within distance R of some sites in $B(0)$. There are at most N_BS such sites, and each of them has $||[A(t),\mathcal{H}_j]||\leqslant G^{k-1}(t)$, where $k=l/2R$. Here, we take k to be the ceiling of $l/2R$, the smallest integer greater than or equal to $l/2R$; we obtain this value of k since each such site is at least a distance $l-R$ from all sites in $A(0)$, and so we need $l-R>[2(k-1)-1]R$. Then, $||[A(t),B(0)]||\leqslant 2N_B||A||||B||\,|2JSt|^{l/2R}e^{2JS|t|}/(l/2R)!$.

Define $f(t)\equiv\langle[A(t),B(0)]\rangle$. Since $f(t)\leqslant||[A(t),B(0)]||$,

$$f(t)\leqslant\frac{2N_B||A||||B||\,|2JSt|^{l/2R}e^{2JS|t|}}{(l/2R)!}\equiv N_B||A||||B||g(t,l). \tag{4}$$

For $t=c_1l$, the large l behavior of $g(t,l)\sim\exp[(l/R)(2JSc_1R+1/2+(1/2)\ln(4JSc_1R))]$. If we choose a sufficiently small c_1, then $g(c_1l,l)$ decays exponentially in l for large l. Numerically, we find that the zero of $2JSc_1R+1/2+(1/2)\ln(4JSc_1R)$, is at $c_1\approx 0.139\,232/(2JSR)$. Any c_1 smaller than this value [for example, $c_1=0.1/(2JSR)$ will work] will cause $g(c_1l,l)$ to be exponentially decaying for large l. The velocity at which correlations spread in the system is of order c_1^{-1}.

B. Spectral decomposition

Now, we use a spectral decomposition of $f(t)$ to relate $f(t)$ to the desired correlation function, $\langle A(0)B(0)\rangle$. Without loss of generality, let us set the ground-state energy E_0 to 0. The spectral decomposition of $f(t)$ gives

$$f(t)=\sum_i\,(e^{-iE_it}A_{0i}B_{i0}-e^{iE_it}B_{0i}A_{i0}), \tag{5}$$

where A_{i0} is the matrix element of operator A between the ground state 0 and the eigenstate i, with energy $E_i\geqslant\Delta E$ above the ground-state energy, and similarly for the other A_{0i},B_{i0},B_{0i}. There are no terms in Eq. (5) involving A_{00},B_{00} since we have assumed $\langle A\rangle=\langle B\rangle=0$.

Let us define a function $f^+(t)=\Sigma_ie^{-E_it}A_{0i}B_{i0}$, which thus contains only the negative frequency (positive energy) terms in $f(t)$. The significance of $f^+(t)$ is that $f^+(t)=\langle A(t)B(0)\rangle$, so that the positive energy part of $f(t)$ contains the desired correlation function in it. In this subsection, we combine the bound (4) on $f(t)$ with the existence of a gap to bound $f^+(0)$.

Define $\tilde{f}(t)=f(t)e^{-t^2\Delta E^2/(2q)}$, with q to be chosen later. We have two bounds on $\tilde{f}(t)$. First, we have the bound (4) on $f(t)$ which gives us the bound $\tilde{f}(t)\leqslant N_B||A||||B||g(t,l)$. We also have

$$\tilde{f}(t)\leqslant 2||A||||B||e^{-t^2\Delta E^2/(2q)}. \tag{6}$$

We will use the first of these bounds for times $|t|<c_1 l$, and the second for long times $|t|>c_1 l$. Finally, we define $\tilde{f}^+(t)$ to contain only the negative frequency terms in $\tilde{f}(t)$.

Now, the desired expectation value $\langle A(0)B(0)\rangle = f^+(0)$. To bound $f^+(0)$, we first bound $\tilde{f}^+(0)$, and then bound $\tilde{f}(0)^+ - f^+(0)$. To bound $\tilde{f}^+(0)$, we use the bounds on $\tilde{f}$ and an integral representation of the positive energy part,[15,16]

$$
\begin{aligned}
|\tilde{f}^+(0)| &= \frac{1}{2\pi}\left|\int_{-\infty}^{\infty} dt\tilde{f}(t)/(-it+\epsilon)\right| \\
&= \frac{1}{2\pi}\left(\left|\int_{|t|<c_1 l}\tilde{f}(t)/(-it+\epsilon)\right| \right. \\
&\quad \left. + \left|\int_{|t|>c_1 l}\tilde{f}(t)/(-it+\epsilon)\right|\right) \\
&\leqslant \frac{1}{2\pi}||A||||B||\left(2N_B g(c_1 l,l) \right. \\
&\quad \left. + 2\frac{\sqrt{2\pi q}}{\Delta E c_1 l}e^{-c_1^2 l^2 \Delta E^2/(2q)}\right). \qquad (7)
\end{aligned}
$$

In Eq. (7), to bound the integral over $|t|<c_1 l$, we used $|\tilde{f}(t)|\leqslant|f(t)|\leqslant N_B||A||||B||g(t,l)$ $\leqslant N_B||A||||B||[|t|/(c_1 l)]g(c_1 l,l)$. To derive this inequality we have assumed that $l>0$ so that taking the ceiling of $l/2R$ above gives a $k\geqslant 1$. Then,

$$
\begin{aligned}
&\int_{|t|<c_1 l} dt\tilde{f}(t)/(-it+\epsilon) \\
&\quad \leqslant N_B||A||||B||\int_{|t|<c_1 l} dt g(c_1 l,l)/(c_1 l) \\
&\quad = 2N_B||A||||B||g(c_1 l,l). \qquad (8)
\end{aligned}
$$

To bound the integral over $|t|>c_1 l$ in Eq. (7) we have used Eq. (6) to show $|\int_{|t|>c_1 l}\tilde{f}(t)/(-it+\epsilon)|\leqslant 2||A||||B||\times\int_{|t|>c_1 l} dt\exp[-t^2\Delta E^2/(2q)]/(c_1 l)\leqslant 2||A||||B||(\sqrt{2\pi q}/\Delta E c_1 l)\times\exp[-c_1^2 l^2\Delta E^2/(2q)]$.

To bound $|\tilde{f}(0)^+ - f(0)^+|$, we start with the definition of $\tilde{f}$. Expressed as a convolution in Fourier space,[17] this is

$$
\tilde{f}(\omega') = (\sqrt{2\pi q}/\Delta E)\int d\omega f(\omega)e^{-q(\omega-\omega')^2/(2\Delta E^2)}. \qquad (9)
$$

Now is where the existence of an energy gap becomes essential. For motivation, let us first pictorially (see Fig. 1) describe how the gap enables us to bound $|\tilde{f}^+(0)-f^+(0)|$ and then present it more mathematically. By definition $\tilde{f}(0) = f(0)$; this follows in Fourier space from $\int_{-\infty}^{\infty} d\omega f(\omega) = \int_{-\infty}^{\infty} d\omega\tilde{f}(\omega)$. The convolution (9) means that a given Fourier component in f which is, for example, negative frequency, will produce both positive and negative frequency Fourier components in $\tilde{f}$. So, consider a δ function spike in

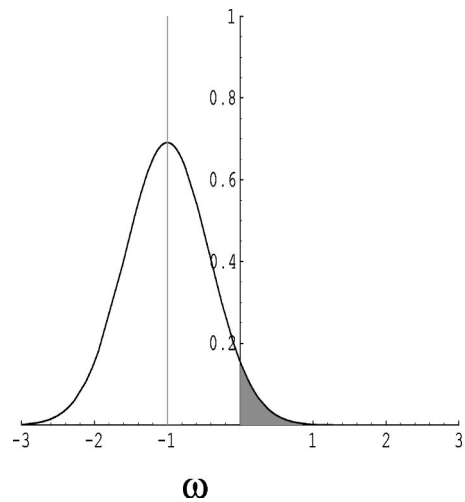

FIG. 1. Illustration of the bound on $|\tilde{f}^+(0)-f^+(0)|$, as described in the text. The vertical line describes a δ-function spike in $f(\omega)$. This produces a Gaussian in $\tilde{f}(\omega)$. The integral over ω of the Gaussian is the same as the integral of the δ function; however, the shaded region of the curve of the Gaussian falls above $\omega=0$, and hence does not contribute to $f^+(0)$. This leads to the difference between $\tilde{f}^+(0)$ and $f^+(0)$.

$f(\omega)$, produced by an intermediate state i with energy $E_i = -\omega>0$. This produces a Gaussian in $\tilde{f}(\omega)$, as shown. The integral over all ω of the Gaussian is the same as the integral of the δ function; however, the shaded portion of the curve has $\omega>0$. Since $\tilde{f}^+(0)=(2\pi)^{-1}\int_{-\infty}^{0} d\omega\tilde{f}(\omega)$ and $f^+(0) = (2\pi)^{-1}\int_{-\infty}^{0} d\omega f(\omega)$, we find a difference between $\tilde{f}^+(0)$ and $f^+(0)$ equal to the integral of the shaded portion of the curve. At $\omega=0$ the height of the Gaussian is reduced by a factor $e^{-q\omega^2/(2\Delta E^2)}$. However, since $E_i\geqslant\Delta E$, this factor is bounded by $e^{-q/2}$.

Now, let us do the calculation more directly: $\tilde{f}^+(0) = (2\pi)^{-1}\int_{-\infty}^{\infty} d\omega f(\omega)\int_{-\infty}^{0} d\omega'(\sqrt{2\pi q}/\Delta E)\exp[-q(\omega-\omega')^2/(2\Delta E^2)]$, while $f^+(0)=(2\pi)^{-1}\int_{-\infty}^{0} d\omega f(\omega)$. Then

$$
\tilde{f}^+(0) - f^+(0) = (2\pi)^{-1}\int_{-\infty}^{\infty} d\omega f(\omega)[\Theta_q(-\omega)-\Theta(-\omega)]. \qquad (10)
$$

Here $\Theta(\omega)$ is a step function: $\Theta(\omega)=1$ for $\omega>0$ and $\Theta(\omega)=0$ for $\omega<0$. We have defined

$$
\Theta_q(\omega) = \int_0^{\infty} d\omega'(\sqrt{2\pi q}/\Delta E)\exp[-q(\omega-\omega')^2/(2\Delta E^2)]. \qquad (11)
$$

Since the system has a gap, the integral in Eq. (10) vanishes for $|\omega|<\Delta E$. However, for $|\omega|\geqslant\Delta E$, we have $|\Theta_q(-\omega) - \Theta(-\omega)|\leqslant e^{-q/2}$. Thus, Eq. (10) is bounded by

$$
(2\pi)^{-1}e^{-q/2}\int_{-\infty}^{\infty} d\omega|f(\omega)|\leqslant 2||A||||B||e^{-q/2}. \qquad (12)
$$

Thus, combining Eqs. (7) and (12), $|f^+(0)| \leqslant |\tilde{f}^+(0) - f^+(0)| + |\tilde{f}^+(0)| \leqslant 1/2\pi ||A||||B||(2N_B g(c_1 l, l) + 2(\sqrt{2\pi q}/\Delta E c_1 l) e^{-c_1^2 l^2 \Delta E^2/(2q)}) + 2||A||||B|| e^{-q/2}$. We finally choose $q = c_1 l \Delta E$ to get

$$|\langle A(0)B(0)\rangle_c| \leqslant \frac{1}{2\pi} 2N_B ||A||||B|| g(c_1 l, l) + 2\left(1 + \frac{1}{\sqrt{2\pi c_1 l \Delta E}}\right) \times ||A||||B|| e^{-c_1 l \Delta E/2}, \tag{13}$$

giving the desired bound. The first term in Eq. (13) decays as $\exp[-\mathcal{O}(l/R)]$, while the second term decays as $\exp[-c_1 l \Delta E/2] = \exp\{-\mathcal{O}[\Delta E l/(JSR)]\}$; here, by $\mathcal{O}(l/R)$, we mean some quantity of order l/R.

In what follows in the next three sections, the first term in Eq. (13) will be negligible: we will be considering operators separated by a distance l which is of order L, so that the first term in Eq. (13) will lead to only exponentially small (in L) contributions to the correlation functions. The second term will be more important: since we will consider gaps $\Delta E \propto \ln(L)/L$, the second term will lead to terms which are suppressed only by powers of L when considering correlation functions of operators separated by a distance of order L.

C. Operators at different times

It is possible to extend the result (13) to correlation functions $\langle A(0)B(i\tau)\rangle$, with τ real and $\tau > 0$. Then, in Eq. (7), we must evaluate $|\tilde{f}^+(-i\tau)|$, so that the denominator $(-it + \epsilon)$ is replaced by $-it + \tau$. In this case, we are still able to find just as tight a bound on $|\tilde{f}^+(-i\tau)|$ as we previously found for $|\tilde{f}^+(0)|$: $|\tilde{f}^+(-i\tau)| \leqslant (1/2\pi)||A||||B||(2N_B g(c_1 l, l) + 2(\sqrt{2\pi q}/\Delta E c_1 l) e^{-c_1^2 l^2 \Delta E^2/(2q)})$.

Of course, for $\tau \geqslant q/\Delta E$ there is the trivial bound $|\langle A(0)B(i\tau)\rangle_c| \leqslant ||A||||B|| e^{-\tau \Delta E} \leqslant ||A||||B|| e^{-q}$. For $|\tau| \leqslant q/\Delta E$, we claim that $|\tilde{f}^+(-i\tau) - \exp[+\tau^2 \Delta E^2/(2q)] f^+(-i\tau)| \leqslant 2||A||||B|| e^{-q/2}$. To show this,

$$\begin{aligned} &|\tilde{f}^+(-i\tau) - \exp[+\tau^2 \Delta E^2/(2q)] f^+(-i\tau)| \\ &= (2\pi)^{-1} \int_{-\infty}^{\infty} d\omega f(\omega) \\ &\times \left(\int_{-\infty}^{0} d\omega' (\sqrt{2\pi q}/\Delta E) \exp[\omega' \tau] \right. \\ &\times \exp[-q(\omega - \omega')^2/(2\Delta E^2)] \\ &\left. - \exp[+\tau^2 \Delta E^2/(2q)] \Theta(-\omega) \exp[\omega \tau] \right). \end{aligned} \tag{14}$$

The portion of the integral with $\omega < 0$ is equal to

$$(2\pi)^{-1} \int_{-\infty}^{-\Delta E} d\omega f(\omega) \int_0^{\infty} d\omega' (\sqrt{2\pi q}/\Delta E) \exp[\omega' \tau] \times \exp[-q(\omega - \omega')^2/(2\Delta E^2)], \tag{15}$$

where we have used the gap ΔE and the relation $\int_{-\infty}^{\infty} d\omega' (\sqrt{2\pi q}/\Delta E) \exp[\omega' \tau] \exp[-q(\omega - \omega')^2/(2\Delta E^2)] = \exp[\omega \tau] \exp[+\tau^2 \Delta E^2/(2q)]$. Then, for $\tau \leqslant q/\Delta E$, the integral (15) with $\omega < 0$ is bounded in absolute value by $(2\pi)^{-1} \int_{-\infty}^{0} d\omega |f(\omega)| e^{-q/2}$. We can similarly bound the portion of the integral with $\omega > 0$, giving the desired result.

With the given $q = c_1 l \Delta E$ the above bounds show that for $\tau \leqslant c_1 l$,

$$|\langle A(0)B(i\tau)\rangle_c| \leqslant e^{-\tau^2 \Delta E/(2c_1 l)} \left[\frac{1}{2\pi} 2N_B ||A||||B|| g(c_1 l, l) + 2\left(1 + \frac{1}{\sqrt{2\pi c_1 l \Delta E}}\right) ||A||||B|| e^{-c_1 \Delta E l/2} \right]. \tag{16}$$

IV. TWISTED BOUNDARY CONDITIONS

In this section we derive some results on the sensitivity to boundary conditions, as a step towards the the main result, Eq. (1). To derive a contradiction later, we will assume throughout this and the next two sections that there is a gap ΔE that violates Eq. (1), with an appropriately chosen c. In the first subsection, we review the twist of boundary conditions and the topological attempt at proving the LSM theorem. In the second subsection, we show the specific results on the sensitivity to boundary conditions.

A. Topological argument

Here we will define a new twisted Hamiltonian, making use of the coordinates (x,y) introduced previously for lattice sites i. To define the new twisted Hamiltonian $\mathcal{H}_{\theta,\theta'}$, replace all loop operators $\mathrm{tr}((x_1,y_1)(x_2,y_2)(x_3,y_3)\ldots)$ in $\mathcal{H}$ with $\mathrm{tr}((x_1,y_1)R(x_1,x_2)(x_2,y_2)R(x_2,x_3)(x_3,y_3)\ldots)$. Here, the twist operator $R^{\mu\nu}(x_1,x_2) \equiv \exp[\pm i(\theta/2)\sigma_z^{\mu\nu}]$ if the shortest lattice path between x_1, x_2 crosses from $x=0$ to $x=1$, where the sign is positive if the path crosses in the direction of increasing x and negative if it crosses in the opposite direction. Here, $R^{\mu\nu}(x_1,x_2)$ is a 2×2 matrix of numbers, rather than of operators,

$$R^{\mu\nu} = \begin{pmatrix} \exp\left[\pm i \dfrac{\theta}{2}\right] & 0 \\ 0 & \exp\left[\mp i \dfrac{\theta}{2}\right] \end{pmatrix}. \tag{17}$$

Alternately, if the shortest lattice path between x_1, x_2 crosses from $L/2$ to $L/2+1$, $R^{\mu\nu}(x_1,x_2) \equiv \exp[\pm i(\theta'/2)\sigma_z^{\mu\nu}]$. Otherwise, $R(x_1,x_2) = \delta^{\mu\nu}$. In Fig. 2, we show the coordinate system using x,y and show where the two boundary condition twists are inserted.

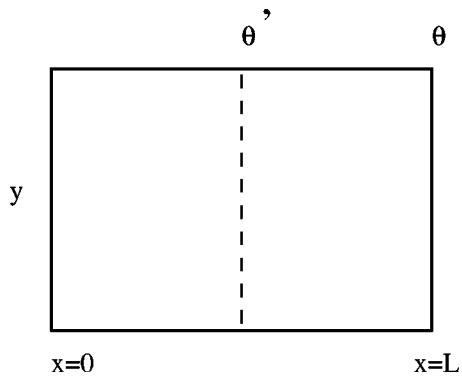

FIG. 2. Plot of the system, showing the x coordinate along the length axis. The x is shown ranging from $x=0$ to $x=L$; due to the periodicity of the system, $x=0$ is identified with $x=L$. The y coordinate specifies the position in the directions normal to the length, as well as specifying the particular site in each unit cell. The twist angles are noted; the twist θ changes the boundary condition near $x=L$, while the twist θ' changes the coupling between sites near $x=L/2$.

Let us see what the effect of this twist is in terms of spin operators. Consider two sites, i,j. Suppose the Hamiltonan $\mathcal{H}_{0,0}$ has a term such as $\mathrm{tr}(ij)=2(S_i^zS_j^z+S_i^xS_j^x+S_i^yS_j^y)+1/2$. Then, let us suppose i has $x=0$ while j has $x=1$. Then, $\mathcal{H}_{\theta,\theta'}$ has a term $\mathrm{tr}(iR(0,1)jR(1,0))$. In terms of spin operators, this is equal to $2[S_i^zS_j^z+\cos(\theta)(S_i^xS_j^x+S_i^yS_j^y)+\sin(\theta)(S_i^xS_j^y-S_i^yS_j^x)]+1/2$. In the untwisted Hamiltonian, we coupled the dot product of the two spin vectors, $\vec{S}_i,\vec{S}_j$; in the twisted Hamiltonian, we couple them after rotating one by an angle θ about the z axis. A good discussion of twists can be found in Ref. 5.

We have considered two different twist angles, θ,θ'. The spectrum of $\mathcal{H}_{\theta,\theta'}$ depends only on the combination $\theta+\theta'$. Further, from any given eigenfunction $\psi(\theta,\theta')$ of $\mathcal{H}_{\theta,\theta'}$, one can find an eigenfunction $\psi(\theta-\delta\theta,\theta'+\delta\theta)$ of $\mathcal{H}_{\theta-\delta\theta,\theta'+\delta\theta}$ by $\psi(\theta-\delta\theta,\theta'+\delta\theta)=\Pi_j e^{i\delta\theta S_j^z}\psi(\theta,\theta')$, where the product extends over all sites $j=(x,y)$ with $0<x\leqslant L/2$.

Given that the spectrum depends only on the combination $\theta+\theta'$, the reader may wonder what the reason is for introducing two twist angles, rather than just one angle. In fact, the second angle is a useful trick, introduced for the following reason: we have previously shown that the existence of a gap causes correlation functions to decay exponentially in the separation of the two operators. However, physically, one expects that the existence of a gap will also imply some insensitivity of the system to boundary conditions, enabling us to bound, for example, the second derivative of the ground-state energy with respect to θ. What we will do in the following subsection is show this insensitivity by using the fact that the spectrum depends only on $\theta+\theta'$ to convert the second derivative (∂_θ^2) of the ground-state energy into a mixed partial derivative $(\partial_\theta\partial_{\theta'})$ of the ground-state energy, and by then evaluating that mixed partial derivative as a correlation function, using the exponential decay of correlation functions. This will be stated more precisely at the start of the following subsection; we mention it here for motivation.

The eigenvalues are of $\mathcal{H}_{\theta,\theta'}$ are invariant under $\theta+\theta'\rightarrow\theta+\theta'+2\pi$, while the wave functions are invariant under $\theta+\theta'\rightarrow\theta+\theta'+4\pi$.[4,5] To motivate the results in this section, we recall the basic idea of the topological attempt[4] at proving the LSM. The idea is that if there is a gap at $\theta+\theta'=0$, and *if the gap remains open* for all $\theta+\theta'$, then under an adiabatic change in the angle θ with θ' fixed at zero, the ground state at $\theta=0$ evolves into the ground state at $\theta=2\pi$. At $\theta=2\pi$, the Hamiltonian is returned to the original Hamiltonian, but, for a system of odd width, the ground state expectation value of the translation operator changes sign, as will be discussed in more detail below. This leads to a contradiction: from the ground state with given expectation value of the translation operator, we construct another ground state with the opposite expectation value. The requirement that the topological attempt requires the gap to remain open for all θ was pointed out in Ref. 5.

What the topological argument actually succeeds in showing is that the gap must close at some value of θ. However, in order to use this argument to obtain any bound on the magnitude of the gap at $\theta=0$, we would have to show that a sufficiently large gap at $\theta=0$ would prevent the gap from closing for all θ; that would then lead to a contradiction, enabling us to bound the gap at $\theta=0$. What we will see is that we can partially show this: for sufficiently large c in Eq. (1), we can show to second order in θ (or indeed, to any finite order) a bound on the change in ground-state energy with respect to θ. However, we will be unable to show that the gap remains open for all θ because to bound the change in ground-state energy for higher orders in θ requires progressively increasing the constant c in Eq. (1), and it is not possible to show the result to all orders. Thus, the topological attempt will ultimately fail, and we will give a physical example of how this can happen. In the following Sec. V, we will give a successful argument.

B. Boundary condition sensitivity

We now show an insensitivity of the ground-state energy $E_0(\theta,\theta')$ to second order[18] in the twist angle $\theta+\theta'$. At $\theta=\theta'=0$, $\partial_\theta E_0(\theta,\theta')=\langle\partial_\theta\mathcal{H}_{\theta,\theta'}\rangle=0$. Indeed, taking any *odd* number of derivatives of $E_0(\theta,\theta')$ leads to a vanishing quantity.[19] To second order in θ,θ', we write a power series: $E_0(\theta,\theta')=E_0(0,0)+a\theta^2/2+a\theta'^2/2+b\theta\theta'$, where $a=\partial_\theta^2E_0=\partial_{\theta'}^2E_0$ and $b=\partial_{\theta\theta'}E_0$. We will show that, for any given negative power of L, we can find a constant c such that if Eq. (1) is violated for that c, then a is bounded by a $\mathcal{H}$-dependent constant times the given negative power of L. We do this by calculating b as a correlation function, and then showing that $b=a$.

Recall linear perturbation theory: suppose a Hamiltonian $\mathcal{H}$ is changed by some $\delta\mathcal{H}$. For a nondegenerate state $|\psi\rangle$ with eigenvalue E, the change $|\delta\psi\rangle$ in $|\psi\rangle$ is given to linear order in $\delta\mathcal{H}$ by $|\delta\psi\rangle=(E-\mathcal{H})^{-1}\delta\mathcal{H}|\psi\rangle$. Since the ground state is the lowest-energy state, all other states have

energies greater than it. Thus, we can write the change in the ground state to linear order as

$$|\delta\psi_0\rangle = -\Sigma_{a\neq 0}\int_0^\infty d\tau e^{(E_0-E_a)\tau}|\psi_a\rangle\langle\psi_a|\delta\mathcal{H}|\psi_0\rangle$$
$$= -\int_0^\infty d\tau\delta\mathcal{H}(i\tau)|\psi_0\rangle,$$

where $|\psi_0\rangle$ is the ground-state wave function, $|\psi_a\rangle$ are a complete set of intermediate states, and where $\delta\mathcal{H}(i\tau) = \exp[-\mathcal{H}\tau]\delta\mathcal{H}\exp[\mathcal{H}\tau]$ is the change in the Hamiltonian operator, taken at imaginary time $i\tau$. Here we have set $E_0(0,0)=0$ without loss of generality.

Specializing to the case of $\delta\mathcal{H}=\partial_\theta\mathcal{H}_{\theta,\theta'}$ and writing the change in ψ_0 in terms of the θ,θ'-dependent ground-state density matrix $\rho^0(\theta,\theta')\equiv|\psi_0(\theta,\theta')\rangle\langle\psi_0(\theta,\theta')|$ we have

$$\partial_\theta\rho^0 = -\int_0^\infty d\tau\partial_\theta\mathcal{H}_{\theta,\theta'}(i\tau)\rho^0 - \int_{-\infty}^0 d\tau\rho^0\partial_\theta\mathcal{H}_{\theta,\theta'}(i\tau),$$

$$\partial_{\theta'}\rho^0 = -\int_0^\infty d\tau\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(i\tau)\rho^0 - \int_{-\infty}^0 d\tau\rho^0\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(i\tau). \tag{18}$$

Note that since $\langle\delta\mathcal{H}\rangle$ vanishes in this case, we do not need to worry about matrix elements of $\delta\mathcal{H}$ from the ground state to the ground state.

Now, we can use the change in the density matrix to compute b by $b=\mathrm{Tr}[(\partial_{\theta'}\mathcal{H}_{\theta,\theta'})(\partial_\theta\rho^0)]$. So,

$$b = -\int_0^\infty d\tau(\langle\partial_\theta\mathcal{H}_{\theta,\theta'}(0)\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(i\tau)\rangle$$
$$-\langle\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(-i\tau)\partial_\theta\mathcal{H}_{\theta,\theta'}(0)\rangle), \tag{19}$$

where the derivatives are evaluated at $\theta=\theta'=0$. The derivative $\partial_\theta\mathcal{H}_i$ is nonvanishing only for sites i which are within distance R of $x=0$; there are at most SV/L such sites. For each i, $||\partial_\theta\mathcal{H}_i||\leqslant JS$, so $||\partial_\theta\mathcal{H}_{\theta,\theta'}||\leqslant JS^2V/L$. We use two bounds for the given correlation functions in Eq. (19). First, each correlation function is bounded by $(JS^2V/L)^2e^{-\tau\Delta E}$. Second, we can use Eq. (16) to bound each correlation function by

$$2S(JS^2V/L)^2(1+1/\sqrt{\pi c_1 L\Delta E})\exp[-c_1\Delta E(L/2)/2], \tag{20}$$

where we neglect the term in $g(c_1 l,l)$ in Eq. (16) as it leads to a correction which is exponentially decaying in L, not in $c_1\Delta EL$, and thus is negligible in what follows. Also, we have used $l=L/2$, ignoring the slight error that in fact $l \geqslant L/2-R$. Finally, we have used $N_B\leqslant S$ in Eq. (20).

Using these two bounds on the correlation function, we arrive at

$$|b|\leqslant 2(JS^2V/L)^2\int_0^\infty d\tau\min\{\exp[-\tau\Delta E],x\}, \tag{21}$$

where $x=2S(1+1/\sqrt{\pi c_1 L\Delta E})\exp[-c_1\Delta E(L/2)/2]$. Thus, $|b|\leqslant 2(JS^2V/L)^2(x/\Delta E)(1+\ln(x))$. The number of sites V is bounded by rL^d, while for ΔE greater than $c\ln(L)/L$, x is bounded by a c-dependent negative power of L. Therefore, we can bound $|b|$ by any desired negative power of L by choosing c sufficiently large.

However, $E_0(\theta,-\theta)=E_0(0,0)$, so $b=a$. Thus, we have also bounded $|a|$ by the same negative power of L. Therefore, at $\theta=\theta'=0$, we find that $|\partial_\theta^2E_0(\theta,\theta')|$ is bounded by a negative power of L. This shows some insensitivity of the ground-state energy to boundary conditions. This realizes the physical idea[5] that a spin liquid state is defined by the lack of response to a twist in boundary conditions to second order in θ.[20]

At fourth order in θ, we must evaluate a correlation function of *four* operators, each of order JS^2V/L; to bound these correlation functions requires a larger c. Each higher order in θ,θ' requires an even larger c, so that it is not possible to bound the change in ground-state energy for arbitrary $\theta+\theta'$. Therefore, the topological attempt[4] to establish the LSM result fails. Indeed, a gap at $\theta+\theta'=0$ must close for $\theta+\theta'\neq 0$.[5]

It is worth giving a specific physical example of this possibility, as the topological argument does show that the gap must close for some $\theta+\theta'$. In many physical examples of spin liquids, the closing of the gap arises because a state which is at some very low energy, of order JL^{-1} or less, above the ground state at $\theta+\theta'=0$ crosses the ground-state energy at a finite $\theta+\theta'$. For example, if the Majumdar-Ghosh Hamiltonian is slightly perturbed, there is a state at an exponentially small energy above the ground state which crosses the ground state at $\theta+\theta'=\pi$.

However, it is also possible for a state which is at some energy JL^0 to cross the ground state: consider a system with two competing phases, one of which is a spin liquid phase while the other is a spin ordered phase with a spiral order. The spiral order is chosen so that the spin ordered phase can be frustrated at $\theta+\theta'=0$, and the spin liquid is the ground state there. At some $\theta+\theta'\neq 0$, however, the spiral phase can take over as the ground state. This taking over as the ground state can happen either via a level crossing (if the two states have different symmetry, for example, or if the spin ordered phase has a nonvanishing net spin), or via an avoided crossing. This provides a specific example of a system in which a state or phase which is at an energy of order JL^0 at $\theta+\theta'=0$ becomes the ground state at some nonvanishing $\theta+\theta'$.

The solution to this problem is simple: it is not necessary to show that there is a gap for all twist angles. Instead, we start with the ground state at vanishing twist and continuously evolve this state, obtaining a state for any twist angle which is an approximate eigenstate of the twisted Hamiltonian, not necessarily the ground state. This approximate eigenstate will be explicitly constructed in the following section, while in the section after that we demonstrate that at a twist of 2π the expectation value of the translation operator has changed sign in the new state compared to the ground state. Thus, this gives a new low-energy state, different from the ground state.

V. TWISTING THE GROUND STATE

A. Constructing the twisted state

Let $\rho(\theta,0)$ be a θ-dependent density matrix that we construct below. Divide the system into two overlapping halves:

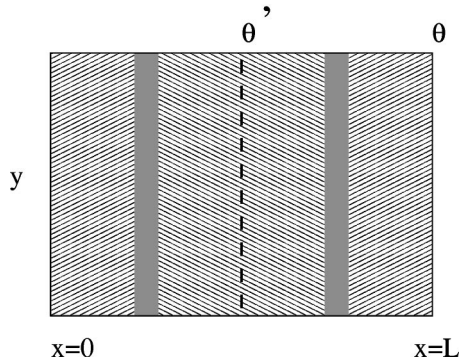

FIG. 3. Plot of the system, showing the twists and coordinates as before. The halves of the system have been shaded in. The shading at the left and right sides of the system (diagonal lines going up and right) denotes sites in half (1), the shading in the middle (diagonal lines going up and left) denotes sites in half (2). The solid shading denotes sites in both halves; the length of the solid region is at least $2R$, so that the Hamiltonian can be written as a sum of terms, each of which is contained in only one half.

half (1) contains sites with $x=3L/4-R,3L/4-R+1,\ldots,L-1,0,1,\ldots,L/4+R$, while half (2) contains sites with $x=L/4-R,L/4-R+1,\ldots,3L/4+R$. That is, half (1) contains all sites from $x=3L/4-R$ up to $x=L-1$, as well as all sites from $x=0$ up to $L/4+R$, while half (2) contains all sites from $x=L/4-R$ up to $x=3L/4+R$. The halves are shown as shaded regions in Fig. 3.

The reason we choose two overlapping halves is that we will be considering density matrices which involve only sites within a given half. These matrices will be defined by tracing over sites outside the given half. Then, to evaluate the expectation value of the energy of the system, we will be able to evaluate the expectation value as a sum of operators which lie completely within one or the other half. That is, by making the two halves overlap, we will deal with the question of the "seam" where the two halves join. This is mentioned here as motivation and will be done in more detail below.

Define $\rho_1(\theta,0)=\mathrm{Tr}_2[\rho(\theta,0)]$, where Tr_2 denotes a trace over all sites *not in half (1)*, and define $\rho_2(\theta,0)=\mathrm{Tr}_1[\rho(\theta,0)]$, the trace over sites not in half (2). Similarly, define $\rho_1^0(\theta,\theta')=\mathrm{Tr}_2[\rho^0(\theta,\theta')]$, and $\rho_2^0(\theta,\theta')=\mathrm{Tr}_1[\rho^0(\theta,\theta')]$. We will assume throughout this section that there is a gap violating Eq. (1). Then, for sufficiently large c, we will construct ρ_1 such that

$$\rho_1(\theta,0)-E_1(\theta)=\left(\prod_j e^{i\theta S_j^z}\right)\rho_1^0(0,0)\left(\prod_j e^{-i\theta S_j^z}\right)=\rho_1^0(\theta,-\theta), \tag{22}$$

where the products extend over all sites $j=(x,y)$ with $0<x\leqslant L/4$ and where $E_1(\theta)$ is an error term such that the trace norm[21] $|E_1(\theta)|$ is bounded by a constant times a negative power of L for all $0\leqslant\theta\leqslant 2\pi$. The particular negative power of L can be determined by choosing the constant c in Eq. (1). As a useful terminology, we will refer to a quantity as "small" if, for any desired negative power of L, we can find sufficiently large c or sufficiently large q (introduced below), such that the given quantity is bounded by a constant times the given negative power of L for all L. Thus, we wish $|E_1(\theta)|$ to be small. Note that, given this definition of small, if a small quantity is multiplied by any fixed power of L, the result is a small quantity. Sometimes, we will indicate that a quantity is made small by choosing c or by choosing q, to specify which of the two needs to be made sufficiently large.

In differential form, we require

$$\partial_\theta\rho_1(\theta,0)=\sum_j i[S_j^z,\rho_1^0(\theta,-\theta)]+e_1(\theta), \tag{23}$$

where $e_1=\partial_\theta E_1$. We will show that the upper[22] derivative $\mathcal{D}_\theta|E_1(\theta)|$ is small, from which Eq. (22) will follow. We will also require $\rho_2(\theta,0)=\rho_2^0(0,0)$, up to a similarly bounded error term $E_2(\theta)$, and $\partial_\theta\rho_1(\theta,0)=e_2(\theta)$, with a similarly bounded $\mathcal{D}_\theta|E_2(\theta)|$.

The physical motivation behind Eq. (22) is to construct a state for the Hamiltonian $\mathcal{H}_{\theta,0}$ that has an energy close to $E_0(0,0)$. The twist θ is along a line that lies completely within half (1) while θ' is along a line that lies completely within half (2). Within half (1), the Hamiltonians $\mathcal{H}_{\theta,0}$ and $\mathcal{H}_{\theta,-\theta}$ are equal, so we construct a density matrix such that *within half (1)* the given density matrix is close to the ground-state density matrix of $\mathcal{H}_{\theta,-\theta}$. Then, the expectation of any operator O which lies completely within half (1) for the density matrix $\rho(\theta,0)$ will be within $||O||||E_1(\theta)|$ of the expectation value of that operator for the density matrix $\rho(\theta,-\theta)$. On the other hand, within half (2), the Hamiltonians $\mathcal{H}_{\theta,0}$ and $\mathcal{H}_{0,0}$ are equal, so we also require that within half (2) the density matrix be close to the ground-state density matrix of $\mathcal{H}_{0,0}$.

Then, the expectation value of the energy in the state defined by $\rho(\theta,0)$ is equal to $\mathrm{Tr}[\rho(\theta,0)\mathcal{H}_{\theta,0}]$. Once we have shown that both Eq. (22) and the bound on $|E_2(\theta)|$ are satisfied, it will follow that this expectation value will be within an amount $||\mathcal{H}_{\theta,0}||\max(|E_1|,|E_2|)$ of $E_0(0,0)$, since the Hamiltonian $\mathcal{H}$ can be written as a sum of operators which are entirely within half (1) or entirely within half (2) (it was for this reason that the halves were chosen to overlap). Therefore, since $||\mathcal{H}||$ is bounded by $VJ<rL^dJ$, if we pick c in Eq. (1) sufficiently large, we will find that $\mathrm{Tr}[\rho(\theta,0)\mathcal{H}_{\theta,0}]-E_0(0,0)$ will also be small at $\theta=2\pi$; this follows from the statement that a small quantity multiplied by a fixed power of L is also small.

Our claim, which we show in this section, is that Eq. (22) is satisfied by a $\rho(\theta,0)$ defined as follows for $0\leqslant\theta\leqslant 2\pi$. We pick

$$\partial_\theta\rho(\theta,0)=-\int_0^{c_1L}d\tau[A^+(i\tau)-A^-(-i\tau),\rho(\theta,0)], \tag{24}$$

where we define

$$A^+(i\tau) = (2\pi)^{-1}\exp[-(\tau\Delta E)^2/(2q\ln(L))]\int_{-\infty}^{\infty} dt\,\partial_\theta\mathcal{H}_{\theta,0}(t) \times\exp[-(t\Delta E)^2/(2q\ln(L))]/(-it+\tau) \tag{25}$$

with q to be chosen later, and $A^-(-i\tau)=(A^+(i\tau))^\dagger$. The time evolution of the operator $\partial_\theta\mathcal{H}_{\theta,\theta'}(t)$ is defined using the Hamiltonian $\mathcal{H}_{\theta,-\theta}$, while the τ dependence of $A^{+,-}$ is defined via Eq. (25).

To give some insight into the definition of A^+,A^-, we note that if q were to be infinite, then they would project onto positive and negative energy parts of $\partial_\theta\mathcal{H}$ at times $\pm i\tau$, respectively. That is, for $q=\infty$, we have $A^+(i\tau)=(2\pi)^{-1}\int_{-\infty}^{\infty}dt\,\partial_\theta\mathcal{H}_{\theta,\theta'}(t)/(-it+\tau)$. Let the matrix elements of the operator $\partial_\theta\mathcal{H}$ in a basis of eigenstates of $\mathcal{H}$ be written $(\partial_\theta\mathcal{H})_{ab}$ where the states have energies E_a,E_b. Let the states have energy difference $-\omega=E_a-E_b$. Then, doing the integral over t we find that $A^+(i\tau)$ has a matrix element between states a,b equal to $\exp[\omega\tau](\partial_\theta\mathcal{H})_{ab}$ for $-\omega>0$ and equal to zero for $-\omega<0$. Similarly, $A^-(-i\tau)$ has a matrix element equal to $\exp[-\omega\tau](\partial_\theta\mathcal{H})_{ab}$ for $-\omega<0$ and equal to zero for $-\omega>0$. Then, for any given time τ, the integrand of Eq. (24) would be the same as that of Eq. (18) for $\theta=0$, since in that case the only nonvanishing terms in Eq. (24) are $-A^+(i\tau)\rho(0,0)-\rho(0,0)A^-(-i\tau)$.

What we will do later is to take a finite q instead. Physically, this means that rather than taking an adiabatic change in θ which keeps us in the ground state, we instead "pass through" the level crossing when the gap closes at some $\theta\neq 0$, going from the ground state to some low-energy excited state.

Equation (24) gives the change in ρ equal to the commutator of ρ with an anti-Hermitian operator, and hence generates an infinitesimal unitary transformation of ρ. Thus, ρ continues to be a density matrix which projects onto a single state, defined to be $\psi(\theta,0)$.

As a first step, we wish to show that for $\theta=\theta'=0$ we can find a c such that $\partial_\theta\rho_1^0(\theta,0)-\Sigma_j i[S_j^z,\rho_1^0(0,0)]$ is small. We have

$$\partial_\theta\rho_1^0(\theta,0)-\sum_j i[S_j^z,\rho_1^0(0,0)]=\partial_\theta\rho_1(\theta,0)-\partial_\theta\rho_1^0(\theta,-\theta) =\mathrm{Tr}_2[\partial_{\theta'}\rho^0(0,\theta')], \tag{26}$$

where all derivatives are evaluated at $\theta=\theta'=0$. To bound the right-hand side of Eq. (26), consider the trace of this term with any operator O with $||O||=1$. This operator must be within half (1), so, using Eq. (18) to compute the derivative of ρ^0 with respect to θ', we obtain the expectation value

$$\mathrm{Tr}[O\partial_{\theta'}\rho^0(0,\theta')]=\partial_{\theta'}\langle O\rangle=-\int_0^\infty d\tau\langle O\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(i\tau) +\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(-i\tau)O\rangle. \tag{27}$$

However, following the arguments from the preceding section and the locality bounds, we can find a c such that Eq. (27) is small. In this case, the distance between O and $\partial_{\theta'}\mathcal{H}$ is at least $L/4-2R$, since $\partial_{\theta'}\mathcal{H}$ includes terms with x down to $L/2-R$, while O is in half (1) so includes x up to $L/4+R$. Note that $\langle\partial_{\theta'}\mathcal{H}_{\theta,\theta'}\rangle=0$ at $\theta=\theta'=0$. Since we have bounded the trace of the right-hand side of Eq. (26) with all operators O with $||O||=1$, we have bounded the trace norm of the right-hand side.

B. Bound on error terms

We now show that we can find a c such that the definition (24) satisfies Eq. (23) in general. We wish to compute $\partial_\theta\rho_1(\theta,0)-\Sigma_j i[S_j^z,\rho_1(\theta,-\theta)]=e_a(\theta)+e_b(\theta)+e_c(\theta)$. Here we define

$$e_a\equiv-\int_0^{c_1L}d\tau\,\mathrm{Tr}_2[[A^+(i\tau)-A^-(-i\tau),\rho(\theta,0)] -[A^+(i\tau)-A^-(-i\tau),\rho^0(\theta,-\theta)]]. \tag{28}$$

$$e_b\equiv-\int_0^{c_1L}\mathrm{Tr}_2[[A^+(i\tau)-A^-(-i\tau),\rho^0(\theta,-\theta)]] -\partial_\theta\rho_1^0(\theta,\theta'), \tag{29}$$

$$e_c\equiv\partial_\theta\rho_1^0(\theta,\theta')-\sum_j i[S_j^z,\rho_1^0(\theta,-\theta)]. \tag{30}$$

In Eq. (29), the derivative of ρ_1^0 is evaluated at $\theta=-\theta'$. We now consider each of these terms e_a,e_b,e_c in turn.

First, consider Eq. (28). In the definition of $A^{+,-}$ as an integral over t, the integral over times $|t|>c_1(L/2-R)$ has an operator norm bounded by $||\partial_\theta\mathcal{H}||\int_{|t|>c_1(L/2-R)}\times\exp[-(t\Delta E)^2/(2q\ln(L))]$. Thus, for any fixed q (to be chosen later) we can find a c such that this integral over times $|t|>c_1(L/2-R)$ has small operator norm, and thus when commuted with $\rho(\theta,0)$ gives a term with small trace norm.

Equation (28) involves an integral of $\partial_\theta\mathcal{H}(t)$ over time t in the definition of $A^{+,-}$; we have shown that the contributions with times $|t|>c_1(L/2-R)$ may be neglected. Then, considering only contributions with $|t|\leq c_1(L/2-R)$, we claim that, up to an error in the operator norm of order $\exp[-\mathcal{O}(L)]$, $\partial_\theta\mathcal{H}(t)$ can be written as an operator involving only terms not in half (2). That is, $||\partial_\theta\mathcal{H}(t)-\mathrm{Tr}_2[\partial_\theta\mathcal{H}(t)]||$ is exponentially small in L. To show this, define U_{12} to be the set of all sites j which lie in both half (1) and half (2); there are at most $2SV/L$ such sites. These sites are shown in the solid regions in Fig. 3. Define operators $O(t=0)=O'(t=0)=\partial_\theta\mathcal{H}$, and define the time evolution of O,O' by $\partial_t O=-i[O,\mathcal{H}]$, while $\partial_t O'=-i\Sigma_{i\notin U_{12}}[O,\mathcal{H}_i]$, i.e., the time evolution of O' includes only the sum over sites i which are either in half (1) or in half (2), but *not* in both halves. Then, using the arguments leading up to Eq. (4), we can show that for $i\in U_{12}$, the operator norm $||[O(t),\mathcal{H}_i]||$ is bounded by $2J||O||g(c_1(L/2-R),L/2-R)$, which is of order $\exp[-\mathcal{O}(L/2-R)]$ for the given range of times t. Then, using the difference in the evolution equations for O,O', we can bound $||O(t)-O'(t)||\leq\Sigma_{i\in U_{12}}\int_0^t dt'\,||[O,\mathcal{H}_i]||\leq 2t(2SV/L)J||O||\times g(c_1(L/2-R),L/2-R)$. This quantity is also of order

$\exp[-\mathcal{O}(L/2-R)]$ for the given range of times t. Finally, we use the fact that $\mathrm{Tr}_2[O']=O'$ to get the desired result.

From the above two paragraphs, it follows that up to small error in the trace norm, $e_a(\theta)=-\int_0^{c_1L}d\tau[\mathrm{Tr}_2[A^+(i\tau)-A^-(-i\tau)],\mathrm{Tr}_2[\rho(\theta,0)-\rho^0(\theta,-\theta)]]$. Then, this is equal to the commutator of $E_1(\theta)$ with an anti-Hermitian operator. It generates an infinitesimal unitary rotation of $E_1(\theta)$ and therefore does not lead to any change in $|E_1(\theta)|$.

Next, consider Eq. (29). First consider the terms in the commutator involving $A^{+,-}$ acting on the left side of ρ^o. As above, the operator $\partial_\theta\mathcal{H}$ can be written in a basis of eigenstates of $\mathcal{H}$ as $(\partial_\theta\mathcal{H})_{ab}$, where the states have energies E_a,E_b. In $\partial_\theta\mathcal{H}\rho^0(\theta,-\theta)$ the only nonvanishing terms involve states with energy difference $-\omega=E_a-E_b\geqslant\Delta E$. Consider a matrix element $(\partial_\theta\mathcal{H})_{ab}$ with given ω. This leads to a matrix element of $A^-(-i\tau)$ equal to $(\partial_\theta\mathcal{H})_{ab}$ times

$$\frac{\sqrt{2\pi q\ln(L)}}{\Delta E}\exp[-(\tau\Delta E)^2/(2q\ln(L))]\times\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\Theta(\omega')\exp[-\omega'\tau]\times\exp[-q\ln(L)(\omega-\omega')^2/(2\Delta E^2)],\qquad(31)$$

where we have converted the time integral to an integral in Fourier space. Since $-\omega\geqslant\Delta E$, Eq. (31) can be made small by choosing q sufficiently large. Thus, the trace norm of $A^-(-i\tau)\rho^0(\theta,-\theta)$ is small, for all $\tau\geqslant0$. Similarly, for A^+, we find that we get a matrix element equal to $(\partial_\theta\mathcal{H})_{ab}$ times

$$\frac{\sqrt{2\pi q\ln(L)}}{\Delta E}\exp[-(\tau\Delta E)^2/(2q\ln(L))]\times\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\Theta(-\omega')\exp[\omega'\tau]\times\exp[-q\ln(L)(\omega-\omega')^2/(2\Delta E^2)].\qquad(32)$$

By choosing q sufficiently large, the integral (32) can be made equal to $\exp[\omega\tau]$, up to small error. Thus, the given matrix element can be made equal to $(\partial_\theta\mathcal{H})_{ab}$ times $\exp[\omega\tau]$, up to small error. Therefore, the trace norm of $-\int_0^{c_1L}d\tau\,\mathrm{Tr}_2\{[A^+(i\tau)-\partial_\theta\mathcal{H}(i\tau)]\rho^0(\theta,-\theta)\}$ is small. These statements amount to saying that, with small error in the operator norm, $A^+(i\tau)$ indeed is equal to the positive energy part of $\partial_\theta\mathcal{H}(i\tau)$, while A^- is equal to the negative energy part.

Now consider $A^{+,-}$ acting to the right side of $\rho^0(\theta,-\theta)$, so that we consider $\rho^0(\theta,-\theta)\partial_\theta\mathcal{H}$. In that case, the only nonvanishing terms in $(\partial_\theta\mathcal{H})_{ab}$ involve $-\omega=E_a-E_b\leqslant-\Delta E$. Repeating the argument above, we find that the trace norm of $\rho^0(\theta,-\theta)A^+(i\tau)$ is small, as is the trace norm of $\int_0^{c_1L}d\tau\,\mathrm{Tr}_2\{\rho^0(\theta,-\theta)[A^-(-i\tau)-\partial_\theta\mathcal{H}(-i\tau)]\}$.

Therefore, up to small error, Eq. (29) is equal to $-\int_0^{c_1L}d\tau\,\mathrm{Tr}_2[\partial_\theta\mathcal{H}(i\tau)\rho^0(\theta,-\theta)+\rho^0(\theta,-\theta)\partial_\theta\mathcal{H}(-i\tau)]-\partial_\theta\rho_1^0(\theta,\theta')$, which equals $-\int_0^{c_1L}d\tau\,\mathrm{Tr}_2[\partial_\theta\mathcal{H}(i\tau)\rho^0(\theta,-\theta)+\rho^0(\theta,-\theta)\partial_\theta\mathcal{H}(-i\tau)]+\int_0^{\infty}d\tau\,\mathrm{Tr}_2[\partial_\theta\mathcal{H}(i\tau)\rho^0(\theta,-\theta)+\rho^0(\theta,-\theta)\partial_\theta\mathcal{H}(-i\tau)]$. This difference is equal to an integral over $\tau\geqslant c_1L$. For sufficiently big c, the trace norm of this integral can be bounded by any desired negative power of L. Thus, e_b has small trace norm.

Finally, consider Eq. (30). This is equal to

$$\partial_\theta\rho_1^0(\theta,\theta')-\partial_\theta\rho_1^0(\theta,-\theta)=\mathrm{Tr}_2[\partial_{\theta'}\rho^0(\theta,\theta')],\qquad(33)$$

where the derivatives are evaluated at $\theta'=-\theta$. The trace norm of the right-hand side of Eq. (33) can be bounded by a negative power of L using the same arguments near Eq. (27), by considering an operator O that is entirely within half (1). The only difference to the arguments near Eq. (27) is that we compute the derivatives and expectation values at $\theta=-\theta'$, rather than at $\theta=\theta'=0$.

Thus, using Eqs. (28)–(30) and picking sufficiently large q we find that $\mathcal{D}_\theta|E_1(\theta)|$ is small for sufficiently large c. A similar sequence of arguments permits one to bound $\mathcal{D}_\theta|E_2(\theta)|$. In the following section, we consider the expectation value of the translation operator on $\rho(2\pi,0)$.

VI. TRANSLATION OPERATOR

Consider the operator $\mathrm{tr}((1,y)(2,y)\ldots(L,y))$, which translates the sites with given y. The translation operator T which translates the entire system by one unit cell is the product of these loop operators over all y (there are an odd number of such loop operators). The ground state of $\mathcal{H}_{0,0}$ is an eigenstate of T. If the ground state is nondegenerate, then it has eigenvalue ±1; without loss of generality we will assume in this section that is has eigenvalue $+1$.

In this section we will show that the expectation value of T for $\rho(2\pi,0)$ is opposite to that for $\rho(0,0)$, up to small error. We note that if $\mathcal{H}_{\theta,0}$ were to have a gap for all $0\leqslant\theta\leqslant2\pi$, then the results in this section would provide the last step in the topological argument discussed above. Instead, the results in this section will complete the argument started in the preceding section: $\rho(2\pi,0)$ gives us a density matrix such that $\mathrm{Tr}[\rho(2\pi,0)\mathcal{H}_{\theta,0}]-E_0(0,0)$ is small, but which, up to small error, has the opposite expectation value for T. Since the difference in the expectation of $\mathcal{H}$ is small, we can find a c such that the difference in expectation value decays faster than $1/L$, and then we can find an L_0 such that for $L>L_0$ the state $\psi(2\pi,0)$ has an energy expectation value which is *less than* $c\ln(L)/L$ above the ground state. However, since the expectation value of T is opposite for $\rho(2\pi,0)$ compared to $\rho(0,0)$, up to small error, this state has an overlap on the ground state which is small. Thus, we will show in this section a contradiction under the assumption that the system had a gap ΔE which violated Eq. (1) and under the assumption that the system was translation symmetric, so that the ground state was an eigenstate of T.

We first define a twisted translation operator,

$$T_{\theta,\theta'}\equiv\Pi_y\mathrm{tr}(\exp[i(\theta/2)\sigma_z](1,y)\times(2,y)\ldots(L/2,y)\exp[i(\theta'/2)\sigma_z]\times(L/2+1,y)\ldots(L,y)).$$

Then, $T_{\theta,\theta'}$ is a unitary operator and a symmetry of $\mathcal{H}_{\theta,\theta'}$. Finally, given that $T\psi_0(0,0)=\psi_0(0,0)$, we have $T_{\theta,-\theta}\psi_0(\theta,-\theta)=\psi_0(\theta,-\theta)$ for all θ.

We will then show that $\partial_\theta|T_{\theta,0}\psi(\theta,0)-\psi(\theta,0)|$ is small for all $0\leqslant\theta\leqslant 2\pi$. It will then follow that, up to a small error, $\mathrm{Tr}[\rho(0,0)T_{0,0}]=\mathrm{Tr}[\rho(2\pi,0)T_{2\pi,0}]=-\mathrm{Tr}[\rho(2\pi,0)T_{0,0}]=1$, thus showing that the twisted state indeed has the opposite expectation value for T. Here we have used the fact that for systems of *odd* width, $T_{2\pi,0}=-T_{0,0}$.

Consider the derivative

$$\begin{aligned}\partial_\theta(T_{\theta,0}\psi(\theta,0)-\psi(\theta,0))&\\ =i\sum_y S^z_{(1,y)}T_{\theta,0}\psi(\theta,0)-\int_0^{c_1L}d\tau[T_{\theta,0}A^+(i\tau)&\\ -A^+(i\tau)-T_{\theta,0}A^-\quad(-i\tau)+A^-(-i\tau)]\psi(\theta,0).&\end{aligned}\tag{34}$$

This is equal to

$$\begin{aligned}&\left\{i\sum_y S^z_{(1,y)}-\int_0^{c_1L}d\tau[T_{\theta,0}A^+(i\tau)T^{-1}_{\theta,0}-A^+(i\tau)-T_{\theta,0}A^-\right.\\ &\left.\times(-i\tau)T^{-1}_{\theta,0}+A^-(-i\tau)]\right\}\psi(\theta,0)+\left\{i\sum_y S^z_{(1,y)}\right.\\ &\left.-\int_0^{c_1L}d\tau[T_{\theta,0}A^+(i\tau)T^{-1}_{\theta,0}-T_{\theta,0}A^-(-i\tau)T^{-1}_{\theta,0}]\right\}\\ &\times\{T_{\theta,0}\psi(\theta,0)-\psi(\theta,0)\}.\end{aligned}\tag{35}$$

The last term of Eq. (35) is equal to an anti-Hermitian operator acting on $T_{\theta,0}\psi(\theta,0)-\psi(\theta,0)$, and thus does not change the norm of this state. Thus, we need to bound the norm of the first term. This term is equal to an anti-Hermitian operator, which we define to be O_1, acting on ψ. The norm square of this term is equal to $\mathrm{Tr}[O_1^\dagger O_1\rho(\theta,0)]$. As shown in the preceding section, up to small error in the operator norm, $A^{+,-}(i\tau)$ can be written entirely as operators in half (1). Therefore, O_1 can be written entirely as an operator in half (1); that is, the operator norm $||O_1-\mathrm{Tr}_2[O_1]||$ is small. Thus, the norm square is, up to small error, $\mathrm{Tr}[O_1^\dagger O_1\rho_1(\theta,0]$, which, again up to small error, is equal to $\mathrm{Tr}[O_1^\dagger O_1\rho_1^0(\theta,-\theta)]$, since $|\rho_1(\theta,0)-\rho_1^0(\theta,-\theta)|$ is small.

We claim, however, that this last expectation value is small. To show this, consider $\partial_\theta(T_{\theta,-\theta}\psi_0(\theta,-\theta)-\psi_0(\theta,-\theta))$. This is equal to zero. However, this derivative can be written as an operator O acting on ψ_0, with $O=i\Sigma_y(S^z_{(1,y)}-S^z_{(L/2+1,y)})-\int_0^\infty d\tau(T_{\theta,-\theta}\partial_\theta\mathcal{H}_{\theta,\theta'}(i\tau)T^{-1}_{\theta,-\theta}-\partial_\theta\mathcal{H}_{\theta,\theta'}(i\tau)-T_{\theta,-\theta}\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(i\tau)T^{-1}_{\theta,-\theta}+\partial_{\theta'}\mathcal{H}_{\theta,\theta'}(i\tau))$, where the derivatives are taken at $\theta=-\theta'$. Since $O\psi_0=0$, it follows that $\mathrm{Tr}[O^\dagger O\rho^0(\theta,-\theta)]=0$. However, up to small error, $O=O_1+O_2$, with O_1 the operator considered above and O_2 defined to be a similar operator acting only in half (2). Then, $2\mathrm{Tr}[O_1^\dagger O_1\rho^0(\theta,-\theta)]+2\mathrm{Tr}[O_1^\dagger O_2\rho^0(\theta,-\theta)]=0$. However, using the locality bounds, the second term can be made small for large enough c,[23] and thus the first term, $\mathrm{Tr}[O_1^\dagger O_1\rho_1^0(\theta,-\theta)]$, is small. Therefore, we have shown the desired result.

VII. DISCUSSION

The main result is Eq. (1), obtaining a bound on the energy gap for spin models in arbitrary dimensions. In order to obtain this result, we have introduced a set of loop operators, and proven a bound on connected correlation functions. This bound on correlation functions did not rely on the system being a spin system; rather, it is valid for any Hamiltonian such that the $\mathcal{H}_i$ have bounded operator norm, and such that the interaction is finite range. Below, we generalize this bound on correlation functions to certain other systems as well.

We note that for the case of higher spin representations of SU(2), Eq. (1) follows automatically from the result for spin-1/2, so long as the total spin within all unit cells is half-odd: the higher spins can be written as various combinations of spin-1/2 spins, and if the total spin in the unit cell is half-odd then there will result an odd number of spin-1/2 spins in each unit cell. Suppose, for example, a unit cell contains one spin-1 spin and one spin-1/2 spin, giving a total spin of 3/2 which is half-odd. Then, the spin-1 can be written as two spin-1/2 spins. Let these two spins be called S_1,S_2 and let the Hamiltonian include only terms symmetric under interchange of S_1,S_2. This new Hamiltonian has three spin-1/2 spins in each unit cell, and hence falls within the class of Hamiltonians considered above. Then, there are two different sectors of the Hilbert space with no terms in the Hamiltonian coupling these two sectors: one sector in which S_1,S_2 form a spin-0, and one in which they form a spin-1. By adding a term coupling S_1 to S_2 to the Hamiltonian with a sufficiently large, negative (ferromagnetic) coefficient, we can ensure that the ground and first excited states lie in the sector in which S_1,S_2 have total spin 1. Then, the existence of a low-lying state satisfying Eq. (1) for the new system with only spin-1/2 implies the existence of such a low-lying state for the original system with both spin-1 and spin-1/2. It would also be interesting to generalize these results to other groups SU(N), as well as to consider the case of even V/L.

We finish with two conjectures. First, we conjecture that the same Eq. (1) holds for systems with an even width, so long as the width V/L is of order L^{d-1} and so long as $d>1$. For $d=1$, this result is of course not true, as Haldane gap behavior is possible.

Second, consider the *thermal* expectation value of T at an inverse temperature β, defined by $\langle T\rangle_\beta\equiv\mathrm{Tr}[\exp(-\beta\mathcal{H})T]/\mathrm{Tr}[\exp(-\beta\mathcal{H})]$. We conjecture that there is a constant c, depending on $\mathcal{H},d$ such that for $\beta>c\ln(L)/L$ the given thermal expectation value $\langle T\rangle_\beta$ vanishes in the limit $L\to\infty$ for systems of odd width. We base this conjecture on the following physical observations: for ferromagnetic systems, there are spin-wave excitations, with dispersion relation $E\propto k^2$. It may be shown that the presence of these excitations causes $\langle T\rangle_\beta$ to vanish for β of order L as $L\to\infty$. For antiferromagnetic systems, the translation symmetry is broken by the antiferromagnetic ordering (in fact,

for these systems, the true ground state has translation symmetry and is a superposition of different broken symmetry states, but there are low-lying states with different expectation values of T so that for β of order L, the expectation value $\langle T\rangle_\beta$ vanishes). Finally, for spin liquid systems, there is a low-lying excited state with the opposite expectation value of T compared to the ground state, as we have found above. We leave a proof of both of these conjectures for future work.

ACKNOWLEDGMENTS

This work was supported by DOE Contract No. W-7405-ENG-36.

APPENDIX: MARKOV PROCESSES AND LOCALITY

Consider a system with a probability p_i of being in state i and a transition matrix T_{ij} so that the equation of motion is $\partial_t p_i = \Sigma_j T_{ij} p_j$. For the total probability to be conserved, we have $\Sigma_i T_{ij} = 0$, which guarantees that T_{ij} has at least one zero eigenvalue. Let us assume, further, that all eigenvalues of T_{ij} are real. This includes all systems for which the stationary state (given by the zero eigenvector of T_{ij}) obeys detailed balance. A typical example of such a process would be the Monte Carlo dynamics of a statistical mechanics system. We will first derive a suitable generalization of the locality result (13) to systems governed by such a Markov process, and then discuss the implication for statistical mechanics systems.

Let us assume that the spectrum of T is such that there is only one zero eigenvalue, with right eigenvector p_i^0, and that all other eigenvalues λ are negative with $\lambda \leqslant -\Delta$, for some $\Delta > 0$. Assume that p_i^0 is normalized by $\Sigma_i p_i^0 = 1$.

Then, introduce various quantities to be measured, $A, B, \ldots$, so that the expectation value of A is given by $\langle A\rangle = \Sigma_i A_i p_i^0$. We can write this slightly differently by introducing for each such quantity a diagonal matrix given by $\hat{A}_{ii} = A_i$ for all i and $\hat{A}_{ij} = 0$ for $i \neq j$. Further, introduce an additional vector I_i, such that $I_i = 1$ for all i. This vector I_i is a left eigenvector of T with zero eigenvalue, since $\Sigma_i T_{ij} = 0$ as mentioned above. Then, $\langle A\rangle = I_i A_{ij} p_j^0$.

We can now consider expectation values of quantities at different times: $\langle A(t)B(0)\rangle \equiv I^\dagger \hat{A}\exp[Tt]\hat{B}p^0 = I^\dagger \exp[-Tt]\hat{A}\exp[Tt]\hat{B}p^0 = I^\dagger \hat{A}(t)\hat{B}(0)p^0$. In these equations, $I^\dagger$ denotes the transpose of the vector I (I is real, so no complex conjugation is necessary), and we have left off all the indices on vectors I, p and matrices $A, B, \exp[\pm Tt]$: the product is evaluated following the usual rules of matrix multiplication. In the sequence of equalities above, the first equality defines the time evolution of the system, the second equality follows since $I\exp[-Tt] = I$, and the last equality follows since we define $\hat{A}(t)$ by the equation of motion: $\partial_t \hat{A}(t) = [A(t), T]$. It is then possible to extend this definition to operators separated by an *imaginary* time separation: $\langle A(it)B(0)\rangle$.

Now, consider a typical physical example: an Ising system, governed by Monte Carlo spin-flip dynamics, with $A(0), B(0)$ representing the value of two different spins which are separated in space. In such a case (as well as in many others), it is possible to obtain a bound similar to Eq. (4). Assume that the matrix T can be written as a sum of matrices T_i, with finite interaction range R and with a bound $||T_i|| \leqslant J$, where i is a site index. Define $f(t) \equiv \langle [A(it), B(0)]\rangle$. Since $f(t) \leqslant ||[A(it), B(0)]||$,

$$f(t) \leqslant \frac{2N_B S||A||||B||2JSt|^{l/2R} e^{2JS|t|}}{(l/2R)!} \equiv N_B ||A||||B|| g(t,l). \tag{A1}$$

At this point, from the existence of a $\Delta > 0$ and a spectral representation with all eigenvalues real[24] follows a result similar to Eq. (13),

$$\begin{aligned} &|\langle A(0)B(0)\rangle_c| \\ &\quad \leqslant \frac{1}{2\pi} 2N_B ||A||||B|| g(c_1 l, l) + 2\left(1 + \frac{1}{\sqrt{2\pi c_1 l \Delta E}}\right) \\ &\quad \times ||A||||B|| e^{-c_1 \Delta E l 2}. \end{aligned} \tag{A2}$$

Therefore, *if* there is a Markov dynamics that gives rise to the equilibrium probability distribution p_i^0 which has a $\Delta > 0$, *then* there is an exponential decay of correlation functions in space. An example is a spin system in the paramagnetic phase with Monte Carlo spin flip dynamics. The converse is not necessarily true: a spin system in the paramagnetic phase with spin exchange dynamics does not have a $\Delta > 0$ but instead has spin correlations which decay with a power law time. However, this dynamics gives rise to the same equilibrium probability distribution as the spin flip dynamics does, and hence has exponentially decaying correlations in space.

*Email address: hastings@cnls.lanl.gov

[1] E.H. Lieb, T.D. Schultz, and D.C. Mattis, Ann. Phys. (N.Y.) **16**, 407 (1961).
[2] F.D.M. Haldane, Phys. Rev. Lett. **50**, 1153 (1983).
[3] I. Affleck, Phys. Rev. B **37**, 5186 (1988).
[4] M. Oshikawa, Phys. Rev. Lett. **84**, 1535 (2000).
[5] G. Misguich, C. Lhuillier, M. Mambrini, and P. Sindzingre, Eur. Phys. J. B **26**, 167 (2002); G. Misguich and C. Lhuillier, cond-mat/0002170 (unpublished).
[6] B. Sutherland, Phys. Rev. B **37**, 3786 (1988); D. Rokhsar and S. Kivelson, Phys. Rev. Lett. **61**, 2376 (1988); N. Read and B. Chakraborty, Phys. Rev. B **40**, 7133 (1989).
[7] N.E. Bonesteel, Phys. Rev. B **40**, 8954 (1989).
[8] It follows automatically that Eq. (1) holds for all L for some c depending on $\mathcal{H}, d, r$.
[9] C.K. Majumdar and D.K. Ghosh, J. Math. Phys. **10**, 1388 (1969).
[10] G. Baskaran and P.W. Anderson, Phys. Rev. B **37**, 580 (1988).
[11] I. Affleck and J.B. Marston, Phys. Rev. B **37**, 3774 (1988); X.G. Wen, *ibid.* **44**, 2664 (1991); N. Read and S. Sachdev, Phys. Rev. Lett. **66**, 1773 (1991).

[12] D.S. Rokhsar and S.A. Kivelson, Phys. Rev. Lett. **61**, 2376 (1988); R. Moessner and S.L. Sondhi, *ibid.* **86**, 1881 (2001).

[13] This norm, written $||O||$, is defined to be the supremum over states ψ, with $|\psi|=1$, of $|O\psi|$, R. Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).

[14] E. Lieb and D. Robinson, Commun. Math. Phys. **28**, 251 (1972).

[15] Following standard physics notation, whenever we write $1/(-it+\epsilon)$ we are taking the limit as $\epsilon\to 0$. The limit is taken outside the integral sign. On the other hand, the use later of $1/(-it+\tau)$ is done at a fixed, nonzero value of τ.

[16] Equation (7) follows from contour integration. Another way to derive it is to use a convolution and Fourier transform. Define the Fourier transform of $f(t)$ to be $f(\omega)=\int_{-\infty}^{\infty}dt f(t)\exp[-i\omega t]$ (see Ref. 17). Then, $f^{+}(t)=(1/2\pi)\int_{-\infty}^{0}d\omega f(\omega)\exp[i\omega t]$. Combining these gives $f^{+}(0)=1/2\pi\int_{-\infty}^{0}d\omega\int_{-\infty}^{\infty}dt f(t)\exp[-i\omega t]$. Doing the integral over t gives Eq. (7). Note that $\tilde{f}$ is rapidly decaying as $t\to\pm\infty$, so that there are no issues with convergence.

[17] Following standard physics notation, in Eq. (9), and throughout, the use of $f(\omega)$ refers to the Fourier transform of $f(t)$ The use of $f(0)$ always refers to $f(t)$ at $t=0$.

[18] We note that since we have assumed that the ground state is unique, with a gap ΔE, then for any finite L, the ground-state energy is analytic in θ,θ' in some neighborhood of $\theta+\theta'=0$.

[19] To show this, consider rotating all the spins by angle π about the x axis. Since the ground state is assumed to be unique and the Hamiltonian is invariant under global SU (2) rotations, the ground state is invariant under global SU (2), so that ground-state expectation values must be unchanged under this rotation. However, if we take an odd number of derivatives of $\mathcal{H}_{\theta,\theta'}$ with respect to θ,θ', the resulting operator changes sign under a rotation by π about the x axis. Therefore, the expectation value of an odd number of derivatives of $\mathcal{H}_{\theta,\theta'}$ must vanish.

[20] Possible low-energy topological excitations do not prevent the use of Eq. (13) for spin liquids as the matrix elements of $\partial_\theta\mathcal{H}_{\theta,\theta'}$ between the ground and topologically excited states are negligible.

[21] The trace norm of an operator O, written $|O|$, is defined to be $\mathrm{Tr}[\sqrt{O^{\dagger}O}]$. We will deal with Hermitian operators, for which it is equal to the sum of the absolute values of the eigenvalues.

[22] In fact, $|E_1(\theta)|$ need not be differentiable. To get around this, we use the upper derivative (Ref. 14), defined as $\mathcal{D}_t f(t)\equiv\limsup_{\epsilon\to 0}(f(t+\epsilon)-f(t))/\epsilon$.

[23] We showed before that up to small error O_1 is entirely within half (1). For larger c, we can in fact ensure that O_1 includes only operators within a distance $L/4$ of $x=0$ so that O_1,O_2 are separated by a distance $L/2$ from each other. Then we can apply the locality bounds to the given correlation function of O_1,O_2.

[24] It turns out that the requirement that the eigenvalues be real is necessary. Although one might have guessed that a similar exponential decay of correlation functions would also hold for complex λ with real part $\mathrm{Re}(\lambda)\leqslant-\Delta$, there seems to be a counterexample to this statement. This will be discussed in a future publication.