Anomalies in the entanglement properties of the square-lattice Heisenberg model

Ann B. Kallin,[1] Matthew B. Hastings,[2,3] Roger G. Melko,[1] and Rajiv R. P. Singh[4]
[1]*Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1 Canada*
[2]*Duke University, Department of Physics, Durham, North Carolina 27708, USA*
[3]*Microsoft Research, Station Q, CNSI Building, University of California, Santa Barbara, California 93106, USA*
[4]*Physics Department, University of California, Davis, California 95616, USA*
(Received 14 July 2011; revised manuscript received 20 September 2011; published 28 October 2011)

We compute the bipartite entanglement properties of the spin-half square-lattice Heisenberg model by a variety of numerical techniques that include valence-bond quantum Monte Carlo (QMC), stochastic series expansion QMC, high-temperature series expansions, and zero-temperature coupling constant expansions around the Ising limit. We find that the area law is always satisfied, but in addition to the entanglement entropy per unit boundary length, there are other terms that depend logarithmically on the subregion size, arising from broken symmetry in the bulk and from the existence of corners at the boundary. We find that the numerical results are anomalous in several ways. First, the bulk term arising from broken symmetry deviates from an exact calculation that can be done for a mean-field Néel state. Second, the corner logs do not agree with the known results for noninteracting Boson modes. And, third, even the finite-temperature mutual information shows an anomalous behavior as T goes to zero, suggesting that the $T \to 0$ and $L \to \infty$ limits do not commute. These calculations show that entanglement entropy demonstrates a very rich behavior in $d > 1$, which deserves further attention.

DOI: 10.1103/PhysRevB.84.165134 PACS number(s): 75.10.Jm, 02.70.Ss, 03.67.Mn

I. INTRODUCTION

The study of bipartite entanglement properties in quantum statistical models is a promising way to understand and classify their topological and universal properties.[1–3] In one spatial dimension, the existence of finite entanglement entropy for gapped systems and a logarithmically divergent entanglement entropy for gapless systems is well understood. The coefficient of the log divergence is universal in that it depends on the central charge and not on the microscopic details of the system. In two dimensions (2D), several gapless systems have been shown to display an area law, where the entropy per unit boundary length is necessarily nonuniversal, reflecting all the microscopic degrees of freedom at the boundary. Thus, looking for universal behavior necessitates looking for subleading terms. The area-law coefficient will have a subleading power-law dependence on the size of the system. However, more interestingly, there can be terms associated with corners on the boundary, with broken symmetry, or with nontrivial topology in the bulk that could be universal and allow one to classify different topological and critical phases.

The spin-half square-lattice Heisenberg model with nearest-neighbor interactions is one of the most studied and best understood models in 2D.[4] The model has long-range order, and spin rotational symmetry is spontaneously broken in the ground state. This broken symmetry is well described by a tower of rotor states,[5] whose energy scales with system size L as $1/L^d$ in d dimensions. For large systems, these states are well separated from excitations around the ground state, which in the long-wavelength limit are spin waves whose energies scale as $1/L$. These spin waves are believed to become noninteracting in the long-wavelength limit as long as one is away from a quantum critical point where the long-range order might go continuously to zero.

In this paper, we have developed a number of different computational methods to calculate entanglement properties, which are valid for arbitrary dimensional quantum statistical models. Stochastic series expansion (SSE), quantum Monte Carlo (QMC),[6,7] and high temperature expansions (HTE) are methods that allow one to calculate thermodynamic properties of the model at finite temperatures and can be used to obtain the Renyi mutual information associated with dividing the system into two regions A and B.[8] This mutual information should reduce to twice the entanglement entropy as $T \to 0$. The valence bond (VB) QMC method is a particularly powerful tool for studying properties of the model directly at $T = 0$.[9–11] We have developed extensions of this method that employ loop updates[12] necessary to accurately calculate the entanglement properties for finite lattices with different bipartite divisions. Finally, Ising expansions at $T = 0$ provide yet another method to calculate properties of the Heisenberg model in its ground state. By approaching the ground state of the Heisenberg model from the ordered side one can calculate a specific ordered state and its entanglement properties.

Our first result is an estimation of the leading nonuniversal area-law coefficient, calculated as $a = 0.097 \pm 0.001$ in VB QMC (Sec. II) and $a = 0.094 \pm 0.001$ in Ising series expansions (Sec. III), which are in good agreement. The small discrepancy shows that there are systematic errors not included in the numerical estimates of error bars. However, these are of order 1%.

The universal contributions to the subleading scaling come from a number of sources. For example, one might expect the universal entanglement properties of the model to be related to the broken symmetry of the Néel state and to the presence of free bosons resulting from noninteracting spin waves. The entanglement properties of free bosons (and free fermions) have been computed in 2D.[13] Furthermore, the entanglement properties of a mean-field Néel state, where all spins on sublattice 1 and sublattice 2 separately form a maximal spin state that is then combined into a singlet, can be calculated exactly. This gives a mean-field bulk entanglement entropy associated with broken symmetry that scales as $c \ln(\ell)$, with $c = 2$, where ℓ is the length of the boundary (Sec. IV). In

contrast, our VB QMC simulations for a bipartite division with no corners show a logarithmic term with $c = 0.74 \pm 0.02$. Recently, the entanglement properties of this model were calculated using modified spin-wave theory,[14] which gave an estimate of $c = 0.92$. Thus, our numerical simulations are much closer to spin-wave theory. This suggests that, in addition to broken symmetry, there are logarithmic contributions in the bulk that come from other sources.

The corner contributions can be obtained in QMC by comparing the log terms in a system using a boundary with corners to a system without boundary corners, giving an estimate of -0.10 ± 0.02. The series estimate, -0.080 ± 0.008, gives reasonable agreement with the QMC results. In contrast, if one takes two free boson modes contributing to the corners, then one gets ~ -0.0496 (Sec. V).

Finally, we have studied the properties of the model at finite temperatures and examined the approach to $T \to 0$. This can be done by SSE QMC and HTE. The HTE extrapolations for the Renyi mutual information agree well with QMC down to $T \approx 1$. Below this temperature, the QMC data shows a sudden decrease and a crossover to a lower saturation value as $T \to 0$. The latter is consistent with the entanglement entropy calculated at $T = 0$. The HTE shows no sharp decrease. The sharp decrease has a size dependence and could imply the limits of $T \to 0$ and $L \to \infty$ do not commute. Such noncommutation is well known for other response functions of the Heisenberg model[15,16] and also suggests a sizable non-mean-field contribution to the area-law term.

II. VALENCE BOND QUANTUM MONTE CARLO ALGORITHMS AND RESULTS

In order to calculate the zero-temperature scaling of the Renyi entanglement entropy in the Heisenberg model on finite-size lattices, we employ the VB QMC method developed by Sandvik.[9] This is a highly efficient method to project out the model's ground state by repeated application of the Hamiltonian to a trial wave function through a Monte Carlo sampling of bond operators. As an improvement on our previous entanglement measurement procedure,[17] we use a modified version of the more efficient loop algorithm.[12] This modification allows for a change to occur in the Monte Carlo weight by modifying the space-time topology of the simulation cell. As in Ref. 17, this modified weight is required so one can measure the *difference* between entanglement entropies of two distinct system subdivisions instead of the absolute entanglement entropy. Then, if the geometries of the regions are chosen properly, the difference can converge faster than the bare entanglement entropy measurement.

A. Valence-bond quantum Monte Carlo

First, we briefly discuss the basic VB QMC algorithm (for more detail see Refs. 9–11) and the more recently developed loop update (see Ref. 12) which significantly improves the scaling of the algorithm. The foundation of the VB QMC technique is to project out the ground state of the system, done by applying a high power of the Hamiltonian $\mathcal{H}^M$ to a trial state. We use the Heisenberg Hamiltonian, rewritten in terms of bond operators ($H_{ab} = \frac{1}{4} - \mathbf{S}_a \cdot \mathbf{S}_b$) acting on pairs of sites (a and b), which are nearest-neighbor pairs in this paper. The Hamiltonian to the power M can be written as a sum of possible arrangements of these bond operators in a list of size M. The Monte Carlo algorithm importance samples terms in this sum, using a weight that depends on the number of off-diagonal operators in the term.[9]

In its original form, the VB QMC scheme can be used to project out one copy of the ground-state wave function (single-projector) or simultaneously project two copies (double projector) that can be used to measure expectation values of observables in the simulation.[9–11] Our previous scheme for measuring Renyi entanglement entropy[17] employed a double-projector method for calculating the expectation value of a SWAP operator. In this paper, we develop a highly efficient loop variation of this measurement algorithm, outlined below.

1. The loop algorithm

The loop update for VB QMC simulations was introduced in Ref. 12 as a highly efficient way of carrying out the sampling procedure. In addition to working in a basis of valence bonds, this scheme also samples over spin states. This combined spin-bond basis is shown to eliminate the need for a rejection step and, thus, samples operators and basis states with high efficiency.

To begin, operators in this case are divided into two classes,

$$H_{ab}(1) = \left(\tfrac{1}{4} - S_a^z S_b^z\right), \tag{1}$$

$$H_{ab}(2) = -\tfrac{1}{2}(S_a^+ S_b^- + S_a^- S_b^+), \tag{2}$$

called diagonal and off-diagonal operators, respectively, where the sum $H_{ab}(1) + H_{ab}(2)$ is equal to the bond operators H_{ab} from the standard VB QMC algorithms mentioned above.[9–11]

The loop algorithm is best visualized using a diagram of the simulation cell showing the placement of valence bonds, spins, and operators, as depicted in Fig. 1. This diagram represents two VB trial states, which are the left and right edges of the figure, each projected "inward" (by $M = 3$ operators). The projected state occurs in the center of the diagram, denoted by the dashed line. Along with the trial VB states, initial spin states are selected at random, with the condition that the two spins in a single VB must be antiparallel. For each trial state, M operators are chosen such that they each act on a pair of antiparallel spins (in the initial step of the algorithm, these are all diagonal operators). There are then two types of updates in this algorithm: spin updates and operator updates.

For the spin update, loops are first constructed by linking the operators and valence bonds (shown in Fig. 1). Then, for each loop that is built, a decision is made to flip all the spins in that loop, with probability 1/2. This update samples possible spin states for the given valence bond configuration. In the second type of update, the operators in the list are changed so diagonal operators are resampled at random, subject to the condition that they remain acting on antiparallel sites. This reconfigures the propagated valence bond states and the topology of the simulation cell loops for future updates. Measurements can be computed as usual[9] using the propagated valence bond states, $|V_L\rangle$ and $|V_R\rangle$, which can be extracted from the simulation cell diagram by following the loops crossing the dotted line in Fig. 1.

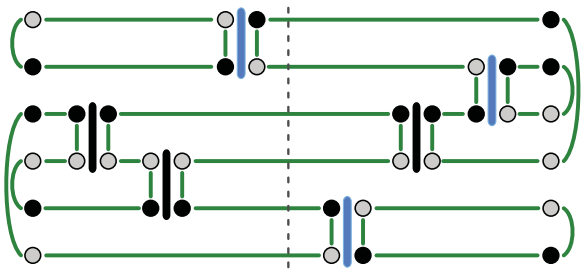

FIG. 1. (Color online) A possible simulation cell diagram for a six-site system, including loops, operators, the initial valence bond states, and the compatible initial spin states. Up (down) spins are shown in gray (black). (Off-) diagonal operators are shown in (blue) black, giving a total of $M = 3$ operators in this example. The dashed line denotes $|V_L\rangle$ (propagated from the trial state on left) and $|V_R\rangle$ (similarly from the right).

In the next section we discuss the measurement of Renyi entanglement entropies with either the double-projector or loop algorithms. In the following section we describe this measurement with the loop algorithm using a modified VB QMC simulation cell.

B. Measuring Renyi entropies with VB QMC

Numerical techniques such as exact diagonalization or density matrix renormalization group (DMRG) simulations are able to directly measure entanglement entropy, since the calculation provides access to the density matrix. In contrast, it is typically a challenge to measure entanglement entropy with QMC, as the density matrix is not sampled in a straightforward way. Over the past few years there have been several proposals of entanglement measures in Monte Carlo simulations.[18–20] Recently, methods to measure Renyi entanglement entropy using a SWAP (or more generally a permutation)[21] operator have been developed and implemented in several types of Monte Carlo methods.[8,17] Next, we outline the basic methodology relevant for VB basis QMC,[17] based on the expectation value of the SWAP operator before detailing current advances for improving the measurement efficiency using a hybrid loop-ratio trick estimator.

1. *Renyi entropies and the* SWAP *operator*

We are interested in the generalized Renyi entropies, which quantify entanglement between a system subdivided into two regions, A and B. They are defined as

$$S_\alpha(A) = \frac{1}{1-\alpha} \ln\left[\mathrm{Tr}\left(\rho_A^\alpha\right)\right], \tag{3}$$

where $\rho_A = \mathrm{Tr}_B(\rho)$ is the reduced density matrix of the total system traced out over region B and the standard von Neumann entanglement entropy is recovered in the limit as $\alpha \to 1$.

Despite the inaccessibility of the full wave function of the system in Monte Carlo techniques, it is possible to sample $\mathrm{Tr}(\rho_A^\alpha)$ for integer $\alpha > 1$. This is accomplished by taking the expectation value of a SWAP operator[17] for $\alpha = 2$ or permutation operator for $\alpha > 2$;[21] e.g.,

$$S_2(A) = -\ln(\langle \mathrm{SWAP}_A \rangle), \tag{4}$$

$$S_3(A) = -\tfrac{1}{2} \ln\left(\left\langle \Pi_3^A \right\rangle\right). \tag{5}$$

To measure the α^{th} entropy, each projected state must be composed of α noninteracting copies of the system. The permutation operators Π_α^A act to cyclicly exchange the state in region A between the α different copies of the system and are constructed such that $\langle \Pi_\alpha^A \rangle = \mathrm{Tr}(\rho_A^\alpha)$. In the case of spin states (which for each MC step in the simulation is a product state) we can simply swap the states within region A between copies of the system. For valence bond states the application of the permutation operator also has a simple result; it acts to exchange the endpoints of valence bonds within region A between copies of the system and as such can create bonds between the noninteracting copies.[17]

The bare measurement of the SWAP operator has been shown[17] to have problems with convergence for large region A, while, in principle, the measurement should be symmetric such that $\langle \mathrm{SWAP}_A \rangle = \langle \mathrm{SWAP}_B \rangle$, since the two density matrices ρ_A and ρ_B have the same eigenvalues. This is because the exchange of a larger region gives a larger number of different states as a result of that SWAP and, thus, a larger range and number of possible values in the expectation value. It simply takes more Monte Carlo steps to converge on the same result. Another way to think of it is that, even though $\mathrm{Tr}(\rho_A^2) = \mathrm{Tr}(\rho_B^2)$, if region A is much larger than region B then $\mathrm{Tr}(\rho_A^2)$ contains exponentially more terms than $\mathrm{Tr}(\rho_B^2)$. Getting the value to converge by a stochastic sampling of these terms takes much longer. However, it is possible to instead sample a series of smaller regions and greatly improve the convergence time. This technique is described in the following section.

2. *The ratio trick*

The convergence issue mentioned above can be addressed by a reweighting of the Monte Carlo sampling scheme. Begin by considering the double-projector method, where one measures the expectation value of SWAP by sampling terms from

$$\langle \mathrm{SWAP}_A \rangle = \frac{\sum_{lr} w_r w_r \langle V_l | V_r \rangle \dfrac{\langle V_l | \mathrm{SWAP}_A | V_r \rangle}{\langle V_l | V_r \rangle}}{\sum_{lr} w_l w_r \langle V_l | V_r \rangle}, \tag{6}$$

where $|V_l\rangle, |V_r\rangle$ are the states obtained by applying lists of bond operators to the trial states and w_l, w_r are the weights accrued by applying those operators. Terms are sampled proportional to the total weight $W = w_l w_r \langle V_l | V_r \rangle$ by accepting a new configuration with probability $W^{\mathrm{new}}/W^{\mathrm{old}}$. Thus one can simply measure $\frac{\langle V_l | \mathrm{SWAP}_A | V_r \rangle}{\langle V_l | V_r \rangle}$ once per Monte Carlo step, and the average value will give us $\langle \mathrm{SWAP}_A \rangle$.

The convergence difficulties can be combatted by using the *ratio trick*.[17] One modifies the sampling weight to include the expectation value of a SWAP operator for a region A that is close in size to the region we intend to measure. One can then measure the ratio of these two operators, e.g.,

$$\frac{\langle \mathrm{SWAP}_A \rangle}{\langle \mathrm{SWAP}_{A'} \rangle} = \frac{\sum_{lr} w_r w_r \langle V_l | \mathrm{SWAP}_{A'} | V_r \rangle \dfrac{\langle V_l | \mathrm{SWAP}_A | V_r \rangle}{\langle V_l | \mathrm{SWAP}_{A'} | V_r \rangle}}{\sum_{lr} w_l w_r \langle V_l | \mathrm{SWAP}_{A'} | V_r \rangle}. \tag{7}$$

This improves the sampling since, if regions A and A' are similar in size, the measurement $\frac{\langle V_l|\mathrm{SWAP}_A|V_r\rangle}{\langle V_l|\mathrm{SWAP}_{A'}|V_r\rangle}$ will have fewer possible values than $\frac{\langle V_l|\mathrm{SWAP}_A|V_r\rangle}{\langle V_l|V_r\rangle}$, and those values will have a smaller variance.

Note, however, that one is only measuring a *ratio* of expectation values. That is, to obtain $\langle \mathrm{SWAP}_A\rangle$, one must know the value of $\langle \mathrm{SWAP}_{A'}\rangle$. If $\langle \mathrm{SWAP}_{A'}\rangle$ was also obtained by a ratio trick simulation, the expectation value for the smaller component of A' must be determined, and so on. Thus, the procedure that we use in this paper is to measure a range of sizes for region A, beginning with a measurement of the bare SWAP for a small region size, and increase the size of regions A and A' over several simulations in sequence. That is, we measure SWAP for a sequence of different region sizes, $A_1, A_2, \ldots, A_n$, where the number of lattice sites in A_{i+1} is greater than the number of sites in region A_i. Then, the Renyi entropy of an arbitrary region A_n is calculated through

$$S_2(A_n) = -\ln\left(\frac{\langle \mathrm{SWAP}_{A_n}\rangle}{\langle \mathrm{SWAP}_{A_{n-1}}\rangle}\right) - \ln\left(\frac{\langle \mathrm{SWAP}_{A_{n-1}}\rangle}{\langle \mathrm{SWAP}_{A_{n-2}}\rangle}\right) - \cdots - \ln\left(\frac{\langle \mathrm{SWAP}_{A_2}\rangle}{\langle \mathrm{SWAP}_{A_1}\rangle}\right) - \ln\left(\langle \mathrm{SWAP}_{A_1}\rangle\right), \quad (8)$$

where each ratio is calculated via Eq. (7), and the last expectation value for A_1 via Eq. (6). Note that each term in the sum requires a different VB QMC simulation, since, although we can measure the entropy for any region A within one simulation, we can use only *one* size of A' per simulation, since it affects the sampling of the valence bond states as described below. The scaling cost of the Ratio trick is therefore n; however, the gain in sampling efficiency is demonstrated to more than compensate for this additional simulation cost.

C. The loop-ratio algorithm

In order to calculate Renyi entanglement entropy with maximal efficiency in VB QMC simulations of the Heisenberg model, an algorithm should be employed that combines the loop update with the ratio trick. When modifying the loop algorithm to use the ratio trick, the same principles as in the double-projector algorithm (above) apply; however, the sampling weight [from Eq. (6)] is not explicit since one instead samples over spin states whose overlap is always unity.

In order to made the necessary modification to the loop algorithm, the system should first be replicated so two noninteracting copies are present, as usual for measurements of the SWAP operator.[17] Links in the simulation cell then are reconnected as if there were a SWAP operator permanently applied to the projected state $|V_r\rangle$, shown in Fig. 2. This causes spins from different noninteracting copies of the system to be connected via loops, which means they can be flipped together, and, thus, the spin states are sampled according to the swapped system $\langle V_l|\mathrm{SWAP}_{A'}|V_r\rangle$. The measurement of $\langle V_l|\mathrm{SWAP}_A|V_r\rangle/\langle V_l|\mathrm{SWAP}_{A'}|V_r\rangle$ is then accomplished by measuring an operator that swaps the states of the sites in region A that were not already swapped in region A', assuming $A' \subset A$.

This method has the same limitation of the double-projector ratio trick, whereby only one value of A' can be used per simulation, so the region to be measured must be built up from a small region A according to Eq. (8). In our results below, we use two geometries for building the region A: "strips" and "squares" (see Fig. 3 or Ref. 22). Strips refer to geometries where region A has one dimension equal to the linear size of the toroidal system itself and is, therefore, without corners. The Renyi entropy $S_2(A)$ is built up through the ratio trick by systematically adding subregions of size $L \times 1$. Squares refer to geometries where the linear size of A is increased symmetrically, starting from size 1×1. Square regions A necessarily have four corners.

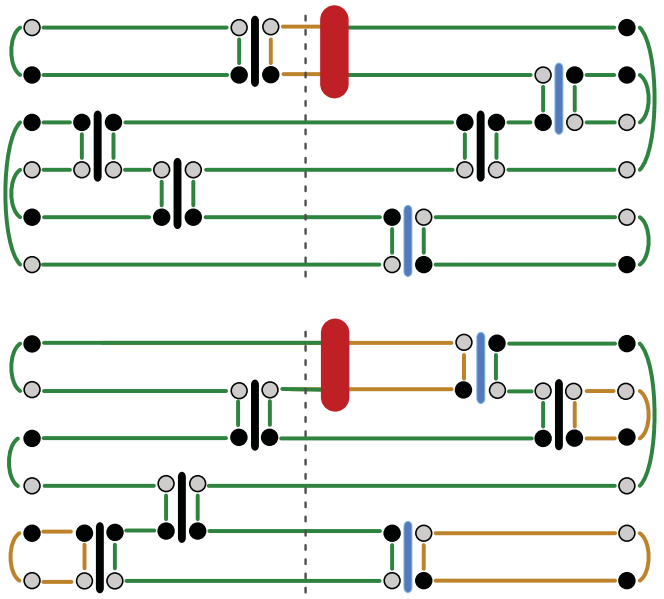

FIG. 2. (Color online) One possible simulation cell configuration for the loop ratio algorithm on a six-site system where $\alpha = 2$ and region A' contains the first two sites of the system. Spins between the usually noninteracting copies are connected through loops via the SWAP operator (red). The green and orange links are used to show the connections between the sheets. The loop on the left side of the SWAP operator on the top sheet is connected to the right side on the bottom sheet and so forth.

D. Results

We begin by testing some basic properties of the Renyi entropy as calculated through the VB basis QMC with the loop-ratio trick outlined above. First, we examine the convergence of the Renyi entropy as a function of operator list length per site, $m = M/N$, as illustrated in Fig. 4. Using several system sizes, both periodic and open boundaries, and strip and square geometries (described above and in Fig. 3), we see very good convergence by $m = 10$. This value of m was used for all the following VB QMC simulations, as the measurements are able to converge with that number of operators, but additional operators would detract from the algorithm's efficiency.

Figure 3 shows examples of $S_2(A)$ for regions A of both strip and square geometries of increasing width for a 20×20 toroidal system. The length of the boundary for all strip regions in this plot is $\ell = 40$, whereas for the square regions $\ell = 4x$. The area law scaling of S_2 is apparent in that the strip and

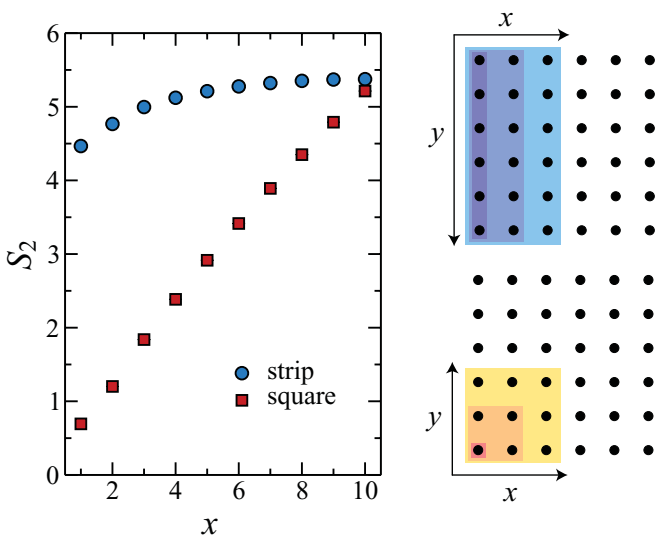

FIG. 3. (Color online) (Left) The Renyi entropy for a 20×20 PBC system versus the width (x) of region A for both the square and strip geometries. The boundary length does not change with the region width for strip geometry (since its height y traverses the periodic lattice), thus the entropy approaches a constant value. For square geometry the boundary length is $4x$. In both geometries the entropy scales with the length of the boundary, with that scaling becoming better as the width of region A approaches half the system size. At right, the "strip" (top) and "square" (bottom) geometries on a 6×6 lattice.

square geometries both approach a straight line with zero and nonzero slope, respectively.

To determine the scaling of entanglement entropy in two dimensions, we examine $L \times L$ systems with periodic boundary conditions, the results of which are shown in Fig. 5. Region A was systematically built up according to the square

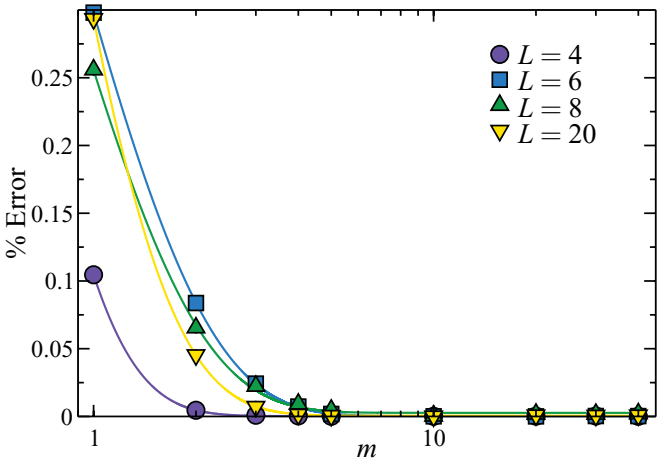

FIG. 4. (Color online) Percentage error in S_2 versus the number of operators per site (m) for different $L \times L$ lattices. For $L = 4,6,8$ the lattices have open boundaries and region A is half the system using the strip geometry. The exact values were found using density matrix renormalization group (DMRG) simulations. The $L = 20$ lattice has periodic boundaries and A is a 2×2 square. The "exact" value is taken from the $m = 10$ simulation. Each data set was fit to an exponential function.

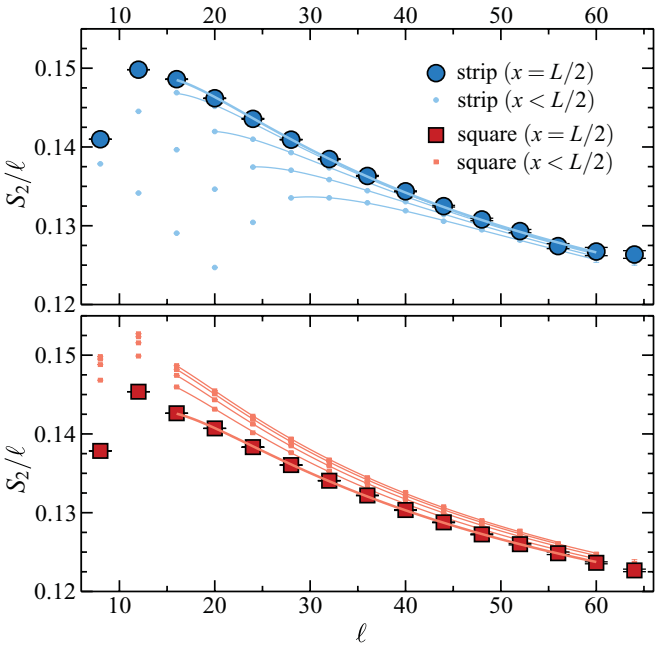

FIG. 5. (Color online) S_2/ℓ vs. ℓ, where ℓ is the boundary length of region A, for regions with square and strip geometry embedded in $L \times L$ systems with periodic boundary conditions. The data sets outlined in black correspond to region A with width $x = L/2$ for both geometries. The other data sets use smaller region A of the same geometries. Fits to Eq. (9) are included for all data, and the coefficients found are listed in Table I.

and strip geometries as defined in Fig. 3. Since each region is built to satisfy Eq. (8), one can perform fits for several sizes of subregion for each lattice size. In this case (as opposed to the plot in Fig. 3), for each set of data in Fig. 5 we use a region A with width proportional to the system size. This is done in an attempt to overcome finite-size effects, and the interaction of boundaries (that can be seen for the strip geometry in Fig. 3). Figure 5 includes data for $S_2(A)$ using regions with width $x = L/2$ as well as smaller regions $x < L/2$. A smaller region A has the advantage that $S_2(A)$ converges faster, but the drawback is it brings one into the regime of finite-size effects, apparent in Fig. 3 with small x, where S_2 is lower than the area-law value found at $x = L/2$. This deviation seems to depend on the fraction of the system contained within region A which is why, for the square case (where both the size of region A and the boundary length are scaled with system size), the different region widths do not change the entropy scaling very much. However, for the strip geometry with a region A width of $x = L/2 - n$, the fraction of the system contained in region A changes with L and approaches $1/2$ as L increases, i.e., when $L/2 \gg n$. This effect is evident in the top panel of Fig. 5, where for $n \neq 0$ the entropy is diminished for smaller system sizes but approaches the $n = 0$ values as system size increases.

As is clear, the data give excellent fits to the function

$$\frac{f(\ell)}{\ell} = a + \frac{c}{\ell}\ln(\ell) + \frac{d}{\ell}, \tag{9}$$

where ℓ is the length of the boundary between regions A and B, provided that the very smallest lattice sizes are excluded.

TABLE I. The coefficients a, c, d found by fitting the data in Fig. 5 to Eq. (9). The fits were done for both strip and square geometries, beginning at $x = L/2$ for both and decreasing the region sizes until $x = L/2$–4, where x is defined in Fig. 3 as the width of region A.

Geometry	$L/2 - x$	a	c	$-d$
Strip	0	0.096(5)	0.7(4)	1.2(2)
	1	0.095(6)	0.7(9)	1.3(6)
	2	0.092(3)	0.9(6)	1.8(8)
	3	0.08(8)	1.2(0)	2.(6)
	4	0.08(2)	1.5(4)	3.(7)
Square	0	0.097(6)	0.64(2)	1.0(6)
	1	0.097(6)	0.62(1)	0.9(5)
	2	0.097(7)	0.61(7)	0.9(1)
	3	0.097(6)	0.62(6)	0.9(3)
	4	0.097(5)	0.6(3)	0.9(4)

The values obtained for coefficients a, c, and d are listed in Table I.

III. SERIES EXPANSIONS

A. Expansion methods

We have developed two different types of series expansions to calculate the entanglement properties of the Heisenberg model. The first is the high temperature expansions. This method was introduced in Ref. 8 for the XXZ model, of which the Heisenberg model is a special case. In this method, the mutual information between two regions A and B associated with their boundaries or corners can be expanded in powers of inverse temperature β. The calculation can be done for Renyi mutual information of index α by introducing α replicas of the system. The coefficient of β^α is a polynomial in α of order $\alpha - 1$ so the limit $\alpha \to 1$ can be readily taken to calculate the von Neumann mutual information as well.

Since there is no finite-temperature phase transition in the Heisenberg model on the square lattice, the expansions in β can, in principle, be extrapolated down to $T = 0$. It is well known that the correlation length of the system grows exponentially at low temperatures as $\exp(C/T)$. Hence, we expect corrections to leading behavior to be exponentially small at low temperature, $\exp(-C/T)$. For this reason, a change of variables $w = \tanh\beta$ is applied before Pade approximants are used. This extrapolation method has been used in the past for other properties of the Heisenberg model[23] and also shows good convergence between different Pade approximants when applied to mutual information.

In addition, we have developed a series expansion directly for the entanglement entropy at $T = 0$ using Ising anisotropy parameter $\lambda = J_{xy}/J_z$ of the XXZ model. When $\lambda = 0$, we pick one of the two Ising states to expand around. That unique state factorizes for any two regions A and B. Therefore, any entanglement entropy vanishes in that limit. For any $\alpha \geqslant 2$, the series expansion for the Renyi entropy can be calculated as a power series in λ using a linked cluster expansion.[24,25] Unlike HTE, the coefficients of the Renyi entropies for different α are not related by a simple polynomial relation and for every α a different calculation is needed. Here, we will restrict ourselves to the calculation of Renyi entropy with $\alpha = 2$.

The Ising series expansions work in the thermodynamic limit, starting with a system that has a short correlation length and long-range order along a particular direction. Since a specific ordered state is picked out, their major limitation is that they cannot be used to study any bulk entanglement properties associated with broken symmetry. In any finite order of perturbation theory, contributions can come only from the boundary between regions A and B. The series are nonsingular as long as the system has a gap. The singular dependence on the size L of the system is replaced in the series expansion studies by a dependence on the correlation length ξ, which diverges as the gapless Heisenberg point is approached. By scaling, the dependence on L should translate into a similar dependence on ξ. An advantage of series expansions is that entanglement associated with different surface manifolds such as surfaces, lines, and corners can be analytically separated and separate expansions can be obtained for the entropy associated with them.

B. Results from series expansions

Pade extrapolation of HTE in the variable $w = \tanh\beta$ are compared with the data from SSE QMC simulations in Fig. 6. The agreement is excellent down to temperatures below J. At low temperatures the QMC data show dramatic finite-size effects, with a maximum and a minimum, followed by a slow rise at lower temperatures. In contrast, the Pade approximants show a steady monotonic rise and saturation at low temperatures. These results suggest that the limit of $T \to 0$ and $L \to \infty$ do not commute for these quantities. We note that such noncommuting limits are well known for other properties of the Heisenberg model.[15,16] However, these noncommuting limits have not been anticipated for the mutual information. In Fig. 7, we plot the crossover temperature, as defined by the temperature of the local maximum in

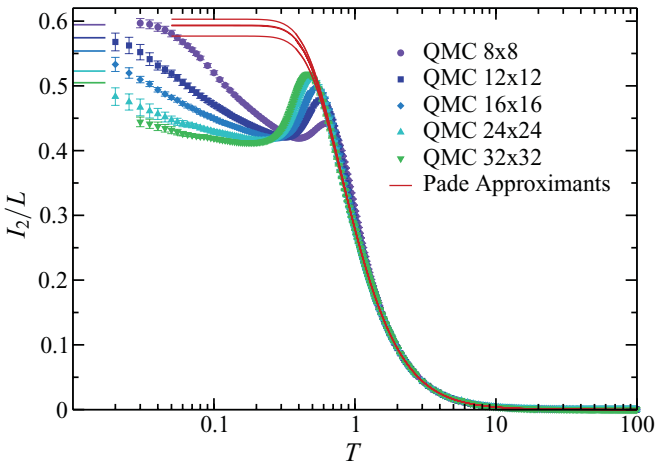

FIG. 6. (Color online) The finite-temperature Renyi mutual information for $L \times L$ periodic systems with strip geometry (note $\ell = 2L$) for region A with SSE QMC results for $L = 8,12,16,24,32$, and [4/6], [5/5], [6/5], and [6/4] Pade approximants for the HTE in the variable $w = \tanh\beta$. The horizontal lines on the left-hand side correspond to the zero temperature results (from VB QMC) for each of the finite-temperature SSE data sets.

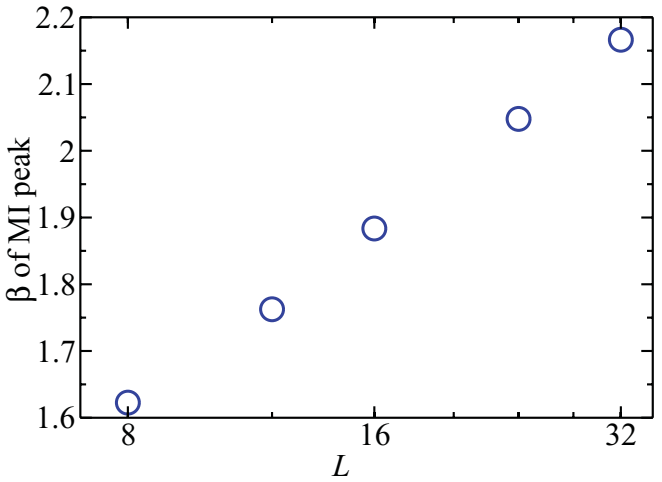

FIG. 7. (Color online) The inverse crossover temperature of Renyi mutual information for the system sizes shown in Fig. 6, where the crossover temperature is taken as the temperature at the highest point of the $T > 0$ mutual information peak. The x axis is logarithmic, showing that the crossover temperature scales as $\sim\log(L)$.

mutual information at $T > 0$ as a function of L. We see that the crossover temperature scales inversely with $\log(L)$ or, equivalently, L scales exponentially in the inverse temperature. Such a length scale agrees with the correlation length in the Heisenberg model, which also scales exponentially at nonzero temperature, and so, at first sight, it is natural to ascribe this peak in the mutual information to the size of the correlation domains. One might imagine the following argument: when the size of the system is larger than the correlation domain size, there are many correlation domains along the boundary. Since the direction of the spin in a correlation domain can be viewed as roughly uncorrelated with that in other correlation domains, this contributes an additional term to the entropy of both regions. However, the correlation domains that cross the boundary produce some correlations between the two regions, leading to a positive contribution to the mutual information. While this argument captures the correct qualitative scaling, suggesting that the mutual information should drop at lower temperatures, it fails on quantitative grounds. The contribution to the entropy of a correlation domain from the random ordering direction of that domain should scale something like the logarithm of the correlation volume (see, for example, the mean-field theory calculation of the next section, where at low temperatures the entire system comprises one correlation domain) and hence should be proportional to β. However, the density of these correlation domains (the number of domains per unit boundary length, which is inversely proportional to the correlation length) is exponentially small in β, and so we should expect that this contribution to the entropy per unit length should become negligible as ℓ gets larger. In contrast, the numerical data shows the difference between the $T > 0$ maximum and the $T = 0$ limit increasing with increasing ℓ, so this difference remains as an unexplained phenomenon.

Let the expansions around the Ising limit for the second Renyi entropy per unit length of the boundary be given by

$$S_{2l}/\ell = \sum_{n=2} a_n \lambda^n,$$

TABLE II. Ising series expansion coefficients for the line and corner terms

n	a_n	b_n
2	0.05555556	0
4	0.00314815	0.00913580
6	0.00558342	−0.00542753
8	0.00353554	−0.00126847
10	0.00220329	−0.00172570
12	0.00186784	−0.00170418
14	0.00144690	−0.00144091

and, for a single corner,

$$S_{2c} = \sum_{n=4} b_n \lambda^n.$$

The nonzero coefficients up to $n = 14$ are given in Table II. Note that all odd order terms are zero, so the expansion is in the variable λ^2. The series for the line term S_{2l} are evaluated by first performing a change of variables $\delta = \lambda^2 - 2\lambda^4$ to remove a square-root singularity at $\lambda = 1$ and then calculating Pade approximants. From these we estimate

$$S_{2l}/\ell = 0.094 \pm 0.001,$$

for the Heisenberg model. Here, the error bars represent spread in values obtained from different Pade approximants.[25] For the corner term, we expect a logarithmic singularity of the form

$$S_{2c} = x \ln \xi \sim -\frac{x}{2} \ln (1 - \lambda^2),$$

where we have used the fact that the correlation length diverges as $(1 - \lambda^2)^{-1/2}$. Given the anticipated logarithmic singularity, we first differentiate the series with respect to the variable λ^2 and then use Pade approximants biased at $\lambda^2 = 1$ to estimate the residue. From these, we estimate

$$x = -0.020 \pm 0.002.$$

Numerical values for S_{2l}/ℓ and x will be compared to VB QMC results and other calculations later in the discussion section.

IV. MEAN-FIELD THEORY AND TOWER-OF-STATES MODES

A. Tower of states and thermodynamics

One starting point for a theoretical treatment of the Heisenberg model is a mean-field theory. This mean-field theory provides a simple framework for understanding some logarithmic bulk corrections associated with spontaneous symmetry breaking, as well as providing some understanding of the low energy "tower-of-states" modes (first identified by Anderson[5]) which are present even in the two-dimensional model.

In mean-field theory we consider the model Hamiltonian

$$H = \frac{J}{N} \sum_{i \in 1, j \in 2} \vec{S}_i \cdot \vec{S}_j, \tag{10}$$

where the sum ranges over spin i in one sublattice and j in the other sublattice (we use 1 and 2 to denote sublattices), and

there are total of N spins, with $N/2$ spins in each sublattice. The ground state of this Hamiltonian corresponds to taking all spins in sublattice 1 in a symmetric state, with total spin $N/4$, and, similarly, taking all spins in sublattice 2 in a symmetric state and then pairing the two spins to form a singlet. We begin by working out thermodynamic properties of the model and we discuss their impact on QMC simulations using SSE and VB projector methods. We then turn to the question of the entanglement entropy.

First, we work out the energy of the aforementioned singlet state as follows: let S_1 denote the total spin operator in sublattice 1 and S_2 denote the same for sublattice 2. The Hamiltonian then can be written as $(J/N)S_1 \cdot S_2$. This equals $(J/2N)[(S_1 + S_2)^2 - S_1^2 - S_2^2]$. The singlet state has $S_1 + S_2 = 0$. Looking at states with spin $N/4$ in each sublattice, we have $S_1^2 = S_2^2 = (N/4)(N/4 + 1)$, so the energy is $-(J/N)(N/4)(N/4 + 1)$, which is of order N.

One can now look at excited states. The excited states where both sublattices have total spin $N/4$, but the whole system is not in a singlet, give the "tower states": we find that in this case S_1+S_2 is not equal to zero. Indeed, the energy difference of this state, compared to the ground state, is simply given by

$$E - E_{\text{ground}} = (J/2N)(S_1 + S_2)^2. \tag{11}$$

That is, the energy is equal to $(J/2N)$ times the total spin squared. In real two-dimensional systems, a similar low-energy structure of states is observed, with the energy of these states proportional to total spin squared divided by N, although the coupling constant J describing the energy of these modes may be renormalized compared to the coupling constant appearing in the lattice Hamiltonian. These low-energy states can be observed in exact diagonalization, and, in fact, they are one of the best checks for the presence of symmetry breaking.[26] Note that these states are much lower energy than the spin-wave states: In a 2D system with linear size L, they have energy of order $1/L^2$, while the lowest energy spin wave has an energy of order $1/L$.

There are also excited states where a given sublattice does not have total spin $N/4$. That is, not all spins in the same sublattice are in a symmetric state. One can check in this case that the energy of these states is increased above the ground-state energy by an amount that is of order 1 (or more, if the spin is reduced a lot compared to $N/4$). These states can be viewed as a mean-field theory analog of the spin-wave states; that is, the mean-field theory raises the energy of the spin waves from order $1/L$ to order 1. Therefore, at a temperature of order 1, it becomes reasonable to ignore those states in mean-field theory [more precisely we need a temp of order $1/\ln(N)$]. In a two-dimensional system, the temperature needs to become of order $1/L$ to ignore the spin-wave modes.

We now consider the effect of the tower of states on the bulk entropy at temperatures sufficiently low that the spin waves can be ignored. We emphasize that this estimate is *not* a calculation of the entanglement entropy but rather a calculation of a bulk thermodynamic entropy. Considering the tower states, the total entropy is not hard to work out: The allowed states are one state (the ground state) which is a singlet, three states with spin 1 and energy increased by $JS(S + 1)/2N = 2J/2N$, five states with spin 2 and energy $6J/2N$, and so on (in this case, the Clebsch-Gordon coefficients work out simply, so the number of states with spin S is exactly $2S+1$). As a rough approximation, at temperature T, we expect to excite states with spin such that JS^2/N is of order T. Therefore, S is of order $\sqrt{TN/J}$ or less. The number of such states is of order S^2 and, hence, of order TN/J. Thus, at a temperature T much larger than $1/N$, but sufficiently small that the spin waves can be ignored, the entropy is equal to

$$\log(TN/J) + \text{const.} \tag{12}$$

B. Effect of tower of states on QMC simulations

In QMC simulations, if we want to access the ground state, it is important to access a temperature sufficiently small that these tower-of-states modes can be ignored. Since the energy of the tower-of-states modes is so small (of order $1/N$) this can require a prohibitively small temperature. However, we will see that this is *not* as big a problem as it might seem.

First, consider the VB projector method. The projector method starts with a trial wave function and then applies a large power of the Hamiltonian to this wave function. The Hamiltonian is appropriately scaled such that its ground state will have the largest eigenvalue, so this high power of the Hamiltonian, acting on the trial wave function, produces a state close to the ground state. This is essentially the power method of finding largest or smallest eigenvalues of a matrix. Given that the Hamiltonian has tower-of-states modes with energy very close to the ground state, it would seem that we would have to apply a very high power of the Hamiltonian, a power that is of order N, in order to produce a final state close to the ground state. However, the trial wave function in the VB projector method has total spin 0, and the Hamiltonian conserves spin. Thus, the wave function produced by acting on the trial wave function with a high power of the Hamiltonian also has total spin 0. However, all the tower-of-states modes have nonzero spin, and, hence, the wave function we produce has no overlap with the tower-of-states modes. This is the reason why it suffices to simply go to high-enough order to project out the spin-wave modes, as once the spin-wave excitations are projected out, the tower of states are also projected out by symmetry (note that we do have excited states with total spin 0 that excite both a spin-wave and a tower mode, but such states have energy of order $1/L$ or higher, not $1/N$).

Consider, finally, the SSE method. As discussed above, the tower-of-states modes do have a noticeable effect on the calculation of the bulk entropy. However, as we will see below, at least in mean-field theory, they have only a small effect on the calculation of the entropy of the reduced density matrix, suggesting that the SSE calculations of the reduced density matrix entropy converge well even without accessing temperature of order $1/N$.

C. Entanglement entropy

We now consider the entanglement entropy in the mean-field model. The entanglement between the 1 and 2 sublattices is the simplest to calculate. Each sublattice has total spin $N/4$ and, hence, has $N/2+1$ states. The ground state, which we

call ψ^0, is maximally entangled and, hence, has entanglement entropy $\ln(N/2+1)$, with all different Renyi entropies equal.

The more interesting entanglement entropy to calculate is to imagine dividing the system into two halves, A and B, with each half having $N/4$ spins in sublattice 1 and $N/4$ spins in sublattice 2. This is, in our opinion, the simplest mean-field model that is relevant to the numerical calculations on the two-dimensional Heisenberg model, as in that case each region A and B contains spins from both sublattices. We will see that this calculation gives a logarithmic dependence on N also, although the result is more complicated.

The ground state ψ^0 can be obtained by taking *any* state that has spin $N/4$ in each sublattice and that has nonzero overlap with the ground state and projecting it into the total spin zero sector. We choose to use a Néel state as our state before projection, where we define this state $\psi^{\text{Néel}}$ as a state in which all spins in sublattice 1 are pointing up and all spins in sublattice 2 are pointing down. We do this projection by averaging over different rotations of the Néel state, so

$$\psi^0 = Z^{-1/2} \int d\theta d\phi R(\theta,\phi)\psi^{\text{Néel}}, \tag{13}$$

where $R(\theta,\phi)$ is the rotation by angles θ and ϕ, and the measure is chosen to be uniform over all rotations (that is, we choose the Haar measure), and where Z is a normalization factor so $|\psi^0|^2 = 1$.

We compute the second Renyi entropy, S_2. The calculation of other Renyi entropies are similar; in this case, in contrast to the previous entropy calculation above, Renyi entropies differ, so S_2 is not equal to the von Neumann entropy. To calculate the Renyi entropy, we must calculate

$$\langle \psi^0 \otimes \psi^0 | \text{SWAP}_A | \psi^0 \otimes \psi^0 \rangle, \tag{14}$$

where SWAP_A is the swap operator used in Ref. 17.

This expectation value is equal to

$$\begin{aligned} Z^{-2} \int & d\theta_1 d\phi_1 d\theta_2 d\phi_2 d\theta_3 d\phi_3 d\theta_4 d\phi_4 \\ & \times \langle \psi^{\text{Néel}} \otimes \psi^{\text{Néel}} | R(\theta_1,\phi_1)^\dagger \otimes R(\theta_2,\phi_2)^\dagger | \text{SWAP}_A | \\ & \times R(\theta_3,\phi_3) \otimes R(\theta_4,\phi_4) | \psi^{\text{Néel}} \otimes \psi^{\text{Néel}} \rangle. \end{aligned} \tag{15}$$

We first estimate Z as follows. We have

$$Z = \int d\theta_1 d\phi_1 d\theta_2 d\phi_2 \langle \psi^{\text{Néel}} | R(\theta_1,\phi_1)^\dagger R(\theta_2,\phi_2) | \psi^{\text{Néel}} \rangle. \tag{16}$$

We can combine the rotations $R(\theta_1,\phi_1)^\dagger R(\theta_2,\phi_2)$ into one rotation by a pair of combined angles, $R(\theta,\phi)$. Equivalently, we note that the integral $\int d\theta_2 d\phi_2 \langle \psi^{\text{Néel}} | R(\theta_1,\phi_1)^\dagger R(\theta_2,\phi_2) | \psi^{\text{Néel}} \rangle$ is independent of θ_1,ϕ_1, so we can fix $\theta_1 = \phi_1 = 0$. Thus, up to constant factors, we have

$$Z = \int d\phi d\theta \langle \psi^{\text{Néel}} | R(\theta,\phi) | \psi^{\text{Néel}} \rangle. \tag{17}$$

The expectation value in the above integral is just the N-th power of $\langle \uparrow | R(\theta,\phi) | \uparrow \rangle$. This is approximated by (for small θ,ϕ) $\exp[-N(\theta^2+\phi^2)]$. Because of the factor of N in the exponent, the restriction to small θ,ϕ is justified as the expectation value is negligible for large θ,ϕ. The integral over θ,ϕ then is Gaussian and the result is that

$$Z \propto N^{-1}. \tag{18}$$

We now estimate the integral in Eq. (15). In this case, we have an integral over four pairs of angles. For the spins not in region A the expression is small unless θ_1 is close to θ_3 and ϕ_1 is close to ϕ_3 and also θ_2 is close to θ_4 and ϕ_2 is close to ϕ_4. Similarly, for the spins in region A we need θ_1 close to θ_4 (and ϕ_1 close to ϕ_4) and θ_2 close to θ_3 (and ϕ_2 close to ϕ_3). Thus, the overlap of spins in region A forces certain pairs of angles to be close and the overlap of spins not in region A forces other pairs of angles to be close. So, in fact, all the angles need to be very similar. So, we get (approximately) a Gaussian integral over three pairs of relative angles and one overall rotation that we can factor out. The result is proportional to $1/N^3$, then, up to constant factors.

Thus, the expectation is $1/N^3/(1/N^2) = 1/N$, giving, again, an entropy that scales as $\ln(N) + \text{const}$ for large N. Finally, we can consider this entropy at a nonzero temperature. The effect of a nonzero temperature (high enough to excite the tower modes but low enough to avoid exciting spin-wave states so the total spin in each sublattice will still equal $N/4$) is to give a global density matrix

$$\begin{aligned} \rho = Z^{-1} \int & d\theta_1 d\phi_1 d\theta_2 d\phi_2 F(\theta_1,\phi_1,\theta_2,\phi_2) \\ & \times R(\theta_1,\phi_1) | \psi^{\text{Néel}} \rangle \langle \psi^{\text{Néel}} | R(\theta_2,\phi_2)^\dagger, \end{aligned} \tag{19}$$

for some function F, which depends only on the relative angles between the two rotations. That is, at higher temperatures the rotation angles in the bra and ket vectors become coupled, while at zero temperature F is a constant. One can check that this still leaves us with an entropy of region A which is equal to $\ln(N) + \text{const}$, so, at least in mean-field theory, the $\ln(N)$ term in the entropy of region A does not depend on whether or not the tower modes are excited.

V. SPIN-WAVE THEORY

In understanding the area law for the Heisenberg model, an immediate question arises: what is the entanglement entropy in spin-wave theory? After all, spin waves are gapless, and gapless modes might make one concerned whether an area law holds. In fact, the entanglement entropy of gapless modes depends strongly on dimension. Free bosons with a linear dispersion relation have a logarithmically divergent entanglement entropy in one dimension, following conformal field theory,[1,27] but in two or more dimensions they obey an area law.[28]

In Ref. 29, it was shown numerically that the Heisenberg model itself obeyed an area law using density matrix renormalization group and a spin-wave calculation also led to an area law. In Ref. 14, a spin-wave calculation was carried out for a finite-size system and was shown to roughly match the qualitative behavior from a quantum Monte Carlo simulation.[17]

While gapless bosons with linear dispersion do produce an area law in two dimensions, they also produce nontrivial exponents associated with corners.[30] The entanglement

entropy of a region A has a term equal to the log of the length scale of A, multiplied by the sum over corners of a scaling function of angle of each corner. For the entropy S_2, this term is equal to

$$\approx -0.0062 \ln(\ell) \tag{20}$$

for each 90° corner for a real scalar field with linear dispersion.[30] Note that the sign of this correction is negative.

For the system we are concerned with, we must multiply this result by 2. There are two ways to understand this counting of modes to see why the result must be multiplied by 2. On the one hand, we can consider an O(3) nonlinear σ model. In 2+1 dimensions, this model has a symmetry broken phase, and the Heisenberg model ground state corresponds to this phase. In the symmetry broken phase, there are two Goldstone modes, corresponding to two different transverse directions in which the order parameter can move. In the Hamiltonian spin-wave language, this factor of 2 again arises, but for a reason that initially might seem to differ. Suppose we use a spin-wave representation in which the operator $b_i^\dagger$ always creates an excitation on site i (we are choosing to follow the notation of Ref. 14, though, of course, the Hamiltonian spin-wave calculation is a textbook calculation); that is, if we do a spin-wave expansion about a state with spins up on the 1 sublattice and down on the 2 sublattice, then this operator $b_i^\dagger$ corresponds to a lowering operator on the 1 sublattice and a raising operator on the 2 sublattice. We then find gapless modes at momenta near (0,0) and (π,π) on a two-dimensional square lattice. Thus, we again see a factor of 2 arising from the existence of two different gapless points.

We can clarify the relation between the factor of 2 in these two different approaches. These gapless modes in the Hamiltonian model correspond to the following states, respectively. Acting with the operator $\sum_i b_i^\dagger$ on the ground state, which creates a zero-energy excitation with momentum (0,0), produces a state that is a superposition of all possible ways of flipping one spin. Acting with the operator $\sum_i (-1)^i b_i^\dagger$, which creates a zero-energy excitation with momentum (π,π), produces a state that again is a superposition of all possible ways of flipping one spin but with a plus sign for flipping a spin in the 1 sublattice and a minus sign for flipping a spin in the 2 sublattice. The first of these states corresponds to acting with the operator $\sum_i \sigma_i^x$ on the ground state and the second to acting with the operator $\sum_i \sigma_i^y$ on the ground state. Thus, they correspond to two different ways of rotating the symmetry broken ground state, either in the YZ plane or the XZ plane, matching the two different Goldstone modes above.

Therefore, for a square region, which has four such corners, we expect a correction of

$$\approx -0.0496 \ln(\ell). \tag{21}$$

As we saw above, numerical results disagree with this, suggesting some nontrivial effects that are not accounted for by this framework.

VI. SUMMARY AND DISCUSSION

In this paper, we have presented several computational methods to calculate the entanglement properties of lattice statistical models, in dimensionality greater than 1, in a systematic manner. The stochastic series expansion QMC and high-temperature expansions are finite-temperature methods. Since the former works with finite systems, it is possible to calculate the ground-state properties by going to sufficiently low temperatures. The latter is a series expansion, defined in the thermodynamic limit. It can, in principle, be extrapolated to $T \to 0$ limit if there is no finite-temperature phase transition. An interesting finding of our paper is that the limits $T \to 0$ and $L \to \infty$ need not commute in these calculations.

We have also developed computational methods that work directly at $T = 0$. The valence bond QMC is an exact ground-state projection method for a finite system. The Ising series expansions represent an expansion in exchange anisotropy around a given classical state. Using these methods we have calculated the area law associated with the boundary, as well as subleading logarithmic terms associated with the corners and the bulk. Whenever quantities are calculated by two different methods, there is good quantitative agreement.

We have also discussed a mean-field calculation of the entanglement properties as well as the results expected from noninteracting bosons. Here, we find a surprise insofar as the numerical results disagree with simple expectations. The mean-field state, where all spins on one sublattice are equally entangled with spins on the other sublattice, predicts a bulk log term of $2 \ln \ell$, whereas our valence bond QMC results give $c \ln \ell$, with $c = 0.74 \pm 0.02$. The latter is closer to a recent spin-wave calculation, where an estimate of $c = 0.92$ was obtained.[14] The discrepancy is even more puzzling for the corner terms. If long-wavelength spin waves act as noninteracting bosons, they should contribute a log term with a coefficient of -0.0496. That number should be compared with -0.080 ± 0.008 obtained in Ising series expansions and ≈ -0.1 in the QMC studies. These disagreements suggest the need for further theoretical study.

ACKNOWLEDGMENTS

We acknowledge collaboration related to the development of the VB QMC loop-ratio algorithm by H. G. Evertz. We are indebted to I. Gonzalez for providing DMRG results for Fig. 4. Support was provided by NSF Grant NO. DMR-1004231 and NSERC of Canada. The simulations were made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca) and Compute/Calcul Canada.

[1] P. Calabrese and J. Cardy, J. Stat. Mech.: Theor. Exp. (2004) P06002.

[2] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).

[3] L. Tagliacozzo, G. Evenbly, and G. Vidal, Phys. Rev. B **80**, 235127 (2009).

[4] S. Chakravarty, B. I. Halperin, and D. R. Nelson, Phys. Rev. B **39**, 2344 (1989).

[5]P. W. Anderson, Phys. Rev. **86**, 694 (1952).
[6]A. W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991).
[7]A. W. Sandvik, J. Phys. A **25**, 3667 (1992).
[8]R. G. Melko, A. B. Kallin, and M. B. Hastings, Phys. Rev. B **82**, 100409 (2010).
[9]A. W. Sandvik, Phys. Rev. Lett. **95**, 207203 (2005).
[10]A. W. Sandvik and K. S. D. Beach, *Computer Simulation Studies in Condensed-Matter Physics XX* (Springer, Berlin, 2008), Chap. 22.
[11]K. S. D. Beach and A. W. Sandvik, Nucl. Phys. B **750**, 142 (2006).
[12]A. W. Sandvik and H. G. Evertz, Phys. Rev. B **82**, 024407 (2010).
[13]M. Cramer, J. Eisert, and M. B. Plenio, Phys. Rev. Lett. **98**, 220603 (2007).
[14]H. F. Song, N. Laflorencie, S. Rachel, and K. Le Hur, Phys. Rev. B **83**, 224410 (2011).
[15]D. S. Fisher, Phys. Rev. B **39**, 11783 (1989).
[16]P. Hasenfratz and F. Niedermayer, Z. Phys. B Cond. Matt. **92**, 91 (1993).
[17]M. B. Hastings, I. González, A. B. Kallin, and R. G. Melko, Phys. Rev. Lett. **104**, 157201 (2010).
[18]F. Alet, S. Capponi, N. Laflorencie, and M. Mambrini, Phys. Rev. Lett. **99**, 117204 (2007).
[19]H. F. Song, S. Rachel, and K. Le Hur, Phys. Rev. B **82**, 012405 (2010).
[20]Y.-C. Lin and A. W. Sandvik, Phys. Rev. B **82**, 224414 (2010).
[21]J. Cardy, Phys. Rev. Lett. **106**, 150404 (2011).
[22]R. R. P. Singh, M. B. Hastings, A. B. Kallin, and R. G. Melko, Phys. Rev. Lett. **106**, 135701 (2011).
[23]R. R. P. Singh and M. P. Gelfand, Phys. Rev. B **42**, 996 (1990).
[24]M. P. Gelfand, R. R. P. Singh, and D. A. Huse, J. Stat. Phys. **59**, 1093 (1990).
[25]J. Oitmaa, C. J. Hamer, and W. Zheng, *Series Expansion Methods for Strongly Interacting Lattice Models* (Cambridge University Press, Cambridge, UK, 2006).
[26]B. Bernu, C. Lhuillier, and L. Pierre, Phys. Rev. Lett. **69**, 2590 (1992).
[27]H.-Q. Zhou, T. Barthel, J. O. Fjærestad, and U. Schollwöck, Phys. Rev. A **74**, 050305 (2006).
[28]M. Cramer, J. Eisert, M. B. Plenio, and J. Dreißig, Phys. Rev. A **73**, 012309 (2006).
[29]A. B. Kallin, I. González, M. B. Hastings, and R. G. Melko, Phys. Rev. Lett. **103**, 117203 (2009).
[30]H. Casini and M. Huerta, Nucl. Phys. B **764**, 183 (2007).