Variational treatment of the Shastry-Sutherland antiferromagnet using projected entangled pair states

A. Isacsson and Olav F. Syljuåsen
NORDITA, Blegdamsvej 17, Copenhagen Ø, DK-2100, Denmark
(Received 10 April 2006; published 1 August 2006)

We have applied a variational algorithm based on projected entangled pair states (PEPS) to a two dimensional frustrated spin system, the spin-1/2 antiferromagnetic Heisenberg model on the Shastry-Sutherland lattice. We use the class of PEPS with internal tensor dimension $D=2$, the first step beyond product states ($D=1$ PEPS). We have found that the $D=2$ variational PEPS algorithm is able to capture the physics in both the valence-bond crystal and the Néel ordered state. Also the spin textures giving rise to the magnetization plateaus seen in experiments on $SrCu_2(BO_3)_2$ are well reproduced. This shows that PEPS with the smallest nontrivial internal dimension, $D=2$, can provide valuable insights into frustrated spin systems.

DOI: 10.1103/PhysRevE.74.026701 PACS number(s): 75.40.Mg

I. INTRODUCTION

Spins with antiferromagnetic interactions prefer opposite alignment. However, in many materials the lattice structure does not allow all antiferromagnetic bonds to be satisfied simultaneously. These are known as frustrated antiferromagnets and display a variety of different phases ranging from rather well-known Néel-ordered phases to much less understood exotic phases such as valence bond crystals and spin liquids. There is no general theory of frustrated antiferromagnets, thus the different lattice structures are usually studied as separate models. Especially difficult are two-dimensional (2D) models. Finding the phase diagram of any of these models is made difficult by the lack of effective numerical tools that go beyond exact diagonalization of the Hamiltonian.

It is an unfortunate fact that the most powerful numerical methods such as the density matrix renormalization group (DMRG) and quantum Monte Carlo (QMC) that are very effective for studying general quantum magnets do not work well when applied to frustrated 2D antiferromagnets. DMRG is mainly restricted to 1D systems, and QMC suffers from the sign problem. However, there are promising variational methods [1–5] that perform energy minimization in a large class of states known as tensor product states (TPSs). Recently Verstraete and Cirac suggested an alternative minimization strategy in this space of states, there termed projected entangled pair states (PEPSs), which promises to be very efficient [6] and deserves further study. The PEPSs or TPSs have a natural "refinement" parameter, the internal dimension D of the tensors. This parameter determines how well the particular class of states covers the full Hilbert space. The lowest level $D=1$ corresponds to product states, thus yielding results on the level of mean field theory. The aim of the present paper is to investigate how well the next level in the hierarchy, $D=2$, can describe a 2D frustrated antiferromagnet of real physical interest.

An interesting frustrated antiferromagnet, which exhibits both a Néel ordered phase and a valence bond crystal phase, is the spin-1/2 Heisenberg antiferromagnet on the Shastry-Sutherland lattice, see Fig. 1. This model was initially proposed as a toy model possessing an exact dimerized eigenstate known as a valence bond crystal [7]. However, the interest in this model is more than academic as it is believed that the material compound $SrCu_2(BO_3)_2$ is reasonably well described by this model for particular values of the antiferromagnetic couplings [8]. Although extensively studied, the zero temperature phase diagram of the Shastry-Sutherland antiferromagnet remains elusive. While two of the phases are known, the possible existence of an intermediate phase and its nature are still unresolved issues. In addition, experiments on $SrCu_2(BO_3)_2$ in a magnetic field show the appearance of magnetization plateaus [9] with rather peculiar spin structures [10]. Several theoretical approaches based on the Shastry-Sutherland model have attempted to explain these steps [11–19]. The approaches used so far have ranged from exact diagonalization [10,11,19] to perturbative analysis [12–17] and mean field theory calculations [18].

The PEPSs or TPSs are higher dimensional generalizations of matrix product states (MPSs) [20,21] which are known to be particularly useful variational states in one dimension [22,23]. In contrast to the variational algorithm proposed in Ref. [6] the variational calculations using TPSs carried out in Refs. [1–5] build in translational invariance at the outset of the minimization procedure by using site-independent tensors or take into account the positional dependence only partially [24]. While this reduces the number of variational parameters it is often desirable not to assume this when dealing with a spin system where the *a priori*

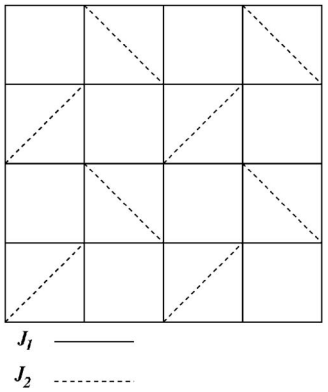

FIG. 1. Bond configuration for the Shastry-Sutherland model. All bonds have antiferromagnetic couplings. Vertical and horizontal bonds have coupling strength J_1 and diagonal bonds J_2.

unknown magnetic unit cell can be bigger than the unit cell of the lattice. There are also systems for which translational symmetry is explicitly broken by, for instance, impurities or boundaries. Thus it is desirable to have a method that is capable of treating also these situations. As a relevant example here, the NMR experiment on the 1/8 magnetization plateau in $SrCu_2(BO_3)_2$ showed that the results were best explained in terms of a state that breaks translational symmetry [10].

From the viewpoint of 1D variational calculations where the MPSs have internal matrix dimensions $D \sim 32-128$ [25], it would at first sight seem inadequate to restrict the 2D calculations to $D=2$. However, as argued in Ref. [26] even the $D=2$ class of states is very rich, a fact that is supported by our findings. We find that the $D=2$ PEPSs capture most of the known physics of the Shastry-Sutherland model such as the valence bond crystal phase, the Néel-ordered phase, and the magnetization steps.

The outline of this paper is as follows: in Sec. II we give a detailed outline of the variational method and in Sec. III we introduce the Shastry-Sutherland model and give a brief account of what is known about the ground state and the connection to the experimental results on $SrCu_2(BO_3)_2$. In Sec. IV we comment on the application of variational PEPSs to the Shastry-Sutherland model and in Sec. V and VI, we look at the ground state in absence and in presence of an external magnetic field respectively, examining the phase transition and magnetization plateaus. Finally, in Sec. VII we address the performance of the algorithm.

II. VARIATIONAL METHOD USING PEPS

Although the algorithm is described in Ref. [6] we reiterate it here in detail for completeness. As any variational algorithm, the aim is to minimize the expectation value of the Hamiltonian within a given class of trial states. The class of states used here are projected entangled pair states (PEPSs) represented by an array of complex tensors A_i, each tensor associated with a physical spin. To define a PEPS trial wave function an auxiliary lattice, the *computational lattice*, is introduced. While the sites on the computational lattice coincide with the sites on the physical lattice the bonds need not. However, it is important that the dimensionality of the computational lattice is the same as the dimensionality of the physical lattice. The bonds in the computational lattice determine the index structure of the tensors A_i. A tensor A_i will have one index for each bond in the computational lattice emanating out from site i. Note that different choices of the underlying computational lattice lead to different classes of variational PEPS-wave functions. As an example, consider Fig. 2 where a computational lattice in the form of a simple 3×3 square lattice with open boundary conditions is shown. With each lattice site i we then associate two tensors A_i^s, $s=\uparrow,\downarrow$ corresponding to spin up and spin down, respectively. Each tensor has a rank determined by the number of bonds in the computational lattice connecting the site and a dimension D. Hence on site 5 in the lattice in Fig. 2 we have two $(s_5=\uparrow,\downarrow)$ D-dimensional rank 4 tensors $[A_5^{s_5}]_{l,r}^{d,u}$ with indices for the bonds going down, up, left and right, whereas

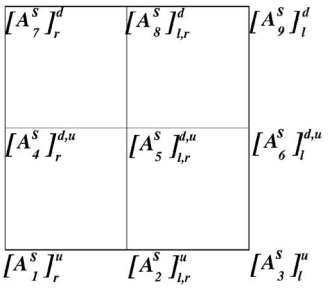

FIG. 2. Computational lattice and associated tensors. Shown here is an example of a 3×3 computational lattice. Two tensors A_i^s, $s=\uparrow,\downarrow$ are associated with each site i. Each tensor has indices u,d,l,r corresponding to bonds connecting the site i to neighboring sites.

on site 6 we have two rank 3 tensors $[A_6^{s_5}]_l^{d,u}$. While for PEPSs one associates the tensor indices with the bonds in the computational lattice one could alternatively associate them with plaquettes as in the interaction-round-face TPSs [4,27,28].

For a system with M sites we have the following form of the trial wave function:

$$|\Psi> = \sum_{s_1=\uparrow,\downarrow} \cdots \sum_{s_M=\uparrow,\downarrow} \mathrm{Tr}(A_1^{s_1}\cdots A_M^{s_M})|s_1\rangle\cdots|s_M\rangle.$$

The symbol $\mathrm{Tr}(\cdot)$ means here that one should trace over all indices (bonds) in the computational lattice. As an example, for the 3×3 lattice in Fig. 2 this operation becomes

$$\begin{aligned}\mathrm{Tr}(A_1^{s_1}A_2^{s_2}\cdots A_9^{s_9}) = {} & [A_1^{s_1}]_{r_1}^{u_1}[A_2^{s_2}]_{r_1,r_2}^{u_2}[A_3^{s_3}]_{r_2}^{u_3} \\ & \times [A_4^{s_4}]_{r_4}^{u_1u_4}[A_5^{s_5}]_{r_4,r_5}^{u_2,u_5}[A_6^{s_6}]_{r_5}^{u_3,u_6} \\ & \times [A_7^{s_7}]_{r_7}^{u_4}[A_8^{s_8}]_{r_7,r_8}^{u_5}[A_9^{s_9}]_{r_8}^{u_6}, \end{aligned} \quad (1)$$

where repeated indices should be summed over.

For $D=1$ the A_i^s's are complex scalars and the trial wave function is a simple product state ansatz similar to a mean field,

$$|\Psi_{D=1}\rangle = \prod_{i=1}^{M} \sum_{s_i=\uparrow,\downarrow} A_i^{s_i}|s_i\rangle.$$

For $D=2$ each index takes on two values. Although we will not make explicit use of it in the following, each $A_i^{s_i}$ can for $D=2$ be represented as a vertex with arrows, one for each index, each pointing either in or out. The contraction of all indices corresponds then to evaluating the partition function of a particular vertex model where $A_i^{s_i}$ represent the vertex weights [29].

To minimize the energy (or to even calculate it) we need to evaluate

$$\langle H\rangle = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}.$$

To see how this is done in practice we consider first the normalization $N=\langle\Psi|\Psi\rangle$ with our 3×3 example above which explicitly gives

$$\langle\Psi|\Psi\rangle = \sum_{\{s_i'\}}\sum_{\{s_i\}} \langle\{s'\}|([A_1^{s_1'}]_{r_1'}^{u_1'})^*([A_2^{s_2'}]_{r_1',r_2'}^{u_2'})^*([A_3^{s_3'}]_{r_2'}^{u_3'})^*$$
$$\times([A_4^{s_4'}]_{r_4'}^{u_1'u_4'})^*([A_5^{s_5'}]_{r_4',r_5'}^{u_2',u_5'})^*([A_6^{s_6'}]_{r_5'}^{u_3',u_6'})^*$$
$$\times([A_7^{s_7'}]_{r_7'}^{u_4'})^*([A_8^{s_8'}]_{r_7',r_8'}^{u_5'})^*([A_9^{s_9'}]_{r_8'}^{u_6'})^*$$
$$\times[A_1^{s_1}]_{r_1}^{u_1}[A_2^{s_2}]_{r_1,r_2}^{u_2}[A_3^{s_3}]_{r_2}^{u_3}$$
$$\times[A_4^{s_4}]_{r_4}^{u_1u_4}[A_5^{s_5}]_{r_4,r_5}^{u_2,u_5}[A_6^{s_6}]_{r_5}^{u_3,u_6}$$
$$\times[A_7^{s_7}]_{r_7}^{u_4}[A_8^{s_8}]_{r_7,r_8}^{u_5}[A_9^{s_9}]_{r_8}^{u_6}|\{s\}\rangle. \quad (2)$$

We now single out a specific site, say $k=5$, and construct the D^2 dimensional tensors E_i, $i\neq 5$,

$$E_i = \sum_s (A_i^s)^* \otimes (A_i^s).$$

Here the tensor product acts on all indices in the tensor, i.e., the tensors E_j have composite indices

$$[E_j]_{\tilde{l},\tilde{r}}^{\tilde{d},\tilde{u}} = [E_j]_{(l'l),(r'r)}^{(d'd),(u'u)}.$$

The normalization can now be written as

$$\langle\Psi|\Psi\rangle = \sum_{s_5}([A_5^{s_5}]_{r_4',r_5'}^{u_2',u_5'})^*[E_1]_{(r_1'r_1)}^{(u_1'u_1)}[E_2]_{(r_1'r_1),(r_2'r_2)}^{(u_1'u_2)}[E_3]_{(r_2'r_2)}^{(u_3'u_3)}$$
$$\times[E_4]_{(r_4'r_4)}^{(u_1'u_1),(u_4'u_4)}[E_6]_{(r_5'r_5)}^{(u_3'u_3),(u_6'u_6)}[E_7]_{(r_7'r_7)}^{(u_4'u_4)}[E_8]_{(r_7'r_7),(r_8'r_8)}^{(u_5'u_5)}$$
$$\times[E_9]_{(r_8'r_8)}^{(u_6'u_6)}[A_5^{s_5}]_{r_4,r_5}^{u_2,u_5}.$$

Contracting all indices except those connecting site $k=5$ we get

$$\langle\Psi|\Psi\rangle = \sum_{s_5,s_5'}([A_5^{s_5'}]_{r_4',r_5'}^{u_2',u_5'})^*\delta_{s_5',s_5}[N_5]_{(r_4'r_4),(r_5'r_5)}^{(u_2'u_2),(u_5'u_5)}[A_5^{s_5}]_{r_4,r_5}^{u_2,u_5}. \quad (3)$$

By treating the $2D^4$ components of A_5 as a $2D^4$ dimensional vector $\mathbf{A}_5$ and the $4D^8$ components of $\delta_{s_5',s_5}N_5$ as a $(2D^4)\times(2D^4)$ matrix $\mathcal{N}_5^{\text{eff}}$ Eq. (3) becomes

$$\langle\Psi|\Psi\rangle = \mathbf{A}_5^\dagger\mathcal{N}_5^{\text{eff}}\mathbf{A}_5.$$

The evaluation of $\langle\Psi|H|\Psi\rangle$ can be done in a similar fashion if we treat each term in the Hamiltonian individually. For a Hamiltonian H consisting of $\mathcal{M}$ terms we can always decompose H into individual terms $H^{(n)}$ as $H=\Sigma_{n=1}^{\mathcal{M}}H^{(n)}$, where each $H^{(n)}$ can be written as a product of on-site operators $H^{(n)}=\Pi_{j=1}^{M}\hat{O}_j^{(n)}$. For a Hamiltonian with only nearest neighbor interactions the majority of the onsite operators $\hat{O}_j^{(n)}$ will be the identity operators $\mathbf{1}_j$. Again we form D^2 dimensional tensors $E_j^{(n)}$ from the D dimensional tensors $A_i^{s_i}$ by

$$E_i^{(n)} = \sum_{s,s'}(A_i^{s'})^*\otimes(A_i^s)\langle s'|\hat{O}_i^{(n)}|s\rangle.$$

For each term $H^{(n)}$ we can now again write this in vector form if we single out a particular site k,

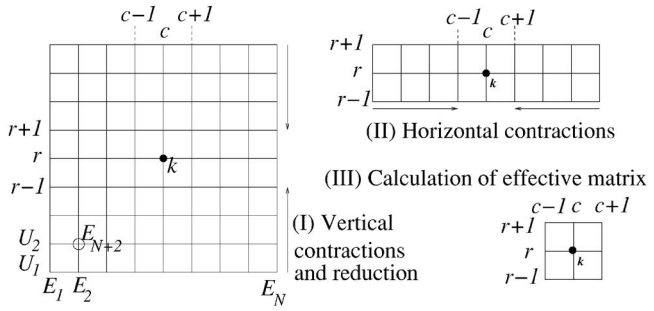

FIG. 3. Steps in obtaining contribution to the effective operator matrices $\mathcal{N}_k^{\text{eff}}$ and $\mathcal{H}_k^{\text{eff}}$ in Eq. (4). For a site k located at row r and column c we first contract row-wise from the top down to row $r+1$ and upwards from the bottom to row $r-1$ all the E tensors contributing to the operator. After each row contraction the left-right dimension of the tensors are reduced to a dimension D_f before the next row is contracted. When only rows r and $r\pm1$ remain we contract vertically to columns $c\pm1$ after which the effective matrices can be obtained.

$$\langle\Psi|H^{(n)}|\Psi\rangle = \mathbf{A}_k^\dagger\mathcal{H}_k^{(n)}\mathbf{A}_k,$$

and sum the matrices $\mathcal{H}_k^{(n)}$ to obtain an effective Hamiltonian matrix $\mathcal{H}_k^{\text{eff}}=\Sigma_n\mathcal{H}_k^{(n)}$ for site k.

The overall structure of the optimization algorithm is now the following. We pick a site k and calculate $\mathcal{N}_k^{\text{eff}}$ and $\mathcal{H}_k^{\text{eff}}$ by contracting all indices of the E tensors surrounding it. Then we solve the generalized eigenvalue problem

$$\mathcal{H}_k^{\text{eff}}\mathbf{A}_k = \lambda\mathcal{N}_k^{\text{eff}}\mathbf{A}_k \quad (4)$$

from which a new A_k^s with lower energy can be determined. While it is in principle possible to chose this new A_k^s to be the eigenvector corresponding to the smallest eigenvalue in Eq. (4) this occasionally leads to problems with convergence. Instead we only gradually project out the high energy eigenvectors from $\mathbf{A}_k$ in the optimization. We then continue in this vein sweeping over all sites $1\leq k\leq M$ until no further reduction in energy can be achieved.

While it is no problem to obtain $\mathcal{N}_k^{\text{eff}}$ and $\mathcal{H}_k^{\text{eff}}$ in the small 3×3 example above it becomes a problem as we move to larger systems. For a general set of tensors E_i the trace $\text{Tr}(\Pi_iE_i)$ is in its most general form, an NP-complete problem [26] and cannot be evaluated exactly for large systems but approximate strategies have to be used. We have employed the strategy suggested in Ref. [6] doing this in a row-wise fashion. To calculate either $\mathcal{N}_k^{\text{eff}}$ or one of the contributions $\mathcal{H}_k^{(n)}$ the three steps in Fig. 3 are performed. For a site k located at row r and column c we first contract vertically from the top down to row $r+1$ and upwards from the bottom to row $r-1$. When only rows r and $r\pm1$ remain we contract vertically to columns $c\pm1$ after which the effective matrices can be obtained.

When contracting vertically the left-right dimension of the E tensors will increase. For instance, contracting the D^2 dimensional tensors $[E_2]_{l,r}^u$ with $[E_{N+2}]_{l,r}^{d,u}$ in Fig. 3 generates a new tensor

$$[F_2]^u_{(l'l),(r'r)} = \sum_{x=1}^{D^2} [E_2]^x_{l',r'} [E_{N+2}]^{xu}_{l,r} \tag{5}$$

with D^4 dimensional left-right indices. This leads to an exponential growth of the left-right index dimensions with each row contraction. To handle this we use the approximation technique suggested for approximating an MPS with dimension D_i with another MPS with a lower dimension $D_f \leq D_i$ described in Ref. [6], which has also been successfully used to simulate time evolution in 1D systems [30]. Below we give an example of how this is done for the contraction of row 1 with row 2 in Fig. 3.

Any row in Fig. 3 can be viewed as a matrix product operator (MPO). For the bottom row this corresponds formally to a vector

$$U_1 = \sum_{\{u_i\}} E_1^{u_1} E_2^{u_2} \cdots E_N^{u_N} |u_1, \ldots, u_N\rangle$$

represented by a set of ND^2 $D^2 \times D^2$ matrices E_i^u. One of the middle rows, for instance the second one, can be formally viewed as a matrix

$$U_2 = \sum_{\{d_i\},\{u_i\}} E_{N+1}^{d_1,u_1} \cdots E_{N+N}^{d_N,u_N} |u_1, \ldots, u_N\rangle\langle d_1, \ldots, d_N|.$$

Contracting row 1 with row 2 is thus formally equivalent to a vector-matrix multiplication giving rise to a new ND^2 dimensional vector $U_{21} = U_2 U_1$ represented by ND^2 $D^4 \times D^4$ matrices F_i^u [cf. Eq. (5)]. We now seek a new vector

$$\tilde{U}_{21} = \sum_{\{u_i\}} \tilde{F}_1^{u_1} \cdots \tilde{F}_N^{u_N} |u_1, \ldots, u_N\rangle \tag{6}$$

represented by ND^2 $D_f \times D_f$ matrices ($D_f \leq D^4$) $\tilde{F}_i^{u_i}$ such that

$$\kappa = |U_{21} - \tilde{U}_{21}|^2$$

is minimal. One does this in an iterative way starting with an ansatz for the solution and then optimizes the matrices $\tilde{F}_i^{u_i}$ one by one until convergence is reached. In practice we do this by first forming the $D_f^2 \times D_f^2$ matrices

$$G_i = \sum_u (\tilde{F}_i^u)^* \otimes \tilde{F}_i^u,$$

the $D_f D^4 \times D_f D^4$ matrices

$$H_i = \sum_u (\tilde{F}_i^u)^* \otimes F_i^u,$$

and the $D^8 \times D^8$ matrices

$$J_i = \sum_u (F_i^u)^* \otimes F_i^u$$

in terms of which κ can be written

$$\kappa = \prod_i G_i - 2\,\mathrm{Re} \prod_i H_i + \prod_i J_i.$$

For a given site $1 \leq k \leq N$ along the row we can now get a linear equation for $\tilde{F}_k^u$ that will locally minimize κ. To see this we differentiate with respect to $([\tilde{F}_k]^u_{l',r'})^*$,

$$\frac{\partial \kappa}{\partial ([\tilde{F}_k]^u_{l',r'})^*} = \left[\prod_{i<k} G_i\right]_{l'l} [\tilde{F}_k]^u_{l,r} \left[\prod_{i>k} G_i\right]_{r'r} - \left[\prod_{i<k} H_i\right]_{l'l} [F_k]^u_{l,r} \left[\prod_{i>k} H_i\right]_{r'r} = 0,$$

and treat the left-right indices of F_k^u and $\tilde{F}_k^u$ as the indices of vectors $\mathbf{F}_k^u$ and $\tilde{\mathbf{F}}_k^u$ which leads to the system of equations

$$G\tilde{\mathbf{F}}_k^u = H\mathbf{F}_k^u.$$

Although the matrices J_i are not needed to actually do the minimization we still calculate them to keep control of the error. If the error, after the $\tilde{F}_k^u$'s have converged is too large we increase D_f to obtain a better approximation.

For an $N \times N$ system we need to calculate of the order of N^2 contributions $\mathcal{H}^{(n)}$ to the effective Hamiltonian $\mathcal{H}^{\mathrm{eff}}$. For each contribution the contractions and approximations of MPO's (cf. Fig. 3) need to be calculated. This is the most computationally costly part of the algorithm and to avoid unnecessary calculations we optimize the tensors A_i^s in the computational lattice row-wise and store all calculated MPOs which can be reused. This means that the memory needed for storage of MPOs scales as $N^4 D^2 D_f^2$.

Once the algorithm has converged it is straightforward to use the obtained state to calculate other expectation values. This is done in the same way as the calculation of the energy. The operator one wishes to evaluate is divided into a sum of terms, i.e., $O = \Sigma_{n=1}^{\mathcal{M}} O^{(n)}$, where each term can be written as a product of on-site operators, $O^{(n)} = \Pi_{j=1}^{M} \hat{O}_j^{(n)}$. One then chooses a site k in the lattice and calculates effective matrices $\mathcal{O}_k^{(n)}$ for each of the $\mathcal{M}$ terms from which an effective matrix $\mathcal{O}_k^{\mathrm{eff}} = \Sigma_{n=1}^{\mathcal{M}} \mathcal{O}_k^{(n)}$ is obtained. The desired expectation value is then evaluated as

$$\langle O \rangle = \frac{\mathbf{A}_k^\dagger \mathcal{O}_k^{\mathrm{eff}} \mathbf{A}_k}{\mathbf{A}_k^\dagger \mathcal{N}_k^{\mathrm{eff}} \mathbf{A}_k}.$$

Finally we would like to point out that in this implementation we make no use whatsoever of any symmetries of the Hamiltonian, either in algorithm or in the trial states. This means that our program can treat very general Hamiltonians with nonuniform ground states.

III. SHASTRY-SUTHERLAND MODEL

The Shastry-Sutherland model was originally introduced as an example of a model with an exact dimerized ground state [7]. The model is a frustrated spin-1/2 antiferromagnet with a bond configuration shown in Fig. 1,

$$H = J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + J_2 \sum_{\langle i,j \rangle'} \mathbf{S}_i \cdot \mathbf{S}_j. \tag{7}$$

Although the model was introduced for reasons of purely theoretical nature, interest was renewed along with experiments on $SrCu_2(BO_3)_2$ [9]. In $SrCu_2(BO_3)_2$ the crystal structure is layered with alternating planes of $CuBO_3$ and Sr and the magnetic properties stem from the $CuBO_3$ layers. It has

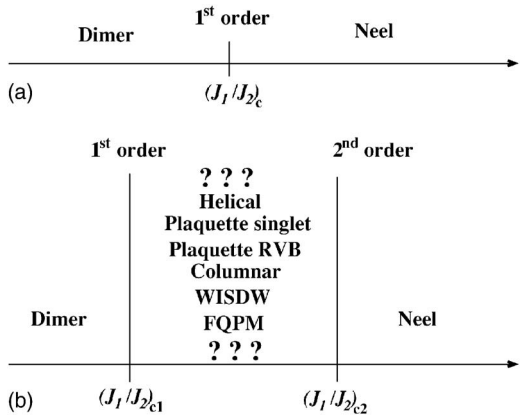

FIG. 4. Proposed phase diagrams for the Shastry-Sutherland model. Two possible scenarios for the phase diagram have been proposed. Either a direct transition between the dimer state and the Néel state (a) or a transition via an intermediate phase as in (b). The nature of this intermediate phase has not been established.

been argued that these layers are well modeled by the Shastry-Sutherland model [8].

In the limit $J_2 \gg J_1$ the ground state, which is separated from the excited states by a gap, is a dimer state with localized spin singlets on the diagonal bonds, the ground state energy per spin being $E_{\text{dimer}}=-3/8J_2$ per diagonal bond. In the other limit $J_1 \gg J_2$ the model reverts to the ordinary antiferromagnetic Heisenberg model with a Néel ordered ground state and gapless spectrum. From high temperature series expansion and exact diagonalization [11,31,32] a possible direct transition between the dimer phase to the Néel phase has been estimated to lie at $(J_1/J_2)_c=0.7\pm0.01$.

Other works point to the existence of an intermediate phase between the antiferromagnet and the dimer phase. A sketch of the phase diagram is shown in Fig. 4. Most estimates agree that below $(J_1/J_2)_{c1}>0.6$ the ground state is the dimer state and above $(J_1/J_2)_{c2}<0.9$ the ground state is the Néel state. The nature of this intermediate state has been addressed in several publications. In Ref. [33] Albrecht and Mila used Schwinger Boson mean field theory to argue in favor of a first order transition between the dimer state into a helical state and a second order transition to a Néel state. Another possible intermediate state, the plaquette singlet phase, was discussed by Koga and Kawakimi in Ref. [34]. Both plaquette states and Helical states were considered in Ref. [35]. Arguments against both plaquette and helical phases was put forward in Ref. [36] based on extensive series expansions around both the helical and plaquette phases and even columnar phases. This is in contrast to the findings in Ref. [37] that lend support to either a plaquette phase or a columnar phase. Other suggestions for the intermediate phase are weakly incommensurate spin-density waves (WISDWs) or fractionalized quantum paramagnet (FQPM) [38]. Finally a resonant valence bond plaquette phase was suggested as the intermediate state in Ref. [39]. Thus neither the existence of an intermediate phase nor its exact nature are presently known.

Experiments on $SrCu_2(BO_3)_2$ in strong external fields show the existence of magnetization plateaus [8,9]. While the ground state in absence of an external field is believed to be the dimer state, the magnetization steps were originally thought to be formed by strongly localized triplets forming periodic patterns which may spontaneously break the translational symmetry [12–17] (for an alternative explanation hypothesis see Ref. [18]). Subsequent NMR experiments [10] at the 1/8 plateau revealed a more complex structure, inconsistent with the simple triplet-singlet picture. By including coupling to phonon degrees of freedom, with the sole purpose of breaking the translational symmetry, exact diagonalization studies of small systems [10,19] revealed more complex spin textures. Again, at all steps (except the 1/2) translational symmetry is broken and larger unit cells are formed.

IV. APPLICATION OF VARIATIONAL PEPS TO THE SHASTRY-SUTHERLAND MODEL

Applying variational PEPSs to the Shastry-Sutherland model is straight forward. Two issues should be noted. First, we stress again that the computational lattice does not need to have the same bond configuration as the underlying Hamiltonian. For the purpose of studying the Shastry-Sutherland model it is sufficient to use an ordinary square lattice as depicted in Fig. 2. It is easy to show that already with a low tensor dimension $D=2$ it is possible to represent exactly the dimerized ground state with singlets on all diagonal bonds.

The second issue regards boundary conditions. In implementing the algorithm we have used a computational lattice with open boundary conditions. The reason for this is twofold. First, using a computational lattice with periodic boundary conditions severely reduces the performance of the algorithm, the difference being that between matrix-vector multiplications rather than matrix-matrix multiplications. Second, our program suffers from stability problems arising due to ill-conditioning and round-off errors in the case of computational lattices with periodic boundary conditions.

Although the computational lattice does not have periodic boundary conditions it is not necessary to adopt the same boundary conditions to the Hamiltonian. In this study we have used two different physical boundary conditions, open and periodic. For the open BC we have adopted the geometry shown in Fig. 5(a) while for the periodic the geometry in Fig. 5(b). The interpretation of using periodic boundary conditions (BCs) in the physical problem while using open BC in the computational lattice is reminiscent of using a self-consistent field on the boundary. The bonds across the boundary have only a tensor dimension $D_{\text{boundary}}=1$ which in the limit of large lattices implies that the contribution from such a bond will approach the product form $\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle \rightarrow \langle \mathbf{S}_i \rangle \cdot \langle \mathbf{S} \rangle_j$ for limited D.

We have mainly restricted ourselves to using a tensor dimension $D=2$ for which a usual work station with 1 GB of internal memory suffices. Although our program can in principle handle $D>2$ (see Sec. VII), it is in its present incarnation too slow and unstable for $D>2$. For the dimension of effective MPOs when calculating effective operators we have used a variable $16 \leq D_f \leq 24$ for all simulations except for

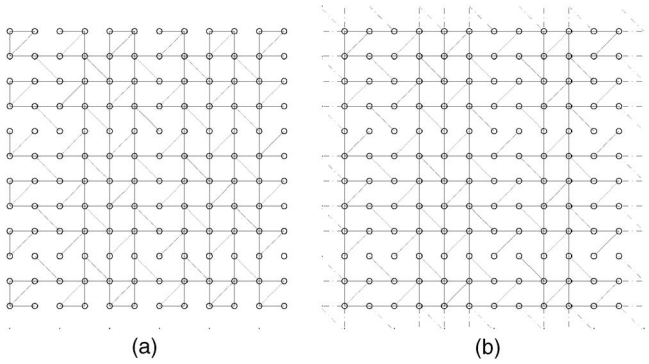

FIG. 5. Boundary conditions used in the simulations. (a) Open boundary condition. (b) Periodic boundary condition.

the largest (12×12) systems where the 1-GB memory limit restricts us to $16\leq D_f\leq 18$.

V. GROUND STATE IN ZERO FIELD

In Fig. 6 the lowest energies obtained by the algorithm for $D=2$ are shown for system sizes of 6×6, 8×8, 10×10, and 12×12 with both open (main panel) and periodic boundary conditions (inset). The total energy has been scaled by the number of internal diagonal bonds N_s and the coupling energy J_2. As can be seen, for low J_1/J_2 the ground state energy approaches $-3/8J_2$ for the system with open boundary conditions confirming the convergence to the dimerized ground state. From the graph obtained using open boundary conditions it can be seen that already for $D=2$ we find a phase transition from the dimerized state, the transition point being located at 0.69 ± 0.02 which is in agreement with estimates for the transition point of the direct dimer-Néel transition. Further, in the energy a clear finite size effect is seen as the transition point is approached.

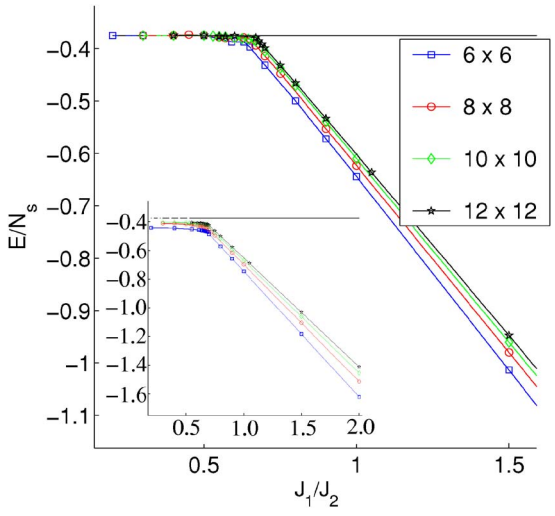

FIG. 6. (Color online) Variational minimum energy as a function of J_1/J_2 for system sizes 6×6 (blue squares), 8×8 (red circles), 10×10 (green diamonds), and 12×12 (black stars). For open boundary conditions (main figure) the transition from the dimer state to the Néel state is clearly visible in the energy which has been scaled to the number of diagonal bonds. The inset shows energies for periodic boundary conditions using the same energy scaling.

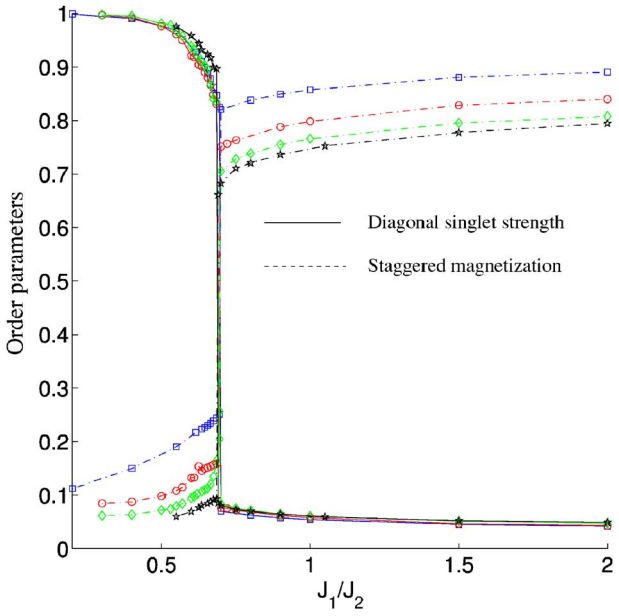

FIG. 7. (Color online) Order parameters, singlet mixing on diagonal bonds (solid lines), and staggered magnetization (dashed lines) for periodic boundary conditions for system sizes 6×6 (blue squares), 8×8 (red circles), 10×10 (green diamonds), and 12×12 (black stars). Results obtained with internal tensor dimension $D=2$ suggest a direct first order transition from the dimer state to the Néel state.

From the inset showing the energies obtained using periodic boundary conditions the location and nature of the transition is less clear. Here we have again divided the total energy by the number of internal diagonal bonds. Since the diagonal bonds across the boundary are not counted this gives rise to energies lower than $-3/8J_2$. The transition is revealed by looking at the strength of the diagonal singlets and the staggered magnetization, as shown in Fig. 7. The singlet mixing is calculated by projecting the diagonal bonds on to the singlet state, i.e., 1 indicates a singlet while 0 indicates a triplet. The staggered magnetization displayed is calculated as $2<[\Sigma_i \mathbf{S}_i(-1)^i]^2>^{1/2}$. Note that for large J_1/J_2 the staggered magnetization is higher than the value expected for a Heisenberg antiferromagnet. The reason for this can be twofold. First, finite size effects explain a part of the discrepancy as can be seen from the graph. Second, while $D=2$ gives a good value for the energies involved, being only a few percent off the exact values, observables may differ by more (see Sec. VII).

As stated in Sec. III, there are good reasons to believe that an intermediate phase exists between the dimer phase and the Néel phase. However, it is hardly surprising that we do not see this intermediate phase in our $D=2$ variational calculation. First of all, with tensor dimension $D=2$ we typically overshoot the true ground state energy by a few percent, thus higher D is likely needed to capture any additional phase that may differ by a percent or less in energy. Second, the influence of boundary conditions scales as $1/N$ which implies that boundary effects can have a big impact even for the

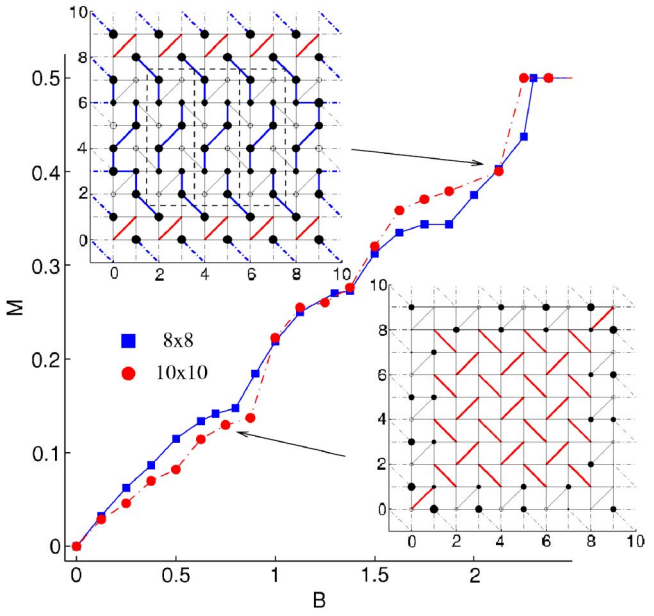

FIG. 8. (Color online) Magnetization curve at $J_1/J_2=0.6$ for system sizes 8×8 (blue squares) and 10×10 (red circles). A clear steplike structure corresponding to fillings 1/4, 1/3, and 1/2 is seen. The insets show the distribution of singlets (red lines) and triplets (blue lines). The S_z component of the spin on each site is illustrated by a circle with radius proportional to $|\langle S_z\rangle|$. Filled circles are aligned with the field while open circles represent spins pointing opposite to the field.

largest systems (12×12) in cases where the energy splittings between ground state candidates are small.

VI. MAGNETIZATION PLATEAUS

To study the magnetization curve we have deliberately chosen a somewhat smaller coupling constant $J_1/J_2=0.6$ rather than the experimental value $(J_1/J_2)_{SrCu_2(BO_3)_2}=0.635$. This makes the dimer state more stable and the algorithm converges faster but should not significantly affect the physics. In Fig. 8 we show the results for finite magnetic fields for systems with periodic boundary conditions of sizes 8×8 and 10×10 ($D=2$). Although our square geometry and (periodic) boundary conditions are inconsistent with the unit cells proposed in Refs. [10] and [19]a clear steplike structure is nevertheless visible in Fig. 8. A closer inspection reveals that *locally* the spin configurations we have obtained match those in Ref. [19] very well.

To study the spin configurations at the plateaus we have visualized the wave functions by coloring the bonds according to the amount of triplet or singlet mixing (see Fig. 8). Localized singlets are drawn in red whereas triplets are blue. To determine the color of a bond we use the following criterion:

$$\begin{cases} \text{red} & \langle\Psi|\mathbf{S}_i\cdot\mathbf{S}_j|\Psi\rangle < 0.1\epsilon_{\text{triplet}}+0.9\epsilon_{\text{singlet}} \\ \text{blue} & \langle\Psi|\mathbf{S}_i\cdot\mathbf{S}_j|\Psi\rangle > 0.9\epsilon_{\text{triplet}}+0.1\epsilon_{\text{singlet}}. \\ \text{black} & \text{otherwise} \end{cases} \quad (8)$$

Furthermore, we have measured the spin component parallel to the direction of the applied field ($\mathbf{B}=B\hat{\mathbf{z}}$) and visualized $\langle S_z\rangle$ by circles with radii proportional to $|\langle S_z\rangle|$. Spins aligned (antialigned) with the field are drawn as filled (open) circles.

For small fields, $B<1$, we see a finite magnetization where one would expect a spin gap. This is due to the inability of our trial states to form singlets across the boundary. As can be seen in the inset, for $B<1$ the interior of the system is still in the dimer phase while only spins on the boundary have aligned with the field. Thus by looking at when the magnetization in the interior of the system becomes finite we estimate the spin gap to be roughly $B=1$ corresponding to 33 T where we have used coupling constants $J=85$ K, $g=2.28$ (with $J=71$ K the corresponding number is 28 T).

For $B>1$ three steps can be distinguished, 1/4, 1/3, and 1/2. While the spin texture at 1/2 matches that of earlier predictions, half of the diagonal bonds being triplets while the other half being singlets, the spin textures for other points are more elaborate and only agree with earlier predictions locally. An example is shown in the top left inset of Fig. 8. Here, three of the 1/3 unit cells obtained by Miyahara *et al.* [19] are reproduced in the interior of the 10×10 system. Note that the total magnetization is larger than 1/3 at this point due to the spin configurations on the boundary.

The spin textures obtained at the steps can only be found when translational symmetry is broken. In the variational method employed here translational symmetry is broken partly because of our choice of computational lattice and boundary conditions, and partly because a $D=2$ PEPS cannot represent the coherent superposition of degenerate plateau states connected by global symmetry transformations.

VII. ALGORITHM PERFORMANCE

So far we have only concerned ourselves with $D=2$ which is the first step beyond a simple on-site factorizable wave function. Restricting ourselves to $D=2$ and sizes up to 12×12 allows the program to run on an ordinary workstation with 1-GB internal memory without using any swapping to disk.

Because the method is based on a sequence of approximations, i.e., for a 12×12 system there are over 700 contributions to the effective Hamiltonian on any given site, each contribution being obtained in a series of up to ten consecutive approximations, one has to ask whether or not the precision is compromised. Another important factor to consider is how much more accuracy (how much closer to the true ground state energy we can come) can be obtained by increasing D, and how the computational effort scales with increasing D.

In Fig. 9 the energy as a function of number of optimizations is shown for a 10×10 system with $J_1=J_2=1$. While smooth on a large scale, the errors accumulated in the successive approximations are clearly visible in the inset which shows a closeup of the final convergence. In this simulation and others we have used a final dimension in the approximation of E tensors (see Sec. II) $16\leq D_f\leq24$ [for 12×12 we have been restricted to $16\leq D_f\leq18$ due to the limited memory (1 GB) of the workstation]. As can be seen, despite the heavy reduction of the state space in the calculations, we have a precision of the order of four digits. This can also be

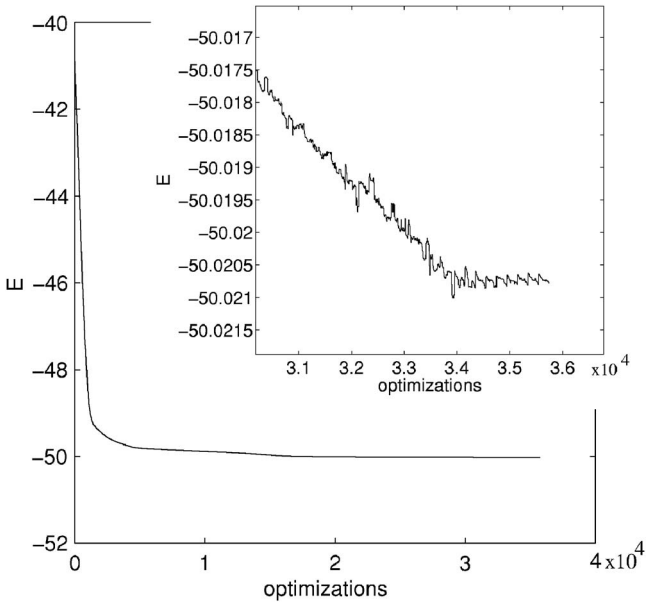

FIG. 9. Convergence of energy in a 10×10 system with periodic boundary conditions for $D=2$. The energy is shown as a function of the number of diagonalizations of the generalized eigenvalue problem in Eq. (4). The inset shows a closeup of the final part of the optimization where fluctuations due to the approximation strategy used to calculate expectation values are visible.

seen by looking at the norm of the wave function. In Fig. 10 the deviation of the norm from the nominal value 1 is shown for the simulation in Fig. 9.

To estimate the accuracy, i.e., how close to the true ground state energy the variational PEPS algorithm can get we have compared it to QMC on an ordinary Heisenberg antiferromagnet (QMC being unable to handle the Shastry-Sutherland (SS) model), $H=J\Sigma_{\langle i,j\rangle}\mathbf{S}_i\cdot\mathbf{S}_j$ with open boundary conditions. The QMC was run at an inverse temperature of $\beta=256/J$ and the results are shown in Fig. 11 where the quantity $1-E_{\mathrm{PEPS}}/E_{\mathrm{QMC}}$ is shown for different system sizes

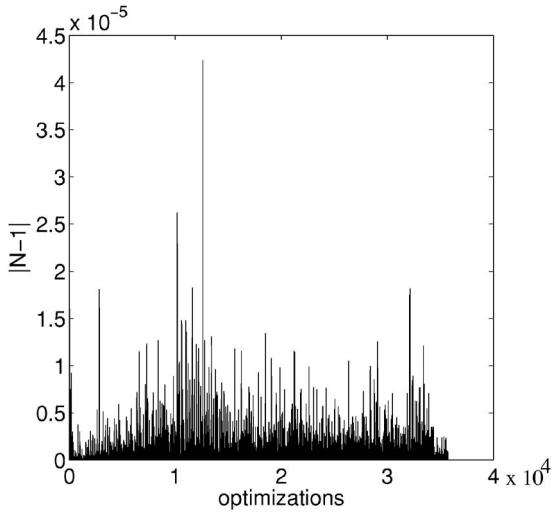

FIG. 10. Deviation of the norm from unity during the optimization of a 10×10 system with $J_1=J_2$.

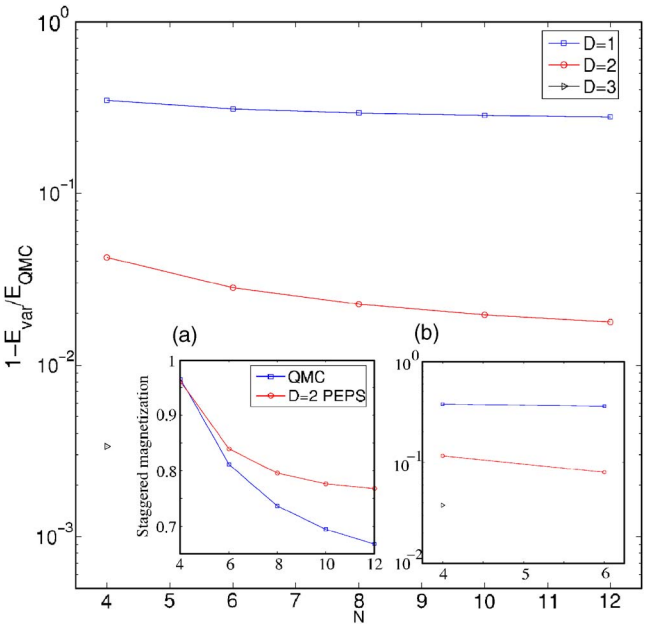

FIG. 11. (Color online) Comparison between ground state energies obtained by variational PEPS and quantum Monte Carlo for a Heisenberg antiferromagnet with open boundary conditions for system sizes 4×4 to 12×12 and tensor dimensions ranging from $D=1$ to $D=3$. Inset (a) shows a comparison between QMC and $D=2$ PEPS for the staggered magnetization for the Heisenberg model. Inset (b) shows a comparison between exact diagonalization of the Shastry-Sutherland model with periodic boundary conditions and PEPS $D=1-D=3$.

and tensor dimensions D. For the smallest system 4×4 sites we find good agreement with the figures reported in Ref. [6] obtained using imaginary time evolution. We further note that as the system size is increased the relative error decreases slightly, and that a linear increase in D seems to give an exponential increase in accuracy.

We want to point out that although the energies obtained by PEPS are in good agreement with exact results, observables may deviate more. In the left inset of Fig. 11 the staggered magnetization for the Heisenberg model is compared with the staggered magnetization obtained using $D=2$ PEPS. Compared to the accuracy in energy which is of the order 1% (see main panel) the error in the staggered magnetization is an order of magnitude larger.

Although the above comparison for the Heisenberg model does not, in a strict sense, tell us anything about the accuracy obtained for the SS model away from the limit $J_2/J_1\gg1$ it still serves as a good indication on the general behavior as one varies N and D. For the SS model we have compared with exact diagonalization results obtained using SPINPACK [40,41] at $J_1=J_2$ using periodic boundary conditions. The comparison is shown in the right inset of Fig. 11.

The internal tensor dimension D plays an important role in how faithfully a PEPS can represent ground states. The algorithm scales very badly with increasing D, the bottleneck being the contraction of two rows [see Eq. (6)]. To form an MPO by contracting two rows requires of the order $ND^6D_f^2$ ($D_f>D^2$) operations and scaling with D is at best D^{10}. This then sets a limit to the maximum internal dimensions that

can be practically used and the large values of $D \sim 10^2$ used in 1D variational MPSs cannot be reached. However, as we have seen here (see also Ref. [26]), for 2D PEPSs we expect small $D \sim 2-5$ to be able capture the essential physics for many problems with short range interactions. The scaling of the algorithm with linear system size is N^4 and is less severe.

All simulations were run on Linux workstations with a 2.0 GHz AMD Athlon processor and 1 GB of internal memory. On such a machine the graph in Fig. 9 took 96 h to produce.

Finally we comment on the stability of the algorithm. We have found that the technique used to optimize tensors one by one often becomes unstable and may not be the optimal way to find the ground state. It may well be that using imaginary time evolution is a more effective way.

VIII. CONCLUSIONS

We have applied a variational procedure based on projected entangled pair states (PEPSs) to study the ground state properties of a frustrated spin system, the Shastry-Sutherland model. Using the smallest nontrivial dimension on the tensors $D=2$ a direct phase transition between the dimer state and the Néel state can be observed, the location being well in agreement with other theoretical estimates for a direct transition. Within $D=2$ we see no clear indication of an intermediate phase which may require higher D, larger systems, or proper handling of periodic boundary conditions. We also find that already with PEPS $D=2$, magnetization plateaus are possible to reproduce, and that the nontrivial spin textures associated with these plateaus can be seen.

Furthermore, we have examined the performance of the algorithm and conclude that it degrades rather severely for intermediate to large values of the internal dimension D. However, this scaling of the performance degradation might not be so restrictive as already the $D=2$ class of PEPSs is well suited for studies of frustrated spin systems at a level beyond mean field theory. The method can readily be extended to other 2D frustrated spin models.

[1] T. Nishino, K. Okunishi, Y. Hieida, N. Maeshima, and Y. Akutsu, Nucl. Phys. B **575**, 504 (2000).
[2] T. Nishino, Y. Hieida, K. Okunishi, N. Maeshima, Y. Akutsu, and A. Gendiar, Prog. Theor. Phys. **105**, 409 (2001).
[3] N. Maeshima, Y. Hieida, Y. Akutsu, T. Nishino, and K. Okunishi, Phys. Rev. E **64**, 016705 (2001).
[4] A. Gendiar, N. Maeshima, and T. Nishino, Prog. Theor. Phys. **110**, 691 (2003).
[5] N. Maeshima, J. Phys. Soc. Jpn. **73**, 60 (2004).
[6] F. Verstraete and J. I. Cirac, e-print cond-mat/0407066.
[7] B. S. Shastry and B. Sutherland, Physica B & C **108**, 1069 (1981).
[8] S. Miyahara and K. Ueda, J. Phys.: Condens. Matter **15**, R327 (2003).
[9] H. Kageyama *et al.*, Phys. Rev. Lett. **82**, 3168 (1999).
[10] K. Kodama *et al.*, Science **298**, 395 (2002).
[11] S. Miyahara and K. Ueda, Phys. Rev. Lett. **82**, 3701 (1999).
[12] T. Momoi and K. Totsuka, Phys. Rev. B **61**, 3231 (2000).
[13] T. Momoi and K. Totsuka, Phys. Rev. B **62**, 15067 (2000).
[14] S. Miyahara and K. Ueda, Phys. Rev. B **61**, 3417 (2000).
[15] S. Miyahara and K. Ueda, Physica B **281/282**, 661 (2000).
[16] Y. Fukumoto and A. Oguchi, J. Phys. Soc. Jpn. **69**, 1286 (2000).
[17] Y. Fukumoto, J. Phys. Soc. Jpn. **70**, 1397 (2001).
[18] G. Misguich, Th. Jolicoeur, and S. M. Girvin, Phys. Rev. Lett. **87**, 097203 (2001).
[19] S. Miyahara, F. Becca, and F. Mila, Phys. Rev. B **68**, 024401 (2003).
[20] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Phys. Rev. Lett. **59**, 799 (1987).
[21] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).
[22] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).
[23] S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997).
[24] A. Gendiar and T. Nishino, Phys. Rev. B **71**, 024404 (2005).
[25] F. Verstraete, A. Weichselbaum, U. Schollwöck, J. I. Cirac, and J. von Delft, e-print cond-mat/0504305.
[26] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006).
[27] G. Sierra and M. A. Martin-Delgado, *Proceedings of the Workshop on the Exact Renormalization Group, Faro (Portugal)*, 10-12 September 1998 (World Scientific, Singapore, 1998).
[28] A. Gendiar, T. Nishino, and R. Derian, Acta Phys. Slov. **55**, 141 (2005).
[29] H. Niggemann, A. Klümper, and J. Zittatz, Z. Phys. B: Condens. Matter **104**, 103 (1997).
[30] J. J. Garcia-Ripoll, e-print cond-mat/0602305.
[31] Z. Weihong, C. J. Hamer, and J. Oitmaa, Phys. Rev. B **60**, 6608 (1999).
[32] E. Müller-Hartmann, R. R. P. Singh, C. Knetter, and G. S. Uhrig, Phys. Rev. Lett. **84**, 1808 (2000).
[33] M. Albrecht and F. Mila, Europhys. Lett. **34**, 145 (1996).
[34] A. Koga and N. Kawakami, Phys. Rev. Lett. **84**, 4461 (2000).
[35] C. H. Chung, J. B. Marston, and S. Sachdev, Phys. Rev. B **64**, 134407 (2001).
[36] W. Zheng, J. Oitmaa, and C. J. Hamer, Phys. Rev. B **65**, 014408 (2002).
[37] M. Al. Hajj and J.-P. Malrieu, Phys. Rev. B **72**, 094436 (2005).
[38] D. Carpentier and L. Balents, Phys. Rev. B **65**, 024427 (2001).
[39] A. Lauchli, S. Wessel, and M. Sigrist, Phys. Rev. B **66**, 014401 (2002).
[40] J. Schulenburg, SPINPACK 2.24, http://wase.urz.uni-magdeburg.de/jschulen/spin/index.html
[41] J. Richter, J. Schulenburg, and A. Honecker, in *Quantum Magnetism*, Lecture Notes in Physics, No. 645, edited by U. Schollwöck, J. Richter, D. J. J. Farell, R. F. Bishop (Springer, Berlin, 2004), Chap. 2, p. 85.