Variational optimization with infinite projected entangled-pair states

Corboz, P.

DOI
10.1103/PhysRevB.94.035133
Publication date
2016
Document Version
Final published version
Published in
Physical Review B

PHYSICAL REVIEW B **94**, 035133 (2016)

Variational optimization with infinite projected entangled-pair states

Philippe Corboz
Institute for Theoretical Physics, University of Amsterdam, Science Park 904 Postbus 94485, 1090 GL Amsterdam, Netherlands
(Received 11 May 2016; published 14 July 2016)

We present a scheme to perform an iterative variational optimization with infinite projected entangled-pair states, a tensor network ansatz for a two-dimensional wave function in the thermodynamic limit, to compute the ground state of a local Hamiltonian. The method is based on a systematic summation of Hamiltonian contributions using the corner-transfer-matrix method. Benchmark results for challenging problems are presented, including the two-dimensional Heisenberg model, the Shastry-Sutherland model, and the t-J model, which show that the variational scheme yields considerably more accurate results than the previously best imaginary-time evolution algorithm, with a similar computational cost and with a faster convergence towards the ground state.

DOI: 10.1103/PhysRevB.94.035133

I. INTRODUCTION

Understanding the collective phenomena in strongly correlated quantum many-body systems is one of the grand challenges in modern physics. For one-dimensional problems tremendous progress has been made thanks to the well-known density matrix renormalization group (DMRG) method [1]. DMRG has an underlying variational ansatz, called matrix product states (MPS), in which the wave function is efficiently represented by a product of matrices [2,3]. The accuracy of the ansatz can be systematically controlled by the bond dimension D, corresponding to the linear size of the matrices. For two-dimensional systems, however, DMRG suffers from an exponential scaling of the computational cost with the linear system size [4].

Progress in quantum information theory, in particular a better understanding of entanglement in quantum many-body systems, has led to the generalization of MPS to higher dimensions, called projected entangled-pair states [2,5,6] or tensor product states [7–9]. These states reproduce an area law scaling of the entanglement entropy which typical ground states of local Hamiltonians fulfill [10] and thus provide an efficient variational ansatz overcoming the exponential scaling of the two-dimensional (2D) DMRG with system size. A particularly powerful ansatz is infinite projected entangled-pair states [11] (iPEPS), which represents a wave function directly in the thermodynamic limit and thereby minimizes finite-size and boundary effects [12].

In recent years 2D tensor network methods have become more and more powerful tools for the study of 2D fermionic and frustrated systems (see, e.g., Refs. [13–26] and references therein). For example, it has been demonstrated that iPEPS provides better variational energies than state-of-the-art variational methods for the t-J and the 2D Hubbard models [21,25]. Thanks to largely unbiased simulations iPEPS also played the key role in understanding the magnetization process in $SrCu_2(BO_3)_2$ [22], which has been an intriguing open problem for many years. Thus, already today, 2D tensor networks are very powerful; however, it is important to realize that the algorithms can still be further improved, such that even higher accuracies can be reached.

The biggest challenge in a 2D tensor network algorithm is the *optimization* of the tensors, i.e., finding the best variational parameters stored in the tensors to have the best approximation of the ground state of a given input Hamiltonian. For iPEPS this is commonly done by performing an imaginary-time evolution (ITE) [11,27–29], where a random initial iPEPS gets projected onto the ground state. The ITE cannot be done in an exact way and requires truncating the bond indices of the iPEPS at each time step. Besides the truncation error there is an additional error coming from the Trotter-Suzuki decomposition of the ITE operator. In order to reduce the error one has to use very small time steps τ, which is computationally not efficient because many steps are required to reach convergence.

In this paper we present an alternative optimization method for iPEPS based on a variational energy minimization. As in the DMRG, the idea is to perform sweeps over all tensors in the ansatz, and at each step one minimizes the energy with respect to a tensor while keeping the other tensors fixed. We show that this optimization method not only converges faster to the ground state but, surprisingly, also yields considerably more accurate results than the best ITE algorithm (the so-called full update), even when taking the limit $\tau \to 0$.

A related scheme was already applied to finite PEPS previously [2,5,6]; however, for iPEPS there are two additional complications: (1) in order to perform the energy minimization one needs to take into account an infinite number of Hamiltonian contributions. In this paper we show how a systematic summation of all these contributions can be achieved using the corner-transfer-matrix (CTM) method [30–32]. [Alternatively, the summation can also be done based on "channel environments" [33] or by representing the Hamiltonian as a projected entangled-pair operator [34], which is closely related to variational tensor network schemes which have been developed for three-dimensional (3D) classical systems [7,35].] (2) Because in iPEPS each tensor appears an infinite number of times in the ansatz (instead of only once as in finite PEPS), the energy optimization for a given tensor is a highly nonlinear problem. We present a practical scheme to deal with this issue, which provides a good convergence to the ground state.

This paper is organized as follows: in the next section we provide a short introduction to the iPEPS ansatz, the CTM method, and the optimization based on ITE. In Sec. III we first present the main idea of the iterative variational optimization, then explain how to use the CTM method for the systematic summation of Hamiltonian contributions, and finally discuss practical schemes for the optimization. In Sec. IV we present

benchmark results for the 2D Heisenberg model, the Shastry-Sutherland model, and the t-J model to demonstrate the performance of the variational approach compared to results based on ITE. Finally, we discuss and summarize our findings in Sec. V. In addition, in the Appendix we explain how to implement a two-site variational optimization which can be used complementary to the one-site update discussed in the main text.

II. INTRODUCTION TO iPEPS

A. iPEPS ansatz

An iPEPS is an efficient variational tensor network ansatz for 2D ground states of local Hamiltonians in the thermodynamic limit [5,7,8,11,36] which obey an area law of the entanglement entropy [10]. It consists of a rectangular unit cell of tensors with one tensor per lattice site $A^{[x,y]}$, where $[x,y]$ label the coordinates of a tensor relative to the unit cell of size $L_x \times L_y = N_T$, shown in Fig. 1(a). Each tensor has one physical index carrying the local Hilbert space of a lattice site and four auxiliary indices which connect to the nearest-neighbor tensors on a square lattice (more generally, a PEPS has z auxiliary indices, where z is the coordination number of the lattice). The accuracy of the ansatz can be systematically controlled by the bond dimension D of the auxiliary indices.

For a translationally invariant state an ansatz with a single-tensor unit cell can be chosen. However, if translational symmetry is spontaneously broken, a larger unit-cell size compatible with the periodicity of the ground state is required (for example, for an antiferromagnetic state two different tensors for the two sublattices are needed). Since the periodicity of the ground state is typically not known in advance, one has to perform simulations with different unit cell sizes to determine which cell size leads to the lowest variational energy. Using different unit cells also offers the possibility to find different competing low-energy states (see, e.g., Ref. [21]).

B. Contraction of an iPEPS

In order to compute an expectation value of an observable $\hat{O}$ with respect to an iPEPS wave function $|\Psi\rangle$, the corresponding 2D tensor network representing $\langle\Psi|\hat{O}|\Psi\rangle$ has to be contracted in a controlled, approximate way. In this work we use a variant of the CTM renormalization group method [30–32] for arbitrary unit-cell sizes [21,36], which is summarized in the following.

Consider the problem of computing the norm of an iPEPS $\langle\Psi|\Psi\rangle$, which boils down to contracting the infinite 2D square-lattice network of the reduced tensors $a^{[x,y]}$, shown in Fig. 1(c), where each $a^{[x,y]}$ is obtained from contracting $A^{[x,y]}$ with its conjugate tensor $A^{\dagger[x,y]}$ [see Fig. 1(b)]. The goal of the CTM approach is to compute the four corner tensors C_1, C_2, C_3, C_4 and the four edge tensors T_1, T_2, T_3, T_4 for each coordinate $[x,y]$ in the unit cell, where each corner tensor represents a quadrant and the edge tensors represent a half row (or half column) of the infinite 2D network. All these tensors together form the so-called environment, representing the infinite system surrounding a bulk site (or several bulk sites), as shown in Fig. 1(c). Once the environment has been

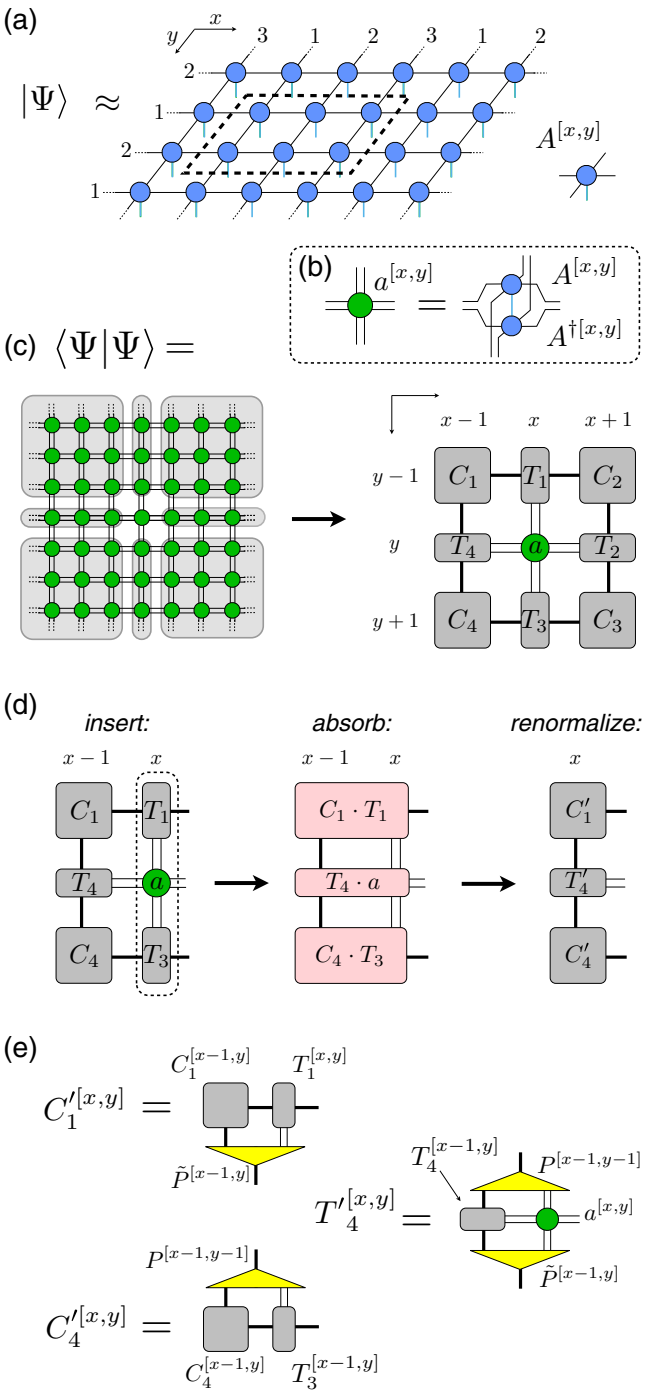

FIG. 1. (a) iPEPS ansatz with a 3×2 unit cell of tensors which is periodically repeated in the lattice. (b) The reduced tensor $a^{[x,y]}$ is obtained from contracting an iPEPS tensor $A^{[x,y]}$ with its conjugate $A^{\dagger[x,y]}$ along the physical leg. (c) The norm $\langle\Psi|\Psi\rangle$ is represented as an infinite square-lattice network of reduced tensors. The CTM approach yields the *environment* tensors surrounding a bulk tensor $a^{[x,y]}$ where the corner tensors C_1, C_2, C_3, C_4 take into account a quarter-infinite system and the edge tensors T_1, T_2, T_3, T_4 take into account an infinite half row or half column of the system. (d) A left move is done by inserting a new column of tensors, multiplying the tensors to the left, and performing a renormalization step (this is done for all coordinates y). (e) Diagrams to compute the updated corner and edge tensors C'_1, C'_4, T'_4 at coordinate x (for all y coordinates). Note that the coordinates are always taken modulo the unit cell size.

computed, one can easily evaluate expectation values of local observables by introducing the corresponding operators in between the physical legs of the iPEPS tensors.

The environment tensors are computed iteratively by letting the system grow in all directions. One starts from an initial guess for the boundary tensors, either by initializing them randomly or, alternatively, by initializing them with the bulk tensors (by tracing out the auxiliary bonds on the edges). In the directional CTM approach [32] one first performs a growth step on, e.g., the left side of the system (called a left move) by introducing a new column of tensors, multiplying them onto the left boundary tensors, followed by a renormalization step [see Fig. 1(d)].

In the renormalization step a bond dimension χ is kept at the boundary which controls the accuracy of the approximate contraction. There are different ways to perform this renormalization step. Here we use a set of projectors P and $\tilde{P}$, introduced in Refs. [37,38] and first applied in the CTM method in Ref. [21], to project from the enlarged space χD^2 down to a dimension χ. These projectors are then used to compute the renormalized corner and edge tensors, C_1', C_4', and T_4', as shown in Fig. 1(e).

For a unit cell of size $L_x \times L_y$ one proceeds in the following way for a full left move (i.e., an absorption of the entire unit cell into the left boundary):

(1) Do for all $x \in [1, L_x]$.
  (a) Do for all $y \in [1, L_y]$.
    i. Compute the projectors $P^{[x-1,y]}$ and $\tilde{P}^{[x-1,y]}$ (see Ref. [21] for details).
  (b) Do for all $y \in [1, L_y]$.
    i. Compute the new renormalized corner tensors $C_1'^{[x,y]}$, $C_4'^{[x,y]}$ and edge tensor $T_4'^{[x,y]}$, as shown in Fig. 1(e).

After a full left move one proceeds with full right, top, and bottom moves in a similar way and reiterates until convergence is reached (e.g., by checking the convergence of the energy with CTM iterations).

C. Optimization based on imaginary-time evolution

In order to get an approximate representation of the ground state of a given Hamiltonian $\hat{H}$, the tensors need to be *optimized*; that is, one needs to find the best variational parameters stored in the tensors. In previous iPEPS simulations this has been done based on an ITE of an initial (e.g., random) state. Using a Trotter-Suzuki decomposition, the imaginary-time-evolution operator is split into a product of two-site operators,

$$e^{-\beta \hat{H}} = e^{-\beta \sum_b \hat{H}_b} \approx \left(\prod_b \hat{U}_b \right)^n, \quad \hat{U}_b = e^{-\tau \hat{H}_b}, \qquad (1)$$

where the product goes over all nearest-neighbor bonds b in the unit cell (assuming a Hamiltonian with only nearest-neighbor terms), $\hat{H}_b$ is the Hamiltonian term on bond b, and $\tau = \beta/n$ is a small imaginary time step. The error of the Trotter-Suzuki decomposition decreases with the size of the time step τ [39]. The ITE is then performed by sequentially multiplying the two-site operators $\hat{U}_b$ onto the iPEPS and representing the resulting wave function again as an iPEPS with the same bond dimension until convergence is reached. There exist different schemes to truncate a bond. In the so-called simple-update scheme the truncation is done based on a local singular-value decomposition (SVD) [27,28,40], whereas in the full update [11,28] (or fast full update [29]) the entire 2D wave function is taken into account for the truncation of a bond index. The simple update is computationally cheaper, but less accurate, than the full update.

III. VARIATIONAL OPTIMIZATION

A. Basic idea

Iterative variational optimization schemes are commonly used for MPS and finite PEPS [1–3,5,6] but so far have not been used for iPEPS (except for 3D classical systems [7,35]). The main idea is to iteratively optimize one tensor after the other until convergence is reached. Optimizing a single tensor A (while keeping all other tensors fixed) boils down to minimizing the energy with respect to tensor A,

$$\min_A E(A) = \min_A \frac{\langle \Psi(A)|\hat{H}|\Psi(A)\rangle}{\langle \Psi(A)|\Psi(A)\rangle} = \min_{\vec{A}} \frac{\vec{A}^\dagger \mathbf{H} \vec{A}}{\vec{A}^\dagger \mathbf{N} \vec{A}}, \qquad (2)$$

where tensor A and its conjugate have been reshaped into vectors. The matrices $\mathbf{N}$ and $\mathbf{H}$ correspond to the (reshaped) tensor network representing the norm and the expectation value of $\hat{H}$ excluding tensor A and its conjugate $A^\dagger$, respectively (see Fig. 2). Minimizing with respect to $A^\dagger$ yields a generalized eigenvalue problem,

$$\frac{\partial}{\partial \vec{A}^\dagger} \left(\frac{\vec{A}^\dagger \mathbf{H} \vec{A}}{\vec{A}^\dagger \mathbf{N} \vec{A}} \right) = 0 \quad \rightarrow \quad \mathbf{H}\vec{A} = E\mathbf{N}\vec{A}. \qquad (3)$$

The eigenvector $\vec{\tilde{A}}$ with the lowest eigenvalue $\tilde{E}$ provides the solution to the local minimization problem, and the updated tensor A' is obtained by reshaping $\vec{\tilde{A}}$ back to a tensor.

The main challenge of such a scheme for iPEPS is the computation of the matrix $\mathbf{H}$, which consists of an infinite sum of the expectation values of all Hamiltonian terms. In the following we explain how to obtain $\mathbf{H}$ using the CTM method in a way similar to how we computed the environment for the norm $\mathbf{N}$ discussed in Sec. II B. The second complication comes from the fact that in iPEPS tensor A appears not only once in the ansatz (unlike in finite PEPS), but actually, $\mathbf{H}$ and $\mathbf{N}$ also depend on A, making each step a highly nonlinear

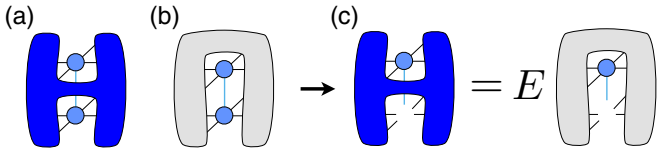

FIG. 2. (a) The H-shaped tensor is obtained by contracting the entire network representing $\langle \Psi|\hat{H}|\Psi\rangle$ except tensors A and $A^\dagger$. (b) Similarly, the N-shaped tensor contains all tensors representing the norm $\langle \Psi|\Psi\rangle$ except tensors A and $A^\dagger$ and can be obtained by contracting all environment tensors in Fig. 1(c) together. Minimizing the energy with respect to tensor A boils down to solving the generalized eigenvalue problem shown in (c) by reshaping the H- and N-shaped tensors into matrices and A into a vector.

optimization problem. We present a practical scheme dealing with this issue in Sec. III C.

B. Systematic summation of Hamiltonian terms with the CTM method

The CTM method discussed in Sec. II B provides a convenient way to compute the norm (and local expectation values) by using the environment tensors, as shown in Fig. 1(c). The expectation value $\langle\Psi|\hat{H}|\Psi\rangle$, which is an infinite sum, can be computed in a similar way by introducing a new type of environment tensors which we call H-environment tensors, shown in dark blue in Fig. 3.

Each H-environment tensor consists of a sum of Hamiltonian contributions. For example, the corner tensor $\tilde{C}_1$ contains all contributions from Hamiltonian terms acting on the infinite upper left part of the system (see bottom panel in Fig. 3). Similarly, $\tilde{T}_4$ contains all Hamiltonian terms acting on the corresponding infinite half row. We further introduce horizontal and vertical corner tensors, denoted by $\tilde{C}_{h1}$ and $\tilde{C}_{v1}$, respectively, for the upper left corner. These tensors take into account Hamiltonian terms which connect sites located in the corner C_1 and edges T_1 or T_4, respectively (see bottom of Fig. 3). Similar tensors are also defined for the other corners. Finally, we also have to sum up the local Hamiltonian terms connecting the center site with its four nearest neighbors (located on the four edge tensors). With this, the sum represented in Fig. 3 takes into account all Hamiltonian terms.

The H-environment tensors can be computed in a systematic way within the regular CTM method, as shown in Fig. 4 for a left move. Importantly, the H-environment tensors are renormalized in the same way as the norm-environment tensors, i.e., using the same projectors P and $\tilde{P}$. In this way the indices of the H-environment tensors match the ones from the norm-environment tensors, and thus, different diagrams, as shown in Fig. 4, can simply be added [41].

Note that the $\tilde{T}\tilde{T}$ tensors, which include Hamiltonian contributions between two edge tensors, do not appear in the expectation value of the Hamiltonian shown in Fig. 3. However, it is crucial to keep track of these tensors since they add contributions to the $\tilde{C}_h$ and $\tilde{C}_v$ tensors, as, for example, shown in the second row in Fig. 3, where the contributions in the $T\tilde{T}_1$ tensor are added to the $\tilde{C}'_{h1}$ tensor.

We end this section with three additional remarks: (1) It is convenient to store also the edge tensors where the physical legs of the outermost site are kept open, e.g., $T_4'^{o}$ shown in Fig. 4. These tensors can then be used to compute the local Hamiltonian terms (connecting to the center site) shown in Fig. 3. (2) The computation of the $\tilde{T}\tilde{T}$ edge terms has a relatively large computational cost of $O(\chi^3 D^6)$ compared to the other terms. This is the same complexity [42] as for the computation of the projectors P and $\tilde{P}$. [21] One way to reduce the complexity of the $\tilde{T}\tilde{T}$ term is to split it in the middle into two parts using an SVD, keeping only a bond dimension of $O(\chi)$ between the two parts. (3) In some implementations of the CTM algorithm one normalizes the environment tensors in a certain way after each step (e.g., division by the largest element of a tensor) in order to keep the numbers in the tensors bounded. In this case one has to make sure for consistency that the same normalization is used also for the H-environment tensors (i.e., the same normalization factor has to be used, e.g., for C_1 and $\tilde{C}_1$).

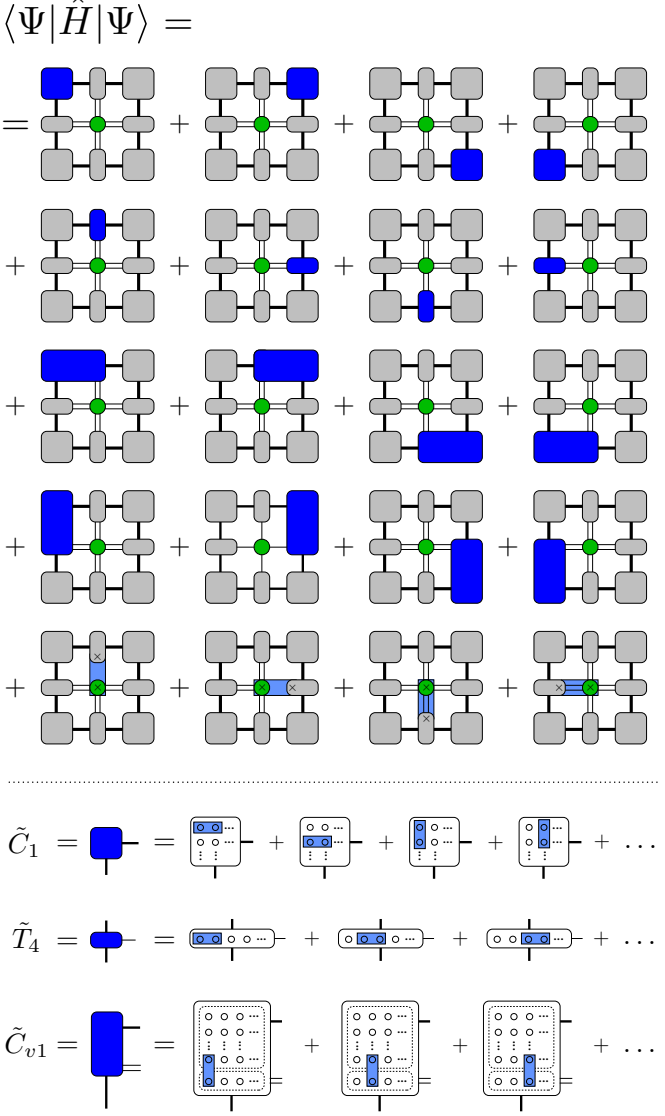

FIG. 3. Representation of the expectation value of the Hamiltonian, where the blue tensors contain sums of local Hamiltonian terms, as illustrated in the bottom part of the figure. For example, the corner tensor $\tilde{C}_1$ contains all contributions of local Hamiltonian terms in the upper left corner of the infinite system, whereas the edge tensor $\tilde{T}_4$ contains all contributions from an infinite half row, as depicted in the bottom part of the figure. The vertical corner tensor $\tilde{C}_{v1}$ takes into account all Hamiltonian terms located between the corner C_1 and the edge tensor T_4 (see bottom image; a similar definition holds for the horizontal corner tensors $\tilde{C}_{h1}$). All the other dark blue tensors on the other corners/edges are defined in a similar way. Finally, there are four remaining Hamiltonian terms (light blue) between the center site and its nearest neighbors. The cross on top of a tensor indicates that the Hamiltonian term is connected to the corresponding physical legs which are not shown in this top view.

C. Practical schemes

With the CTM approach discussed in the previous sections we can compute the **H** and the **N** matrices and solve the

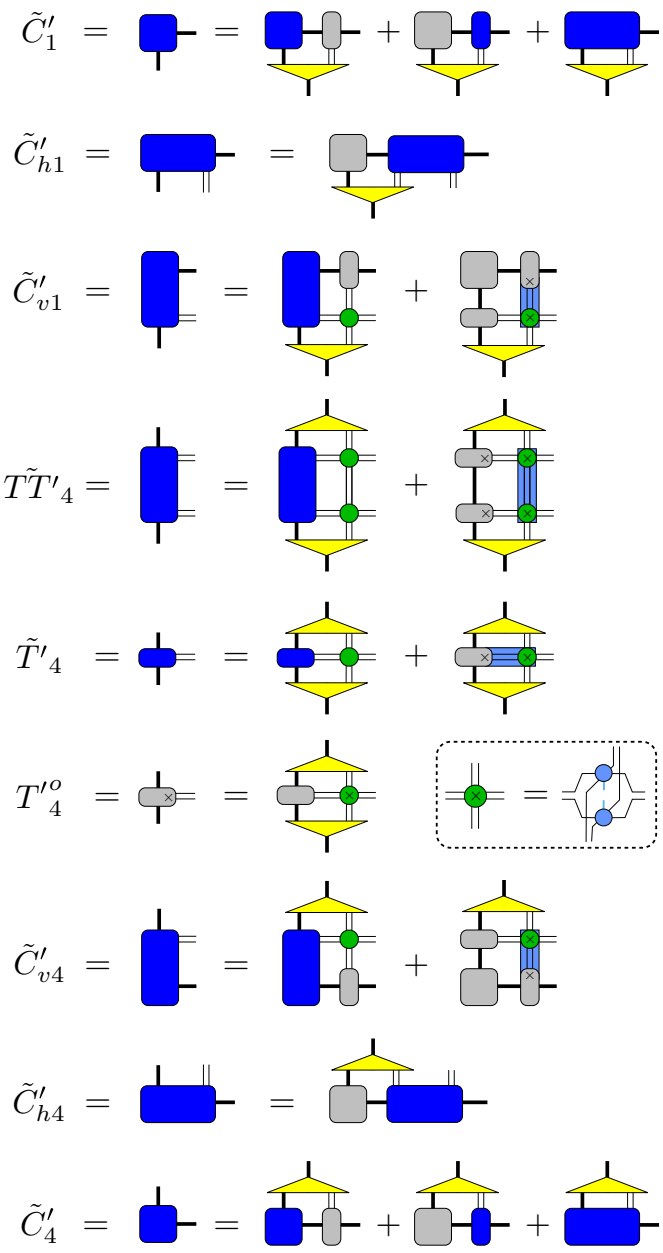

FIG. 4. All relevant diagrams to perform a left move to update the H-environment tensors in the CTM method. Coordinates of the tensors relative to the unit cell have been omitted for simplicity. The projectors to perform the renormalization step (yellow triangles) are the same as the ones computed for the norm. The tensor $\tilde{T}T_4$ contains Hamiltonian terms connecting the sites between two edges T_4. The tensor T_4^o is an edge where the physical legs of the right outermost bulk tensors are left open (such that a Hamiltonian term can be connected to it). Similar diagrams are defined for a right move, top move, and bottom move. In this way one keeps track of all nearest-neighbor Hamiltonian terms in a systematic way.

generalized eigenvalue problem (3) for the eigenstate $\tilde{A}^{[x,y]}$ with lowest-energy eigenvalue. In finite PEPS, where each tensor appears only once, this provides the best solution at the current iteration. In iPEPS, however, each tensor $A^{[x,y]}$ appears infinitely many times, and thus, replacing each tensor $A^{[x,y]}$ by the solution $\tilde{A}^{[x,y]}$ might not be the optimal choice. This is because both $\mathbf{H}$ and $\mathbf{N}$ also depend on $A^{[x,y]}$, making Eq. (2) a highly nonlinear problem (instead of a quadratic one).

One could solve the minimization problem (2), e.g., by a conjugate-gradient method. Here we use a different strategy, which turns out to work well in practice: we solve the generalized eigenvalue problem, but instead of using the solution $\tilde{A}^{[x,y]}$, we take a linear combination with the previous tensor $A^{[x,y]}$,

$$A'(\lambda)^{[x,y]} = \tilde{A}^{[x,y]} \sin\lambda\pi - A^{[x,y]} \cos\lambda\pi. \tag{4}$$

We then optimize the energy $E(\lambda)$ with respect to the single parameter $\lambda \in [0.5, 1.5]$ [43], which, in principle, can be done by standard minimization solvers. For each evaluation of $E(\lambda)$ one has to recompute the environment for the norm (typically, a few iterations starting from the previous environment is accurate enough) and evaluate all local Hamiltonian terms. For this reason it is desirable to keep the number of function evaluations of $E(\lambda)$ low. We had a good experience with the following scheme:

(1) Compute $E(1)$ (corresponding to the previous energy with the old tensor $A' = A$) and $E(0.5)$ (corresponding to the energy with $A' = \tilde{A}$).

(2) If $E(0.5) < E(1)$, take $A' = \tilde{A}$ as the solution and exit.

(3) Define an initial step size Δ_0 (e.g., $\Delta_0 = 0.1$) and a tiny step size h (e.g., $h = 10^{-4}$).

(4) If $E(1+h) < E(1)$, set $\Delta = \Delta_0$, else $\Delta = -\Delta_0$.

(5) For *iter* $= 1$ to *maxiter*

(a) If $E(1+\Delta) < E(1)$, accept solution [44] with $\lambda = 1 + \Delta$ and exit.

(b) Else $\Delta = \Delta/2$.

With this scheme, typically, only a few evaluations of the energy are required. The algorithm stops as soon as a lower-energy solution is found. This does not provide the optimal λ at each iteration, but in practice this does not seem to matter since in the end we are interested in the global minimum after many sweeps and not the "local" optimum at each iteration.

Finally, we repeat the minimization for each tensor in the unit cell, i.e., for all coordinates $[x,y]$, and reiterate until the desired convergence in the energy is reached.

For computations of both the H-environment and the norm-environment we can start from the environment from the previous iteration (like in the fast full update [29]), so that only a few additional CTM iterations are needed at each step [45].

IV. BENCHMARK RESULTS

In this section we present a series of benchmark results, ranging from a standard problem (the Heisenberg model) to challenging cases, including the Shastry-Sutherland and the t-J models. In all cases we show that the iPEPS results for each bond dimension can be considerably improved with the variational optimization, the energy, and order parameters. For the larger-D simulations we have exploited the $U(1)$ symmetry of the models in order to increase the efficiency of the calculation [46,47].

A. Heisenberg model

As a first example we consider the two-dimensional $S = 1/2$ Heisenberg model on a square lattice with Hamiltonian

$$\hat{H} = J \sum_{\langle i,j \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j, \tag{5}$$

where the sum goes over nearest-neighbor sites and $\mathbf{S_i}$ is a spin-1/2 operator on site i. Reference values are taken from Ref. [48] and are based on state-of-the-art quantum Monte Carlo calculations. We use an iPEPS ansatz with two different tensors (one for each sublattice) to represent the ground state with antiferromagnetic order.

Figure 5(a) shows a comparison of the relative error of the energy as a function of the bond dimension D, obtained with the three optimization methods. One can clearly see that there is a substantial improvement when going from the simple update to the full update calculations, as previously found. But interestingly, the results obtained with the variational optimization are even better, by roughly a factor 2! Also the order parameter, the staggered magnetization m shown in Fig. 5(b), is considerably improved.

It is important to point out that the difference between the full and the variational update is not due to the Trotter error in the imaginary-time evolution. In Fig. 5(c) the dependence of the result for $D = 4$ as a function of the Trotter step τ shows that even in the limit $\tau \to 0$ the full update does not yield the same accuracy as the variational approach. We believe that this has mainly to do with the fact that at every step a two-site operator $\hat{U}_b$ (see Sec II C) is applied only to a *single* bond in the middle of the system, which is then truncated in an optimal way, resulting in updated tensors A', B'. These tensors provide a locally optimal solution for that single bond; however, it is not guaranteed that they also provide a globally optimal solution when replacing the tensors everywhere in the ansatz. A globally optimal solution could be found by applying $\hat{U}_b$ to all bonds in the ansatz and then truncating all bonds simultaneously. However, such a scheme would be computationally much more expensive because it would require us to contract the time-evolved iPEPS with an enlarged bond dimension, and thus, it would not be very useful in practice.

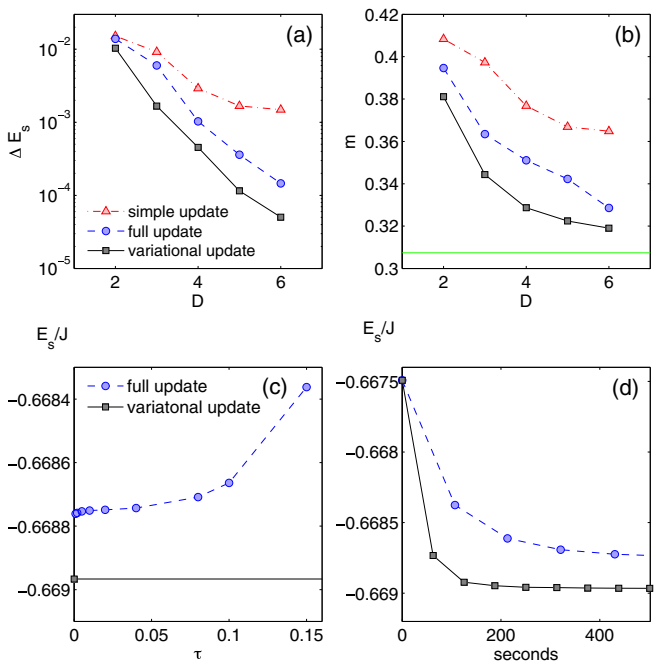

FIG. 5. (a) Relative error of the energy per site of the 2D $S = 1/2$ Heisenberg model as a function of the bond dimension D obtained with three different optimization methods for iPEPS. (b) The order parameter (staggered magnetization) as a function of D, compared to the extrapolated QMC result in the thermodynamic limit (horizontal line). (c) Full update result as a function of the Trotter imaginary time step τ ($D = 4$) compared to the variational result. (d) Evolution of the energy as a function of simulation runtime on a MacBook Pro laptop for $D = 4$ and $\chi = 50$. The results from the variational optimization are shown at each sweep (squares), whereas the energy obtained with the full update is computed after every 30 time steps with $\tau = 0.02$.

The variational optimization is not only more accurate but typically also converges faster to the lowest-energy state. This is illustrated in Fig. 5(d), where we compare the performance of the full update with the variational update on a MacBook Pro laptop for $D = 4$, $\chi = 50$. Starting from the simple update result already after one step with the variational scheme (taking 1 min), one has a better result than with the full update in the long-run-time limit.

As a side remark on the performance we note that it is best to start with an initial state which is already close to the ground state, particularly for the large-D simulations. For example, we can take the solution from the simple update (which is computationally very cheap) as an initial state for the full- or variational-update scheme [49] or use the full-update solution (e.g., obtained with a large τ) as a starting point for a variational optimization. One can also use a converged solution with bond dimension $D - 1$ as an initial state for a simulation with bond dimension D. In this case one can first perform a few full update steps (to increase the bond dimension from $D - 1$ to D) and then continue with the variational update. Alternatively, the bond dimension can also be dynamically increased by using a two-site update (see the Appendix).

Finally, it is interesting to compare our best result with 2D DMRG on cylinders from Ref. [4]. Our lowest variational energy per site for $D = 6$ ($\chi = 250$) is $-0.669408J$, which is very close to the exact Monte Carlo result, $-0.6694421(4)J$ [48], with a relative error of only 5×10^{-5}. This precision is comparable to a 2D DMRG calculation on a width-10 cylinder with $m = 3000$ states. However, due to the exponential scaling of the required bond dimension with the cylinder width in DMRG, already a width-12 cylinder with the same number of states is an order of magnitude worse than the iPEPS result. Furthermore, the iPEPS accuracy is obtained for the infinite system (infinite cylinder width), and we would actually expect an even higher accuracy with iPEPS on a finite cylinder. It is also remarkable to compare the number of variational parameters: A $D = 6$ iPEPS has roughly 2.6×10^3 variational parameters per tensor (with two different tensors in the entire ansatz), whereas an $m = 3000$ state has of the order of 1.8×10^7 parameters, i.e., a difference of four orders of magnitude per tensor [50]. This illustrates that (i)PEPS offers a much more efficient representation of a 2D wave function than MPS, even for modest cylinder widths.

In summary, the results from this section demonstrate that (1) the full update actually *fails* to reproduce the most optimal result for a given bond dimension D, even in the limit $\tau \to 0$, and (2) more accurate results can be obtained with the variational optimization, with typically a faster convergence towards the ground state.

B. Shastry-Sutherland model

We next move to a more challenging case: the Shastry-Sutherland model, [51] which is a frustrated spin system given by the Hamiltonian

$$H = J \sum_{\langle i,j \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j + J' \sum_{\langle\langle i,j \rangle\rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j - h \sum_i S_i^z, \qquad (6)$$

where the $\langle i,j \rangle$ bonds with coupling strength J build an array of orthogonal dimers while the bonds with coupling J' denote interdimer couplings and h is the strength of an external magnetic field. This model is realized in the material $SrCu_2(BO_3)_2$ in which experiments have found a series of different magnetization plateaus [52–61]. Recently, iPEPS has played the key role in understanding the structure of the states realized in these plateaus [22].

At zero field, $h = 0$, there exists a nonmagnetic plaquette phase [62–65] in a narrow parameter range in between a dimer phase for $J'/J < 0.675(1)$ and an antiferromagnetic phase for $J'/J > 0.765(1)$ [18]. Previous iPEPS simulations confirmed the existence of the plaquette phase [18]; however, it was observed that due to the close adjacency of the other two phases it is difficult to converge into the plaquette phase with the full update (depending on the simulation setup and initial state used). For example, when starting from an antiferromagnetic state (using a simulation setup with one tensor per dimer), the full update fails to reproduce the plaquette state and remains stuck in an antiferromagnetic state.

Here we use this challenging case as a benchmark for the variational optimization. Encouragingly, we find that even if the simulation is initialized in the wrong state, the variational optimization converges to the correct plaquette state. A comparison between the full update and variational results is presented in Fig. 6 for $D = 4$ and $J'/J = 0.7$. Both simulations have been initialized with the same antiferromagnetic state obtained with the simple update (for $J'/J = 0.78$). The resulting energy per site with the variational optimization, $E_s = -0.3862J$, is considerably lower than the full update result, $E_s = -0.3843J$.

In order to test the variational optimization method for larger unit cells we consider the Shastry-Sutherland model in an external magnetic field, i.e., with $h > 0$. It has previously been found with iPEPS that the magnetization plateaus correspond to crystals of bound states of triplet excitations [22], which have a typical spin structure reminiscent of a pinwheel, shown in Fig. 7(a) for a 4×4 unit cell of dimers (i.e., 16 different tensors). For details on the physics we refer to Ref. [22] and references therein. The bound state is symmetric under rotations by 90°; however, an inaccurate optimization scheme like the simple update fails to create a nicely symmetric state, as shown, for example, in Fig. 7(b). The full update creates a more symmetric state; however, here we show that the variational optimization even provides a substantially better result.

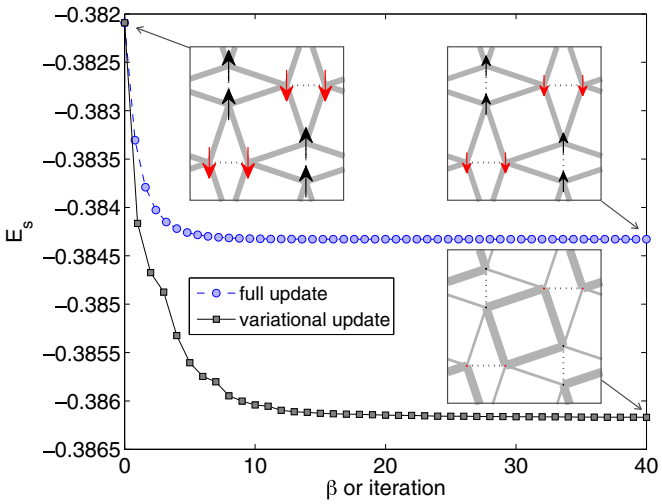

FIG. 6. Evolution of the energy as a function of imaginary time (full update) or iteration step (variational update) for $J'/J = 0.7$ (in the plaquette phase) for $D = 4$. Both simulations have been initialized from an initial antiferromagnetic state which was obtained using the simple update (for $J'/J = 0.78$). The initial state is shown in the top left inset, where the lengths of the arrows are proportional to the magnitude of the local spin, and the thickness of the lines on the J' bonds scales with the magnitude of the energy on that bond (i.e., the thicker a line is, the lower the energy is). The variational update (squares) correctly reproduces a plaquette state, shown by the bottom right inset, where the bond energies around a plaquette are lower than on the other bonds, and the values of the local spins vanish. The full update (circles) fails to reproduce the state and yields an antiferromagnetic state with higher energy and slightly suppressed antiferromagnetic order compared to the initial state (top right inset).

To quantify the rotational symmetry breaking we define a "symmetry error" in the following way: On each site we determine the expectation value of the spin and compute the standard deviation over all sites which are equivalent by symmetry. The symmetry error is then defined as the mean value of all the standard deviations. Thus, in a perfectly rotationally invariant state, this error vanishes. Figure 7(d) shows the results for the symmetry error as a function of D obtained with the three optimization schemes. Again, we find that the error is considerably lowered with the variational optimization compared to the simple and full updates. Consistently, the corresponding variational energies are also lower, as shown in Fig. 7(c).

The results further demonstrate that with the variational optimization not only is a better variational energy obtained, but the states (order parameters) are more accurately reproduced, even if they exhibit extended features requiring large unit cells.

C. The t-J model

Finally, we test the variational optimization also for a challenging fermionic system which requires a fermionic iPEPS ansatz. The formalism to apply 2D tensor networks to fermionic systems has been developed in several works [28,66–73], which all describe the same fermionic

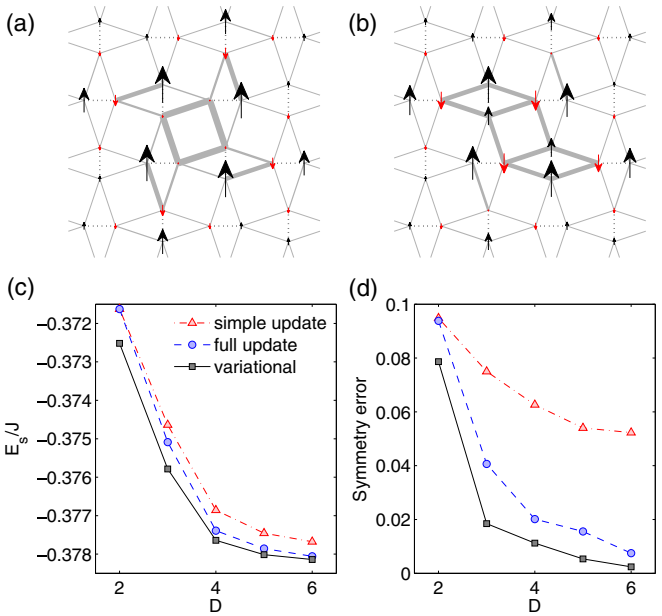

FIG. 7. Results for the Shastry-Sutherland model in a finite magnetic field at magnetization $m = 1/8$, obtained in a 4×4 unit cell of dimers. (a) Spin structure of a bound state of two triplet excitations exhibiting a 90° rotational symmetry. (b) The rotational symmetry of a bound state gets strongly broken when using an inaccurate optimization such as the simple update (here for $D = 5$). (c) Comparison of the variational energies as a function of D obtained with the three optimization methods. (d) Comparison of the symmetry error, quantifying the degree of the broken rotational symmetry.

tensor network ansatz. Here we follow the formalism from Ref. [28] to take into account fermionic exchange statistics.

As a challenging example we consider the t-J model, which is an effective model of the Hubbard model in the strongly interacting limit, given by the Hamiltonian

$$\hat{H} = -t \sum_{\langle i,j\rangle\sigma} (\tilde{c}^\dagger_{i\sigma}\tilde{c}_{j\sigma} + \text{H.c.}) + J \sum_{\langle i,j\rangle} \left(\hat{S}_i \hat{S}_j - \frac{1}{4}\hat{n}_i\hat{n}_j \right), \quad (7)$$

with $\sigma = \{\uparrow, \downarrow\}$ being the spin index, $\hat{n}_i = \sum_\sigma \hat{c}^\dagger_{i\sigma}\hat{c}_{i\sigma}$ being the electron density, $\hat{S}_i$ being the spin-1/2 operator on site i, and $\tilde{c}_{i\sigma} = \hat{c}_{i\sigma}(1 - \hat{c}^\dagger_{i\bar{\sigma}}\hat{c}_{i\bar{\sigma}})$ being hopping operators forbidding doubly occupied sites.

The t-J model has been previously studied with iPEPS using the simple [36] and full updates [21]. Here we show that we can substantially improve the variational energies by just adding a few (three to five) variational update steps on top of the full-update results from Ref. [21]. The results are presented in Fig. 8 for $J/t = 0.4$ and hole density $\delta = 0.12$.

Thus, this example further illustrates the usefulness of the variational optimization to increase the accuracy of iPEPS simulations for challenging problems. Similar improvements can also be obtained for the 2D Hubbard model (not shown) [25].

V. SUMMARY AND OUTLOOK

We have presented a variational optimization scheme for iPEPS based on a systematic summation of Hamiltonian terms

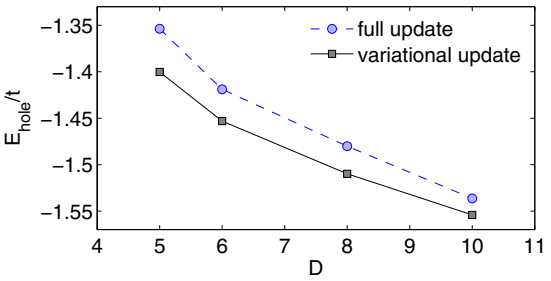

FIG. 8. Energy per hole in the t-J model for $J/t = 0.4$ and doping $\delta = 0.12$. For each value of D the energies are substantially improved by the variational optimization.

using the corner-transfer-matrix method. For a given bond dimension D the scheme yields a higher accuracy for the energy and order parameters than the previously best results based on an imaginary-time evolution using the full update. For example, for the Heisenberg model with bond dimension $D = 6$ the relative error of the energy is only 5×10^{-5} (without using any extrapolation) compared to the extrapolated quantum Monte Carlo result. The variational approach not only yields a higher accuracy but can also help to speed up the convergence towards the ground state. We have also presented a challenging case where the full update failed to reproduce the correct ground state in the Shastry-Sutherland model, whereas the variational optimization converged to the right plaquette state.

We note that a similar approach can also be used for Hamiltonians with next-nearest-neighbor terms with only minor modifications. In principle, also longer-ranged Hamiltonians can be simulated, which requires keeping track of additional H-environment tensors.

The results for the t-J model provide further evidence that with the variational scheme more accurate results can be obtained not only for (frustrated) spin models but also for strongly correlated electron systems. We thus believe that the scheme will play a key role for future state-of-the-art simulations of challenging problems, like the single- and multiband 2D Hubbard models.

Finally, we point out that there are also ways to perform a variational optimization with iPEPS other than the scheme presented here. A systematic summation of Hamiltonian contributions can also be achieved using the channel environments introduced in Ref. [33]. As an alternative to solving a generalized eigenvalue problem at each step, one can also use other approaches to minimize the energy, e.g., a conjugate gradient method [74]. Thus, there are more promising options which are interesting to explore in order to further improve the performance of iPEPS ground-state simulations.

ACKNOWLEDGMENTS

The author acknowledges insightful discussions with G. Vidal on how to use the CTM method to systematically summing up Hamiltonian terms and with L. Vanderstraeten, F. Verstraete, A. Gendiar, and T. Nishino about other variational optimization schemes. This work is part of the D-ITP

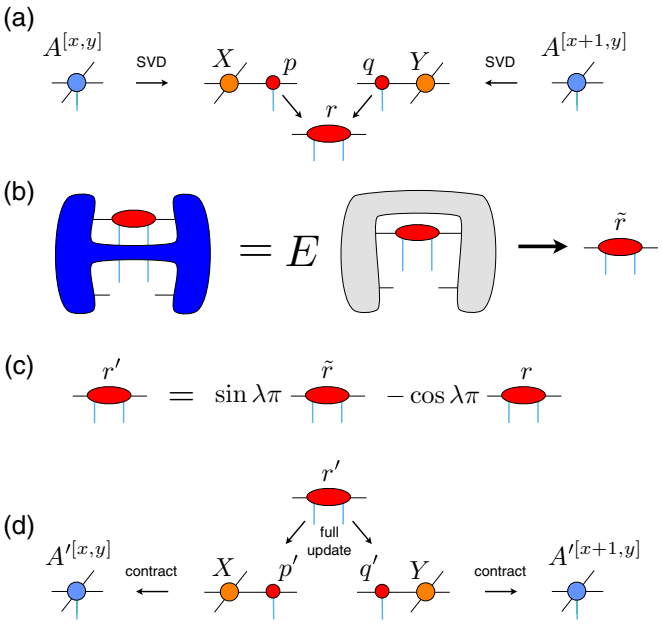

FIG. 9. (a) In order to update the horizontal bond between coordinates $[x,y]$ and $[x+1,y]$, we first split the involved tensors into smaller pieces using an SVD (the singular values are absorbed in the p and q tensors). (b) Solving the generalized eigenvalue problem for tensor r yields the solution $\tilde{r}$. (c) Ansatz for the solution r' made of a linear combination of $\tilde{r}$ and the initial tensor r, depending on a parameter λ. (d) The updated tensors $A'^{[x,y]}$ and $A'^{[x+1,y]}$ are obtained from recombining p' with X and q' with Y, respectively, where p' and q' are determined using the full update.

consortium, a program of the Netherlands Organization for Scientific Research (NWO) that is funded by the Dutch Ministry of Education, Culture and Science (OCW). This research was supported in part by the Perimeter Institute for Theoretical Physics. Research at the Perimeter Institute is supported by the government of Canada through Industry Canada and by the province of Ontario through the Ministry of Research and Innovation.

APPENDIX: TWO-SITE VARIATIONAL OPTIMIZATION

In the main part of the paper we discussed how to perform a one-site update, i.e., where one tensor in the unit cell after the other gets iteratively optimized. Like for MPS in one dimension, one can also perform a two-site update, where two tensors are updated at once. One advantage of this approach is that the bond dimension can be dynamically adjusted (which is particularly important also when making use of global Abelian symmetries [46,47]). In the following we describe the main idea of how to update, e.g., a horizontal bond involving tensors $A^{[x,y]}$ and $A^{[x+1,y]}$.

Since updating two full tensors at once is computationally expensive, we first split the two tensors $A^{[x,y]}$ and $A^{[x+1,y]}$ into smaller pieces using an SVD, shown in Fig. 9(a). A similar splitting is also used in the simple update and full update to increase the efficiency [28,29]. The contributions p and q are then combined into a single tensor r. We can now formulate the variational optimization with respect to this tensor r by constructing the corresponding **H** and **N** matrices using the H- and norm-environment tensors, like in the one-site update. Solving the generalized eigenvalue problem for the lowest-energy eigenvalue yields the tensor $\tilde{r}$ in Fig. 9(b). We again make an ansatz for the solution r' by a linear combination of $\tilde{r}$ and the previous tensor r [Fig. 9(c)] and determine the best mixing parameter λ, e.g., with the scheme from Sec. III C. To evaluate the energy one has to decompose r' into two pieces, p' and q' (with bond dimension D between them), which can be done using the full update, [28] and then use p' and q' to compute the updated tensors $A'^{[x,y]}$ and $A'^{[x+1,y]}$, respectively [see Fig. 9(d)].

This completes the update of a horizontal bond, and one can now proceed with updating all the other horizontal and vertical bonds in the unit cell and perform several sweeps until convergence of the energy is reached.

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[2] F. Verstraete, V. Murg, and J. I. Cirac, Adv. Phys. **57**, 143 (2008).
[3] U. Schollwöck, Ann. Phys. (NY) **326**, 96 (2011).
[4] E. Stoudenmire and S. R. White, Annu. Rev. Condens. Matter Phys. **3**, 111 (2012).
[5] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.
[6] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. A **75**, 033605 (2007).
[7] T. Nishino, Y. Hieida, K. Okunishi, N. Maeshima, Y. Akutsu, and A. Gendiar, Prog. Theor. Phys. **105**, 409 (2001).
[8] Y. Nishio, N. Maeshima, A. Gendiar, and T. Nishino, arXiv:cond-mat/0401115.
[9] M. Daniška and A. Gendiar, J. Phys. A **48**, 435002 (2015).
[10] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[11] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **101**, 250602 (2008).
[12] Another class of efficient tensor network states for 1D and 2D systems is the multiscale entanglement renormalization ansatz [75,76].
[13] P. Corboz, A. M. Läuchli, K. Penc, M. Troyer, and F. Mila, Phys. Rev. Lett. **107**, 215301 (2011).
[14] L. Wang, Z.-C. Gu, F. Verstraete, and X.-G. Wen, arXiv:1112.3331.
[15] H. H. Zhao, C. Xu, Q. N. Chen, Z. C. Wei, M. P. Qin, G. M. Zhang, and T. Xiang, Phys. Rev. B **85**, 134416 (2012).
[16] P. Corboz, M. Lajkó, A. M. Läuchli, K. Penc, and F. Mila, Phys. Rev. X **2**, 041013 (2012).
[17] Z. Y. Xie, J. Chen, J. F. Yu, X. Kong, B. Normand, and T. Xiang, Phys. Rev. X **4**, 011025 (2014).
[18] P. Corboz and F. Mila, Phys. Rev. B **87**, 115144 (2013).
[19] Z.-C. Gu, H.-C. Jiang, D. N. Sheng, H. Yao, L. Balents, and X.-G. Wen, Phys. Rev. B **88**, 155112 (2013).
[20] J. Osorio Iregui, P. Corboz, and M. Troyer, Phys. Rev. B **90**, 195102 (2014).

[21] P. Corboz, T. M. Rice, and M. Troyer, Phys. Rev. Lett. **113**, 046402 (2014).

[22] P. Corboz and F. Mila, Phys. Rev. Lett. **112**, 147203 (2014).

[23] T. Picot and D. Poilblanc, Phys. Rev. B **91**, 064415 (2015).

[24] T. Picot, M. Ziegler, R. Orús, and D. Poilblanc, Phys. Rev. B **93**, 060407(R) (2016).

[25] P. Corboz, Phys. Rev. B **93**, 045116 (2016).

[26] P. Nataf, M. Lajkó, P. Corboz, A. M. Läuchli, K. Penc, and F. Mila, Phys. Rev. B **93**, 201113(R) (2016).

[27] H. C. Jiang, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **101**, 090603 (2008).

[28] P. Corboz, R. Orus, B. Bauer, and G. Vidal, Phys. Rev. B **81**, 165104 (2010).

[29] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, and R. Orus, Phys. Rev. B **92**, 035142 (2015).

[30] T. Nishino and K. Okunishi, J. Phys. Soc. Jpn. **65**, 891 (1996).

[31] T. Nishino and K. Okunishi, J. Phys. Soc. Jpn. **66**, 3040 (1997).

[32] R. Orús and G. Vidal, Phys. Rev. B **80**, 094403 (2009).

[33] L. Vanderstraeten, M. Mariën, F. Verstraete, and J. Haegeman, Phys. Rev. B **92**, 201111(R) (2015).

[34] F. Fröwis, V. Nebendahl, and W. Dür, Phys. Rev. A **81**, 062337 (2010).

[35] A. Gendiar, N. Maeshima, and T. Nishino, Prog. Theor. Phys. **110**, 691 (2003).

[36] P. Corboz, S. R. White, G. Vidal, and M. Troyer, Phys. Rev. B **84**, 041108 (2011).

[37] L. Wang, I. Pižorn, and F. Verstraete, Phys. Rev. B **83**, 134421 (2011).

[38] Y.-K. Huang, P. Chen, and Y.-J. Kao, Phys. Rev. B **86**, 235102 (2012).

[39] We usually use a second-order decomposition in our simulations.

[40] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

[41] If one were to compute a separate set of projectors to renormalize the bond indices adjacent to an H-environment tensor, one would also lose translational invariance in the environment since, e.g., the T-environment tensors adjacent to a $\tilde{T}$ tensor would also get altered, and the number of different tensors would diverge with increasing CTM iterations, which would make the scheme inefficient.

[42] By introducing further approximations the projectors can also be computed with a computational cost of $O(\chi^3 D^4)$.

[43] A similar mixing was also used in Refs. [7,35] in the variational optimization schemes for 3D classical systems, but without performing a second minimization step with respect to the mixing parameter.

[44] Some (large) values of Δ may result in a "pathological" state where the CTM scheme does not properly converge anymore and where the norm of the state becomes very small. Such solutions should not be accepted since they may lead to stability problems (see also discussion in Ref. [35]).

[45] After many iterations it can be useful to reinitialize the environments in some cases and to recompute the environments from scratch.

[46] S. Singh, R. N. C. Pfeifer, and G. Vidal, Phys. Rev. B **83**, 115125 (2011).

[47] B. Bauer, P. Corboz, R. Orús, and M. Troyer, Phys. Rev. B **83**, 125106 (2011).

[48] A. W. Sandvik, in *Lectures on the Physics of Strongly Correlated Systems XIV: Fourteenth Training Course in the Physics of Strongly Correlated Systems*, AIP Conf. Proc. No. 1297 (AIP, New York, 2010), pp. 135–338.

[49] There are cases where simple update fails to provide the correct state, e.g., the spin-liquid state [20], and does not provide a good initial state for these cases. It is thus important to cross-check results also starting from random initial states or from full-update solutions.

[50] In addition, a finite-size DMRG algorithm uses $L_x \times L_y$ different tensors, e.g., with $L_x = 20$ and $L_y = 10$ in Ref. [4], which increases the number of variational parameters by an additional two orders of magnitude. We note that the effective number of parameters is reduced in both cases by exploiting the global $U(1)$ symmetry in the model.

[51] B. Sriram Shastry and B. Sutherland, Physica B+C (Amsterdam) **108**, 1069 (1981).

[52] H. Kageyama, K. Yoshimura, R. Stern, N. V. Mushnikov, K. Onizuka, M. Kato, K. Kosuge, C. P. Slichter, T. Goto, and Y. Ueda, Phys. Rev. Lett. **82**, 3168 (1999).

[53] K. Onizuka, H. Kageyama, Y. Narumi, K. Kindo, Y. Ueda, and T. Goto, J. Phys. Soc. Jpn. **69**, 1016 (2000).

[54] H. Kageyama, M. Nishi, N. Aso, K. Onizuka, T. Yosihama, K. Nukui, K. Kodama, K. Kakurai, and Y. Ueda, Phys. Rev. Lett. **84**, 5876 (2000).

[55] K. Kodama, M. Takigawa, M. Horvatić, C. Berthier, H. Kageyama, Y. Ueda, S. Miyahara, F. Becca, and F. Mila, Science **298**, 395 (2002).

[56] M. Takigawa, K. Kodama, M. Horvatić, C. Berthier, H. Kageyama, Y. Ueda, S. Miyahara, F. Becca, and F. Mila, Physica B (Amsterdam, Neth.) **346-347**, 27 (2004).

[57] F. Levy, I. Sheikin, C. Berthier, M. Horvatić, M. Takigawa, H. Kageyama, T. Waki, and Y. Ueda, Europhys. Lett. **81**, 67004 (2008).

[58] S. E. Sebastian, N. Harrison, P. Sengupta, C. D. Batista, S. Francoual, E. Palm, T. Murphy, N. Marcano, H. A. Dabkowska, and B. D. Gaulin, Proc. Natl. Acad. Sci. USA **105**, 20157 (2008).

[59] M. Jaime, R. Daou, S. A. Crooker, F. Weickert, A. Uchida, A. E. Feiguin, C. D. Batista, H. A. Dabkowska, and B. D. Gaulin, Proc. Natl. Acad. Sci. USA **109**, 12404 (2012).

[60] M. Takigawa, M. Horvatić, T. Waki, S. Krämer, C. Berthier, F. Lévy-Bertrand, I. Sheikin, H. Kageyama, Y. Ueda, and F. Mila, Phys. Rev. Lett. **110**, 067210 (2013).

[61] Y. H. Matsuda, N. Abe, S. Takeyama, H. Kageyama, P. Corboz, A. Honecker, S. R. Manmana, G. R. Foltin, K. P. Schmidt, and F. Mila, Phys. Rev. Lett. **111**, 137204 (2013).

[62] A. Koga and N. Kawakami, Phys. Rev. Lett. **84**, 4461 (2000).

[63] Y. Takushima, A. Koga, and N. Kawakami, J. Phys. Soc. Jpn. **70**, 1369 (2001).

[64] C. H. Chung, J. B. Marston, and S. Sachdev, Phys. Rev. B **64**, 134407 (2001).

[65] A. Läuchli, S. Wessel, and M. Sigrist, Phys. Rev. B **66**, 014401 (2002).

[66] P. Corboz, G. Evenbly, F. Verstraete, and G. Vidal, Phys. Rev. A **81**, 010303(R) (2010).

[67] C. V. Kraus, N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. A **81**, 052338 (2010).

[68] C. Pineda, T. Barthel, and J. Eisert, Phys. Rev. A **81**, 050303 (2010).

[69] T. Barthel, C. Pineda, and J. Eisert, Phys. Rev. A **80**, 042333 (2009).
[70] Q.-Q. Shi, S.-H. Li, J.-H. Zhao, and H.-Q. Zhou, arXiv:0907.5520.
[71] P. Corboz and G. Vidal, Phys. Rev. B **80**, 165129 (2009).
[72] I. Pižorn and F. Verstraete, Phys. Rev. B **81**, 245110 (2010).
[73] Z.-C. Gu, F. Verstraete, and X.-G. Wen, arXiv:1004.2563.
[74] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, arXiv:1606.09170.
[75] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).
[76] G. Evenbly and G. Vidal, Phys. Rev. Lett. **102**, 180406 (2009).