# Simulating strongly correlated quantum systems with tree tensor networks

V. Murg and F. Verstraete

*Fakultät für Physik, Universität Wien, Boltzmanngasse 3, A-1090 Vienna, Austria*

Ö. Legeza* and R. M. Noack

*Fachbereich Physik, Philipps-Universität Marburg, 35032 Marburg, Germany*

We present a tree-tensor-network-based method to study strongly correlated systems with nonlocal interactions in higher dimensions. Although the momentum-space and quantum-chemistry versions of the density-matrix renormalization group (DMRG) method have long been applied to such systems, the spatial topology of DMRG-based methods allows efficient optimizations to be carried out with respect to one spatial dimension only. Extending the matrix-product-state picture, we formulate a more general approach by allowing the local sites to be coupled to more than two neighboring auxiliary subspaces. Following [Y. Shi, L. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006)], we treat a treelike network ansatz with arbitrary coordination number $z$, where the $z=2$ case corresponds to the one-dimensional (1D) scheme. For this ansatz, the long-range correlation deviates from the mean-field value polynomially with distance, in contrast to the matrix-product ansatz, which deviates exponentially. The computational cost of the tree-tensor-network method is significantly smaller than that of previous DMRG-based attempts, which renormalize several blocks into a single block. In addition, we investigate the effect of unitary transformations on the local basis states and present a method for optimizing such transformations. For the 1D interacting spinless fermion model, the optimized transformation interpolates smoothly between real space and momentum space. Calculations carried out on small quantum chemical systems support our approach.

## I. INTRODUCTION

Understanding and simulating strongly correlated systems has long been a major challenge in theoretical physics and in theoretical chemistry. In the past two decades, the density-matrix renormalization group (DMRG) method has been applied effectively to study problems in these fields.[1,2] In particular, it has been widely used to study fermionic and spin-chain problems in one dimension for models with both local and long-range interactions. Application to systems with long-range interactions gained impetus when the method was reformulated to treat models defined in momentum space[3–5] (MS-DMRG), quantum chemistry calculations[6–16] (QC-DMRG), and on random graphs.[17] Common characteristics of these approaches are that a higher dimensional system is mapped to a one-dimensional (1D) chain and that variational approximations to the eigenstates are obtained by an iterative diagonalization procedure. Introduction of various quantum information entropies[5,18–20] and the reformulation of the problem in terms of matrix product states[21–23] (MPS) has clarified the mathematical underpinnings of the method.

Recently, extensions to two dimensions based on the controlled manipulation of entanglement between subsystems has led to alternate methods that can be viewed as generalizations of matrix-product-state-based methods.[24] These methods include projected entangled pair states[21,25–27] (PEPS), the multiscale-entanglement-renormalization ansatz[28] (MERA), and correlator product states[29] or complete-graph tensor network states.[30] The PEPS and MERA methods have already shown considerable promise for frustrated and fermionic problems; they do not suffer from the fermion sign problem that appears in quantum Monte Carlo simulations.[31,32] These new methods, however, have primarily been restricted to the treatment of local Hamiltonians. Thus, developing effective algorithms to treat higher dimensional systems in which the interactions are nonlocal remains an important problem.

In order to efficiently treat quantum chemical systems, it has become evident in the past few years that methods must take into account more general spatial topology. It has been shown that quantum information entropy[5,15,33] can be used to determine the entanglement or quantum correlation among sites or orbitals in a pairwise way. Such a two-dimensional (2D) entanglement matrix leads to a picture of the topology of orbitals that corresponds to a multiply connected network. Two possible approaches to optimize computational efficiency are to work with an appropriate networklike structure that reflects the entanglement topology and to vary the single-particle basis[34] so that entanglement is reduced. For states based on a one-dimensional topology, the only possibility to take entanglement topology into account is to reorder the orbitals. The next step in generalizing the topology is to use a tree network. This was first pointed out in Ref. 35. A tree network makes it possible to reduce the distance between highly entangled orbitals when they are multiply connected. In addition, the coordination number $z$ can be varied from orbital to orbital to adapt the variational state to a particular entanglement topology. For a tree network that is bipartite, it is possible to adapt methods and optimizations developed for matrix-product-state-based algorithms such as the DMRG (Ref. 36) to tensor-product states on the tree network. This is because the tree network method can be related to a generalized DMRG with $z$ blocks instead of two. In addition since an effective Hamiltonian can be formed, real time evolution, like that done in time-dependent density-matrix renormalization group (t-DMRG) (Refs. 19 and 37–

39) and in PEPS,[26] can be carried out. Real time evolution could be used, for example, to calculate spectral functions for quantum chemical systems.

The use of more general spatial topologies is also potentially important for quantum impurity problems. In these systems, the impurity subsystem, which can consist of either a single site or a strongly correlated cluster, is coupled to a bath of free fermions. The classical method to treat quantum impurity problems numerically is the numerical renormalization group (NRG).[40] The starting point of the NRG method is the mapping of the problem onto a one-dimensional semi-infinite lattice in which the first site (or set of sites) describes the impurity subsystem and the remainder of the chain represents the logarithmically discretized conduction band. In the past few years, the NRG has been extended by dividing the chain into a system and an environment as in the DMRG;[41] the resulting method, the DM-NRG, has undergone significant development recently.[42–46] While these extensions have led to significant improvements, they nevertheless have focused primarily on optimizing algorithms based on a one-dimensional topology. In general, the conduction bands are entangled through the impurity subsystem only. Therefore, it is a natural choice to describe the problem as a treelike structure in which the impurity subsystem at the center is surrounded by shells of the conduction bath. Such a treelike topology has been utilized to treat the quantum impurity problem that occur with dynamic mean-field theory.[47,48]

The first attempts to apply matrix-product states on tree networks were carried out by Otsuka[49] and by Friedman *et al.*,[50] who calculated ground-state properties of the spin-1/2 XXZ and Heisenberg chains using the DMRG. Subsequently, Lepetit *et al.*[51] studied the half-filled Hubbard model on a Bethe lattice (also known as a Cayley tree). Recently, Martin-Delgado *et al.*[52] applied the DMRG to a model of dendritic polymers, a system that can also be represented as a Cayley tree. All of this work uses a Bethe lattice, whose characteristics are that the number of nearest neighbors at each node is $z$, i.e., the coordination number, and that closed loops do not occur. Since there is only one path between any pair of sites, a DMRG-based solution of the problem is possible. In this approach, however, $z$ systems blocks must be renormalized to a single block at each iteration step, leading to computational cost which increases exponentially with $z$, hindering systematic DMRG studies for large systems and/or large $z$; up to now only $z=3$ has been treated. In contrast, the tree tensor network (TTNS) methods we introduce here have a much lower computational cost because the topology consists of a single site and $z$ blocks whereas the superblock configuration included $2z$ blocks or $z+1$ in previous DMRG attempts. In this work, we present calculations on small systems in order to demonstrate the viability of the tree tensor network method and to compare its accuracy and efficiency with existing DMRG methods.

Another approach to take advantage of the benefits of the tree network was formulated in Ref. 53. In this work, the tree tensor network is formed by placing physical sites on the boundary sites only. The remaining interior sites are virtual; they are only used to transfer entanglement up the tree. This tree tensor product state is designed to treat models in which
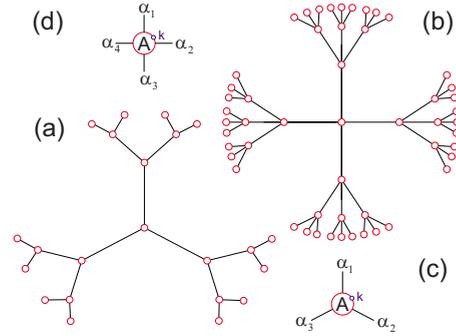


FIG. 1. (Color online) Tree tensor network with (a) $z=3$ and (b) $z=4$. The structure of the tensors is shown in (c) and (d). The bonds indicate the virtual indices $\alpha_1, \ldots, \alpha_z$ and the circle the physical index $k$.

sites contribute the same amount of entanglement, i.e., have the same value of single-site entropy.

In this work, we form a tree tensor network in which all sites in the tree represent physical sites and in which entanglement is transferred via the virtual bonds that connect the sites. Our motivation is to treat models in which physical sites have varying degrees of entanglement; positions closer to the center of the tree should be better suited to represent more entangled sites. An additional motivation is to take advantage of the property of the tree tensor network ansatz that the long-range correlations differ from the mean-field value polynomially with distance rather than exponentially with distance as for MPS. In our algorithmic approach to optimize the tree tensor network, we use tools similar to those used in Refs. 35 and 52–55, and optimize the network site-by-site as in the DMRG. In addition, we explicitly describe how to deal with fermions and long-range interactions.

The paper is organized as follows. In Sec. II, we describe the theoretical background for the TTNS algorithm and the orbital optimization used during the iterative procedure. Section III is devoted to the analysis of the error in the ground-state energy for various spin and fermion models as a function of bond dimension for TTNS with different coordination numbers. These include the 2D Heisenberg model, the 2D spinless fermion model, the 1D spinless fermion model in momentum space, and the Beryllium atom as an application in quantum chemistry. In the latter case, we compare the results to those from DMRG calculations. We conclude and discuss future prospects in Sec. IV.

## II. THEORETICAL BACKGROUND OF THE TREE-NETWORK APPROACH

We approach the problem of finding the ground state of strongly correlated systems on a finite lattice of $M$ sites with long-range interactions using a TTNS ansatz of the form

$$|\Psi\rangle = \sum_{k_1,\ldots,k_M} C_{k_1,\ldots,k_M} |k_1, \ldots, k_M\rangle.$$

Here the coefficients $C_{k_1,\ldots,k_M}$ describe a tree tensor network, i.e., they emerge from contractions of a set of tensors $\{A_1, \ldots, A_M\}$ according to a tree network, as shown in Fig. 1.

We associate a tensor with $z+1$ indices,

$$[A_m]^k_{\alpha_1,\ldots,\alpha_z}$$

with each vertex $m$ of the network, that is, each tensor has $z$ virtual indices $\alpha_1,\ldots,\alpha_z$ of dimension $D$ and one physical index $k$ of dimension $d$, with $z$ being the coordination number of that site. In the following, $D$ will be referred to as the *virtual dimension*, and $d$ as the *physical dimension*. The coefficients $C_{k_1,\ldots,k_M}$ are obtained by contracting the virtual indices of the tensors according to the scheme of a tree tensor network (see Fig. 1). The structure of the network can be arbitrary and the coordination number can vary from site to site. The only condition is that the network is bipartite, i.e., by cutting one bond, the network separates into two disjoint parts. In the special case $z=2$, the one-dimensional MPS-ansatz used in DMRG is recovered.

In a tensor network, entanglement is transferred via the virtual bonds that connect the sites. Thus, it is preferable to put strongly correlated sites close together, i.e., to minimize the number of bonds between them. Evidently, the diversity of arranging the sites in the network increases drastically with increasing coordination number $z$. Also, with a coordination number $z>2$ the number of virtual bonds required to connect two arbitrary sites scales logarithmically with the number of sites $M$ whereas the scaling is linear in $M$ for $z=2$.[35] This can be seen by considering a Cayley-tree of depth $\Delta$, as shown in Fig. 1. The number of sites in the tree is

$$M = 1 + z\sum_{j=1}^{\Delta}(z-1)^{j-1} = \frac{z(z-1)^{\Delta}-2}{z-2}$$

and thus, the maximal distance between two sites, $2\Delta$, scales logarithmically with $M$ for $z>2$. This logarithmic scaling is fundamental because, with a MPS ansatz, the expectation value of a long-range correlation differs from the mean-field result only by a quantity that decays exponentially with the distance:

$$\langle \tau_n \tau_{n+\Delta}\rangle - \langle \tau_n\rangle\langle \tau_{n+\Delta}\rangle \propto c^{-\Delta}.$$

With a TTNS ansatz, the logarithmic scaling counteracts this exponential decay, so that the difference from the mean-field result only scales polynomially with the number of sites $M$.

The way to arrange the physical sites on the network is determined by the choice of the basis $|k_1,\ldots,k_M\rangle$. Obviously, the precision of the ansatz depends critically on the choice of the basis. For example, the noninteracting Fermi gas has a ground state that is a direct product in the momentum space representation. Such a state corresponds to a tree tensor network state with virtual dimension $D=1$ at all bonds. In position representation, however, a bond dimension that increases exponentially with the number of sites would be required. Thus, it will be favorable to optimize not only over the tensors in the coefficients $C_{k_1,\ldots,k_M}$ but also over the basis $|k_1,\ldots,k_M\rangle$.

In second quantization, $|k_1,\ldots,k_M\rangle$ denotes the basis in occupation number representation,

$$|k_1,\ldots,k_M\rangle = (a_1^{\dagger})^{k_1}\cdots(a_M^{\dagger})^{k_M}|0\rangle$$

and a basis transformation is obtained from the canonical transformation

$$b_j^{\dagger}(U) = \sum_{r=1}^{M}U_{jr}a_r^{\dagger}$$

defined by the $M\times M$ unitary matrix $U$. The whole variational ansatz then depends on two sets of parameters: the tensors $\{A_1,\ldots,A_M\}$ and the unitary $U$. The goal of the algorithm is to find the minimum of the energy with respect to these two sets of parameters, i.e., to calculate

$$E = \min_{A_1,\ldots,A_M U}\langle\Psi(A_1,\ldots,A_M,U)|H|\Psi(A_1,\ldots,A_M,U)\rangle.$$

The idea is to perform the optimizations over the parameter sets $\{A_1,\ldots,A_M\}$ and $U$ consecutively and repeatedly until convergence is reached. In the following two sections, we describe the two optimization procedures in more detail.

### A. Network optimization

First, let us sketch how to optimize the tensors $\{A_1,\ldots,A_M\}$ while keeping the basis fixed. The minimization of the energy with the constraint that the norm remains constant is equivalent to minimizing the functional

$$F = \langle\Psi|H|\Psi\rangle - E(\langle\Psi|\Psi\rangle - 1).$$

This functional is nonconvex with respect to all parameters $\{A_1,\ldots,A_M\}$. However, due to the tensor network structure of the ansatz, it is quadratic in the parameters $A_m$ associated with one lattice site $m$. Because of this, the optimal parameters $A_m$ can simply be found by solving a generalized eigenvalue problem $\mathcal{H}_m\vec{A}_m = E\mathcal{N}_m\vec{A}_m$. For a bipartite network, it is always possible to assume a gauge condition so that $\mathcal{N}_m = \mathbb{I}$, and thus reduce the generalized eigenvalue problem to an ordinary one.[35] We will discuss this in more detail later in this section. The concept of the algorithm is to do this one-site optimization site-by-site until convergence is reached.

The challenge that remains is to calculate the effective Hamiltonian $\mathcal{H}_m$ of the eigenvalue problem. In principle, this is done by contracting all indices in the expression for the expectation value $\langle\Psi|H|\Psi\rangle$ except those that connect to $A_m$. By interpreting the tensor $A_m$ as a $dD^z$-dimensional vector $\vec{A}_m$, this expression can be written as

$$\langle\Psi|H|\Psi\rangle = \vec{A}_m^{\dagger}\mathcal{H}_m\vec{A}_m. \tag{1}$$

Since

$$\langle\Psi|\Psi\rangle = \vec{A}_m^{\dagger}\mathcal{N}_m\vec{A}_m \tag{2}$$

and $\mathcal{N}_m = \mathbb{I}$, the functional $F$ attains its minimum when

$$\mathcal{H}_m\vec{A}_m = E\vec{A}_m.$$

Due to the bipartite structure of the tensor network, the calculation of $\mathcal{H}_m$ can be performed efficiently, i.e., on a time that scales polynomially with $M$ and $D$. Assuming that the coordination number $z$ is the same everywhere, the scaling will be $MdD^{z+1}$.
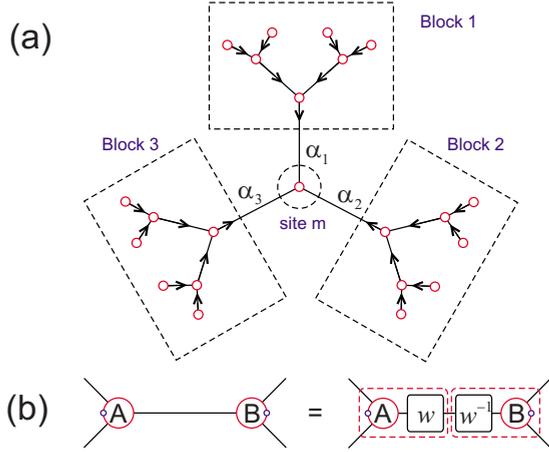
FIG. 2. (Color online) (a) Separation of the state into $z$ blocks plus the site under optimization, as described by Eq. (3). (b) Natural freedom in the tensor network: insertion of a matrix $w$ and $w^{-1}$ at one bond leaves the state invariant. The contraction of $A$ with $w$ forms the new tensor $A'$ on the left-hand side; the contraction of $B$ with $w^{-1}$ forms the new tensor $B'$ at the right-hand side.

This procedure is similar to a DMRG calculation with $z$ blocks instead of two, where a block consists of all of the sites within one of the branches emerging from site $m$ [see Fig. 2(a)]. The wave function is then formed as

$$|\Psi\rangle = \sum_{\alpha_1,\ldots,\alpha_z=1}^{D} |\varphi_{\alpha_1,\ldots,\alpha_z}\rangle \otimes |\phi_{\alpha_1}^1\rangle \otimes \cdots \otimes |\phi_{\alpha_z}^z\rangle, \quad (3)$$

where $|\phi_\alpha^i\rangle$ ($\alpha=1,\ldots,D$) is the basis in environment block $i$ ($i=1,\ldots,z$) and $|\varphi_{\alpha_1,\ldots,\alpha_z}\rangle$ is the state of site $m$. Since $\mathcal{N}_m$ is obtained from the norm $\langle\Psi|\Psi\rangle$ by contracting all tensors except $A_m$ [see Eq. (2)], it factorizes into a tensor product of $z$ matrices $\mathcal{N}_m^i$,

$$\mathcal{N}_m = \mathcal{N}_m^1 \otimes \cdots \otimes \mathcal{N}_m^z,$$

where each matrix $\mathcal{N}_m^i$ is formed by taking the overlap of the basis states in environment block $i$

$$[\mathcal{N}_m^i]_{\alpha\beta} = \langle\phi_\alpha^i|\phi_\beta^i\rangle.$$

Obviously, in order to guarantee that $\mathcal{N}_m=\mathbb{I}$, the basis in each environment block must be orthonormal. This is a similar requirement as in the DMRG. In the tree tensor network, this may be achieved by assuming an appropriate gauge condition. The purpose of this gauge condition is to fix the natural freedom in the tensor network that a matrix and its inverse can be inserted at any bond, with the matrix being contracted with the first attached tensor and the inverse being contracted with the second attached tensor, leaving the network invariant [see Fig. 2(b)]. The gauge condition for all sites $n$ ($n \neq m$) that assures that $\mathcal{N}_m=\mathbb{I}$ is

$$\sum_{k\beta_2,\ldots,\beta_z} [A_n^*]_{\alpha'\beta_2,\ldots,\beta_z}^k [A_n]_{\alpha\beta_2,\ldots,\beta_z}^k = \delta_{\alpha\alpha'}. \quad (4)$$

Here we take the index $\alpha$ to be outgoing, i.e., the branch attached to this index contains site $m$, and the indices

$\beta_2,\ldots,\beta_z$ to be incoming, i.e., the attached branches *do not* contain site $m$ [see Fig. 2(a)].

A stable way to apply this gauge condition is to "orthonormalize" the tensors from outside to inside. That is, starting with the tensors on the leaves of the tree, we take into account condition (4) by QR-decomposing $[A_n]_\alpha^k$, i.e., by writing it as

$$[A_n]_\alpha^k = \sum_{\alpha'} [Q_n]_{\alpha'}^k R_{\alpha'\alpha}.$$

The unitary matrix $[Q_n]_{\alpha'}^k$ is the new orthonormalized tensor at site $n$. In order to keep the tensor network invariant, $R_{\alpha'\alpha}$ must be contracted with the tensor on the first inner shell that is connected to the tensor at site $n$. This procedure continues iteratively with the tensors on the first inner shell, the second inner shell, and so on until site $m$ is reached.

Thus, by assuring that the gauge condition is always satisfied in the course of the algorithm, the only term that must be calculated is the effective Hamiltonian $\mathcal{H}_m$. This term is obtained, as can be gathered from Eq. (1), by contracting all tensors except $A_m$ in the expectation value $\langle\Psi|H|\Psi\rangle$. Since the Hamiltonian is a sum of interaction terms

$$H = \sum_{r=1}^{R} h_r$$

with $h_r$ being a tensor product of matrices, e.g., $h_r = \tau^{(7)} \otimes \tau^{(15)}$ for a two-body interaction acting on sites 7 and 15, the effective Hamiltonian splits into a sum of effective Hamiltonians $\mathfrak{h}_{m,r}$ with respect to a single interaction term $h_r$

$$\langle\Psi|h_r|\Psi\rangle = \vec{A}_m^\dagger \mathfrak{h}_{m,r} \vec{A}_m.$$

Due to the structure, Eq. (3), of the TTNS, each effective Hamiltonian $\mathfrak{h}_{m,r}$ factorizes into a tensor product of $z$ matrices

$$\mathfrak{h}_{m,r} = \mathfrak{h}_{m,r}^1 \otimes \cdots \otimes \mathfrak{h}_{m,r}^z,$$

where each matrix $\mathfrak{h}_{m,r}^i$ corresponds to the matrix elements of $h_r$ with respect to the basis in environment block $i$

$$[\mathfrak{h}_{m,r}^i]_{\alpha\beta} = \langle\phi_\alpha^i|h_r|\phi_\beta^i\rangle.$$

Graphically, the evaluation of $\langle\phi_\alpha^i|h_r|\phi_\beta^i\rangle$ corresponds to the contraction of a three-layered tensor network according to the structure of the branch in block $i$, as depicted in Fig. 3. This network can be contracted efficiently by starting from the leaves and working in the inward direction. The numerical effort for contracting one additional site into the network scales as $dD^{z+1}$, so that the total effort scales as $dD^{z+1}$ times the number of sites in the block.

Of course, if $h_r$ has no support in environment block $i$, $\mathfrak{h}_{m,r}^i$ is equal to the identity because the basis is chosen to be orthogonal in each environment block. Because of this, the calculation simplifies significantly. For example, each two-site interaction has support in at most two blocks. This means that at most two effective Hamiltonians $\mathfrak{h}_{m,r}^i$ have to be calculated (for each interaction term); all others are equal to the identity. Since the orthonormalization of the state is applied iteratively from the leaves inwards to the optimized
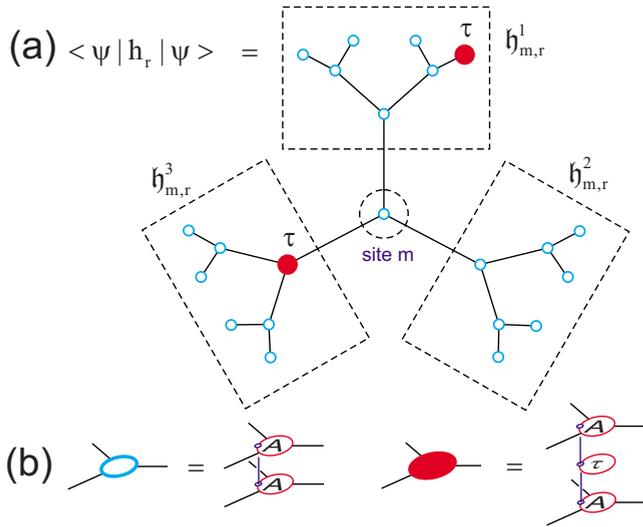
FIG. 3. (Color online) (a) Formation of the effective Hamiltonian $\mathfrak{h}_{m,r} = \mathfrak{h}^1_{m,r} \otimes \mathfrak{h}^2_{m,r} \otimes \mathfrak{h}^3_{m,r}$ with respect to the interaction $h_r = \tau^{(7)} \otimes \tau^{(15)}$. The sites on which the interaction has support are marked by a filled circle. Each open (filled) circle in the tensor network corresponds to the contraction of the layered structure of tensors shown in (b).

site $m$, the calculation of $\mathfrak{h}^i_{m,r}$ simplifies substantially, as well. Within each block, the contraction of all subbranches on which $\mathfrak{h}^i_{m,r}$ has no support automatically yields the identity. Thus, for the example of a two-site interaction, it is sufficient to take into account the sites on the path connecting the two sites (see Fig. 4). The treatment of long-range interactions on a tree is therefore not significantly more complicated than the treatment of long-range interaction on a chain in the DMRG.

At first glance, it might seem that the advantage of "cropping" all subbranches with no support is lost when switching to fermions. This is because fermionic interactions with local support are turned into interactions with nonlocal support in the spin picture via the Jordan-Wigner transformation. The two-particle fermionic interaction $a^\dagger_7 a_{15}$, for example, is turned into the spin interaction
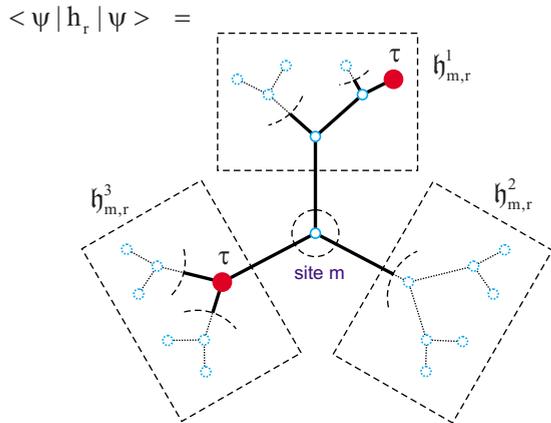


FIG. 4. (Color online) Branches with no support (marked by dotted lines and circles), which yield the identity when contracted and therefore do not have to be taken into account in calculating the effective Hamiltonian.
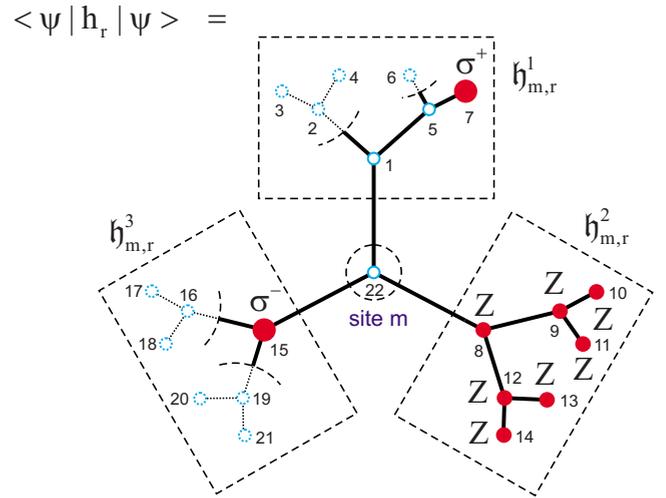


FIG. 5. (Color online) Formation of the effective Hamiltonian $\mathfrak{h}_{m,r} = \mathfrak{h}^1_{m,r} \otimes \mathfrak{h}^2_{m,r} \otimes \mathfrak{h}^3_{m,r}$ with respect to the fermionic interaction $h_r = a^\dagger_7 a_{15}$. The sites on which the interaction has support are marked in red.

$$a^\dagger_7 a_{15} = \sigma^{(7)}_+ \left[ \prod_{7<n<15} Z^{(n)} \right] \sigma^{(15)}_-$$

that has support on 9 sites, where $Z = -\sigma_z$ and $\sigma_+$, $\sigma_-$, $\sigma_z$ denote the Pauli operators (see Fig. 5). However, as we will show, local fermionic interactions can be treated in the tree network with the same effort as local spin interactions by including the $Z_2$ symmetry in the ansatz. This is the fermion number parity in the fermionic language.

The $Z_2$ symmetry can be incorporated into the ansatz by making the tensors $[A_m]^k_{\alpha_1,\dots,\alpha_z}$ block diagonal. That is, each virtual index is split into an index pair $(\alpha_i, p_i)$, with $p_i$ carrying the parity information. By stipulating that $p_1 \oplus \cdots \oplus p_z \oplus k = 1$ (where $\oplus$ denotes summation modulo 2), it is possible to "move" the operator $Z$ that acts on the physical index of $A_m$ to the virtual parity indices $p_i$

$$Z_{k\tilde{k}}[A_m]^{\tilde{k}}_{\alpha_1 p_1 \cdots \alpha_z p_z} = [A_m]^k_{\alpha_1 \tilde{p}_1 \cdots \alpha_z \tilde{p}_z} Z_{\tilde{p}_1 p_1} \cdots Z_{\tilde{p}_z p_z}. \quad (5)$$

From this relation, it immediately follows that:

$$Z \otimes \cdots \otimes Z |\Psi\rangle = |\Psi\rangle \quad (6)$$

and thus that the state is $Z_2$ symmetric. This is because the left-hand side of Eq. (6) corresponds to the contraction of $Z$'s to the physical indices of all tensors $A_m$. After moving all $Z$'s to the virtual bonds, the operator $Z$ appears twice on each bond. Thus since $Z^2 = \mathbb{I}$, the state $|\Psi\rangle$ remains unchanged.

Using the same idea, we can immediately enforce a more restrictive symmetry that is fulfilled by all fermionic Hamiltonians: the charge symmetry $U(1)$, i.e. the conservation of the number of particles. For this, the tree graph has to be made directed (see Fig. 6), so that all sites (except one) have $z-1$ incoming and one outgoing bond. The exception is the sink site, which has only incoming bonds. Thus, each virtual index of a tensor $A_m$ is equipped with the additional information of whether it is "incoming" or "outgoing." We assume that the index 1 is always the outgoing index in the
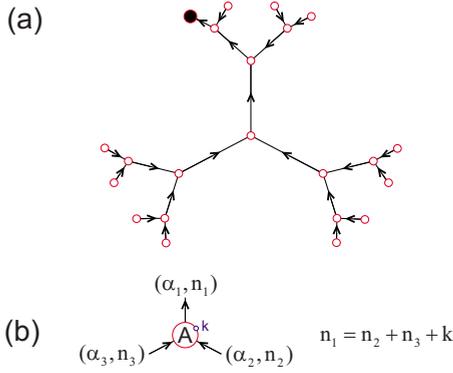
FIG. 6. (Color online) (a) Possible choice of an ordered tree graph with each site having two incoming and one outgoing bond. The sink site is marked in black. The tensors $A_m$ associated with each site with virtual indices consisting of pairs $(\alpha_i, n_i)$ have the structure shown in (b).

following. As before, each virtual index is split into an index pair $(\alpha_i, n_i)$. In addition, we require that $n_1 = n_2 + \cdots + n_z + k$. Thus, for $i = 2, \ldots, z$, $n_i$ counts the number of particles within the branch attached to index $i$. The index $n_1$, on the other hand, is equal to the number of particles in all the branches plus the number of particles at site $m$. Since the parity can be immediately derived from the particle-number information, a similar relation as Eq. (5) holds, namely,

$$Z_{k\tilde{k}}[A_m]^{\tilde{k}}_{\alpha_1 n_1 \cdots \alpha_z n_z} = [A_m]^{k}_{\alpha_1 \tilde{n}_1 \cdots \alpha_z \tilde{n}_z} \widetilde{Z}_{\tilde{n}_1 n_1} \cdots \widetilde{Z}_{\tilde{n}_z n_z} \qquad (7)$$

with $\widetilde{Z}_{\tilde{n}n} = \delta_{\tilde{n}n}(-1)^n$. Thus, as before, $Z$ acting on the physical index can be moved to the virtual particle-number bonds and since $\widetilde{Z}^2 = \mathbb{I}$, the state is $Z_2$ symmetric.

The advantage of taking into account the particle number is twofold: on the one hand, the block structure of the tensors reduces the numerical effort considerably: a bond dimension of $\chi = ND$ can be treated with an effort of order $N^z D^{z+1}$ instead of $\chi^{z+1}$ for a nonsymmetric ansatz. Here $\chi$ is the full bond dimension, so it is equivalent to the number of block states kept in a DMRG procedure. On the other hand, the calculation of the effective Hamiltonians $\mathfrak{h}_{m,r}$ with respect to an interaction with support on a few sites only, e.g., a two-particle or four-particle interaction, is simplified. The main idea is depicted in Fig. 7 for the interaction $a_7^\dagger a_{15}$: with an appropriately chosen numbering of the fermions, each sub-branch that has no fermionic support either has only identities acting on the sites or only operators $Z$. The subbranches including only identities simplify to the identity because of the orthonormalization of the state. The $Z$ operators can be moved, using relation (7), to the virtual bonds, and all of them except one cancel (see Figs. 5 and 7). What remains is a subbranch that includes only identities, which reduces to the identity because of the orthonormalization of the state. Thus, for a fermionic two-site interaction, it is sufficient to take into account the path connecting the two sites. In this way, the treatment of long-range fermionic interactions is feasible with the same numerical effort as the treatment of long-range spin interactions.
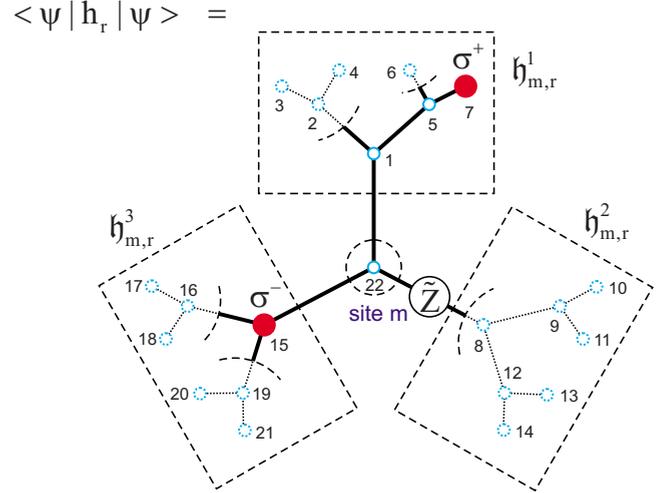


FIG. 7. (Color online) Formation of the effective Hamiltonian $\mathfrak{h}_{m,r} = \mathfrak{h}^1_{m,r} \otimes \mathfrak{h}^2_{m,r} \otimes \mathfrak{h}^3_{m,r}$ with respect to the fermionic interaction $h_r = a_7^\dagger a_{15}$ with parity symmetry taken into account. The sites on which the interaction has support are marked in red. All branches marked by dotted lines and circles yield the identity when contracted. The parity operator $\widetilde{Z}$ is contracted to the virtual bond.

Using the same scheme as for minimizing the energy, the efficient simulation of time evolution is also possible. For this, the time evolution is split up into small steps of duration $\delta t$. For $\delta t \ll 1$ and a starting state $|\Psi_0\rangle$ of the form of a TTNS, the time-evolved state $e^{-iH\delta t}|\Psi_0\rangle$ is a TTNS as well. However, the virtual dimension is multiplied by a factor $\kappa > 1$. In order to prevent the bond dimension from increasing exponentially with time, the TTNS has to be approximated by a TTNS with a reduced virtual dimension $|\Psi\rangle$ after every time step, i.e., the functional

$$K = \| e^{-iHt}|\Psi_0\rangle - |\Psi\rangle \|^2$$

must be minimized. The optimization with respect to a single site $m$ leads to the system of linear equations $\mathcal{N}_m \vec{A}_m = \vec{w}_m$ (where $\mathcal{N}_m$ can again be set equal to identity by using the appropriate gauge) with

$$\langle \Psi | e^{-iHt} | \Psi_0 \rangle = \vec{A}_m^\dagger \vec{w}_m$$

and thus can be performed efficiently with the scaling $MD^{z+1}$, as before. Clearly, all previous considerations regarding particle-number symmetry can also be adapted to the case of time evolution.

### B. Orbital optimization

As mentioned previously, the entanglement properties of the system depend critically on the choice of the basis. Our goal is to find a basis in which entanglement is localized as much as possible at the sites of the tree network. Such a choice of basis guarantees that a given precision can be attained with a smaller virtual dimension $D$, and thus with less computational effort. In QC-DMRG the optimization of the basis is fundamental and has been used in much previous work.[34,56–59]

In contrast to previous work, our approach aims to find the optimal basis that can be obtained by a canonical transformation of the fermionic modes. The canonical transformation is defined by a $M \times M$ unitary matrix $U$. Since the number of parameters is relatively small, a gradient search applied to the expectation value of the energy,

$$E(U) = \langle \Psi(U) | H | \Psi(U) \rangle$$

is certainly feasible. Since $E(U)$ is a nonconvex function of the parameters $U$, it is a highly nontrivial problem to find the absolute minimum. The idea is to perform the gradient search multiple times in the course of the algorithm, e.g., after each optimization sweep of the tensor network, and improve the energy at each gradient search by only a small amount. This will eventually adapt the orbitals optimally to the tree tensor network.

There are two ways to implement the basis transformation: one based on the state and the other based on the Hamiltonian. We have applied the basis transformation to the Hamiltonian. For the fermionic Hamiltonian with long-range interactions,

$$H = \sum_{ij} T_{ij} a_i^\dagger a_j + \sum_{ijkl} V_{ijkl} a_i^\dagger a_j^\dagger a_k a_l$$

that appears, e.g., in quantum chemistry, in momentum-space descriptions of the Hubbard model, or in descriptions of the Hubbard model on higher dimensional lattices, the function $E(U)$ can be expressed as

$$E(U) = \sum_{ij} \widetilde{T}(U)_{ij} \langle a_i^\dagger a_j \rangle + \sum_{ijkl} \widetilde{V}(U)_{ijkl} \langle a_i^\dagger a_j^\dagger a_k a_l \rangle$$

with

$$\widetilde{T}(U) = UTU^\dagger,$$

$$\widetilde{V}(U) = (U \otimes U) V (U \otimes U)^\dagger.$$

The correlation functions $\langle a_i^\dagger a_j \rangle$ and $\langle a_i^\dagger a_j^\dagger a_k a_l \rangle$ are calculated with respect to the original state and are not dependent on the parameters in $U$. With the function $E(U)$ in this form, its gradient can be calculated explicitly. Both quantities can be evaluated efficiently for different parameter sets $U$, which makes the gradient search feasible.

In order to guarantee that $U$ remains unitary in the course of the gradient search, we express $U$ as the exponential of an $M \times M$ hermitian matrix $X$, $U = e^{iX}$, and perform the gradient search with respect to the $M^2$ real parameters that parameterize $X$, compactly described as a vector $\vec{x}$ of length $M^2$. The gradient search that is used for the orbital optimization typically finds the local minima in the vicinity of the starting point of the optimization. In order to avoid falling into the same minimum, we shift the starting point in a random direction by an appropriately chosen amount, i.e., we add a random vector to the starting vector $\vec{x}_0$. This has the consequence that the gradient search generally falls into another local minimum that may have a higher energy. However, it assures that the orbitals change considerably and that since
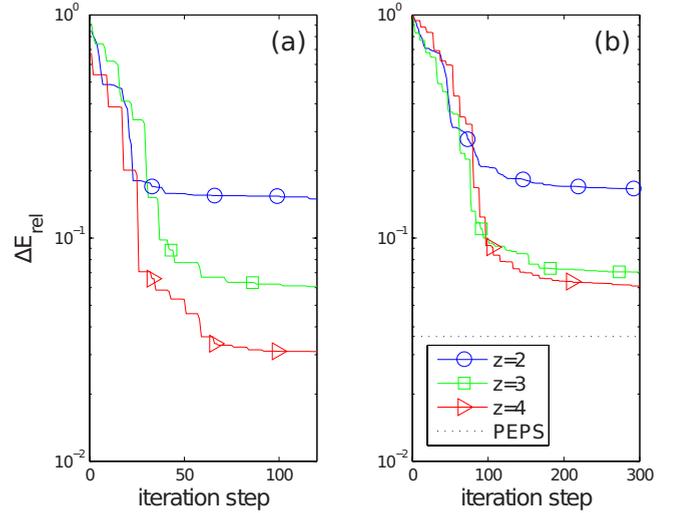


FIG. 8. (Color online) Relative error $\Delta E_{rel} = |(E - E_{exact})/E_{exact}|$ in the energy for Heisenberg model on (a) a $4 \times 4$ lattice and (b) a $6 \times 6$ lattice as a function of the optimization steps. The reference energy $E_{exact}$ was obtained by exact diagonalization in the $4 \times 4$ case and by Monte Carlo simulations using Algorithms and Libraries for Physics Simulations (ALPS) [60] in the $6 \times 6$ case. The calculations were performed with a fixed virtual dimension of $D = 4$ and tree tensor networks with coordination numbers $z=2$ (blue curve), $z=3$ (green curve), and $z=4$ (red curve).

the orbital optimization is performed repeatedly interspersed with network optimization, the energy decreases notably in the course of the algorithm.

## III. NUMERICAL RESULTS

We have applied the algorithm consisting of two optimization procedures described above, the optimization of the tensor network and the optimization of the basis according to this network, to several models. In this section, we discuss the results.

### A. 2D Heisenberg model

First, let us consider the tree tensor network optimization only and show using the 2D Heisenberg model with open boundary conditions how results improve with a TTNS ansatz compared to a one-dimensional MPS ansatz. The relative error in the energy of a system consisting of $4 \times 4$ spins as a function of the optimization step is shown in the left hand side of Fig. 8 for the MPS ansatz and the TTNS ansatz with $z=3$ and $z=4$. In all calculations, a fixed virtual dimension of $D=4$ is used. The manner in which the spins have been assigned to the nodes in the tree network is shown in Fig. 9. As can be seen, the precision increases considerably with increasing coordination number. As a reference we used results obtained by exact diagonalization and by Monte Carlo simulations using Algorithms and Libraries for Physics Simulations (ALPS).[60]

For larger systems such as for $6 \times 6$ spins, shown in the right-hand side of Fig. 8, the energy is plotted as a function of the optimization steps for $z=2$, $z=3$, and $z=4$. The virtual
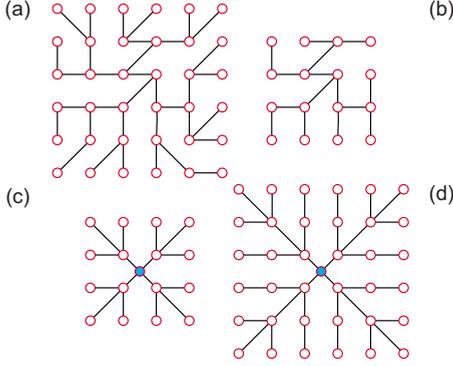
FIG. 9. (Color online) Tree structures used for the study of models on a rectangular lattice. Structures with coordination number $z=3$ mapped on a rectangular lattices of size $4 \times 4$ and $6 \times 6$ are shown in (a) and (b), the structures with $z=4$ is depicted in (c) and (d). The sites marked in blue are virtual sites with no corresponding particle or mode.

dimension is fixed to $D=4$. As before, the energy decreases considerably and approaches the PEPS result for $D=4$, indicated by the dotted line which, however, is obtained with a much larger numerical effort.[26]

### B. 2D interacting spinless fermions

The TTNS Ansatz also performs well on two-dimensional fermionic models. We have applied the TTNS algorithm to a system of interacting fermions on a two-dimensional lattice, described by the Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} a_i^\dagger a_j + U \sum_{\langle i,j \rangle} \hat{n}_i \hat{n}_j$$

with $\hat{n}_i = a_i^\dagger a_i$. The boundary conditions are assumed to be periodic. As can be gathered from Fig. 10, the ground-state energy improves with increasing coordination number $z$. The figure shows the relative error in the energy as a function of the iteration step for $z=2$ and $z=3$ for a $4 \times 4$-lattice in the left part, and the ground-state energy as a function of the iteration step for a $6 \times 6$ lattice in the right part. For the calculations, we have used the virtual dimension $D=4$. We have chosen the parameters $J=1$ and $U=0.5$ and have fixed the number of fermions to $N=3$.

Thus, a TTNS Ansatz might be useful for the study of higher dimensional models of small size because the effective long-range interactions are represented better in a tree than in a chain and the numerical effort is relatively low compared to a PEPS calculation.

### C. 1D interacting spinless fermions

In order to assess the effectiveness of the orbital optimization, we have applied the algorithm to a simple fermionic model, the one-dimensional interacting spinless fermion model
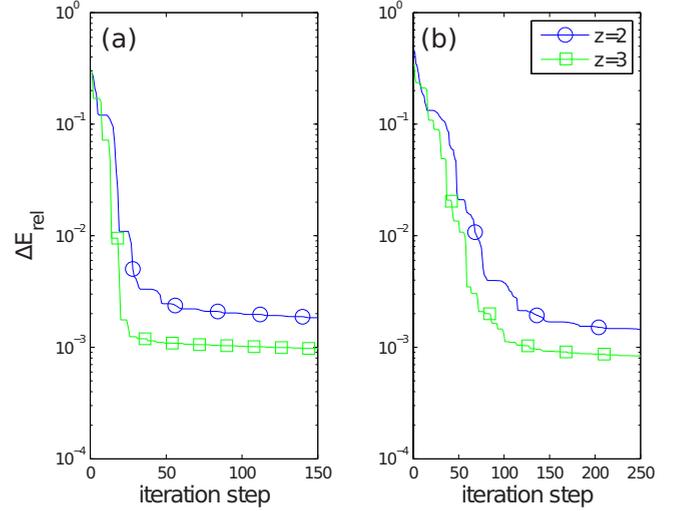


FIG. 10. (Color online) Relative error $\Delta E_{rel} = |(E - E_{exact})/E_{exact}|$ in the energy for the interacting spinless fermion model on (a) a $4 \times 4$-lattice and (b) a $6 \times 6$-lattice as a function of the optimization steps. The calculations were performed with a fixed virtual dimension of $D=4$ and tree tensor networks with coordination numbers $z=2$ (blue curve) and $z=3$ (green curve). The chosen parameters are $J=1$, $U=0.5$ and the number of fermions is fixed to $N=3$.

$$H = -J \sum_{i=1}^{M} (a_i^\dagger a_{i+1} + \text{h.c.}) + U \sum_{i=1}^{M} \hat{n}_i \hat{n}_{i+1},$$

where $\hat{n}_i = a_i^\dagger a_i$. We assume periodic boundary conditions, i.e., $a_{M+1} = a_1$. This model can be mapped to the XXZ spin chain via a Jordan-Wigner transformation. In this model, it is known that the choice of the basis has a big effect on the precision of the DMRG or TTNS calculation. For $U \to \infty$, the ground state is a product state in the position representation, and thus optimally represented by a tensor network with $D=1$ in this basis. For $U=0$, on the other hand, the ground state can be represented as a direct product in momentum space. Thus, the momentum-space basis is clearly best suited in this limit. For intermediate $U$, one might expect a basis intermediate between real and momentum space to represent the entanglement properties optimally; our goal is to find such a basis automatically by carrying out orbital optimization.

The results of calculations incorporating the optimization procedure are displayed in Fig. 11 for $z=2$ and $z=3$, both on systems of $M=7$ sites and $N=3$ particles with a fixed virtual dimension of $D=2$ and parameters $U=1$. As before, the relative error in the energy as a function of the optimization step is plotted. For comparison, calculations performed in the position representation and in the momentum representation are also shown. As can be seen, the energy improves significantly in the course of the optimization; the improvement is three orders of magnitude for the $z=3$ case.

### D. Quantum chemical systems

In this section, we compare numerical results for quantum chemical systems obtained using the QC-DMRG and TTNS
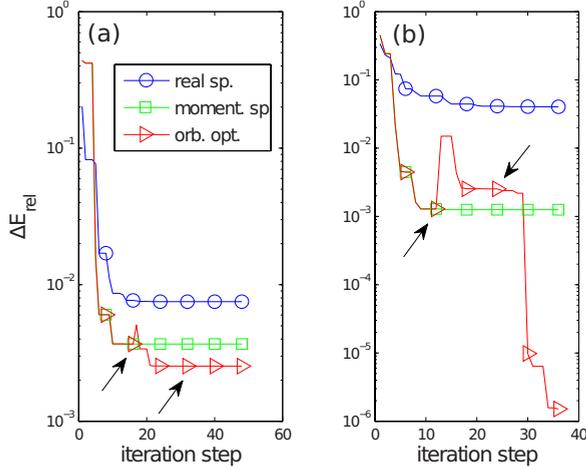
FIG. 11. (Color online) Relative error $\Delta E_{rel}=|(E-E_{exact})/E_{exact}|$ in the energy for interacting spinless fermions with $M=7$ sites as a function of the optimization steps for the TTNS ansatz with (a) $z=2$ and (b) $z=3$. Here $U=1$, the particle number is fixed to $N=3$, and the virtual dimension is assumed to be $D=2$. Results with orbital optimization are compared to the results obtained in the real-space basis and in the momentum-space basis. The steps at which orbital optimizations are performed are marked with arrows.

methods. In these applications, the electron-electron correlation is taken into account by an iterative procedure that minimizes the Rayleigh quotient corresponding to the Hamiltonian describing the electronic structure of the molecule, given by

$$H = \sum_{ij\sigma} T_{ij} a_{i\sigma}^{\dagger} a_{j\sigma} + \sum_{ijkl\sigma\sigma'} V_{ijkl} a_{i\sigma}^{\dagger} a_{j\sigma'}^{\dagger} a_{k\sigma'} a_{l\sigma} \quad (8)$$

and thus determines the full-CI wave function. In Eq. (8), $T_{ij}$ denotes the matrix elements of the one-particle Hamiltonian, which is comprised of the kinetic energy and the external electric field of the nuclei, and $V_{ijkl}$ stands for the matrix elements of the electron repulsion operator, defined as

$$V_{ijkl} = \int d^3 x_1 d^3 x_2 \Phi_i^*(\vec{x}_1) \Phi_j^*(\vec{x}_2) \frac{1}{\vec{x}_1 - \vec{x}_2} \Phi_k(\vec{x}_2) \Phi_l(\vec{x}_1).$$

We obtain the Hartree-Fock orbitals in a given basis of Gaussian orbitals and transform the matrix elements $T_{ij}$ and $V_{ijkl}$ to the Hartree-Fock basis using the MOLPRO program package,[61] which we also use to obtain the full-CI energies used as a benchmark.

In the QC-DMRG, a one-dimensional chain is built up from the atomic or molecular orbitals obtained from a suitable mean-field or multiconfiguration self-consistent-field calculation. The tree network is constructed similarly but there is greater freedom to form the proper structure of the network. The two-orbital mutual information[15] provides a good starting configuration. A general approach to reduce entanglement is to form the network by placing the highly correlated orbitals at or near the center of the tree and less correlated orbitals at the boundary sites.
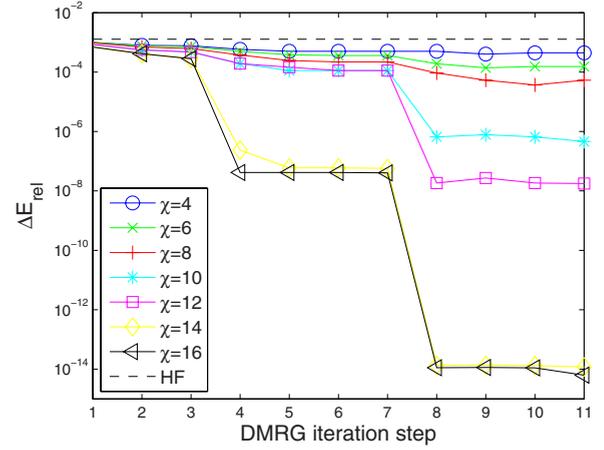


FIG. 12. (Color online) The relative error $\Delta E_{rel}=|(E-E_{exact})/E_{exact}|$ in the ground-state energy for the Beryllium atom as a function of DMRG iteration steps for various values of the number of DMRG block states. The dashed line corresponds to the Hartree-Fock energy.

In Fig. 12, we plot the relative error in the ground-state energy for the Beryllium atom as a function of DMRG iteration steps for various fixed values of the DMRG block states. Corresponding data gathered after the fourth DMRG sweep is shown in Fig. 14. In this calculation, four electrons have been placed on eight orbitals. The Hartree-Fock energy is $-14.351880250000$ while the full-CI energy is $-14.37016558404629$. Figure 13 depicts the relative error as a function of the optimization step for the Be atom, calculated using a TTNS ansatz with coordination number $z=3$. The cases with and without orbital optimization are considered separately and the $z=2$ results are included for comparison. As can be seen, the relative error is considerably smaller
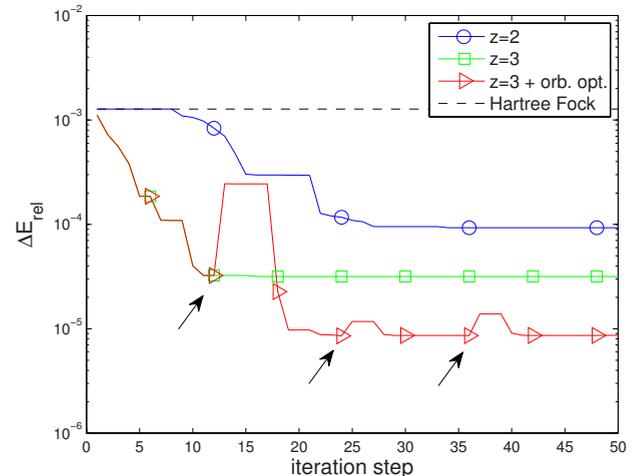


FIG. 13. (Color online) Relative error $\Delta E_{rel}=|(E-E_{exact})/E_{exact}|$ in the ground-state energy for the Be atom as a function of the optimization step for $D=2$. The results obtained using a TTNS ansatz with $z=3$ plus orbital optimization are compared to the TTNS-results without orbital optimization For comparison, the results with a MPS ansatz (corresponding to $z=2$) are also included. The steps at which orbital optimizations are performed are marked with arrows.
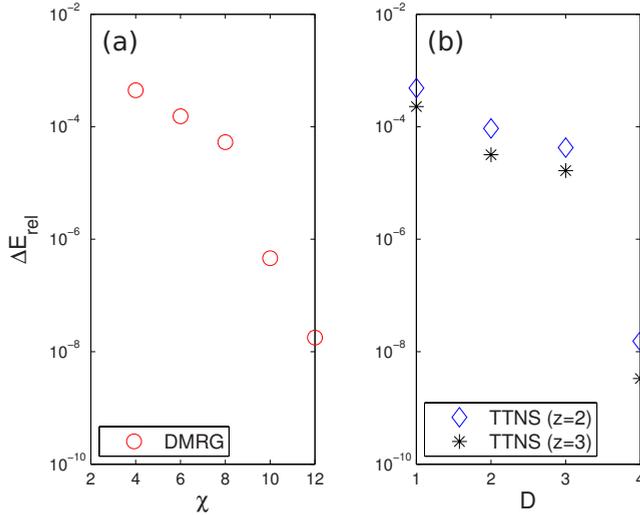
FIG. 14. (Color online) Relative error $\Delta E_{rel} = |(E - E_{exact})/E_{exact}|$ in the ground-state energy for the Beryllium atom as a function of the full bond dimension $\chi$ obtained using (a) the DMRG and (b) as a function of the virtual dimension $D$ obtained from a TTNS with $z=2$ and $z=3$.

for $z=3$ than for $z=2$, and there is a significant gain in precision when orbital optimization is taken into account. In Fig. 14, we display the dependence of the $z=2$ and 3 calculations on bond dimension $D$. As can be seen, the DMRG and TTNS calculations with $z=2$ yield similar accuracy while the $z=3$ TTNS calculation is significantly more accurate for a given bond dimension.

It should be noted that there is currently a discrepancy between the QC-DMRG and the TTNS calculations in the speed of convergence with optimization step. Since our QC-DMRG code is highly optimized, the calculation converges faster than that in the present version of the TTNS method. As can be seen in Fig. 12, for example, the QC-DMRG ground-state energies are always below the Hartree-Fock energy, and the active space is extended dynamically using CI-based expansion techniques (CI-DEAS) (Ref. 5) (DEAS). In general, there is no fundamental difficulty in incorporating optimizations such as the CI-DEAS into the TTNS method.

## IV. CONCLUSION

In this paper, we have described and applied a method, the tree tensor network state method, to treat strongly correlated systems with long-range interactions on a tree network with arbitrary coordination number $z$. Our approach is based on a tensor product state ansatz that generalizes a DMRG-like matrix product state to $z$ rather than two blocks. The number of virtual bonds required to connect two arbitrary sites scales logarithmically with the number of sites in TTNS, in contrast to the linear scaling of a one-dimensional topology.[35] In this sense, our TTNS method has a lower computational cost than currently used DMRG-based methods. We have also incorporated optimization of the single-particle orbitals in our method, treating the case of a procedure for transforming the basis that smoothly interpolates between real space and mo-
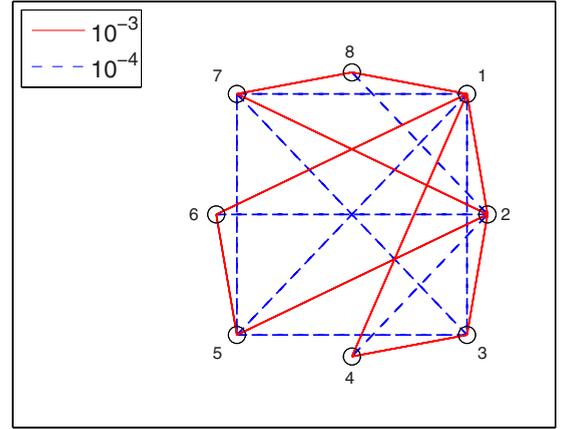


FIG. 15. (Color online) Graphical representation of the magnitude of the components of the two-dimensional entanglement matrix obtained from the two-site mutual information for the beryllium atom. For clarity, the one-dimensional chain of orbitals used in the DMRG calculation is plotted on a circle. Numbers next to the symbols label the ordering of the molecular orbitals within the chain.

mentum space. We have tested our method using numerical calculations on various systems with local and nonlocal interactions, including the two-dimensional Heisenberg lattice, the momentum-space version of the 1D interacting spinless fermion model, and small quantum chemical systems. For the quantum chemical systems, we have compared TTNS results to those of DMRG calculations.

Although we have performed calculations using fixed values of the coordination number, we expect that varying $z$ over the sites of the lattice will give rise to a better optimized lattice topology for quantum chemical systems. As an example, in Fig. 15 we display the components of the two-dimensional entanglement matrix for the beryllium atom, which we have obtained from the two-site mutual information calculated with the QC-DMRG method according to the recipe given in Ref. 15. As can be seen, sites 1 and 2 are highly entangled with the other sites while sites 4 and 6 are less entangled. The site-dependent coordination number can be optimized to reflect such an entanglement structure.

Since the TTNS approach is defined on a bipartite network, previous algorithmic developments and optimizations procedures developed in the context of the quantum chemistry version of DMRG can also be integrated into the TTNS method. Such optimizations include dynamic adjustment of the bond dimension,[5] orbital optimization,[34,56–59,62,63] and initialization procedures based on CI expansions,[11,14,64] among others. Incorporation of these elements into the TTNS approach will be carried out in future work. In light of the promising features of this method, we expect it to provide a viable alternate means of treating atoms and molecules efficiently in the near future.

---

*Also at Research Institute for Solid-State Physics and Optics, Hungarian Academy of Sciences, H-1121 Budapest, Hungary.

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[2] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[3] T. Xiang, Phys. Rev. B **53**, R10445 (1996).

[4] S. Nishimoto, E. Jeckelmann, F. Gebhard, and R. M. Noack, Phys. Rev. B **65**, 165114 (2002).

[5] Ö. Legeza and J. Solyom, Phys. Rev. B **68**, 195116 (2003).

[6] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1999).

[7] S. Daul, I. Ciofini, C. Daul, and S. R. White, Int. J. Quantum Chem. **79**, 331 (2000).

[8] A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, J. Chem. Phys. **115**, 6815 (2001).

[9] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).

[10] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **118**, 8551 (2003).

[11] Ö. Legeza, J. Röder, and B. Hess, Phys. Rev. B **67**, 125114 (2003).

[12] Ö. Legeza, J. Röder, and B. Hess, Mol. Phys. **101**, 2019 (2003).

[13] G. Moritz, B. A. Hess, and M. Reiher, J. Chem. Phys. **122**, 024107 (2005).

[14] G. Moritz, B. A. Hess, and M. Reiher, J. Chem. Phys. **124**, 034103 (2006).

[15] J. Rissler, R. Noack, and S. White, Chem. Phys. **323**, 519 (2006).

[16] K. Marti, I. M. Ondik, G. Moritz, and M. Reiher, J. Chem. Phys. **128**, 014104 (2008).

[17] J. Rodriguez-Laguna, J. Phys. A **40**, 12043 (2007).

[18] G. Vidal, J. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).

[19] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003); **93**, 040502 (2004).

[20] Ö. Legeza and J. Solyom, Phys. Rev. B **70**, 205118 (2004).

[21] F. Verstraete, J. I. Cirac, and V. Murg, Adv. Phys. **57**, 143 (2008).

[22] F. Verstraete and J. Cirac, Phys. Rev. B **73**, 094423 (2006).

[23] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 030504 (2008).

[24] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).

[25] F. Verstraete and J. Cirac, arXiv:cond-mat/0407066 (unpublished).

[26] V. Murg, F. Verstraete, and J. Cirac, Phys. Rev. A **75**, 033605 (2007).

[27] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. B **79**, 195119 (2009).

[28] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).

[29] H. Changlani, J. Kinder, C. Umrigar, and Garnet kin-lic Chan, Phys. Rev. B **80**, 245116 (2009).

[30] K. Marti, B. Bauer, M. Reiher, M. Troyer, and F. Verstraete, New J. Phys. **12**, 103008 (2010).

[31] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[32] M. Troyer and U. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[33] G. Barcza, Ö. Legeza, K. Marti, and M. Reiher, arXiv:1008.4607 (unpublished).

[34] H. G. Luo, M. P. Qin, and T. Xiang, arXiv:1002.1287 (unpublished).

[35] Y. Y. Shi, L. M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).

[36] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[37] M. A. Cazalilla and J. B. Marston, Phys. Rev. Lett. **88**, 256403 (2002).

[38] S. White and A. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).

[39] A. Daley, C. Kollath, U. Schollwoeck, and G. Vidal, J. Stat. Mech.: Theory Exp. 2004, P04005.

[40] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[41] W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (2000).

[42] R. Peters, T. Pruschke, and F. B. Anders, Phys. Rev. B **74**, 245114 (2006).

[43] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005).

[44] A. I. Toth, C. P. Moca, Ö. Legeza, and G. Zarand, Phys. Rev. B **78**, 245109 (2008).

[45] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. **99**, 076402 (2007).

[46] A. Holzner, A. Weichselbaum, and J. von Delft, Phys. Rev. B **81**, 125126 (2010).

[47] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[48] M. Eckstein, M. Kollar, K. Byczuk, and D. Vollhardt, Phys. Rev. B **71**, 235119 (2005).

[49] H. Otsuka, Phys. Rev. B **53**, 14004 (1996).

[50] B. Friedman, J. Phys.: Condens. Matter **9**, 9021 (1997).

[51] M. B. Lepetit, M. Cousy, and G. Pastor, Eur. Phys. J. B **13**, 421 (2000).

[52] M. A. Martin-Delgado, J. Rodriguez-Laguna, and G. Sierra, Phys. Rev. B **65**, 155116 (2002).

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

[54] P. Corboz and G. Vidal, Phys. Rev. B **80**, 165129 (2009).

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

[56] S. R. White, J. Chem. Phys. **117**, 7472 (2002).

[57] D. Zgid and M. Nooijen, J. Chem. Phys. **128**, 144116 (2008).

[58] D. Zgid and M. Nooijen, J. Chem. Phys. **128**, 014107 (2008).

[59] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **128**, 144117 (2008).

[60] F. Albuquerque *et al.*, Journal of Magnetism and Magnetic Materials **310**, 1187 (2007).

[61] MOLPRO, a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, version 2002.1, R. D. Amos, A. Bernhardsson, A. Berning *et al.*, see http://www.molpro.net.

[62] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).

[63] T. Yanai, Y. Kurashige, E. Neuscamman, and G. K.-L. Chan, J. Chem. Phys. **132**, 024105 (2010).

[64] G. Moritz and M. Reiher, J. Chem. Phys. **126**, 244109 (2007).