Automatic structural optimization of tree tensor networks

Toshiya Hikihara [1,*] Hiroshi Ueda [2,3,4] Kouichi Okunishi [5] Kenji Harada [6] and Tomotoshi Nishino [7]

[1] *Graduate School of Science and Technology, Gunma University, Kiryu, Gunma 376-8515, Japan*
[2] *Center for Quantum Information and Quantum Biology, Osaka University, Toyonaka 560-0043, Japan*
[3] *Precursory Research for Embryonic Science and Technology (PRESTO),*
Japan Science and Technology Agency (JST), Kawaguchi 332-0012, Japan
[4] *Computational Materials Science Research Team, RIKEN Center for Computational Science (R-CCS), Kobe 650-0047, Japan*
[5] *Department of Physics, Niigata University, Niigata 950-2181, Japan*
[6] *Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan*
[7] *Department of Physics, Graduate School of Science, Kobe University, Kobe 657-8501, Japan*

The tree tensor network (TTN) provides an essential theoretical framework for the practical simulation of quantum many-body systems, where the network structure defined by the connectivity of the isometry tensors plays a crucial role in improving its approximation accuracy. In this paper, we propose a TTN algorithm that enables us to automatically optimize the network structure by local reconnections of isometries to suppress the bipartite entanglement entropy on their legs. The algorithm can be seamlessly implemented to such a conventional TTN approach as the density-matrix renormalization group. We apply the algorithm to the inhomogeneous antiferromagnetic Heisenberg spin chain, having a hierarchical spatial distribution of the interactions. We then demonstrate that the entanglement structure embedded in the ground state of the system can be efficiently visualized as a perfect binary tree in the optimized TTN. Possible improvements and applications of the algorithm are also discussed.

I. INTRODUCTION

Tensor networks have been attracting growing interest in a variety of research fields, including condensed matter physics, quantum information, quantum cosmology, and data science [1–4]. Under the context of quantum many-body physics, tensor networks have provided flexible representations for low-energy states in correlated systems and have been employed in various theoretical approaches [5]. A tensor network framework had been used earlier in the field of statistical physics [6], but recent rapid developments in this field chiefly originated from the density matrix renormalization group (DMRG) method [7,8]. It turned out that the matrix product state (MPS) is variationally improved through the numerical sweeping processes in the DMRG method [9,10].

A couple of fundamental frameworks of the tensor network have been established during the first decade of this century, in which generalization of the network structure plays a key role. For instance, the tensor-product state [11,12], equivalently, the projected entangled-pair state [13,14], has been applied successfully to two-dimensional (2D) quantum lattice models

and three-dimensional statistical ones. The multiscale entanglement renormalization ansatz (MERA) revealed the intrinsic network structure for properly describing one-dimensional (1D) quantum critical states [15,16]. Moreover, the tree tensor network (TTN) framework has been established as a solid basis for practical numerical analyses in quantum chemical problems [17–22], equilibrium phase and ground state studies [23–32], quantum many-body systems with randomness [33–38], and information science [39–44]. Also, we note that the DMRG applied to Bethe lattice systems invokes the TTN framework [45–47].

Suppose that we have a quantum state described by a tensor network, which is a contraction of tensors. In general, a tensor leg that connects a pair of tensors, which is often mentioned as an auxiliary bond, with a finite dimension χ can be capable of representing the entanglement entropy (EE) up to $\ln \chi$. Thus, if n auxiliary bonds cross the boundary between a subsystem and the rest of the entire network, the bipartition EE can be $n \ln \chi$ at most. This upper bound controls the precision of the tensor network when it is used as an approximation of a certain quantum state. For practical tensor network approaches, the network structure is thus crucial in improving the accuracy.

For spatially uniform 1D quantum systems, the network structure of the MPS, i.e., the linear arrangement of three-leg tensors along the chain direction, seems to fit the lattice structure of their ground state wave functions. However, there is no guarantee, in general, that such an intuitive network structure is really suitable to the accurate representation of the

target state. Under the presence of long-range interactions or position-dependent interactions such as random couplings, it is even difficult to imagine a good network structure in advance. How can we resolve the situation? Formulations in the strong-disorder renormalization group (SDRG) [48,49] and its tensor network generalization (tSDRG) [33,34,36] provide us an insight into this fundamental question. In the tSDRG, the network structure of tensors is determined by the energy scale contained in the effective Hamiltonians, which are sequentially generated by successive real-space renormalization group (RG) transformations. Another variant of tSDRG that utilizes the EE in the construction of the TTN has also been clarified to generate slightly different network structure compared with the above tSDRG based on the energy scale [37]. Recently, Roy *et al.* [50–52] and Santalla *et al.* [53] introduced the concept of *emergent geometry*, that is, the entanglement structure visualized from the EEs for all the possible bipartitions of the system. The relation between the structure of entanglement in a quantum state and the geometry of the tensor network has also been explored [54,55].

As a step toward the construction of better tensor network structures, we focus on TTNs, which do not contain any loop in the diagrammatic representation. An important feature of the TTN is that the bipartition of any TTN is possible just by cutting an auxiliary bond between three-leg tensors, which we call isometries following the convention [15,16]. This means that each auxiliary bond bears the EE for the corresponding bipartition. If there is no limitation in the dimension of auxiliary bonds, every bond can carry any amount of EE so that the TTN with an arbitrary tree structure can, in principle, represent any quantum state exactly. However, in the practical situation where the bond dimension is bounded by χ, a large amount of EE exceeding $\ln \chi$ causes a loss of accuracy in representing the quantum state. This fact leads us to a guiding principle that one should select a TTN structure in which the EE at every bond is as small as possible so that one can minimize the loss of accuracy due to the restriction of the bond dimension. We call a TTN that satisfies this requirement of the least-EE principle the *optimal* one in the following. We note that the preclusion of the bonds with large EE can be partially achieved by reordering of the sites [56–60] or the readjustment of the network structure [18,19,61], based on the mutual information between sites. Larsson [22] introduced a scheme to reconstruct the local tree structure by monitoring the cutoff bond dimension. Very recently, Li *et al.* [62] examined the efficiency of the site reordering by the successive minimization of the bond EE and/or truncation error within the MPS. An entanglement bipartition approach to construct the optimal TTN for a given exact wave function of a small system has also been proposed [63].

In this paper, we propose an iterative algorithm that generates the optimal TTN structures automatically. More precisely, in the DMRG-like updating step of tensors, we determine the locally optimal connectivity of the isometry pair among three possible candidates to reduce the EE on the bond under consideration. During sweeps over the variational TTN wave function, a better network structure is thereby generated successively toward the optimal one through local reconnections of auxiliary bonds. We apply the algorithm to the inhomogeneous antiferromagnetic Heisenberg spin chain which
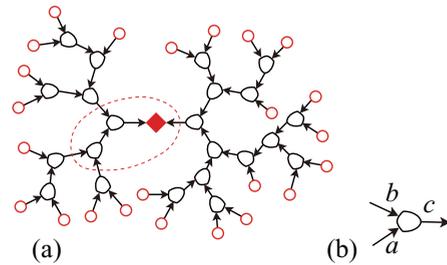
FIG. 1. (a) Schematic picture of the binary tree tensor network (TTN). The ovals (black) and circles (red) represent isometries and bare spin degrees of freedom, respectively. The arrows are the tensor legs, which we call the (auxiliary) bonds. The rhombus (red) represents the singular values. The part surrounded by the dotted curve (red) is the central area. (b) The isometry V_{ab}^c. Directions of the bonds are shown by arrows.

possesses hierarchical spatial modulation of the interaction constants and then confirm that the perfect binary tree network emerges automatically. Note that the sweeping process of the algorithm can be interpreted as a visualization of the entanglement structure embedded in the TTN.

The rest of this paper is organized as follows. We introduce the basic notations in the conventional variational TTN formulation in Sec. II. We explain the automatic optimization scheme of the TTN structure in Sec. III. In Sec. IV, we apply the proposed algorithm to the trial system and check the numerical validity of the algorithm. Section V is devoted to the summary and concluding remarks.

II. TTN

Before explaining the optimization scheme on the TTN structure, we shortly review fundamental concepts and notations in the conventional formulation of the variational TTN. Let us consider the binary TTN. Figure 1(a) shows the schematic diagram of a TTN, which can be also viewed as the TTN representation of the ground state wave function of a quantum spin model. In the figure, the oval and circle symbols represent isometries and bare spin degrees of freedom, respectively, and they are connected to other isometries or circles through arrows representing auxiliary bonds. As depicted in the diagram, the direction of the bond arrows flows to the *center* indicated by the rhombus symbol representing singular values. As a result, the network contains no loop. Also, we define the *central area* by the dotted curve surrounding the center of the network for later convenience.

As shown in Fig. 1(b), the isometry has two incoming bonds (arrows) and an outgoing one. Let us write tensor elements of the isometry as V_{ab}^c, where a and b are indices for incoming bonds, and c is that for the outgoing one. We distinguish tensors by the letters put on the bonds, if necessary. The isometry satisfies the orthonormal condition:

$$\sum_{ab} \bar{V}_{ab}^c V_{ab}^{c'} = \delta_{cc'}, \tag{1}$$

where $\bar{V}_{ab}^c$ is the complex conjugate of V_{ab}^c, and $\delta_{cc'}$ is the Kronecker delta. Throughout this paper, we assume that the number of bond degrees of freedom, i.e., the bond dimension
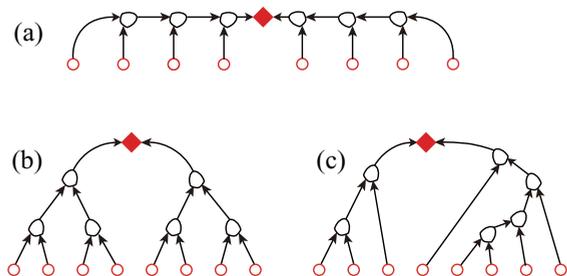
FIG. 2. A variety of tree tensor network (TTN) geometries.
(a) The matrix product network (MPN). (b) The perfect binary TTN.
(c) An example of the nonuniform TTN.

FIG. 3. Local improvement process. (a) The current central bond a and two isometries connected to the new central bond e. (b) The renormalized ground-state wave function $\tilde{\Psi}_{abcd}$. (c) Updated isometries connected to the new central bond e.

for a, b, and c, is bounded by a finite number χ. For the *central bond* attached with the rhombus, meanwhile, we assign k to specify its index for a moment. We then assume that singular values D_k are arranged in the descending order:

$$D_1 \geqslant D_2 \geqslant D_3 \geqslant \cdots \geqslant D_\chi (\geqslant 0). \tag{2}$$

Since the TTN state $|\Psi\rangle$ is written in the canonical form, the norm is expressed as

$$\langle \Psi | \Psi \rangle = \sum_{k=1}^{\chi} (D_k)^2 \tag{3}$$

because of the orthonormal condition of the isometry [64]. Here, we note that the position of the center can be moved to an arbitrary bond by means of fusion processes of tensors [64].

Given the Hamiltonian $\hat{H}$ of a system, we evaluate the variational energy:

$$E_{\text{var}} = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \tag{4}$$

where $|\Psi\rangle$ denotes a trial quantum state in terms of TTN. Here, the TTN representation of $|\Psi\rangle$ is not unique. Consider TTNs for a system with eight spins as typical examples. Figure 2(a) shows the TTN of the MPS form, which is a TTN established for uniform 1D quantum systems. We call this type of TTN the matrix product network (MPN) [65]. Note that, although the MPN has a simple chain structure, there remains room for various ordering of the site index [56–60,62,66,67]. Figure 2(b) is the perfect binary TTN with a hierarchical structure, which is often used in the framework of real-space RGs. Figure 2(c) shows a nonuniform TTN structure, which often appears in tSDRGs for random systems. These examples suggest that we have several options in choosing an appropriate TTN structure for obtaining better variational energy E_{var} of Eq. (4).

If the network structure of the TTN is fixed, the variational minimum of E_{var} can be obtained by the successive improvement of isometries, which is quite similar to that of the finite-system DMRG algorithm [18,19,61,64,68,69]. We briefly explain the local improvement procedure of the isometries at the center of the TTN, which is also essential for constructing the local reconnection algorithm of isometries discussed in Sec. III. Figure 3(a) shows the central area in the TTN to move the central bond from a to e. At first, the singular value D_a is located at the central bond a, where two
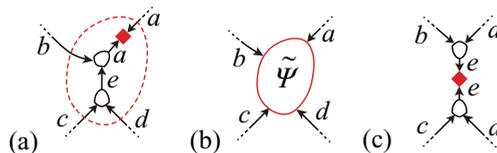
isometries V_{be}^a and V_{cd}^e are connected by the bond e, which will be a new central bond. The bonds a, b, c, and d have arrows incoming to the central area surrounded by the dotted curve, and then the isometries outside the central area define the corresponding environment. We can then construct the renormalized Hamiltonian $\tilde{H}$ by recursively applying the isometries in the environment to $\hat{H}$ from the boundary of the TTN toward the center. In practical computations, this RG transformation can be quickly completed with the reuse of the environments in the previous steps. Note that $\tilde{H}$ corresponds to the so-called *super-block* Hamiltonian in the DMRG. Diagonalizing $\tilde{H}$, we obtain the lowest eigenvalue and the corresponding eigenvector $\tilde{\Psi}_{abcd}$, which can be called the renormalized ground state wave function.

Figure 3(b) shows the diagram for $\tilde{\Psi}_{abcd}$. Rearranging the leg indices into ab and cd in the matrix form, we next perform the singular-value decomposition (SVD) to obtain the new tensors:

$$\tilde{\Psi}_{abcd} = \sum_e V_{ab}^e D_e V_{cd}^e, \tag{5}$$

where the right-hand side is diagrammatically illustrated in Fig. 3(c). Discarding tiny singular values, we keep the number of degrees of freedom for the new central bond e within the upper bound χ. If there is degeneracy in D_e, we keep all the degenerating singular values or discard all of them to maintain the symmetry of the state. Note that any basis truncation is not performed near the boundary of the TTN where the number of positive D_e does not exceed χ.

After the decomposition in Eq. (5), the tensor V_{be}^a in Fig. 3(a) is replaced by V_{ab}^e in Fig. 3(c), where the position of the singular value is also shifted to the bond e. Also the isometry V_{cd}^e can be updated. Using the above local update accompanying the position shift of the central area, we sweep the entire TTN with improving the isometries. After repeating this procedure several times, we finally obtain a good variational wave function.

III. TREE STRUCTURE OPTIMIZATION

Now we explain the variational scheme of automatically searching for the optimal TTN structure based on the conventional variational TTN algorithm above. An essential idea is that the local reconnection of bonds can be implemented in every improvement step of isometries inside the central area. Let us see $\tilde{\Psi}_{abcd}$ in Fig. 3(b) again. In Eq. (5), we arranged the collective indices of ab and cd according to the original network structure and then performed SVD, as shown
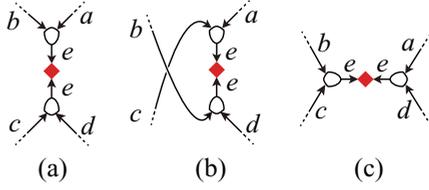
FIG. 4. Three possible applications of singular-value decomposition (SVD) to $\tilde{\Psi}_{abcd}$, as candidates for the local reconnection.

in Fig. 3(c). We draw this diagram again in Fig. 4(a). As in Figs. 4(b) and 4(c), however, there are the other two possible arrangements of the collective indices [64]. If we adopt ac and bd, the SVD yields

$$\tilde{\Psi}_{abcd} = \sum_e V^e_{ac} D'_e V^e_{bd}, \tag{6}$$

as shown in Fig. 4(b). If we choose ad and bc, we obtain

$$\tilde{\Psi}_{abcd} = \sum_e V^e_{ad} D''_e V^e_{bc}, \tag{7}$$

which is depicted as Fig. 4(c). These diagrams provide three options of the local connectivity of isometries in the TTN. Figures 4(b) and 4(c) accompany the reconnection of the network. Here, note that D_e in Eq. (5), D'_e in Eq. (6), and D''_e in Eq. (7) are different, but for all cases, the bond dimension of e is at most χ after the basis truncation.

To quantitatively evaluate the quality of the above three connections, we employ the EEs associated with their singular values. Let us write the arrangement in Eq. (5) as $(ab|cd)$. Then the corresponding EE is given by

$$\mathcal{S}^{(ab|cd)} = -\sum_e (D_e)^2 \ln(D_e)^2, \tag{8}$$

where we have assumed the normalization of the TTN. Similarly, we have the EEs:

$$\mathcal{S}^{(ac|bd)} = -\sum_e (D'_e)^2 \ln(D'_e)^2, \tag{9}$$

$$\mathcal{S}^{(ad|bc)} = -\sum_e (D''_e)^2 \ln(D''_e)^2, \tag{10}$$

for $(ac|bd)$ and $(ad|bc)$, respectively. We then select the connection having the least EE as the locally optimal connectivity of the isometries. Recall that, in the TTN, cutting a certain bond always has the corresponding bipartitioning of the entire wave function. The choice of the reconnection, which locally satisfies the least EE principle discussed in Sec. I, is thus expected to reduce the loss of accuracy caused by the limitation of finite bond dimension. This reconnection process is the heart of the proposed algorithm.

Once the new connectivity inside the central area is determined, we also update the corresponding isometry. Suppose that the connection of $(ad|bc)$ is selected, and the central bond is shifted to c in the next step, as in Fig. 5. Following Eq. (7), we first update the isometry V^e_{ad}, which is on the outside of the new central area illustrated as the dotted curve in Fig. 5. We next recalculate the block Hamiltonian and the
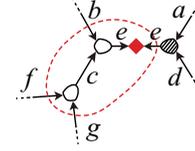
FIG. 5. The new central area in the case that Fig. 4(c) is chosen as the optimal connection and that c is treated as the central bond of the new step. The shaded oval represents V^e_{ad}, which has been updated.

spin operators belonging to V^e_{ad}, which turns out to be a part of $\tilde{H}$ in the next step. Meanwhile, we may skip the update of V^e_{bc} since it is included in the new central area that will be treated in the next step. Now we can perform the same updating processes by shifting the central area. The iterative algorithm summarized in Table I enables us to simultaneously optimize the variational energy E_{var} and the TTN structure by sweeping the central area over the entire TTN.

Here, it is worth mentioning an acceleration algorithm effective for the numerical diagonalization of $\tilde{H}$. So far, we have stored the wave function $\tilde{\Psi}_{abcd}$ and the updated isometry V^e_{ad} in computer memory. Also, we assume that the isometry V^c_{fg} was obtained previously. Then we calculate the contraction $\sum_{acd} V^c_{fg} \bar{V}^e_{ad} \tilde{\Psi}_{abcd}$, which can be a good initial vector for $\tilde{\Psi}_{bfge}$. Alternatively, if we have V^e_{bc} and D''_e explicitly, the initial vector can be constructed as $\sum_c V^c_{fg} V^e_{bc} D''_e$. These initial vectors may substantially reduce the computational cost of the Lanczos or the Davidson method to diagonalize $\tilde{H}$ in the next step [70].

In the following, we describe several comments on the technical aspect of the algorithm in order. The initial setup of the TTN is flexible since both the tree structure and the isometries will be optimized afterward. A MPN generated by the finite-system DMRG is one of the realistic candidates. It is even possible, in principle, to use a naïve product state, which can be regarded as a TTN with $\chi = 1$. However, a good initial setup may decrease the number of sweeps required for convergence. For random spin systems, the TTN generated by the tSDRG protocol [33–38] will be a reasonable initial TTN. Also, we note that the central area must move around all auxiliary bonds and isometries in the sweeping process of the network. In Appendix A, we explain the computational detail

TABLE I. Algorithm of automatic structural optimization for the TTN.

(1)	Construct the super-block Hamiltonian $\tilde{H}$ according to the current central area.
(2)	Diagonalize $\tilde{H}$ to obtain the renormalized ground state wave function $\tilde{\Psi}$.
(3)	Perform the SVD on $\tilde{\Psi}$ in three different manners, as shown in Fig. 4.
(4)	Choose the grouping with the smallest EE and update the tensors in the central area accordingly.
(5)	Move the central bond to the appropriate direction (see Appendix A).
(6)	Iterate the above steps until the entire TTN is updated.
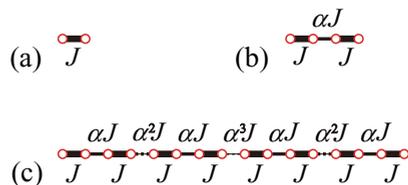
FIG. 6. Inhomogeneous interactions on the hierarchical chain when (a) $N = 2$, (b) $N = 4$, and (c) $N = 16$.

of this technical issue about the sweeping path of iterative computation.

While we employ the bond EE for the evaluation of the local network structure in the proposed algorithm, one may use other quantities such as the truncation error, i.e., the sum of the eigenvalues of the reduced density matrix (DM) for the discarded states, or the Renyi entropy [62]. The truncation error is directly related to the variational energy, while the bond EE can access the entanglement structure of the target state.

Since the reconnection of the isometries in the central area is always local, the optimization process of the TTN structure may be trapped at local minima of the distribution landscape of the EEs on the network. A possible device to escape from such trapping is to combine a stochastic method with the algorithm. More precisely, we can select one of the diagrams in Fig. 4 according to the relative probabilities based on the heat-bath method:

$$P^{(ab|cd)} \propto \exp[-\beta \mathcal{S}^{(ab|cd)}], \tag{11a}$$

$$P^{(ac|bd)} \propto \exp[-\beta \mathcal{S}^{(ac|bd)}], \tag{11b}$$

$$P^{(ad|bc)} \propto \exp[-\beta \mathcal{S}^{(ad|bc)}], \tag{11c}$$

where β is an effective inverse temperature. If β is gradually increased during the sweeping process, the TTN structure converges to the one where the sum of EE on all the bonds is minimized, which is a candidate for the optimal TTN. If one prefers to suppress the EE on each bond further, the square or higher power of $\mathcal{S}^{(ab|cd)}$, $\mathcal{S}^{(ac|bd)}$, and $\mathcal{S}^{(ad|bc)}$ can be used in the exponent of the right-hand side of Eq. (11). Note that the stochastic sampling of the tree structure can be parallelized.

IV. NUMERICAL RESULTS

We check the validity of the proposed optimization scheme in Table I by applying it to the inhomogeneous Heisenberg spin chain, whose Hamiltonian is formally written as

$$\hat{H} = \sum_{i=1}^{N-1} J_i \mathbf{S}_i \cdot \mathbf{S}_{i+1}, \tag{12}$$

where $\mathbf{S}_i = (S_i^x, S_i^y, S_i^z)$ represents the $S = \frac{1}{2}$ quantum spin on the ith site. We consider the case where the position-dependent exchange coupling $J_i > 0$ is recursively determined, as shown in Fig. 6. We call the model the hierarchical chain in the following. The minimum unit is a two-site system ($N = 2$) in Fig. 6(a), where the coupling constant between $\mathbf{S}_1$ and $\mathbf{S}_2$ is $J > 0$, which defines the unit of the energy scale. Joining two two-site units with the coupling αJ, we obtain a

four-site unit ($N = 4$) in Fig. 6(b). The parameter $0 < \alpha \leqslant 1$ controls the decay rate of the coupling. In general, joining the 2^n-site units ($N = 2^n$) with the coupling $\alpha^n J$, we obtain the 2^{n+1}-site system. Figure 6(c) shows the 2^4-site system as an example. We treat the system size $N = 64$ (for several α's) and $N = 128$ (for $\alpha = 0.50$) in the following numerical calculations.

A significant feature of the hierarchical chain is that, when α is sufficiently small, one can deduce its optimal TTN with the perturbative RG scheme. In the hierarchical chain, the largest exchange coupling is J, with which two spins $\mathbf{S}_{2l-1}$ and $\mathbf{S}_{2l}$ are entangled most strongly. Therefore, it is natural to connect the two spins coupled with J by an isometry to form the spin blocks at the first step of the RG. In the second step, two blocks connected via αJ are entangled most strongly, and thus, one should put an isometry to merge the blocks to form a new block. A recursive application of the above RG process eventually gives rise to the perfect binary TTN, which is expected to be the optimal one if α is small enough. As α increases to unity, the system approaches the uniform Heisenberg chain, where the critical ground state is realized. Thus, the hierarchical chain may provide a good platform to visualize the crossover of the entanglement structure between the ones of the perfect binary tree and of the uniform wave function with critical fluctuation.

To demonstrate that the optimal TTN can be automatically obtained with the proposed algorithm, we start the calculation from the MPN prepared by the following processes. We first focus on two spins at an open end of the chain. We then diagonalize the block Hamiltonian, truncate the high-energy eigenstates, and include the neighboring spin into the new block. We perform this recursive RG process from both ends of the chain and increase the block size one by one. When the left and right blocks meet at the center, we have an MPN including all the isometries, from which we start the first sweep of the structural optimization algorithm. The upper bound of the bond dimension in the initial MPN is $\chi = 40$.

During the iterative sweeps of the algorithm in Table I, the isometries and their connectivity are successively modified. We perform the diagonalization of the effective Hamiltonian $\tilde{H}$ within the subspace of zero total magnetization $\sum_i S_i^z = 0$ and take the lowest energy state of the subspace as the ground state. After the numerical convergence, the variational state with optimal TTN structure is automatically obtained. For the hierarchical chain, we do not use the stochastic choice of reconnection in Eq. (11), but always choose the smallest EE connection in each local update. We perform the SVD of the renormalized ground state wave function $\tilde{\Psi}$ by the diagonalization of the reduced DM, whose eigenvalues are equal to the square of the corresponding singular values. We note that the SVD can also be directly carried out by using a linear algebra package. For comparison, we performed additional calculation with forbidding the reconstruction of the TTN structure, which remains MPN throughout the variational calculation. Note that this calculation is equivalent to the finite-system DMRG method.

We set the maximum number of the bond dimension to be $\chi = 40$. We discard the block states whose DM eigenvalues are smaller than 10^{-12} since their contribution to the ground state wave function is negligible. Due to this

FIG. 7. Optimization process when $\alpha = 0.5$ and $N = 64$. (a) The initial matrix product network (MPN). (b) After the first sweep. (c) Any time after the second sweep. Only the left half of the tree tensor network (TTN) is presented, while the right half is symmetric with respect to the center of the system.

TABLE II. Maximum and average of the bond EE obtained for optimized TTN and fixed MPN when $N = 64$ and 128 with $\alpha = 0.5$. The EEs on the boundary bonds are not included in the analysis.

Type	Maximum	Average
Optimized TTN ($N = 64$)	0.1110	0.0640
Optimized TTN ($N = 128$)	0.1110	0.0625
MPN ($N = 64$)	0.6935	0.3697
MPN ($N = 128$)	0.6935	0.3719

average of the bond EEs are lessened by the tree structure optimization, with the suppression ratio about $\frac{1}{6}$.

Figure 8 shows the expectation values of the nearest neighbor spin-correlation functions $\langle \mathbf{S}_i \cdot \mathbf{S}_{i+1} \rangle$ when $\alpha = 0.5$ and $N = 128$. As seen in the figure, the correlation functions obtained from the MPN coincide with those from the optimal TTN. This is because the ground state for $\alpha = 0.5$ is close to the product state consisting of the singlet dimers, and accordingly, both the optimized TTN and MPN obtained have the sufficient potential of representing the ground state accurately enough. Indeed, the strong dimerization in every two sites can be confirmed in the figure where $\langle \mathbf{S}_{2\ell-1} \cdot \mathbf{S}_{2\ell} \rangle$ is distributed within the range -0.732 to -0.735, while the spin correlations across neighboring dimers are weak. Also, we note that the ground state energy calculated by the MPN agrees with that of the optimal TTN up to 11 digits.

Distinctions between the optimized TTN and the MPN can be found in the spatial distribution of the EE on the auxiliary bonds, although both approaches provide accurate results at the level of the correlation functions. In Fig. 9, we compare the EE distributions on the two different networks of the optimized TTN and the MPN for $\alpha = 0.5$ and $N = 128$. In the figure, we have introduced the horizontal coordinate x (for both the optimized TTN and the MPN) and the height coordinate y (only for the optimized TTN) to specify the positions of the auxiliary bonds; the definitions of the coordinates are presented in Appendix C. The EEs on the boundary bonds directly connected to the bare spins that take the value $\mathcal{S} = \ln 2$ are not included in Fig. 9.

cutoff condition, the number of block states kept in the practical calculation can be smaller than χ. For example, we kept only 26 states, at most, for $\alpha = 0.5$ and $N = 64$ and 128, while 40 states were kept for $\alpha = 1.0$ and $N = 64$. The truncation error is less than 1.3×10^{-11} for $(\alpha, N) = (0.50, 64), (0.50, 128), (0.75, 64),$ and $(0.80, 64)$ and 2.5×10^{-8} at the worst case of $(\alpha, N) = (1.00, 64)$.

Figure 7 shows the transition process of the TTN structure during the sweeps for $\alpha = 0.5$ and $N = 64$. The calculation starts with the initial MPN shown in Fig. 7(a). After the first sweep, a structure like a perfect binary tree appears for every four sites, as shown in Fig. 7(b), where the MPN structure remains in the upper part of the diagram. After the second sweep, the perfect binary tree structure shown in Fig. 7(c) expands up to the top. This is the optimal TTN structure, which is unchanged afterward. We emphasize that the optimal TTN structure obtained by the calculation agrees with the one derived from the perturbative RG scheme discussed above. We observed a similar transition process for $\alpha = 0.5$ and $N = 128$, where the perfect binary TTN emerges after three sweeps. Here, we note that the intermediate TTN structure may depend on how the center bond moves during a sweep; the structure in Fig. 7(b) is the one obtained with the sweep procedure explained in Appendix A. The point is that, as the calculation proceeds, the TTN structure smoothly changes toward the optimal one shown in Fig. 7(c).

Table II shows the maximum and average of the bond EE obtained in the optimized TTN for $\alpha = 0.5$ and $N = 64$ and 128. For comparison, the same set of data obtained in the fixed MPN is shown. Here, the bonds directly connected to the bare spins are not included in the analysis since they always carry the EE of the amount $\ln 2$ irrespective of the network structure, reflecting the fact that the two states $|\uparrow\rangle$ and $|\downarrow\rangle$ of each bare spin have the same reduced DM weight $\frac{1}{2}$. It is clearly seen in the table that both the maximum and

FIG. 8. Nearest neighbor spin correlation function when $\alpha = 0.5$ and $N = 128$. Open circles and solid squares, respectively, represent the obtained result by the matrix product network (MPN) and the optimal tree tensor network (TTN).

FIG. 9. Bond entanglement entropy (EE) calculated when $\alpha = 0.5$ and $N = 128$ as a function of the horizontal position of the bonds x, defined in Appendix C. Open circles represent the EE calculated by the matrix product network (MPN). Other plots are calculated by the optimal tree tensor network (TTN). y denotes the height of the bond in the optimal TTN.

FIG. 10. Optimal structures for (a) $\alpha = 1.00$, (b) $\alpha = 0.80$, and (c) $\alpha = 0.75$. The system size is $N = 64$. Only the left half of the tree tensor network (TTN) is presented, while the right half is symmetric with respect to the center of the system.

Let us discuss the distribution of the EE in Fig. 9 from the bottom to the top. First, the same smallest value of $\mathcal{S} = 1.7 \times 10^{-4}$ at $x = 64.5$ is found in the data of the optimized TTN and the MPN since the corresponding bipartition of the system into the left half $1 \leqslant i \leqslant 64$ and the right one $65 \leqslant i \leqslant 128$ is captured by both networks. Also, both of them capture the second smallest $\mathcal{S} = 6.1 \times 10^{-4}$ at $x = 24.5$ and 104.5 for the optimized TTN and $x \simeq 31.5$ and 97.5 for the MPN, corresponding to the bipartition into $1 \leqslant i \leqslant 32$ and $33 \leqslant i \leqslant 128$ and into $1 \leqslant i \leqslant 96$ and $97 \leqslant i \leqslant 128$. Meanwhile, only the optimized TTN captures a slightly larger value of $\mathcal{S} = 7.7 \times 10^{-4}$ at $x = 40.5$ and 88.5, which corresponds to the bipartitions into $33 \leqslant i \leqslant 64$ and the rest and into $65 \leqslant i \leqslant 96$ and the rest. Note that these bipartitions separate an inner branch of the perfect binary tree network from the rest. The MPN does not have the bond corresponding to such bipartitions and thus cannot capture the EE of $\mathcal{S} = 7.7 \times 10^{-4}$. As the height coordinate y in the optimized TTN decreases, the number of the bipartitions with small EEs that cannot be captured by the MPN increases. In the upper part of the figure, the EEs on almost half of the auxiliary bonds in the MPN take values slightly larger than $\ln 2$ since the bonds carry the EE of the strongly dimerized pairs sitting in every two sites. Such large EEs never appear in the optimized TTN, indicating that the automatic structural optimization algorithm successfully avoids the bipartitions with the large EEs.

Finally, we discuss the α dependence of the optimal tree structures. Figure 10(a) shows the optimized TTN for the chain with $\alpha = 1$ and $N = 64$, which is nothing but a uniform Heisenberg chain. In the bulk limit, the ground state of the chain is usually described by the uniform MPS. In the finite-size system with open boundaries, however, the strong dimer instability induces a sizable staggered component both in the nearest neighbor spin correlations and the EEs [71]. Reflecting this dimerization effect, the optimized TTN for $\alpha = 1$ and $N = 64$ has the MPN structure in the unit of the isometries merging every pair of neighboring bare spins. As α decreases,
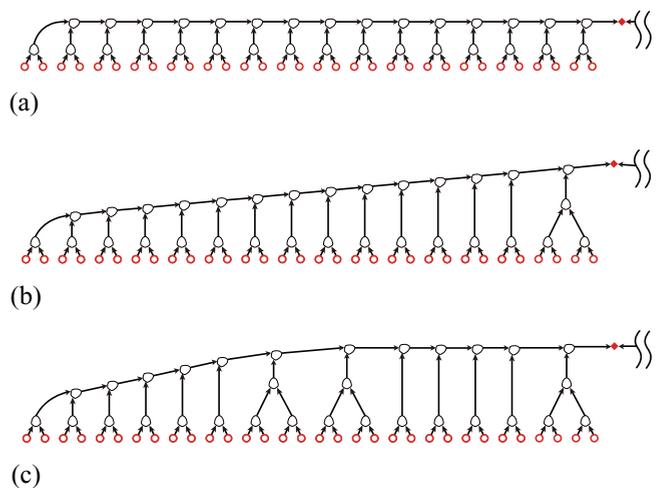
the hierarchical distribution of the EE gradually enhances between the dimerized MPN for $\alpha = 1$ and the perfect binary tree for $\alpha = 0.5$ [Fig. 7(c)]. In Figs. 10(b) and 10(c), we respectively show the optimal TTN structures for $\alpha = 0.8$ and 0.75 with $N = 64$. Here, recall that, in the hierarchical chain, the weakest coupling is placed at the center of the chain and the second weakest at the quarter from the open boundary. Figures 10(b) and 10(c) clearly illustrate that a structure like a perfect binary tree grows from the regions around the weaker interactions.

V. CONCLUDING REMARKS

We have proposed the automatic structural optimization algorithm for the TTN, where the auxiliary bonds between isometry tensors are locally reconnected based on the least EE principle during the variational sweeping on the TTN. We have then demonstrated that, for the hierarchical chain, the algorithm generates a perfect binary tree as the optimal network structure, where the emergence of the bonds with a large EE is actually suppressed. The resulting network structure can be regarded as a fingerprint of the entanglement structure contained in the quantum state.

In this paper, we have focused on the analysis of the distribution of the EE in the optimal TTN to verify that the proposed algorithm indeed works well in terms of the least EE principle. In practical situations, of course, it is also an important issue to clarify how the variational energy E_{var} in Eq. (4) can be improved with the structural optimization. Preliminary calculations for random quantum spin systems with long-range interactions suggest that the optimal TTN realized in the proposed algorithm provides better accuracy of E_{var} compared with the MPN in the DMRG and a TTN obtained by the tSDRG. For such complex systems, the annealing process according to Eq. (11) is essential to escape from the local minimum in the EE landscape. The detailed analysis for the random spin system will be published elsewhere.

The structural optimization based on the local reconnection scheme is promising to be effective in a wide variety of TTN applications to, e.g., 2D or higher dimensional systems and systems with complicated interactions. Random spin systems, for which the optimal TTN is expected to exhibit nontrivial structure, are interesting subjects for the application of the present algorithm. Quantum chemistry systems containing complex configuration interactions among electrons [17–22,62,66,67] are another intriguing candidate to be studied. How the structural optimization of the TTN improves the computational accuracy of such systems is an interesting future problem. In addition, the real-time evolution of quantum many-body states is often described with the framework of the TTN [64,69,72]. Emulation of quantum circuits can also be viewed as a sort of time-evolved quantum state generated by gate operations and measurements [44]. The dynamical reconnection of the TTN for such time evolution problems is also a potential application target.

For the practical application of the algorithm, the estimation of computational costs is another important issue. In the variational TTN calculation, the most time-consuming part is the diagonalization of the renormalized Hamiltonian $\tilde{H}$ to obtain the renormalized ground state wave function $\tilde{\Psi}$. The cost scales as $O(N_{\text{int}} \chi^5)$, where N_{int} is a numerical factor that depends on the boundary area or volume of subsystems connected to the central area and the range of interaction. (See Appendix B for the evaluation of the cost.) Another computational hotspot with the cost $O(\chi^6)$ is the full diagonalization of the reduced DM. The cost for this part may be reduced to $O(\chi^5)$ or $O(\chi^4 \log \chi)$ by use of the partial SVD or randomized SVD [73]. These computational costs in the proposed algorithm based on the TTN are much heavier than those in the DMRG based on the MPN since, in the DMRG, two of the four bonds entering the central area always have the dimension ν, the degree of freedom of a bare spin, and the relation $\chi \gg \nu$ usually holds.

A possible scheme to reduce the computational cost is to carry out the structural optimization with a small χ and then improve isometries in the obtained optimal TTN structure with a large χ. This two-step optimization is effective if the convergence of the TTN structure with respect to χ is fast. Variational improvement of the isometries can also be accelerated by employing the single-site update algorithm [74,75], in which the central area contains only a single isometry. The dimension of $\tilde{H}$ is thus suppressed to be of the order of χ^3 so that the cost of diagonalizing $\tilde{H}$ becomes $O(N_{\text{int}} \chi^4)$. If the system possesses high symmetry, the use of the symmetric bases [76] would be further efficient. The techniques in the real-space parallel DMRG [77] that are effective in accelerating the sweep process for large systems can also be incorporated with the present algorithm.

Finally, we would like to comment on structural optimization for tensor networks other than the TTN. In general, the philosophy of optimal network structure with the least EE principle can be relevant to a wider class of tensor networks containing loop structure. How to clarify the optimal MERA structure may be an interesting problem in connection with this paper since the MERA network is constructed by inserting disentanglers into the TTN.

APPENDIX A: SWEEP PROCEDURE

The automatic optimization scheme we have explained performs local reconnection of isometries when it is necessary; therefore, computational care must be taken so that all the isometries are updated at least once during a sweep. We introduce the *flags* for all the bonds and isometries to indicate whether they have been updated (flag on) or not (flag off). At the beginning of each sweep, all the boundary bonds that are directly connected to bare spins are set to be flag on, while other bonds and all the isometries are flag off. We may choose an arbitrary bond, excluding the boundary ones, as the *origin* and start the sweeping by selecting a bond which is flag off and neighboring to the origin as the central one for the first update step.

After the optimal local connection in the central area is determined in each step, we select a bond from flag-off ones entering the current central area as the central bond in the next step. If there are multiple choices, we select the one which is most distant from the origin. If there is still more than one choice, one may choose any of them. For each isometry in the current central area, if both the entering bonds are flag on, we turn on the flag of the isometry and that of the outgoing bond, which is the current central bond. An exception applies to the step when the central bond comes to the origin; after the update of the isometry and local connection, we keep the flags off for the origin bond and the isometries connected to it even if the condition for turning on the flags is satisfied.

By means of these flag operations, the central bond maneuvers around in the manner that it moves toward the boundary of TTN, if possible, along the bonds which are flag off. If the central bond hits a dead end of flag-on bonds, it steps back toward the origin with turning on the flags of the bond and isometry of the dead end branch. We finish the sweep when all the bonds neighboring the central bond are flag on; this situation occurs in the step where the central bond is at the origin. (We note that, even when the central bond comes to the origin, we continue the sweep if there remains a flag-off bond neighboring the origin.) At the end of the sweep, we update the isometries and local network connection around the origin. We perform the next sweep in the same manner.

We note that the procedure discussed above is not the only way to realize the adequate sweep. The sweeping path of

the central bond should affect the convergence speed of the calculation. Finding the best way of sweeping in practice will be a future problem.

APPENDIX B: COMPUTATIONAL COST

The computational cost for the diagonalization of the renormalized Hamiltonian $\tilde{H}$ to obtain the renormalized ground state wave function $\tilde{\Psi}$ scales with the cost to multiply $\tilde{H}$ to a vector $\tilde{\psi}$ in the Hilbert space of the super-block, which is expanded by the four bonds entering the central area. Here, let us consider the general Heisenberg model in which any pair of spins may have the exchange interaction. The configuration of a conventional DMRG-like super-block Hamiltonian, described for example in Fig. 3, is given by

$$\tilde{H} = \tilde{H}_0^{(ab)} + \tilde{H}_0^{(cd)} + \sum_{i\in(ab)} \sum_{i'\in(cd)} J_{i,i'} \mathbf{S}_i \cdot \mathbf{S}_{i'}, \qquad \text{(B1)}$$

where $\tilde{H}_0^{(ab)}$ and $\tilde{H}_0^{(cd)}$ are renormalized Hamiltonians for the subsystems (ab) and (cd), respectively, and $J_{i,i'}$ is the exchange constant between spins in ith and i'th sites. If one employs this decomposition, a computational cost of $O(\tilde{N}_{\text{int}}\chi^6)$ is required for the operation of $\tilde{H}\tilde{\psi}$, where $\tilde{N}_{\text{int}}$ is the number of operations to multiply the exchange interaction terms in Eq. (B1) to $\tilde{\psi}$. For the systems with short-range interactions, only $J_{i,i'}$ for the spins close to the boundary between the subsystems (ab) and (cd) is nonzero so that $\tilde{N}_{\text{int}}$ scales with the boundary area between the subsystems. For the systems with long-range interactions, all $J_{i,i'}$ are nonzero in general. In this case, one can rewrite the exchange interaction part in Eq. (B1) as $\sum_{i\in(ab)} \mathbf{S}_i \cdot [\sum_{i'\in(cd)} J_{i,i'} \mathbf{S}_{i'}]$ or $\sum_{i'\in(cd)} \mathbf{S}_{i'} \cdot [\sum_{i\in(ab)} J_{i,i'} \mathbf{S}_i]$ so that $\tilde{N}_{\text{int}}$ scales with the volume of the smaller subsystem. Furthermore, one can reduce the order of the computational cost with respect to χ by dividing the super-block into four subsystems:

$$\tilde{H} = \sum_p \tilde{H}_0^{(p)} + \sum_p \sum_{p'(>p)} \sum_{i\in p} \sum_{i'\in p'} J_{i,i'} \mathbf{S}_i \cdot \mathbf{S}_{i'}, \qquad \text{(B2)}$$

where $p, p' \in \{a, b, c, d\}$, and $\tilde{H}_0^{(p)}$ is a renormalized Hamiltonian for the subsystem p. In this decomposition, the cost to multiply $\tilde{H}$ to $\tilde{\psi}$ is $O(N_{\text{int}}\chi^5)$, where N_{int} is the number of operations to multiply the exchange interaction part in Eq. (B2). Here, N_{int} scales with the boundary area between (the volume of the smaller one of) two out of the four subsystems in the case of short-range (long-range) interactions. Employing the decomposition in Eq. (B2) is thus useful when dealing with χ satisfying $\tilde{N}_{\text{int}}\chi \gg N_{\text{int}}$.

Another computational hotspot with the cost of $O(\chi^6)$ is the full diagonalization of the reduced DM, whose dimension is of the order of χ^2 in the TTN calculation. Here, if we do not need to care about the degeneracy of singular values due to the symmetry of the target state, the partial SVD may effectively reduce the cost down to $O(\chi^5)$ [73]. The randomized SVD may further reduce the cost to $O(\chi^4 \log \chi)$.

The computational costs in the TTN calculation discussed above are much larger than the corresponding ones in the DMRG since, in the MPN employed in the DMRG, two of the four bonds entering the central area, two out of $\{a, b, c, d\}$

(a)

(b)
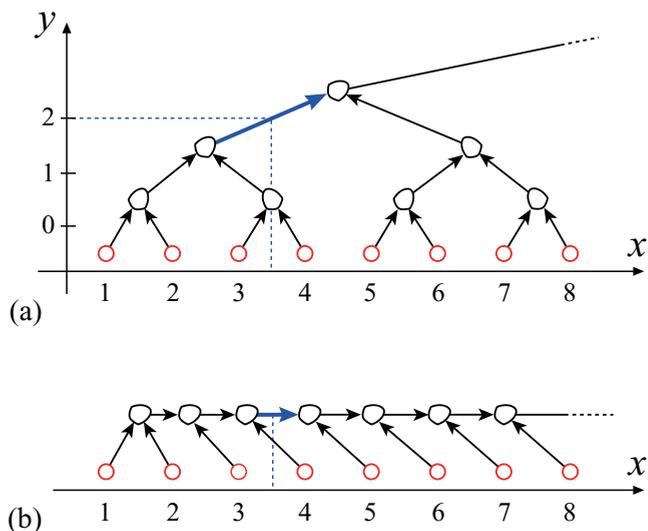
FIG. 11. Schematic pictures for the coordinates indicating the positions of the auxiliary bonds for (a) the optimized tree tensor network (TTN) and (b) the matrix product network (MPN). In (a), the coordinates for the bond denoted by a thick arrow (blue) are $(x, y) = (3.5, 2)$. In (b), the coordinate for the bond is $x = (3.125 + 4.0625)/2 = 3.59375$.

in Fig. 3, have the dimension ν, which is usually much smaller than χ. As a result, the cost for the diagonalization of $\tilde{H}$ represented in the decomposition in Eq. (B2) is $O(N_{\text{int}}\nu^2\chi^3)$ in the DMRG, being much smaller than $O(N_{\text{int}}\chi^5)$ in the TTN algorithm. Similarly, the cost for the full diagonalization of the reduced DM in the DMRG is $O(\nu^3\chi^3)$, which is significantly small compared with $O(\chi^6)$ in the TTN algorithm.

APPENDIX C: COORDINATES OF AUXILIARY BONDS IN FIG. 9

In Fig. 9, we have introduced the coordinates to indicate the position of the auxiliary bonds. The horizontal coordinate x introduced for both the optimized TTN and MPN are defined as follows. We first assign the site index i to the horizontal coordinate of bare spins. The position of an isometry is then given by the average of the horizontal coordinates of bare spins or the isometries connected with the incoming bonds. The position x of the auxiliary bonds is finally defined by the average of the coordinates of the connected isometries.

For the case of the optimized TTN, moreover, we have introduced the height coordinate y, which is defined as the distance of the bond from the nearest bare spin; we assign $y = 0$ for the bonds directly connected to a bare spin, and y increases by one as the position of the bond moves away from the bare spin.

Figure 11 presents the schematic pictures to illustrate the coordinates in the optimized TTN and the MPN. We note that the horizontal coordinate x for the isometries and auxiliary bonds in the MPN is shifted toward the open edges of the system compared with the diagram shown in Fig. 7(a), where we do not use the coordinate x for readability.

[1] K. Okunishi, T. Nishino, and H. Ueda, Developments in the tensor network—From statistical mechanics to quantum entanglement, J. Phys. Soc. Jpn. **91**, 062001 (2022).

[2] A. Milsted and G. Vidal, Geometric interpretation of the multiscale entanglement renormalization ansatz, arXiv:1812.00529 (2018).

[3] R. Orús, Tensor networks for complex quantum systems, Nat. Rev. Phys. **1**, 538 (2019).

[4] E. Stoudenmire and D. J. Schwab, Supervised learning with tensor networks, in *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems, December 5–10, 2016, Barcelona, Spain*, edited by D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (2016), pp. 4799–4807, https://researchr.org/publication/StoudenmireS16.

[5] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, Rev. Mod. Phys. **93**, 045003 (2021).

[6] R. J. Baxter, Dimers on a rectangular lattice, J. Math. Phys. **9**, 650 (1968).

[7] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[8] S. R. White, Density-matrix algorithms for quantum renormalization groups, Phys. Rev. B **48**, 10345 (1993).

[9] S. Östlund and S. Rommer, Thermodynamic Limit of Density Matrix Renormalization, Phys. Rev. Lett. **75**, 3537 (1995).

[10] S. Rommer and S. Östlund, Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group, Phys. Rev. B **55**, 2164 (1997).

[11] T. Nishino, Y. Hieida, K. Okunishi, N. Maeshima, Y. Akutsu, and A. Gendiar, Two-dimensional tensor product variational formulation, Prog. Theor. Phys. **105**, 409 (2001).

[12] A. Gendiar, N. Maeshima, and T. Nishino, Stable optimization of a tensor product variational state, Prog. Theor. Phys. **110**, 691 (2003).

[13] F. Verstraete and J. I. Cirac, Valence-bond states for quantum computation, Phys. Rev. A **70**, 060302(R) (2004).

[14] F. Verstraete, M. M. Wolf, D. Pérez-García, and J. I. Cirac, Criticality, the Area Law, and the Computational Power of Projected Entangled Pair States, Phys. Rev. Lett. **96**, 220601 (2006).

[15] G. Vidal, Entanglement Renormalization, Phys. Rev. Lett. **99**, 220405 (2007).

[16] G. Evenbly and G. Vidal, Algorithms for entanglement renormalization, Phys. Rev. B **79**, 144108 (2009).

[17] N. Nakatani and G. Kin-Lic Chan, Efficient tree tensor network states (TTNS) for quantum chemistry: Generalizations of the density matrix renormalization group algorithm, J. Chem. Phys. **138**, 134113 (2013).

[18] V. Murg, F. Verstraete, Ö. Legeza, and R. M. Noack, Simulating strongly correlated quantum systems with tree tensor networks, Phys. Rev. B **82**, 205105 (2010).

[19] V. Murg, F. Verstraete, R. Schneider, P. R. Nagy, and Ö. Legeza, Tree tensor network state with variable tensor order: An efficient multireference method for strongly correlated systems, J. Chem. Theory Comput. **11**, 1027 (2015).

[20] K. Gunst, F. Verstraete, S. Wouters, Ö. Legeza, and D. Van Neck, T3NS: Three-legged tree tensor network states, J. Chem. Theory Comput. **14**, 2026 (2018).

[21] K. Gunst, F. Verstraete, and D. Van Neck, Three-legged tree tensor networks with $SU(2)$ and molecular point group symmetry, J. Chem. Theory Comput. **15**, 2996 (2019).

[22] H. R. Larsson, Computing vibrational eigenstates with tree tensor network states (TTNS), J. Chem. Phys. **151**, 204102 (2019).

[23] D. Nagaj, E. Farhi, J. Goldstone, P. Shor, and I. Sylvester, Quantum transverse-field Ising model on an infinite tree from matrix product states, Phys. Rev. B **77**, 214431 (2008).

[24] L. Tagliacozzo, G. Evenbly, and G. Vidal, Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law, Phys. Rev. B **80**, 235127 (2009).

[25] P. Silvi, V. Giovannetti, S. Montangero, M. Rizzi, J. I. Cirac, and R. Fazio, Homogeneous binary trees as ground states of quantum critical Hamiltonians, Phys. Rev. A **81**, 062335 (2010).

[26] W. Li, J. von Delft, and T. Xiang, Efficient simulation of infinite tree tensor network states on the Bethe lattice, Phys. Rev. B **86**, 195137 (2012).

[27] H. J. Changlani, S. Ghosh, C. L. Henley, and A. M. Läuchli, Heisenberg antiferromagnet on Cayley trees: Low-energy spectrum and even/odd site imbalance, Phys. Rev. B **87**, 085107 (2013).

[28] I. Pižorn, F. Verstraete, and R. M. Konik, Tree tensor networks and entanglement spectra, Phys. Rev. B **88**, 195102 (2013).

[29] A. Milsted, M. Ganahl, S. Leichenauer, J. Hidary, and G. Vidal, TensorNetwork on TensorFlow: A spin chain application using tree tensor networks, arXiv:1905.01331 (2019).

[30] E. Macaluso, T. Comparin, R.O. Umucalilar, M. Gerster, S. Montangero, M. Rizzi, and I. Carusotto, Charge and statistics of lattice quasiholes from density measurements: A tree tensor network study, Phys. Rev. Res. **2**, 013145 (2020).

[31] P. Lunts, A. Georges, E. M. Stoudenmire, and M. Fishman, Hubbard model on the Bethe lattice via variational uniform tree states: Metal-insulator transition and a Fermi liquid, Phys. Rev. Res. **3**, 023054 (2021).

[32] B. Kloss, R. D. Reichman, and Y. Bar Lev, Studying dynamics in two-dimensional quantum lattices using tree tensor network states, SciPost Phys. **9**, 070 (2020).

[33] T. Hikihara, A. Furusaki, and M. Sigrist, Numerical renormalization-group study of spin correlations in one-dimensional random spin chains, Phys. Rev. B **60**, 12116 (1999).

[34] A. M. Goldsborough and R. A. Römer, Self-assembling tensor networks and holography in disordered spin chains, Phys. Rev. B **89**, 214203 (2014).

[35] Y.-P. Lin, Y.-J. Kao, P. Chen, and Y.-C. Lin, Griffiths singularities in the random quantum Ising antiferromagnet: A tree tensor network renormalization group study, Phys. Rev. B **96**, 064427 (2017).

[36] K. Seki, T. Hikihara, and K. Okunishi, Tensor-network strong-disorder renormalization groups for random quantum spin systems in two dimensions, Phys. Rev. B **102**, 144439 (2020).

[37] K. Seki, T. Hikihara, and K. Okunishi, Entanglement-based tensor-network strong-disorder renormalization group, Phys. Rev. B **104**, 134405 (2021).

[38] G. Ferrari, G. Magnifico, and S. Montangero, Adaptive-weighted tree tensor networks for disordered quantum many-body systems, Phys. Rev. B **105**, 214201 (2022).

[39] E. Dumitrescu, Tree tensor network approach to simulating Shor's algorithm, Phys. Rev. A **96**, 062322 (2017).

[40] S. Cheng, L. Wang, T. Xiang, and P. Zhang, Tree tensor networks for generative modeling, Phys. Rev. B **99**, 155131 (2019).

[41] M. L. Wall and G. D'Aguanno, Tree-tensor-network classifiers for machine learning: From quantum inspired to quantum assisted, Phys. Rev. A **104**, 042408 (2021).

[42] E. Grelier, A. Nouy, and R. Lebrun, Learning high-dimensional probability distributions using tree tensor networks, Int. J. Uncertain. Quantif. **12**, 47 (2022).

[43] H. Chen and T. Barthel, Tensor network states with low-rank tensors, arXiv:2205.15296 (2022).

[44] P. Seitz, I. Medina, E. Cruz, Q. Huang, and C. B. Mendl, Simulating quantum circuits using tree tensor networks, arXiv:2206.01000 (2022).

[45] H. Otsuka, Density-matrix renormalization-group study of the spin-$\frac{1}{2}$XXZ antiferromagnet on the Bethe lattice, Phys. Rev. B **53**, 14004 (1996).

[46] B. Friedman, A density matrix renormalization group approach to interacting quantum systems on Cayley trees, J. Phys.: Condens. Matter **9**, 9021 (1997).

[47] M. Kumar, S. Ramasesha, and Z. G. Soos, Density matrix renormalization group algorithm for Bethe lattices of spin-$\frac{1}{2}$ or spin-1 sites with Heisenberg antiferromagnetic exchange, Phys. Rev. B **85**, 134415 (2012).

[48] S.-k. Ma, C. Dasgupta, and C.-k. Hu, Random Antiferromagnetic Chain, Phys. Rev. Lett. **43**, 1434 (1979).

[49] C. Dasgupta and S.-k. Ma, Low-temperature properties of the random Heisenberg antiferromagnetic chain, Phys. Rev. B **22**, 1305 (1980).

[50] S. S. Roy, S. N. Santalla, J. Rodríguez-Laguna, and G. Sierra, Entanglement as geometry and flow, Phys. Rev. B **101**, 195134 (2020).

[51] S. S. Roy, S. N. Santalla, G. Sierra, and J. Rodríguez-Laguna, Link representation of the entanglement entropies for all bipartitions, J. Phys. A: Math. Theor. **54**, 305301 (2021).

[52] S. S. Roy, S. N. Santalla, J. Rodríguez-Laguna, and G. Sierra, Emergent geometry from entanglement structure, in *Quantum Theory and Symmetries*, edited by M. B. Paranjape, R. MacKenzie, Z. Thomova, P. Winternitz, and W. Witczak-Krempa (Springer International Publishing, Cham, 2021), pp. 347–357.

[53] S. N. Santalla, G. Ramírez, S. S. Roy, G. Sierra, and J. Rodríguez-Laguna, Entanglement links and the quasiparticle picture, arXiv:2208.03766 (2022).

[54] G. Evenbly and G. Vidal, Tensor network states and geometry, J. Stat. Phys. **145**, 891 (2011).

[55] K. Hyatt, J. R. Garrison, and B. Bauer, Extracting Entanglement Geometry from Quantum States, Phys. Rev. Lett. **119**, 140502 (2017).

[56] Ö. Legeza and J. Sólyom, Optimizing the density-matrix renormalization group method using quantum information entropy, Phys. Rev. B **68**, 195116 (2003).

[57] Ö. Legeza, L. Veis, A. Poves, and J. Dukelsky, Advanced density matrix renormalization group method for nuclear structure calculations, Phys. Rev. C **92**, 051303(R) (2015).

[58] J. Brabec, J. Brandejs, K. Kowalski, S. Xantheas, Ö. Legeza, and L. Veis, Massively parallel quantum chemical density matrix renormalization group method, J. Comput. Chem. **42**, 534 (2021).

[59] M. Ali, On the ordering of sites in the density matrix renormalization group using quantum mutual information, arXiv:2103.01111 (2021).

[60] M. Máté, K. Petrov, S. Szalay, and Ö. Legeza, Compressing multireference character of wave functions via fermionic mode optimization, J. Math. Chem. (2022), doi: 10.1007/s10910-022-01379-y.

[61] S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza, Tensor product methods and entanglement optimization for *ab initio* quantum chemistry, Int. J. Quantum Chem. **115**, 1342 (2015).

[62] W. Li, J. Ren, H. Yang, and Z. Shuai, On the fly swapping algorithm for ordering of degrees of freedom in density matrix renormalization group, J. Phys.: Condens. Matter **34**, 254003 (2022).

[63] K. Okunishi, H. Ueda, and T. Nishino, Entanglement bipartitioning and tree tensor networks, arXiv:2210.11741 (2022).

[64] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network, Phys. Rev. A **74**, 022320 (2006).

[65] We use a TTN as a term indicating the amplitudes of a wave function represented in the form of a contraction of tensors on a tree network. Meanwhile, the ket state $|\Psi\rangle$ with a TTN is termed a TTN state. Given a MPS, similarly, the MPN represents the amplitudes of the corresponding wave function.

[66] G. K.-L. Chan and M. Head-Gordon, Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group, J. Chem. Phys. **116**, 4462 (2002).

[67] G. Moritz, B. A. Hess, and M. Reiher, Convergence behavior of the density-matrix renormalization group algorithm for optimized orbital orderings, J. Chem. Phys. **122**, 024107 (2005).

[68] M. Gerster, P. Silvi, M. Rizzi, R. Fazio, T. Calarco, and S. Montangero, Unconstrained tree tensor network: An adaptive gauge picture for enhanced performance, Phys. Rev. B **90**, 125154 (2014).

[69] X. Cao, Y. Lu, P. Hansmann, and M. W. Haverkort, Tree tensor-network real-time multiorbital impurity solver: Spin-orbit coupling and correlation functions in Sr_2RuO_4, Phys. Rev. B **104**, 115119 (2021).

[70] S. R. White, Spin Gaps in a Frustrated Heisenberg Model for CaV_4O_9, Phys. Rev. Lett. **77**, 3633 (1996).

[71] N. Laflorencie, E. S. Sørensen, M.-S. Chang, and I. Affleck, Boundary Effects in the Critical Scaling of Entanglement Entropy in 1D Systems, Phys. Rev. Lett. **96**, 100603 (2006).

[72] D. Bauernfeind and M. Aichhorn, Time dependent variational principle for tree tensor networks, SciPost Phys. **8**, 024 (2020).

[73] N. Halko, P. G. Martinsson, and J. A. Tropp, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, SIAM Rev. **53**, 217 (2011).

[74] H. Takasaki, T. Hikihara, and T. Nishino, Fixed point of the finite system DMRG, J. Phys. Soc. Jpn. **68**, 1537 (1999).

[75] S. R. White, Density matrix renormalization group algorithms with a single center site, Phys. Rev. B **72**, 180403(R) (2005).

[76] I. P. McCulloch and M. Gulácsi, The non-Abelian density matrix renormalization group algorithm, Europhys. Lett. **57**, 852 (2002).

[77] E. M. Stoudenmire and S. R. White, Real-space parallel density matrix renormalization group, Phys. Rev. B **87**, 155137 (2013).