Efficient calculation of three-dimensional tensor networks

Li-Ping Yang [1] Y. F. Fu [2] Z. Y. Xie [2,*] and T. Xiang [3,4,5,†]

[1]*Department of Physics, Chongqing University, Chongqing 401331, China*
[2]*Department of Physics, Renmin University of China, Beijing 100872, China*
[3]*Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[4]*Beijing Academy of Quantum Information Sciences, Beijing 100193, China*
[5]*Department of Physics, University of Chinese Academy of Sciences, Beijing 100190, China*

We have proposed an efficient algorithm to calculate physical quantities in the translational invariant three-dimensional tensor networks, which is particularly relevant to the study of the three-dimensional classical statistical models and the $(2+1)$-dimensional quantum lattice models. In the context of a classical model, we determine the partition function by solving the dominant eigenvalue problem of the transfer matrix, whose left and right dominant eigenvectors are represented by two projected entangled simplex states. These two projected entangled simplex states are not Hermitian conjugate to each other but are appropriately arranged so that their inner product can be computed much more efficiently than in the usual prescription. For the three-dimensional Ising model, the calculated internal energy and spontaneous magnetization agree with the published results in the literature. The possible improvement and extension to other models are also discussed.

I. INTRODUCTION

The many-body problem is one of the central problems in physics, and developing accurate and efficient numerical methods that can effectively handle the exponential growth of the corresponding Hilbert space has always been a great challenge, especially for quantum systems. Based on the idea of renormalization group and the tensor-network representation of partition functions and wave functions, tensor-network methods have evolved progressively to be an important member of many-body computational methods in recent years [1–3]. In fact, due to the absence of sign problem and the ability to deal with two-dimensional systems, tensor networks are drawing increasing attention, and have been applied successfully to strongly-correlated electron systems [4,5], frustrated spin systems [6–10], statistical models [11–13], topological order [14–17], quantum field theory [18–20], machine learning [21,22], and quantum circuit simulation [23], etc. Among the various tensor-network methods, imaginary time evolution is a highly efficient method to determine the ground state of low-dimensional quantum systems [3,24,25].

A central task in the application of tensor network states to $2+1$ dimensional quantum lattice models is to contract a two-dimensional tensor network, which has a double-layer structure and bond dimension D^2, with D the maximal virtual bond dimension of the tensor network representation of a quantum state [3]. This is extremely costly, and thus although D is expected to be larger to produce a more accu-

rate result, it is limited to about 13 in practical calculations [9,17,26]. There have been some efforts in recent years to solve this problem, such as employing symmetries [27,28], combining Monte Carlo sampling [29], and the nested tensor network method [30]. However, apparently the problem is not completely solved, and developing efficient algorithms to resolve this issue is still an important topic for tensor-network communities.

Meanwhile, it is known that, in the formalism of path integral, the equilibrium quantum many-body problem in d dimensions is similar to the classical many-body problem in $d+1$ dimensions [11,31–35]. Thus inspired by the success in one- and two-dimensional quantum systems, there are also some efforts to apply tensor-network methods to three-dimensional (3D) classical models. Actually, both the coarse-graining tensor renormalization group algorithms [11,36–38] and the transfer-matrix-based tensor-network state methods [39–45] have been applied to these models or related previously. Although the transfer-matrix-based methods can work excellently in two-dimensional networks, they seem not so efficient in three dimensions [46]. In the literature, in order to obtain reliable results, the variational optimization procedure is indispensable, thus due to the high computational cost, as will be discussed later, the bond dimension of the involved tensor-network states and the accuracy are rather limited, and some auxiliary techniques are usually employed further to analyze the data. Especially, to our best knowledge, in practice, whether a simple yet efficient algorithm analogous to the simple-update algorithm [47] in two-dimensional quantum lattice models can be developed for 3D classical models is unclear up to now.

In this paper, we are trying to address the above issues. To be specific, following the usual prescription of the

*qingtaoxie@ruc.edu.cn
†txiang@iphy.ac.cn

transfer-matrix-based methods, first we express the partition function of a 3D classical model in terms of some special two-dimensional transfer matrices, and reduce the problem to the dominant eigenvalue problem of the matrices. Then we solve the dominant eigenvalue problem by representing the dominant eigenvector as a special tensor-network state, namely the projected entangled simplex state (PESS) [26], which is efficiently determined through a power iteration procedure analogous to imaginary time evolution, as done for two-dimensional classical models similarly [48]. Furthermore, a simple nesting technique is proposed in which the PESS representations of the left and right dominant eigenvectors are designed appropriately so that their inner product can be expressed as a tensor network with bond dimension D instead of D^2, and thus the contraction can be carried out much more efficiently. Combining this nesting technique with the corner transfer matrix renormalization group (CTMRG) algorithm [4,49,50], we are able to push the bond dimension D to 20 in this paper. For the 3D Ising model, it shows that even if the tensor network state is renormalized by the so-called simple update technique [25,26,47] after each evolution step, the obtained local quantities, such as energy density and spontaneous magnetization, are in good consistent with the previous studies, and the estimated critical temperature T_c is about 4.50984(2), which has only a relative deviation of about 10^{-4} from the best Monte Carlo estimations.

The rest of the paper is organized as follows. In Sec. II, we introduce some details of the algorithm employed in this paper, including the nesting technique. The numerical results for statistical averages, such as energy density E and spontaneous magnetization M, as well as the convergence analysis, are presented in Sec. III. In Sec. IV, we summarize our paper, and discuss the possible improvement and promising extensions briefly.

II. METHOD

A. Tensor-network representation of the partition function

For concreteness, hereinafter, we will focus on the 3D Ising model. The partition function can be written as

$$Z = \sum_{\{s\}} \prod_{\langle ij \rangle} e^{\beta s_i s_j}, \tag{1}$$

where $s_i = \pm 1$ is the spin variable located on the ith lattice site, β is the inverse temperature, $\langle ij \rangle$ means the product is performed over all the nearest-neighboring bonds, and the summation is over all the spin configurations.

For later use, we regroup the product in Eq. (1) in the unit of cubes that are the building block of a cubic lattice, i.e.,

$$Z = \sum_{\{s\}} \prod_{\alpha} T^{(\alpha)}, \tag{2}$$

where $T^{(\alpha)}$ is a rank-8 tensor defined at the αth cube. For example, as shown in the right panel of Fig. 1, for a cube where the spin variables residing on the eight vertices are denoted by $s_1 - s_8$, the local tensor T can be expressed as the following product of twelve Boltzmann weights corresponding to each
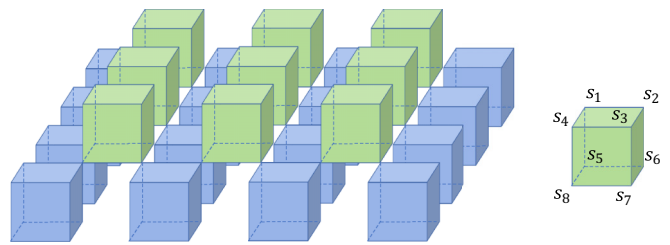

FIG. 1. Special transfer matrices in the expression of the partition function of 3D Ising model. (Left) The green and blue cubes constitute transfer matrices T_2 and T_1, as expressed in Eqs. (4) and (5). (Right) The definition of local tensor T located in each colored cube, as expressed in Eq. (3). As mentioned in the main text, the vertical direction is referred to as z direction for convenience.

edge of the cube, respectively,

$$\begin{aligned} T_{s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8} = \exp[\beta(s_1 s_2 + s_2 s_3 + s_3 s_4 + s_4 s_1 + s_5 s_6 \\ + s_6 s_7 + s_7 s_8 + s_8 s_5 + s_1 s_5 + s_2 s_6 + s_3 s_7 \\ + s_4 s_8)]. \end{aligned} \tag{3}$$

Manifestly, to make Eqs. (2) and (3) consistent, T should be defined only in two kinds of inequivalent cubes, as denoted as green and blue, respectively, in the left panel of Fig. 1. Thus the two kinds of cubes form an alternative or staggered structure in the vertical direction. This is very similar to the case of the imaginary time evolution in quantum lattice models, where the Trotter-Suzuki decomposition of the evolution operator $e^{-\tau H}$ always leads to an alternative structure in imaginary time τ direction. In the following, we will use this special structure extensively. For convenience, the vertical direction will be referred to as z direction hereinafter.

B. Determination of the tensor-network representation of the dominant eigenvector

Following the prescription of the transfer-matrix-based method, one needs to express the partition function in terms of some transfer matrices. To this end, firstly, we introduce two tensor-network operators

$$T_1 = \bigotimes_{\alpha \in b} T^{(\alpha)}, \quad T_2 = \bigotimes_{\alpha \in g} T^{(\alpha)}, \tag{4}$$

where the direct products are performed over Ts defined at *blue* cubes and *green* cubes, respectively, as illustrated in Fig. 1. And then we can identify the following equality:

$$Z = \mathrm{Tr}(T_2 T_1)^{2n}, \tag{5}$$

where the length in the z direction is denoted as $2n$ for convenience. It is worth noting that T_1 and T_2 should be understood as matrices by grouping indices properly in Eqs. (4) and (5), in order to make the operations for matrices therein meaningful.

Once Eq. (5) is established, the calculation of partition function is immediately reduced to the dominant eigenvalue problem of the transfer matrix $T_2 T_1$, which can be solved by the power iteration method. In this paper, we represent the corresponding dominant eigenvector $|\Psi\rangle$ as a PESS form
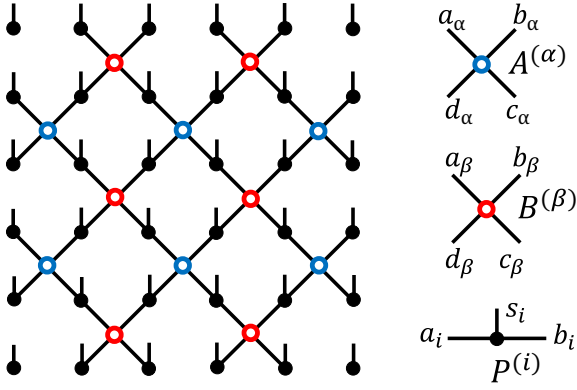
FIG. 2. Tensor network representation of the dominant eigenvector of the transfer matrix $T_2 T_1$ appeared in Eq. (5), as formulated in Eq. (6). The black vertical lines denote the physical configurations $\{s\}$ appeared in Eq. (6). In the right panel, the indices of the local tensors are explicitly shown for clarity.

[26,51], which can be written as

$$|\Psi\rangle = \sum_{\{s\}} \mathrm{Tr}(\dots A^{(\alpha)}_{a_\alpha b_\alpha c_\alpha d_\alpha} B^{(\beta)}_{a_\beta b_\beta c_\beta d_\beta} P^{(i)}_{a_i b_i}[s_i]\dots)|\dots s_i\dots\rangle, \quad (6)$$

as illustrated in Fig. 2. Here α and β denote the coordinates of two inequivalent squares, at the center of which a rank-4 simplex tensor A or B is introduced to characterize the four-spin entanglement in that square. i denotes the coordinates of lattice sites, where a rank-3 projection tensor P is defined with two virtual indices labeled as a, b... and a single physical index labeled as s. Every two virtual indices associated with the same bond take the same values. Tr is over all the repeated virtual indices and $\sum$ is over all the spin configurations $\{s\}$. In this paper, we employed the C_4 rotational symmetry of A and B that corresponds to the symmetry of local tensor T defined in Eq. (3), and the translational symmetry of the 3D lattice, thus there are only one independent P, A, and B present in Eq. (6). More background about the PESS wave function can be found in the Appendix.

As done in time evolution in quantum lattice systems, using the alternative structure of T_1 and T_2, we can take a similar simple update strategy [25,26,47] to determine the variational parameters A, B, and P in $|\Psi\rangle$. To be specific, starting from a random state $|\Psi_0\rangle$, which has the same structure as $|\Psi\rangle$, we apply T_1 and T_2 alternatively to $|\Psi_0\rangle$ and update the parameters accordingly by local decompositions of the related clusters after each projection. This procedure is applied repeatedly until the convergence is reached, and the obtained tensor network state provides an approximate representation of $|\Psi\rangle$ [48]. Taking A as an example, Fig. 3 illustrates how a single projection step is performed. For more details, we suggest referring to Ref. [26].

C. Inner product between the left and right dominant eigenvectors, and its nesting structure

Having described how to obtain the dominant eigenvectors, in this subsection, firstly, we briefly describe how the statistical averages are obtained in the traditional reduced tensor network method (RTN) [3,30]. As long as the dominant
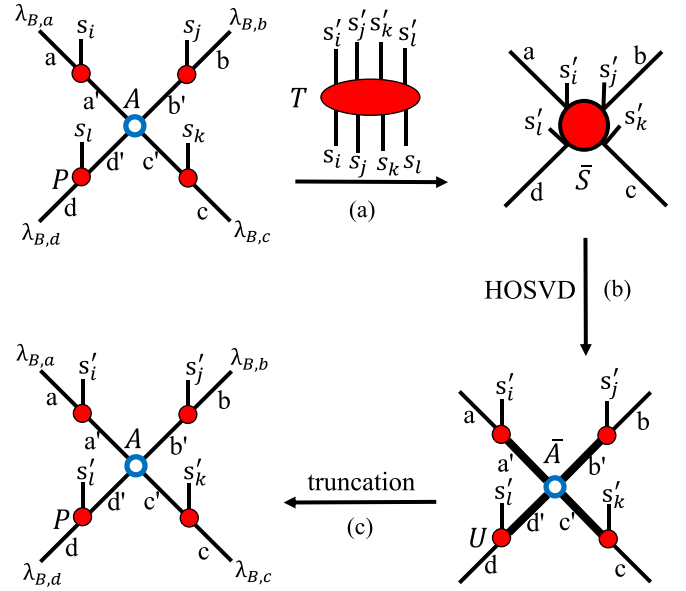

FIG. 3. A single projection step of T on the PESS ansatz (take the update of simplex tensor A for example), as described in the main text. To illustrate the operation, the indices of the local tensor T are labeled explicitly. (a) The indices $\{s_i s_j s_k s_l\}$ in T and A are summed over. A cluster tensor $\bar{S}$ is obtained by further absorbing four mean-field entanglements, i.e., the positive definite vector λ_B obtained from the higher-order singular value decomposition (HOSVD) of the B tensor. (b) The HOSVD of $\bar{S}$ is performed to obtain four unitary matrices Us and the core tensor $\bar{A}$. (c) Truncate $\bar{A}$ to keep bond dimension D, recover the original mean-field entanglement structure, and obtain the updated local tensors A and P. For more details, one can refer to the elaboration in Ref. [26].

eigenvector $|\Psi\rangle$ is obtained, we can use the fundamental formula to calculate the statistical averages of local physical quantities. For example, the average magnetization located at the ith site can be determined by

$$M_i = \frac{\sum_{\{s\}} s_i \prod_\alpha T^{(\alpha)}}{\sum_{\{s\}} \prod_\alpha T^{(\alpha)}} = \frac{\mathrm{Tr}(T_2 T_1)^n s_i (T_2 T_1)^n}{\mathrm{Tr}(T_2 T_1)^n (T_2 T_1)^n}, \quad (7)$$

where we have used the transfer-matrix expression of Z and assumed that s_i sits in the middle in the z direction. In the thermodynamic limit, $n \to \infty$, we reach

$$M_i = \frac{\langle\Psi'|s_i|\Psi\rangle}{\langle\Psi'|\Psi\rangle}, \quad M = \frac{\sum_{i\in\alpha} M_i}{8}, \quad (8)$$

where $\langle\Psi'|$ is the left dominant eigenvector of $T_2 T_1$ and can be derived easily due to the symmetry between T_1 and T_2. The spontaneous magnetization M is obtained eventually as above by averaging over the eight spins in the same cube due to the translational invariance of the tensor networks. Internal energy E of the bonds can be obtained similarly. To evaluate the ratios like Eq. (8), we can follow the widely-used impurity method as discussed in Refs. [27,47,52], and some details can be found in the Appendix.

As indicated explicitly in Eq. (8), the statistical average of a local physical quantity needs to contract a two-dimensional tensor network, which is identical to the expectation value calculation for quantum lattice models in essence. Suppose
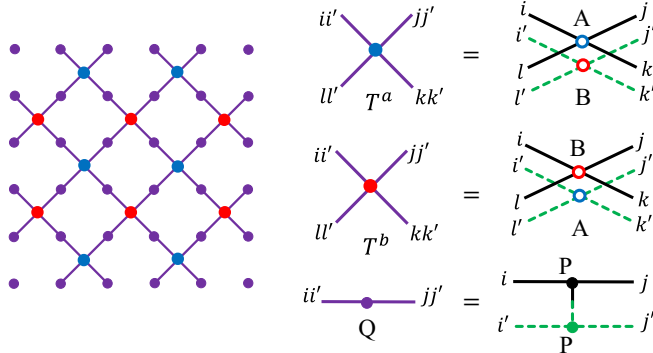
FIG. 4. Sketch of $\langle\Psi'|\Psi\rangle$ appeared in Eq. (8). Here $\langle\Psi'|$ is the left dominant eigenvector of T_2T_1. The relevance to $|\Psi\rangle$, and thus the structure of the local tensors T^a and T^b, come from the symmetry between T_1 and T_2, i.e., they are transformed to each other by a finite translation in the diagonal direction.

both $|\Psi\rangle$ and $\langle\Psi'|$ have bond dimension D, then the generated two-dimensional tensor network has dimension D^2, as shown in Fig. 4. When D is small, this tensor network can be contracted by the CTMRG algorithm [4,49,50] effectively. More introduction of the CTMRG algorithm used in this paper is provided in the Appendix.

The simple-update-like algorithm described in Sec. II B is highly efficient, and the leading computational cost scales as D^5 in our case. Thus PESS eigenvector with large bond dimensions can be obtained without too much cost. However, contracting a two-dimensional tensor network with bond dimension D^2, as shown in Fig. 4, has a leading cost D^{12} empirically, as discussed in Ref. [30], and is extremely costly when D is large. A possible solution is the nested tensor network technique, which can reduce the cost to D^9 [30]. Nevertheless, in this paper, we did not take this strategy. Instead, we proposed a much simpler nesting technique by taking advantage of the special cubic lattice structure. This nesting technique can reduce the computational cost similarly as achieved by the nested tensor network method (also scales as D^9), keeps the symmetry of the ground state properly, and is especially suitable for cubic systems.

To see how it works, we first divide the cubic lattice into three parts, namely two bulks whose contribution to the full partition function can be expressed in terms of transfer matrices, and the surface part, which connects the two bulks and combines as the whole cubic lattice. This means we have rewritten the partition function in another manner, as illustrated in Fig. 5(a),

$$Z = \text{Tr}(T_4T_3)^n YVX(T_2T_1)^n, \qquad (9)$$

where T_1 and T_2 are the transfer matrices corresponding to the lower bulk, T_3 and T_4 correspond to the upper bulk similarly, and X, Y, V are auxiliary matrices to describe the surface part. Here we intentionally decompose the lower and upper parts in different manners, and their relative locations projected in the xy plane are shown in Fig. 5(b). Mathematically,

$$X = \prod_{\langle ij\rangle_t \notin g} e^{\beta s_i s_j}, \quad Y = \prod_{\langle ij\rangle_b \notin o} e^{\beta s_i s_j}, \quad V = \prod_{\langle ij\rangle_{bt}} e^{\beta s_i s_j}, \quad (10)$$
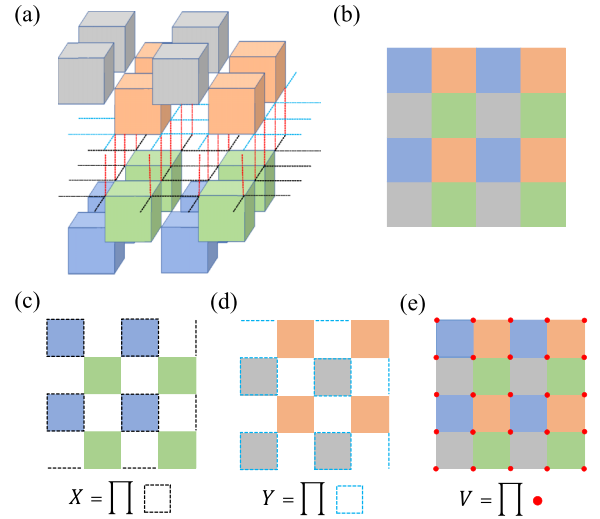

FIG. 5. Decomposition of the cubic lattice into three parts, as described in the nesting technique in Sec. II. (a) The cubes with colors blue, green, orange, and grey, constitute the transfer matrices T_1, T_2, T_3, and T_4, respectively. The dashed lines with color black, blue, and red, constitute the transfer matrices X, Y, and V, respectively. The full partition function can be represented in terms of these transfer matrices, as expressed in Eq. (9). (b) The spatial distribution of the cubes in xy plane, for the sixteen cubes shown in (a). (c) Relative location of X (black dashed lines) to T_1 (blue) and T_2 (green), and this explains how the operation $X|\Psi\rangle$ is performed. (d) Relative location of Y (blue dashed lines) to T_3 (orange) and T_4 (grey), and this explains how the operation $\langle\Phi|Y$ is performed. (e) Relative location of V (on-site red dots) to all four Ts, and this explains how the operation $\langle\tilde{\Phi}|V|\tilde{\Psi}\rangle$ is performed. Note that (b)–(e) display the projection of (a) onto the horizontal plane, therefore the simplex tensors appeared in the PESS ansatz Eq. (6) and updated in Fig. 3, reside exactly in the center of the colored squares.

which are explicitly shown in Figs. 5(a), and 5(c)–5(e). Here X and Y can be understood as on-site diagonal matrices, $\langle ij\rangle_t \notin g$ means the product is over the nearest-spin pairs between green cubes on the top surface of the lower bulk, and $\langle ij\rangle_b \notin o$ means the product is over the nearest-spin pairs between orange cubes at the bottom surface of the upper bulk. $\langle ij\rangle_{bt}$ means V is the product of Boltzmann weights corresponding to all the bonds connecting the two bulks in the z direction. In Fig. 5, the dashed bonds corresponding to X, Y, and V, are denoted as black, blue, and red, respectively. Note in Eq. (9), the size in the z direction is assumed as $2n + 1$ for convenience.

In the thermodynamic limit, the evaluation of Eq. (9) is again reduced to the dominant eigenvalue problem as before, and the statistical averages can be determined similarly. For example, the bond energy in the z direction can be obtained by

$$E_p = \frac{\langle\Phi|YV'X|\Psi\rangle}{\langle\Phi|YVX|\Psi\rangle}, \quad V' = (-s_a s_b e^{\beta s_a s_b}) \prod_{\langle ij\rangle_{bt} \neq p} e^{\beta s_i s_j}, \tag{11}$$

where $\langle\Phi|$ is the left dominant eigenvector of T_4T_3, and V' differs from V by only a single vertical bond denoted as p

(a) (b)

(c) (d)
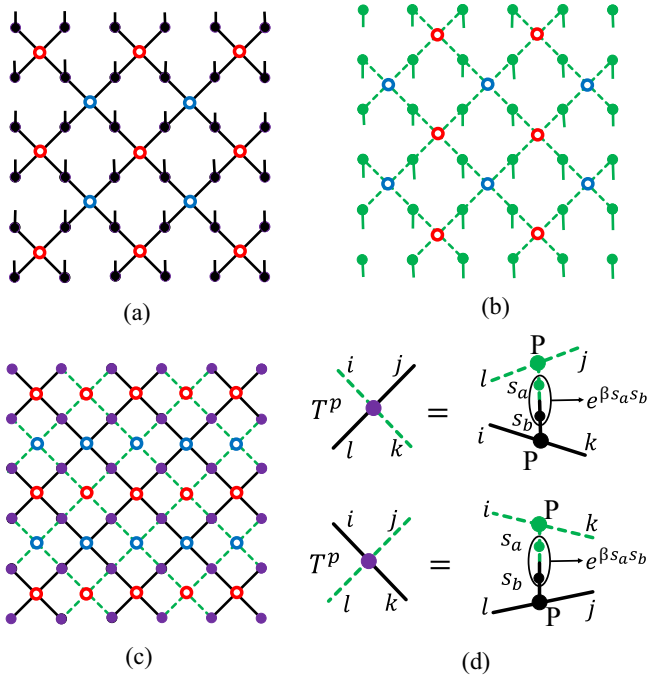
FIG. 6. Illustration of the nesting structure of $\langle\tilde{\Phi}|V|\tilde{\Psi}\rangle$ appeared in Eq. (12). The spatial location corresponds to Fig. 5 exactly. (a) Structure of the ket state $|\tilde{\Psi}\rangle$, whose virtual bonds are indicated as solid black lines. (b) Structure of the bra state $\langle\tilde{\Phi}|$, whose virtual bonds are indicated as dashed green lines. (c) Structure of the inner product $\langle\tilde{\Phi}|V|\tilde{\Psi}\rangle$, where new rank-4 tensor T^p is defined in Eq. (13) and illustrated in (d). Here we have used the symmetry between the two wavefunctions, and both definitions of T^p hold, although it is unnecessary for the nesting method. The local tensors are the parameters in $|\tilde{\Psi}\rangle$ in this context, not to be confused with parameters in $|\Psi\rangle$ in Fig. 2. Clearly, the bond dimension of the resulting tensor network is not squared, which is different from the case in Fig. 4.

with s_a and s_b located at its ends. Finally we have

$$E_p = \frac{\langle\tilde{\Phi}|V'|\tilde{\Psi}\rangle}{\langle\tilde{\Phi}|V|\tilde{\Psi}\rangle}, \quad \langle\tilde{\Phi}| \equiv \langle\Phi|Y, \quad |\tilde{\Psi}\rangle \equiv X|\Psi\rangle. \quad (12)$$

The key observation is that $\langle\tilde{\Phi}|$ and $|\tilde{\Psi}\rangle$ have a similar structure, but the parameters are distributed separately in space, i.e., their simplex tensors are defined separately in space at different sets of squares, which can be seen from Figs. 5(b)–5(d) and Figs. 6(a)–6(c) straightforwardly. And this leads to a great advantage, i.e., if the wavefunctions $\langle\tilde{\Phi}|$ and $|\tilde{\Psi}\rangle$ are approximated by PESS with bond dimension D, then the generated two-dimensional tensor network also has bond dimension D, instead of D^2. To be specific, let us take the denominator as an example. Exploring the symmetry between $\langle\tilde{\Phi}|$ and $|\tilde{\Psi}\rangle$, the resulting tensor network are composed of three parameters, i.e., the local tensors in $|\tilde{\Psi}\rangle$ and $|\tilde{\Phi}\rangle$, and a new local tensor T^p derived from P therein, i.e.,

$$T^p_{ijkl} = \sum_{s_a s_b} P_{lj}[s_a]P_{ki}[s_b]e^{\beta s_a s_b}, \quad (13)$$

as illustrated in Figs. 6(c) and 6(d). Therefore, compared with the reduced method, which generates a tensor network with a squared bond dimension, this nesting technique can greatly
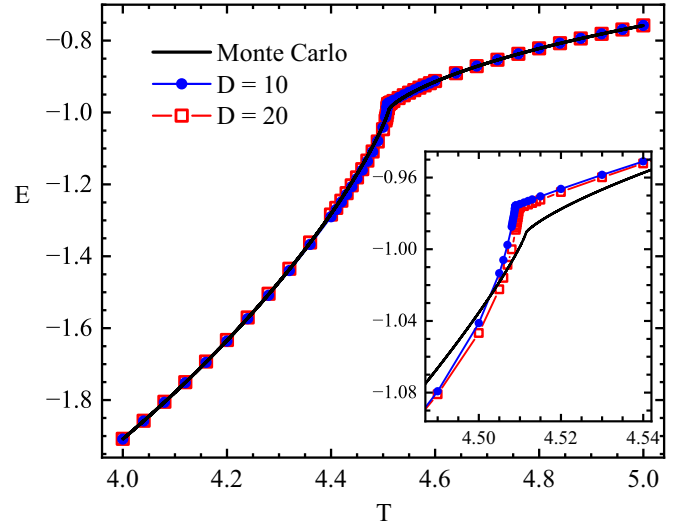

FIG. 7. Energy estimation obtained from the simple update and the nesting method. The Monte Carlo data from Ref. [55] are plotted as a reference.

reduce the cost and leads to much more efficient contraction, as achieved in Ref. [30] similarly. Considering the fact that the environment dimension χ used in the CTMRG iterations has the order D^2 empirically, the nesting technique can reduce the computational cost from D^{12} to D^9. In this paper, we push D to 20 with the help of this technique.

III. RESULTS

In this paper, we focus on the 3D Ising model, which is of long-standing interest in statistical physics and condensed matter physics. Although there is no analytical solution as in two-dimensional case, the higher-order tensor renormalization group (HOTRG) [11,53] and Monte Carlo methods [54–57] have provided very accurate numerical data, which confirms a second-order finite temperature phase transition. Therefore, this model provides a suitable touchstone to test new numerical algorithms in higher dimensions.

In this paper, as described in Sec. II, we use the developed evolution method combined with the simple update technique to determine the tensor-network representation of the dominant eigenvectors of the transfer matrices. Furthermore, to push D to a larger value (up to 20 in this paper), in the expectation value calculations, we always employ the nesting technique described in Sec. II C to improve the efficiency of CTMRG algorithm. Note that when D is small, and the conventional reduced method can be efficiently performed, the results obtained by the CTMRG with and without nesting technique are consistent, as illustrated and discussed in the Appendix.

The result of internal energy E is shown in Fig. 7. The reference curve denotes high-precision Monte Carlo data obtained in Ref. [55]. It shows that the proposed method can give consistent results at the off-critical region, while near the critical point, enlarging the bond dimension D can produce more accurate results as expected. The spontaneous magnetization M shows similar behavior, as shown in Fig. 8, where the reference curve is obtained from Monte Carlo [54]. For both
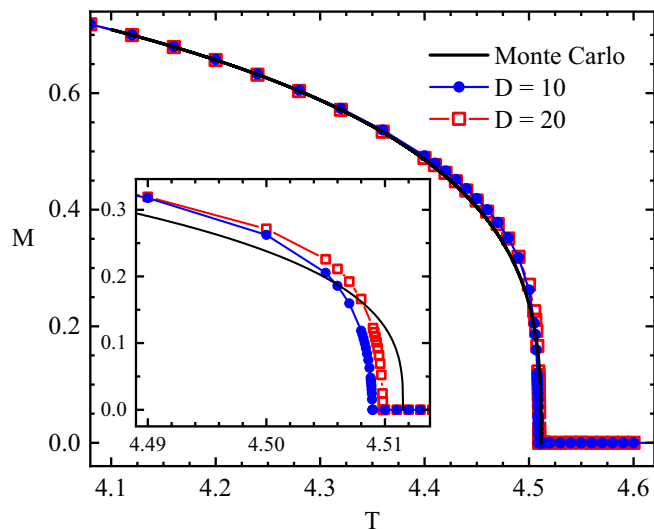
FIG. 8. Magnetization estimation obtained from the simple update and the nesting method. The Monte Carlo data from Ref. [54] are plotted as a reference.

quantities, our result coincides well with previous studies, and the singular behavior can be seen clearly.

From E and M, and the critical temperatures derived as shown in Fig. 9, it seems that the proposed method tends to underestimate the critical temperature a little, which is different from the coarse-graining tensor renormalization group methods as shown, e.g., in Refs. [11,58]. This is probably related to the simple update technique used in the evolution process, since it essentially provides a Bethe lattice approximation of the dominant eigenvector [59,60] and thus tends to be disordered at finite temperature, especially at the temperature close to but lower than critical point where the approximation cannot give a good estimate of the correlation length. When $D = 20$, the estimated critical temperature T_c is about 4.50984(2), which has only about 10^{-4} relative deviation from

TABLE I. Comparison of the critical point T_c for the 3D Ising model obtained from different tensor-network-state-based methods. The Monte Carlo and HOTRG results are included as references. Note that for the methods denoted by (*), variational calculations have been employed and the bond dimensions are rather small ($D \sim 4$). Our results are derived from the magnetization data with $D = 20$ simple-updated wave function.

Method	T_c
Monte Carlo [57]	4.51152326(11)
HOTRG [11,53]	4.51152469(1)
KWA-based TNS* [40]	4.5788
Vertex-type TNS* [41]	4.5392
TPVA* [40,61]	4.5704, 4.554
TNS* [41]	4.504
Algebraic variation* [42]	4.547
TNS data fitting* [44]	4.5118057(41)
TNS data collapse* [45]	4.5104, 4.51170
Simple update (this paper)	4.50984(2)

the best estimation $T_c \sim 4.51152$ [11,53,57]. The comparison with other variational calculations is summarized in Table I. Although obtaining a benchmark result is not our motivation, it shows that our method performs rather well. Compared to the most recent PEPS calculations [44,45], though our result seems slightly less accurate, our method does not need other auxiliary techniques, such as ignoring the long tails of magnetization with careful fitting [44], scaling hypothesis and data collapse [45], bootstrap exponent fixing [45], etc. The accuracy of the critical temperature and the critical exponent can be further improved by further refining the simple-updated wavefunctions through variational calculations and performing these advanced techniques.

Furthermore, as shown in Fig. 10, the critical exponent can be obtained by fitting the magnetization data near critical temperature according to the scaling relation
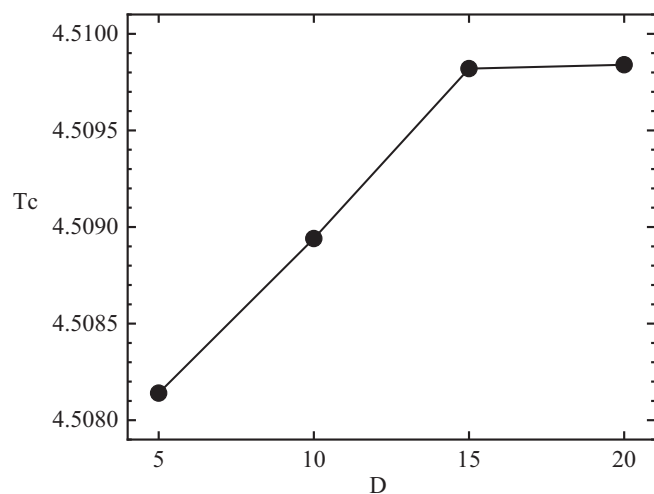
$$M \sim (1 - T/T_c)^{\beta}. \tag{14}$$

FIG. 9. Critical temperature obtained from the magnetization M, with respect to the bond dimension D of the tensor network state representation of the transfer matrices' dominant eigenvectors. The critical temperature is estimated as the lowest temperature, which has almost zero magnetization, as shown in Fig. 8.
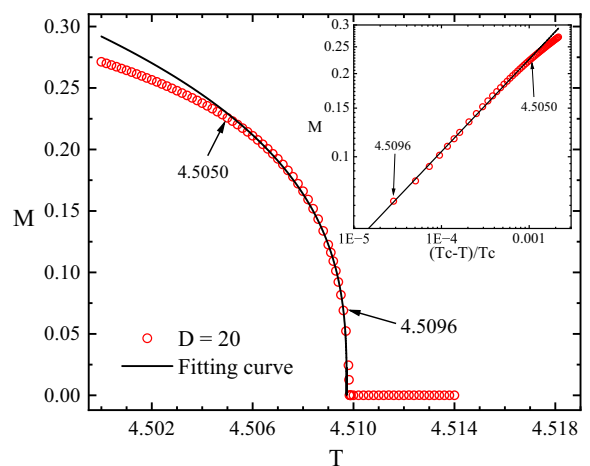

FIG. 10. Critical exponent β for the magnetization obtained from $D = 20$ simple-updated wave function and $\chi = 120$ nesting CTMRG method. β is estimated as 0.335(16), and the data obtained at temperature ranging from 4.5050 to 4.5096 are used for fitting.
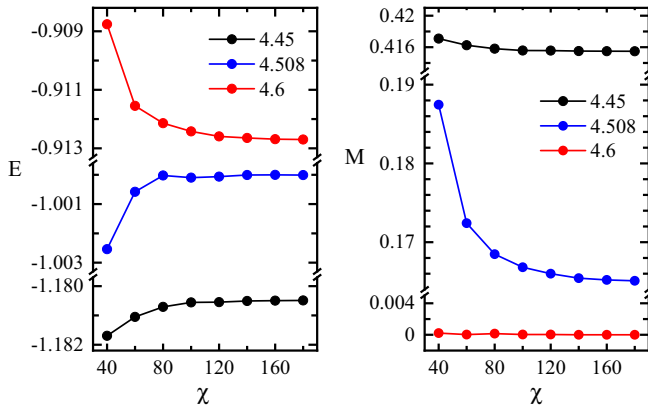
FIG. 11. Convergence analysis about energy and magnetization for $D = 20$. Three typical temperatures are chosen, i.e., temperature in symmetry-breaking phase (4.45), close to critical point (4.508), and in disordered phase (4.6).

We find that the exponent $\beta = 0.335(16)$, close to the Monte Carlo [62] (0.3262), conformal bootstrap [63] (0.326419), and HOTRG [11] (0.3295) results.

As to the nesting technique proposed in this paper, its convergence with respect to the bond dimension χ of the environment tensors used in CTMRG is shown in Fig. 11, for $D = 20$. It shows clearly that for all the three typical temperatures ranging from symmetry-breaking phase to paramagnetic phase, the convergence is at least equally satisfying, compared with the nested tensor network method proposed in Ref. [30]. Particularly, even at the temperature that is very close to the critical value, the convergence is already very acceptable when $\chi = 180$ for $D = 20$, with an error about 10^{-6} for E and 10^{-4} for M. This is a very nice feature for tensor-network methods, especially for large D, as discussed in Ref. [30] in detail.

IV. SUMMARY

In this paper, we propose an efficient numerical method to contract the 3D tensor networks with translational invariance. The result of the contraction is expressed in terms of some transfer matrices, whose dominant eigenvectors are represented as the PESS form and determined by power iterations analogous to the imaginary time evolution algorithm [24,48]. Especially, the PESS representations of the left and right dominant eigenvectors are designed appropriately so that their inner product has a nesting structure and can be expressed as a tensor network with bond dimension D instead of D^2, and thus the contraction can be carried out much more efficiently. As to the 3D Ising model, it can give very consistent results for both energy and magnetization with previous studies, even though only the simple update strategy is employed to update the tensor network states. The convergency of the nesting technique is shown to be at least equally good to that of the nested tensor network method [30]. When $D = 20$, we obtain a critical temperature about 4.50984(2), which is very close to the best-known estimation. As far as we know, this is probably the successful trial of applying the imaginary-time-evolution-like method to 3D classical models without variational update

procedure, thus this extends the application scope of tensor network states. The application to other interesting but unsolvable classical models in three dimensions, such as the Potts model [53], clock model [64], dimer model [44], and lattice gauge models [19], is straightforward.

As mentioned in Sec. III, the simple update strategy employed in this paper is probably a reason why the obtained T_c tends to be underestimated. Besides the scaling hypothesis and data collapse technique [45], this tendency could be eased in two different manners. One is to resort to the more involved but more accurate update methods, such as the cluster update [65] and the full update strategies [66,67], which consider the renormalization effect of the environment better. The other possible way is to utilize more sophisticated ansatz and evolution techniques, such as PESS ansatz with stronger simplex entanglement [26], the recently proposed regularized scheme of time evolution [68], and the so-called minimal canonical form of tensor network states [69], all of which are also expected to produce more accurate representation of the dominant eigenvectors. In both cases, the nesting technique proposed in this paper can still be applied without changing too much, and the improvement in performance can be expected.

At last, it is worth mentioning that the proposed nesting technique in this paper can also be extended to the two-dimensional quantum lattice models. In order to achieve this, one needs to represent the partition function in a different manner, e.g., similarly as we have done in Eq. (9), so that two representations of the ground state with separated distributions in space can be used simultaneously, e.g., as illustrated in Fig. 5(b) and expressed in Eq. (12). This can be advantageous in the study of models, which have square structure and separatable block interactions, such as Shastry-Sutherland model [6,70], checkboard system [71], and so on. Exploring its potential and limits is an interesting and promising project, and we would like to leave it as a future pursuit.

APPENDIX: FURTHER BACKGROUND AND COMPARISONS

In this Appendix, some further background on the concepts and techniques involved in this paper, and some relevant comparisons are provided.

1. PESS wave function ansatz

The PESS ansatz is particularly important in the nesting technique developed in this paper. It was proposed in Ref. [26], and can be regarded as a generalization of the projected entangled pair state (PEPS) ansatz proposed in

Ref. [1]. It extends pair entanglement well characterized by PEPS to simplex entanglement that might be more important in some situations, and has been successfully applied to a series of frustrated spin systems, such as kagome spin liquid systems [8,72], chiral spin liquid systems [17,73], quantum spin-orbital liquid systems [9,74], etc.

As a wavefunction ansatz, the defining feature of a PESS is the existence of simplex tensors in it, each of which connects more than two physical degrees of freedom and captures the many-body entanglement among them. For example, in Eq. (6) and Fig. 2, simplex tensors A^α and B^β are introduced to characterize the entanglement among the four spins defined on the vertices of the αth and βth squares respectively. For any given base ket $|\{s\}\rangle$, the corresponding coefficient in the superposition is given by the Tr operation, which means the contraction of a two-dimensional tensor network and can be accomplished by, e.g., the CTMRG method. For general discussions about tensor-network state, Refs. [1–3] are excellent references covering almost all important topics.

In this paper, the translational symmetry and the C_4 rotational symmetry of the simplex tensors are employed, thus we have only one independent A, B, and P as variational parameters that need to be determined by the simple-update-aided imaginary-time evolution method, as discussed in the main text and Ref. [26].

2. CTMRG algorithm for arbitrary supercell

The CTMRG is a highly-efficient and extensively-used algorithm to contract two-dimensional tensor networks, and in this paper it has been used to calculate the expectation values, since the numerator and denominator in both Eqs. (8) and (12), as well as other terms similar to the inner product of two tensor-network states, can be represented as two-dimensional tensor networks exactly. In this paper, for a given tensor-network state with bond dimension D, the difference in the expectation value calculation between the contraction of a normal network and a network with nesting structure, lies in the fact that the normal one has bond dimension D^2 while the nested one has dimension D. Therefore, though both the two kinds of tensor networks can be contracted by CTMRG algorithm, the leading cost differs greatly, i.e., D^{12} for the normal one and D^9 for the nested one (can be reduced to D^{10} and D^8 respectively by some advanced techniques, e.g., partial SVD.)

In this paper, we used the CTMRG algorithm for arbitrary supercell that was proposed in Ref. [4]. For the normal network depicted in Fig. 4, the supercell size is 2×2 (one can always absorb the Q tensors into T^a and/or T^b), while for the nested network depicted in Fig. 6, the supercell size is 4×4. In the CTMRG algorithm, each independent tensor in the supercell is associated with four corner tensors and four edge tensors, which mimic its effective environment and are used to extract the expectation values. For a supercell with size $m \times n$, we have stored $4mn$ corners and $4mn$ edges in total, which are initialized as random numbers and are updated iteratively until convergence is reached. The detailed iterative relations among these corners and edges are determined according to their relative spatial locations, and have been given clearly in Ref. [4]. For general discussions about the CTMRG algorithm, Ref. [3]
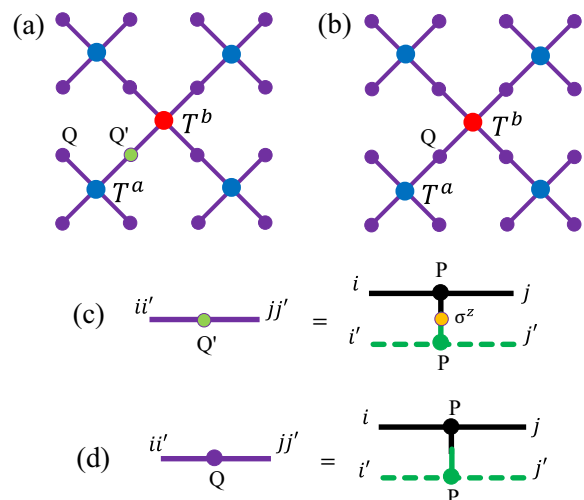

FIG. 12. Sketch of the impurity method and the explanation of Eq. (15). (a) The numerator $\langle\Psi'|s_i|\Psi\rangle$ represented as a 2D tensor network. (b) The denominator $\langle\Psi'|\Psi\rangle$ represented as a 2D tensor network. Note that almost all the tensors in (a) and (b) are the same, as defined in Fig. 4, and the only difference lies in two tensors defined on the ith site, i.e., Q' (impurity tensor, denoted as bright green) as defined in (c) and Q as defined in (d). In the definition of Q', the operator s_i is included, and here its matrix representation σ^z is a diagonal matrix with nonzero elements ± 1.

provide excellent and friendly reviews. For the most recent development, Refs. [75,76] provide promising trials.

3. Impurity method in tensor-network states

The impurity method is frequently used to calculate the expectation values without resorting to the derivatives. It has been successfully applied early in Refs. [47,58], and was elaborated later in Refs. [27,52]. In essence, it aims to calculate the expectation values of some local operators, e.g., the bond
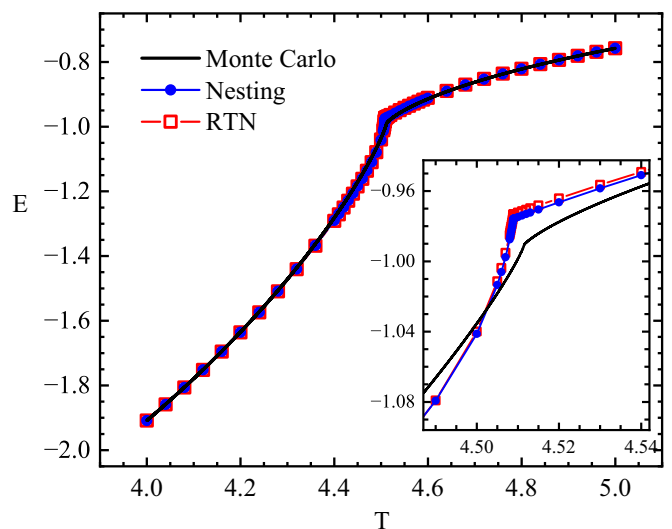

FIG. 13. Comparison between the internal energy E obtained from the conventional RTN method and the nesting technique, for $D = 10$. Note that the PESS wave functions used in these two methods are the same, and in both methods, the network is contracted by the CTMRG algorithm.
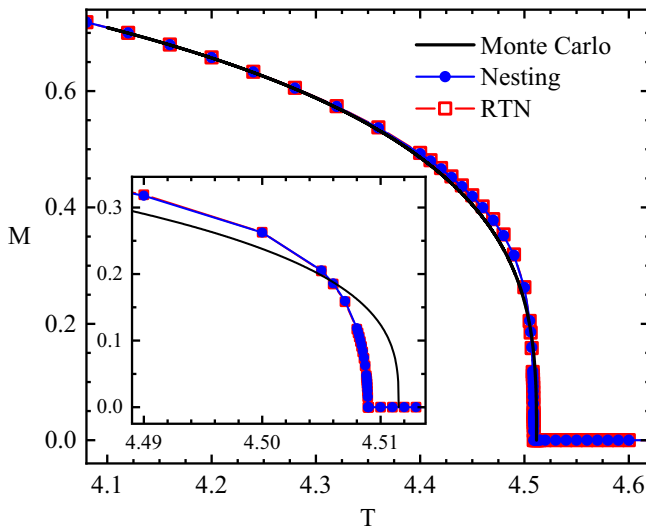
FIG. 14. Comparison between the magnetization M obtained from the conventional RTN method and the nesting technique, for $D = 10$. Note that the PESS wave functions used in these two methods are the same, and in both methods, the network is contracted by the CTMRG algorithm.

energy or the magnetization, which can be expressed as a ratio of two quantities represented as tensor networks and differing from with each other by only very few tensors (dubbed impurity tensors). For simplicity, let us take the on-site magnetization operator M_i as an example, and repeat Eq. (8) in the following:

$$M_i = \frac{\langle \Psi'|s_i|\Psi \rangle}{\langle \Psi'|\Psi \rangle} \quad (15)$$

where $\langle \Psi'|$ and $|\Psi \rangle$ are the left and right dominant eigenvectors of $T_2 T_1$ respectively. Suppose the two eigenvectors have been represented as PESS form, as shown in Fig. 2, then the numerator and the denominator can be represented as Figs. 12(a) and 12(b) respectively. It shows that there is only one special impurity tensor Q' in Fig. 12(a), which constitutes

the only difference from Fig. 12(b). Thus Q and Q' share exactly the same effective environment $N^{(e)}$, and one can evaluate the ratio M_i by determining $N^{(e)}$ through well-developed methods, e.g., the CTMRG method as described above, and obtain the following simple ratio form of M_i,

$$M_i = \frac{\mathrm{Tr}(N^e Q')}{\mathrm{Tr}(N^e Q)} \quad (16)$$

finally. For other local operators like bond energy, the process is very similar but involves two impurity tensors. More details and examples can be found in Ref. [52].

4. The result comparison between the CTMRG with and without nesting technique

In this paper, the data shown in the figures in the main text are all obtained by the CTMRG algorithms combined with the nesting technique. When D is small and the conventional reduced tensor network method, i.e., RTN, can be efficiently performed, we can compare the results obtained by the two methods. The results are shown in Figs. 13 and 14. In both figures, complete convergence of the CTMRG calculations concerning the environment dimensions χ has been reached for both methods. It shows that both the internal energy E and M are very consistent for the two methods. Even though it seems that near critical region, the energy result obtained by the nesting technique is more closer to the Monte Carlo data, they are essentially the same. Actually the tiny difference comes not from the two methods themselves [i.e., whether use Fig. 4 or Fig. 6(c) to represent the overlap of two PESSs, both of which are exact for two given representations], but from the fact that the methods employ different decompositions of the partition function [i.e., Fig. 1 and Eq. (2) in RTN, while Fig. 5 and Eq. (9) in the nesting technique], and that accordingly the two expectation value calculations involve different tensor networks. This tiny difference is expected to vanish gradually as D becomes larger, and should vanish in quantum lattice models where exact Fig. 5 and Eq. (9) are unnecessary when employing the nesting technique.

[1] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066; J. I. Cirac, D. Perez-Garcia, N. Schuch, and F. Verstraete, Rev. Mod. Phys. **93**, 045003 (2021).

[2] S. Montangero, *Introduction to Tensor Network Methods* (Springer, New York, 2018).

[3] R. Orus, Ann. Phys. **349**, 117 (2014); Eur. Phys. J. B **87**, 280 (2014); Nat. Rev. Phys. **1**, 538 (2019).

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

[5] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. Kin-Lic Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos *et al.* (Simons Collaboration on the Many-Electron Problem), Phys. Rev. X **5**, 041041 (2015); **10**, 031016 (2020); B. X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, G. K.-L. Chan, Science **358**, 1155 (2017).

[6] P. Corboz and F. Mila, Phys. Rev. Lett. **112**, 147203 (2014); Phys. Rev. B **87**, 115144 (2013).

[7] L. Wang, Z. C. Gu, F. Verstraete, and X. G. Wen, Phys. Rev. B **94**, 075143 (2016).

[8] H. J. Liao, Z. Y. Xie, J. Chen, Z. Y. Liu, H. D. Xie, R. Z. Huang, B. Normand, and T. Xiang, Phys. Rev. Lett. **118**, 137202 (2017).

[9] Q. Li, H. Li, J. Z. Zhao, H. G. Luo, and Z. Y. Xie, Phys. Rev. B **105**, 184418 (2022).

[10] G. Ferrari, G. Magnifico, and S. Montangero, Phys. Rev. B **105**, 214201 (2022).

[11] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, Phys. Rev. B **86**, 045139 (2012).

[12] C. Wang, S. M. Qin, and H. J. Zhou, Phys. Rev. B **90**, 174201 (2014).

[13] J. G. Liu, L. Wang, and P. Zhang, Phys. Rev. Lett. **126**, 090506 (2021).

[14] C. Wille, O. Buerschaper, and J. Eisert, Phys. Rev. B **95**, 245127 (2017).

[15] W. T. Xu, Q. Zhang, and G. M. Zhang, Phys. Rev. Lett. **124**, 130603 (2020).

[16] D. J. Williamson, N. Bultinck, and F. Verstraete, arXiv:1711.07982.

[17] R. Wang, Z. Y. Xie, B. G. Wang, and T. Sedrakyan, Phys. Rev. B **106**, L121117 (2022).

[18] F. Verstraete and J. I. Cirac, Phys. Rev. Lett. **104**, 190405 (2010).

[19] Y. Z. Liu, Y. Meurice, M. P. Qin, J. Unmuth-Yockey, T. Xiang, Z. Y. Xie, J. F. Yu, and H. Y. Zou, Phys. Rev. D **88**, 056005 (2013).

[20] A. Tilloy and J. I. Cirac, Phys. Rev. X **9**, 021040 (2019).

[21] Z. Y. Han, J. Wang, H. Fan, L. Wang, and P. Zhang, Phys. Rev. X **8**, 031012 (2018).

[22] Z. F. Gao, S. Cheng, R. Q. He, Z. Y. Xie, H. H. Zhao, Z. Y. Lu, and T. Xiang, Phys. Rev. Res. **2**, 023300 (2020).

[23] Y. Meurice, R. Sakai, and J. Unmuth-Yockey, Rev. Mod. Phys. **94**, 025005 (2022).

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

[25] G. Vidal, Phys. Rev. Lett. **98**, 070201 (2007).

[26] Z. Y. Xie, J. Chen, J. F. Yu, X. Kong, B. Normand, and T. Xiang, Phys. Rev. X **4**, 011025 (2014).

[27] H. H. Zhao, Z. Y. Xie, Q. N. Chen, Z. C. Wei, J. W. Cai, and T. Xiang, Phys. Rev. B **81**, 174411 (2010).

[28] P. Schmoll, S. Singh, M. Rizzi, and R. Orus, Ann. Phys. **419**, 168232 (2020); M. Mambrini, R. Orus, and D. Poilblanc, Phys. Rev. B **94**, 205124 (2016).

[29] L. Wang, I. Pizorn, and F. Verstraete, Phys. Rev. B **83**, 134421 (2011); W.-Y. Liu, S.-J. Dong, Y.-J. Han, G.-C. Guo, and L. X. He, *ibid.* **95**, 195154 (2017); H.-H. Zhao, K. Ido, S. Morita, and M. Imada, *ibid.* **96**, 085103 (2017).

[30] Z. Y. Xie, H. J. Liao, R. Z. Huang, H. D. Xie, J. Chen, Z. Y. Liu, and T. Xiang, Phys. Rev. B **96**, 045128 (2017).

[31] S. Sachdev, *Quantum Phase Transitions*, 2nd edition (Cambridge University Press, Cambridge, 2011).

[32] M. Suzuki, Prog. Theor. Phys. **56**, 1454 (1976).

[33] T. Nishino, J. Phys. Soc. Jpn. **64**, 3598 (1995).

[34] R. J. Bursill, T. Xiang, and G. A. Gehring, J. Phys.: Condens. Matter, **8**, L583 (1996).

[35] H. Niggemann, A. Klumper, and J. Zittartz, Zeitschrift für Physik B Condensed Matter **104**, 103 (1997).

[36] P. Y. Teng, Physica A **472**, 117 (2017).

[37] D. Kadoh and K. Nakayama, arXiv:1912.02414.

[38] D. Adachi, T. Okubo, and S. Todo, Phys. Rev. B **102**, 054432 (2020).

[39] T. Nishino and K. Okunishi, J. Phys. Soc. Jpn. **67**, 3066 (1998).

[40] K. Okunishi and T. Nishino, Prog. Theor. Phys. **103**, 541 (2000); T. Nishino, K. Okunishi, Y. Hieida, N. Maeshima, and Y. Akutsu, Nucl. Phys. B **575**, 504 (2000) ; A. Gendiar and T. Nishino, Phys. Rev. E **65**, 046702 (2002).

[41] T. Nishino, Y. Hieida, K. Okunishi, N. Maeshima, Y. Akutsu, and A. Gendiar, Prog. Theor. Phys. **105**, 409 (2001); A. Gendiar, N. Maeshima, and T. Nishino, *ibid.* **110**, 691 (2003).

[42] S. G. Chung, Phys. Lett. A **359**, 707 (2006).

[43] R. Orus, Phys. Rev. B **85**, 205117 (2012).

[44] L. Vanderstraeten, B. Vanhecke, and F. Verstraete, Phys. Rev. E **98**, 042145 (2018).

[45] B. Vanhecke, J. Hasik, F. Verstraete, and L. Vanderstraeten, Phys. Rev. Lett. **129**, 200601 (2022).

[46] K. Okunishi, T. Noshinio, and H. Ueda, J. Phys. Soc. Jpn. **91**, 062001 (2022).

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

[48] Note that essentially the similar idea has been applied to two-dimensional classical models under the help of the canonical form of matrix product state. For example, see R. Orus, and G. Vidal, Phys. Rev. B **78**, 155117 (2008).

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

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

[51] N. Schuch, D. Poilblanc, J. I. Cirac, and D. P. Garcia, Phys. Rev. B **86**, 115108 (2012).

[52] H. H. Zhao, Z. Y. Xie, T. Xiang, and M. Imada, Phys. Rev. B **93**, 125115 (2016).

[53] S. Wang, Z. Y. Xie, J. Chen, B. Normand, and T. Xiang, Chin. Phys. Lett. **31**, 070503 (2014).

[54] A. L. Talapov and H. W. J. Blote, J. Phys. A: Math. Gen. **29**, 5727 (1996).

[55] Y. J. Deng and H. W. J. Blote, Phys. Rev. E **68**, 036125 (2003); X. M. Feng and H. W. J. Blote, *ibid.* **81**, 031103 (2010).

[56] M. Hasenbusch, Phys. Rev. B **82**, 174433 (2010).

[57] A. M. Ferrenberg, J. Xu, and D. P. Landau, Phys. Rev. E **97**, 043301 (2018).

[58] Z. Y. Xie, H. C. Jiang, Q. N. Chen, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **103**, 160601 (2009).

[59] W. Li, J. von Delft, and T. Xiang, Phys. Rev. B, **86**, 195137 (2012).

[60] H. J. Liao, Z. Y. Xie, J. Chen, X. J. Han, H. D. Xie, B. Normand, and T. Xiang, Phys. Rev. B **93**, 075154 (2016).

[61] A. Gendiar and T. Nishino, Phys. Rev. B **71**, 024404 (2005).

[62] M. Hasenbusch, Int. J. Mod. Phys. C **12**, 911 (2001).

[63] F. Kos, D. Poland, D. S. Duffin, and A. Vichi, J. High Energ. Phys. **08**, 36 (2016).

[64] H. Shao, W. Guo, and A. W. Sandvik, Phys. Rev. Lett. **124**, 080602 (2020).

[65] L. Wang, and F. Verstraete, arXiv:1110.4362.

[66] J. Jordan, R. Orus, G. Vidal, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **101**, 250602 (2008).

[67] M. Lubasch, J. I. Cirac, and M. C. Banuls, Phys. Rev. B **90**, 064425 (2014); H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, and R. Orus, *ibid.* **92**, 035142 (2015).

[68] L. X. Cen, arXiv:2208.03436.

[69] A. Acuaviva, V. Makam, H. Nieuwboer, D. P. Garcia, F. Sittner, M. Walter, and F. Witteveen, arXiv:2209.14358.

[70] B. S. Shastry and B. Sutherland, Physica B+C **108**, 1069 (1981); J. Yang, A. W. Sandvik, and L. Wang, Phys. Rev. B **105**, L060409 (2022).

[71] H. Y. Zou, F. Yang, and W. Ku, arXiv:2011.06520.

[72] C. Y. Lee, B. Normand, and Y. J. Kao, Phys. Rev. B **98**, 224414 (2018).

[73] S. Niu, J. Hasik, J. Y. Chen, and D. Poilblanc, Phys. Rev. B **106**, 245119 (2022).

[74] Y. Liu, Z. Y. Xie, H. G. Luo, and J. Z. Zhao, Phys. Rev. B **107**, L041106 (2023).

[75] X. F. Liu, Y. F. Fu, W. Q. Yu, J. F. Yu, and Z. Y. Xie, Chin. Phys. Lett. **39**, 067502 (2022).

[76] I. V. Lukin and A. G. Sotnikov, Phys. Rev. B **107**, 054424 (2023).