

Nuclear norm regularized loop optimization for tensor network

Kenji Homma¹, Tsuyoshi Okubo², and Naoki Kawashima¹

¹*Institute for Solid State Physics, The University of Tokyo, Kashiwa, Chiba 277-8581, Japan*

²*Department of Physics, The University of Tokyo, Tokyo 113-0033, Japan*



(Received 1 July 2024; accepted 16 October 2024; published 5 November 2024)

We propose a loop optimization algorithm based on nuclear norm regularization for the tensor network. The key ingredient of this scheme is to introduce a rank penalty term proposed in the context of data processing. Compared to the standard variational periodic matrix product states method, this algorithm can circumvent the local minima related to short-ranged correlation in a simpler fashion. We demonstrate its performance when used as a part of the tensor network renormalization algorithms [S. Yang, Z.-C. Gu, and X.-G. Wen, *Phys. Rev. Lett.* **118**, 110504 (2017)] for the critical two-dimensional Ising model. The scale invariance of the renormalized tensors is attained with higher accuracy while the higher parts of the scaling dimension spectrum are obtained in a more stable fashion.

DOI: [10.1103/PhysRevResearch.6.043102](https://doi.org/10.1103/PhysRevResearch.6.043102)

I. INTRODUCTION

Tensor network is a convenient representation of quantum and classical many-body problems in that effective truncation of degrees of freedom can be realized in a flexible way. Because of this capability, a number of promising approximation schemes have been proposed, as typified by density matrix renormalization group (DMRG) [1,2] and corner transfer matrix renormalization group (CTMRG) [3,4]. Developments of progressive techniques such as multiscale entanglement renormalization ansatz (MERA) [5] and projected entangled pair states (PEPS) [6–8] would allow us to elucidate the complex many-body systems at finite bond dimension.

However, understanding the expressive power of tensor network for many-body systems is still limited to loop-free network. For the bipartite tensor network ansatz, its gauge freedoms of the loop-free tensor network can be fixed by introducing canonical gauges [9]. Its key idea was to find the center of orthogonality of tensor networks. This led to the success of simulating infinite matrix product states (MPS) [9–12] and generalized to any loop-free tensor networks. Yet, as the authors of Ref. [13] discussed, there did not exist an effective gauge-fixing method for looped tensor networks as the bipartite canonical gauge was ill-defined for looped structure. Thus, developing an efficient and computationally light method for training looped tensor network remains as a challenging task.

This issue can be traced back to short-range correlation problems in tensor renormalization group (TRG) [14]. Although TRG and its modified versions [15,16] have sufficient capability to compute partition functions of statistical models

at moderate accuracies, Gu and Wen pointed out the TRG proposed by Levin-Nave (LN-TRG) does not produce correct RG flow because simple singular value decomposition (SVD) cannot remove the short-range correlations that lives in tensor space [17]. Recently, tensor network renormalization (TNR) [18,19] and its variants [20–22] were proposed to deal with such reduction of short-range correlations and produced more accurate results than TRG. For example, the so-called Loop-TNR approach [20] was formulated as “loop optimization,” which is to decompose the tensor network into a loop-like configuration and variational method on minimizing the distance between two periodic MPSs [23,24]. Typically, it finds the minimum by solving the linearization of cost function. However, such strategies will typically get stuck in local minima. Imposing symmetry restrictions to the search space, and/or trying better initial conditions are common practices to ease the problem. In the case of latter, one can start optimizing periodic MPS from the preconditioned points to accelerate the convergence. In general, a “good” preconditioned point highly depends on the underlying problem characteristics, which requires one to speculate on the nature of local minima. We conjecture that the short-range correlation that lives in tensor space might be the origin of local minima for loop optimization.

In this paper, we propose alternative loop optimization algorithms, which we borrow from tensor ring decomposition techniques developed in the context of general data processing [25,26]. Our strategy is to circumvent the local minima originating from short-range correlations by introducing a rank penalty term using nuclear norm regularization (NNR) [27–29], which is closely related to least absolute shrinkage and selection operator (LASSO) regression but includes much broader applications [30]. It is known that the LASSO regression induces sparsity in solution, and similarly, the NNR induces sparsity of singular values, which corresponds to low-rank solution. Our results suggest that adding these regularization term in the tensor network’s optimization is an

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

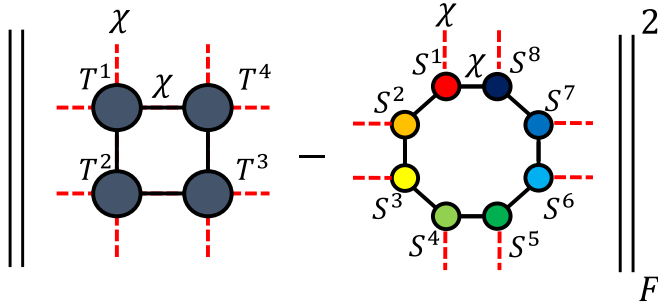


FIG. 1. The cost function of loop optimization. One can regard it as minimizing the distance between two different MPS. The red dotted line corresponds to physical indices and the black solid line are virtual indices of MPS with bond dimension χ .

effective approach for designing optimal entanglement structure. We examine its performance using it as a part of the LN-TRG approach and found that our strategy enables us to obtain the critical fixed-point tensor with high accuracies but less human effort.

This paper is organized as follows. In Sec. II, we explain the target situation of NNR loop optimization and present its algorithmic details. In Sec. III, we compare our proposal against conventional Loop-TNR. Thanks to NNR loop optimizations, short-ranged correlation are efficiently removed even at criticality and we find that our method further improves the scale-invariance of the renormalization group, hence it stabilizes higher parts of the scaling dimension spectrum as a function of renormalization steps compared to Loop-TNR. A discussion and future applications of NNR loop optimization are presented in Sec. IV.

II. LOOP OPTIMIZATION BASED ON NUCLEAR NORM REGULARIZATION

In the case presented below, we consider the following loop optimization problems in Fig. 1 for TNR application. Specifically,

$$\|T^1 \cdot T^2 \cdot T^3 \cdot T^4 - S^1 \cdot S^2 \cdot S^3 \cdot S^4 \cdot S^5 \cdot S^6 \cdot S^7 \cdot S^8\|_F^2, \quad (1)$$

where “ $\cdot$ ” traces out the connecting indices in Fig. 1 and $\|\cdots\|_F$ represents the Frobenius norm. The square configurations of T is now deformed into an octagon consisted of the three-index tensor S . In this paper, we shall call them as target and variational tensors, respectively. Typically, three-index tensors are obtained by LN-TRG and treated independently. In addition, one can regard Eq. (1) as the minimizing the distance of two wave functions with physical indices indicated by red dashed lines in Fig. 1 as $\|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2$. Here, these calligraphic letters, for example, $\mathcal{T}$, refer to the sets of tensors $\{T^1, T^2, \dots\}$ to compose the periodic MPS $|\psi(\mathcal{T})\rangle$. Thus, the problem we are concerned with below is a general MPS optimization problem.

A. Short-range correlations

To illustrate the algorithm, we consider the following special case where the target tensor T in Fig. 2(a) has the

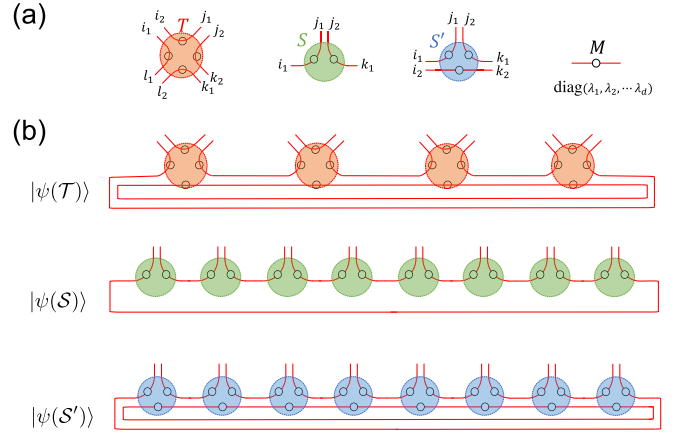


FIG. 2. (a) Target CDL tensor $T_{(i_1 i_2)(j_1 j_2)(k_1 k_2)(l_1 l_2)}$ and two examples of variational tensor $S_{i_1(j_1 j_2)k_1}$, $S'_{(i_1 i_2)(j_1 j_2)(k_1 k_2)}$. (b) periodic MPS wave functions composed of T , S , and S' under translational invariance.

form as

$$T_{(i_1 i_2)(j_1 j_2)(k_1 k_2)(l_1 l_2)} = M_{i_2 j_1} M_{j_2 k_2} M_{k_1 l_2} M_{l_1 i_1}, \quad (2)$$

where $M = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$ and $\lambda_1 > \lambda_2 > \dots > \lambda_d$. Each subscript of T , such as (i_1, i_2) , represents a single natural number, e.g., $i = (i_1, i_2) = i_1 + \eta(i_2 - 1)$ with $i_1, i_2 = 1, 2, \dots, \eta$. The integer η is related to the bond dimension by $\chi = \eta^2$. The tensor T is called corner double line (CDL) tensor, which is a toy model for the short-ranged correlation. In addition, let us define two different variational tensors S and S' for discussion

$$S_{i_1(j_1 j_2)k_1} = M_{i_1 j_1} M_{j_2 k_1}, \quad (3)$$

$$S'_{(i_1 i_2)(j_1 j_2)(k_1 k_2)} = M_{i_1 j_1} M_{j_2 k_1} M_{k_2 i_2}, \quad (4)$$

which follows the index convention we define above. One can construct periodic MPS $|\psi(\mathcal{T})\rangle$, $|\psi(\mathcal{S})\rangle$, and $|\psi(\mathcal{S}')\rangle$ from copies of T , S , and S' under the assumption of translational invariance, as shown in Fig. 2(b). The first thing to note is that the two wave functions $|\psi(\mathcal{S})\rangle$ and $|\psi(\mathcal{S}')\rangle$ are identical to each other up to some normalization factor, as the inner correlation loop formed by $M_{i_2 k_2}$ can be contracted to an overall multiplicative factor. In this sense, the wave function $|\psi(\mathcal{S}')\rangle$ has redundant correlation compared to $|\psi(\mathcal{S})\rangle$. Furthermore, the objective function in Eq. (1) gives $\|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2 = \|\psi(\mathcal{T}) - \psi(\mathcal{S}')\|_F^2 = 0$, which implies that one cannot distinguish them from Eq. (1). The key issues in loop optimizing this periodic MPS can be addressed in the indistinguishability of the relevant correlation and redundant correlation [13], and we speculate that it is the origin of the local minima for loop optimization. In the Loop-TNR approach, there are several methods to implicitly avoid such local minima related to the redundant correlation, such as the so-called “entanglement filtering” and imposing the rotational lattice symmetry during the loop optimization. In this work, we propose an alternative approach to penalize such a redundant correlation via low-rank regularization.

B. NNR loop optimization

In the following, we consider penalizing the high-rank contributions of variational tensors, then gradually relax the penalty term to the original cost function Eq. (1). However, the rank of a tensor is not convenient for the penalty term because it is an integer and also because it is a rather complicated function of the tensor elements. To overcome this problem, one of the commonly used measures in the rank regularization approach is the *nuclear norm*, which is defined as the sum of singular values. The NNR is a concept that has a broad area of possible applications, e.g., the basis pursuit denoising problem in compressed sensing literature, and the LASSO regression popular in machine learning [30]. The fact that the singular values enter as their first power makes the whole algorithm more stable in convergence to the lower rank solution, as we see in the LASSO regression.

While we penalize the high-rank representation through the NNR, there may be several other ways for the same purpose. Among them, we tried local entropy, or mutual information, as a possible alternative to the nuclear norm as we discuss in detail in Appendix B [31]. It turned out that estimation of the central charge and scaling dimensions are noisier and less stable compared to the nuclear norm, while the free-energy precision is comparable or even better in some cases. Though we have not identified the reason of this observations, we suspect that the singular nature of the nuclear norm may play an important role in improving the estimation of critical properties.

The method discussed in this section is a straight-forward application of the NNR loop optimization for the tensor ring decomposition proposed in Refs. [25,26], yet differs in the following two aspects. One is to dynamically adjust the amplitude of the penalty term so that it would not spoil the high precision. This treatment is relevant because, when the redundant correlation loop is only approximately redundant, i.e., when the redundant correlation is still weakly coupled to the other relevant degrees of freedom, which is usually the case in early iterations of the RG procedure, it is harmful to enforce the low-rank representation by the penalty term. The other aspect is that we incorporate constraints on the indices to penalize redundant correlations on the diagonal legs in Fig. 1, which is to enhance consistency with the findings in Ref. [17].

Below, we detail how we define the penalty term using low-rank regularization for periodic MPS. The objective function under nuclear norm regularization in the Fig. 3(a) is given by

$$\min_S \frac{1}{2} \|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2 + \sum_{n=1}^N \mu \|S_{[\alpha(n)]}^n\|_* \quad (5)$$

Here, μ is a nonnegative parameter and $\|\cdots\|_*$ denotes the nuclear norm, i.e., $\|S_{[\alpha(n)]}^n\|_* \equiv \sum_i \sigma_{\alpha(n),i}^n$ where $\sigma_{\alpha(n),i}^n$ are the i th singular value of the α th mode “matricized” tensor $S_{[\alpha(n)]}^n$ at the n th site. For example, $S_{[1]}$ is the matrix such that $S_{[1]} := S_{i,(jk)}$ where the doublet (jk) represents, as before, a single integer that uniquely specifies both j and k [see also Fig. 3(b)]. $\alpha(n)$ is the index corresponding to the diagonal legs in Fig. 1, e.g., for S^1 , $\alpha(n)$ specifies the leg connecting S^1 to S^2 , and for S^8 it specifies the leg connecting to S^7 . To achieve this, we define the index function $\alpha(n)$ as $\alpha(n) = 1$ for n being odd, and $\alpha(n) = 3$ for n being even. Note that we

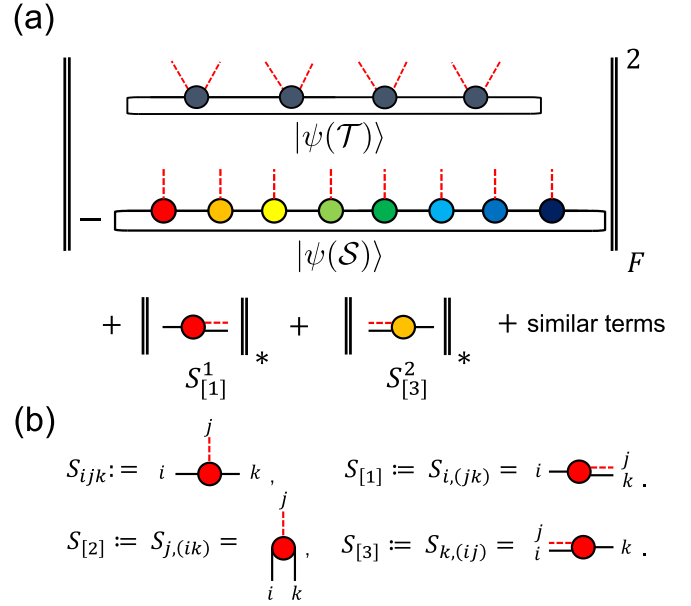


FIG. 3. (a) The cost function of NNR loop optimization. (b) Mode- α matricization. The lower indices $S_{[\alpha]}$ indicate which indices to be grouped.

assume, for later convenience, each target tensor T has been normalized, such that $\langle \psi(\mathcal{T}) | \psi(\mathcal{T}) \rangle = 1$. The initial value of variational tensor S is computed from LN-TRG. The second term penalizes the nuclear norm of matrices $S_{[\alpha(n)]}$, which is supposed to restrict the high-rank solutions of $|\psi(\mathcal{S})\rangle$ and thus favors compact tensor network representations during optimization.

In what follows, for each three-index tensor S^n , we introduce another set of three-index tensors M^n as auxiliary variables to handle the first and second terms in Eq. (5) independently, which is given by

$$\begin{aligned} \min_{S, \mathcal{M}} \frac{1}{2} \|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2 + \sum_{n=1}^N \mu \|M_{[\alpha(n)]}^n\|_* \\ \text{s.t. } M_{[\alpha(n)]}^n = S_{[\alpha(n)]}^n. \end{aligned} \quad (6)$$

As is discussed below, while $M_{[\alpha(n)]}^n = S_{[\alpha(n)]}^n$ will be achieved after sufficient iteration, we let them differ from each other during the optimizations. Incorporating the condition in Eq. (6) in the style of alternating direction method of multipliers (ADMM), the augmented Lagrangian can be written by introducing the Lagrange multipliers Y^n to impose the condition $M_{[\alpha(n)]}^n = S_{[\alpha(n)]}^n$, as

$$\begin{aligned} \mathcal{L}(S, \mathcal{M}, \mathcal{Y}) = \frac{1}{2} \|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2 + \sum_{n=1}^N \mu \|M_{[\alpha(n)]}^n\|_* \\ + \text{tr}[Y_{[\alpha(n)]}^n (M_{[\alpha(n)]}^n - S_{[\alpha(n)]}^n)] \\ + \frac{\xi}{2} \|M_{[\alpha(n)]}^n - S_{[\alpha(n)]}^n\|_F^2, \end{aligned} \quad (7)$$

where $0 < \xi < 1$ is a penalty parameter. The initial conditions of M^n, Y^n are zero-elements tensors. Updated equations of a given augmented Lagrangian can be obtained by the

extremum of Eq. (7) with respect to S^n and M^n for $n = 1, 2, \dots, N$. In what follows, we explain each update equations of S^n , M^n , Y^n for the single iteration.

1. Solving Scheme for S^n

One can simplify the augmented Lagrangian function of S^n as

$$\mathcal{L}(S^n) = \frac{1}{2} \|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2 + \sum_{n=1}^N \text{tr} \left[Y_{[\alpha(n)]}^n (M_{[\alpha(n)]}^n - S_{[\alpha(n)]}^n) + \frac{\xi}{2} \|M_{[\alpha(n)]}^n - S_{[\alpha(n)]}^n\|_F^2 \right]. \quad (8)$$

The first quadratic term in Eq. (8) can be expanded in terms of S^n as

$$\begin{aligned} \|\psi(\mathcal{T}) - \psi(\mathcal{S})\|_F^2 &= \langle \psi(\mathcal{T}) | \psi(\mathcal{T}) \rangle + \langle \psi(\mathcal{S}) | \psi(\mathcal{S}) \rangle - \langle \psi(\mathcal{S}) | \psi(\mathcal{T}) \rangle - \langle \psi(\mathcal{T}) | \psi(\mathcal{S}) \rangle \\ &= 1 + \sum_{ijklm} (S_{ijk}^n)^\dagger N_{iklm}^n S_{ljm}^n - S_{ljm}^n (W_{ljm}^n)^\dagger - (S_{ijk}^n)^\dagger W_{ijk}^n, \end{aligned} \quad (9)$$

or diagrammatically.

Suppose all tensors but S^n are fixed, the minimum of Eq. (8) can be obtained by solving the linear equation

$$(N_{(ik)(lm)}^n + \xi I)(S_{[2]}^n)^\dagger = W_{[2]}^n + Y_{[\alpha(n)]}^n + \xi M_{[\alpha(n)]}^n, \quad (10)$$

where $I \in \mathbb{R}^{\chi^2 \times \chi^2}$ denotes identity matrix. One can update as $S^n \leftarrow \text{tensorize}(S_{[2]}^n)$ for $n = 1, 2, \dots, N$. Here, *tensorize* denotes the tensorization of the given input matrices into the three-index tensor. We used the least-squares method to solve the linear equations here. The cutoff in this case is the machine epsilon.

2. Solving scheme for M^n

For optimization of the auxiliary variable M^n , its augmented Lagrangian can be written by

$$\mathcal{L}(M^n) = \sum_{n=1}^N \xi \left\| M_{[\alpha(n)]}^n - S_{[\alpha(n)]}^n + \frac{1}{\xi} Y_{[\alpha(n)]}^n \right\|_F^2 + \mu \|M_{[\alpha(n)]}^n\|_*. \quad (11)$$

The optimization problem for the augmented Lagrangian in Eq. (11) can be solved by using the singular value shrinkage operator (SVSO) [28] as shown below. For the sake of simplicity, we first consider the following problem under nuclear norm regularization:

$$\min_X \tau \|X\|_* + \frac{1}{2} \|X - Y\|_F^2,$$

where $X, Y \in \mathbb{R}^2$ are unknown and given matrices, respectively. The SVSO $\mathcal{D}_\tau(Y)$ finds the solution X of the above equations as

$$\mathcal{D}_\tau(Y) := U \lfloor (s - \tau) I \rfloor V,$$

where $\lfloor X_{ij} \rfloor \equiv \max(X_{ij}, 0)$ is the element-wise floor function of the matrix X . U, s, V are obtained from the SVD of Y .

Thus, the updating equations for Eq. (11) is given by

$$M^n = \text{tensorize}[\mathcal{D}_{\mu/\xi}(S_{[\alpha(n)]}^n)] \quad (12)$$

for $n = 1, 2, \dots, N$.

3. Updating scheme of Y^n

Lastly, the Lagrangian multiplier Y^n is updated as

$$Y^n \leftarrow Y^n + \xi(M^n - S^n). \quad (13)$$

The aforementioned three steps are applied sequentially to each of the n th site, after which the optimization moves to the next site. Once this sweep has completed a full cycle, it constitutes a single iteration. Moreover, the penalty parameter ξ is updated as $\xi \leftarrow \rho \xi$ at every iterations by $0 < \rho < 1$, such that the regularization term in Eq. (7) is gradually relaxed to the original Eq. (1). We find that by setting $\mu := \xi^2$ for the next iteration it works appropriately for all examples in Sec. III.

In summary, the NNR loop optimization has several parameters for users to tune. The larger ξ , the more our objective functions favor low-rank solutions and ρ controls the penalty schedule. For the examples shown below, we chose the initial value of the parameter ξ , as denoted as ξ^{init} , and the

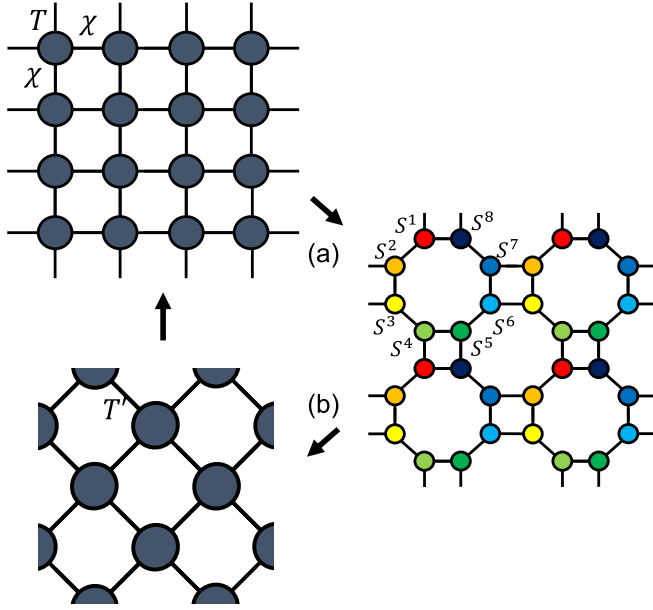


FIG. 4. Single iteration of NNR-TNR for a square lattice. (a) Target tensors T are first decomposed into 3-index tensors using LN-TRG, then NNR loop optimization is applied to the eight tensors on loop $S^1 \cdot S^2 \cdot S^3 \cdot S^4 \cdot S^5 \cdot S^6 \cdot S^7 \cdot S^8$, as introduced in the previous section. (b) The updated four tensors are contracted into single tensor in the same fashion as LN-TRG. New target tensors are recovered from rotating the square lattice by 45° .

parameter ρ within ranges $10^{-7} < \xi^{\text{init}} < 10^{-4}$ and $0.8 < \rho < 1$, respectively.

III. BENCHMARK RESULTS OF NNR LOOP OPTIMIZATION

In this section, we show an example application of NNR loop optimization for TRG. We call the combination of NNR loop optimization and TRG, *NNR-TNR*. One coarse-graining iteration for the square lattice is shown in Fig. 4. At the first step [Fig. 4(a)], we compute the initial value of the variational tensor S from LN-TRG and perform NNR loop optimization as described in the previous section. Note that such a choice for the initial value of S is quite useful for stable convergence, though LN-TRG does not remove redundant correlation. The coarse-graining steps in Fig. 4(b) are the same as LN-TRG, as sets of four updated tensors S are contracted to yield new target tensor T' . As will be discussed later in this section, NNR loop optimization filters out the redundant correlation during optimization, and thus it not only optimizes the approximation given by Eq. (1) but also has an effect similar to the entanglement filtering of Loop-TNR. In this sense, NNR-TNR achieves high-accuracy results with less human efforts, as its minimal programming implementation takes much fewer lines of codes than conventional Loop-TNR (see Appendix A). The overall computational complexity scales as $O(\chi^6)$ if one follows the contraction order given in Ref. [20].

We benchmark NNR-TNR by several physical observables for the critical two-dimensional (2D) Ising model on the square lattice. Note that our formulism does not impose any symmetries, such as Z_2 symmetry, rotational, or reflectional

symmetries of eight octagon tensors for all the results below. Therefore, we compare the results of NNR-TNR with those of Loop-TNR without imposing the rotational symmetries of lattices [32].

In tuning the parameters of NNR-TNR methods for analyzing critical systems, we observe the following trends. When the penalty parameter is set too low, contributions from short-range correlations persist, leading to instability in the RG flow and preventing improvements in accuracy even with increasing bond dimension. Conversely, if the penalty parameter is set too high, not only will the short-range correlations but also the relevant information will be truncated. Consequently, the RG flow rapidly converges to a trivial fixed point. For effective parameter tuning in NNR-TNR, it is recommended to adjust the penalty parameters to maintain the scale-invariant fixed-point tensor up to the largest possible system sizes. The end of RG flow can be checked from the singular value spectrum, as described in the following paragraph. Note that the parameters should ideally be adjusted according to the bond dimension.

In Fig. 5, we present singular value spectrums of coarse-grained tensors, which supposedly monitors the RG flow at the unstable critical fixed-point. For Loop-TNR, it obtains approximately identical singular value spectrums as a function of the RG steps starting from the third iteration, roughly up to the 40th. In contrast, NNR-TNR maintains the almost identical unstable critical fixed-point tensors more than the 40th RG steps. In addition, the RG flow ends in the trivial fixed-point tensor for both results, which implies that there is no footprint of short-range correlation for NNR-TNR, as well as Loop-TNR. Thus, it seems that NNR-TNR successfully truncates the short-range correlation during loop optimization. The bottom panels of Fig. 5 show the approximations to scale invariance measured by $\delta = \|s^{(t)} - s^{(t-1)}\|$, where s^t is the singular values of the normalized tensor $T^{(t)} = U s^{(t)} V^\dagger$ at the t th RG step. It seems that NNR-TNR improves Loop-TNR in terms of fixing the scale-invariant tensor at criticality.

In Fig. 6, we show the results of central charge and the smallest 40 scaling dimensions Δ computed from diagonalizing transfer matrix [17] for Loop-TNR and NNR-TNR. In the Loop-TNR case, the higher levels of scaling dimensions spectrum become unstable as a function of the RG steps and hardly distinguishable, for example, $\Delta \geq 3$ for $\chi = 16$ and $\Delta \geq 4$ for $\chi = 32$, respectively. However, in the case of NNR-TNR, such instabilities of the aforementioned levels of scaling dimensions are not seen. NNR-TNR successfully extracts $\Delta = 3$, $\Delta = 3.125$ scaling dimensions at $\chi = 16$ and $\Delta = 4$ and 4.125 are recognizable at $\chi = 32$. Although there exist some spurious solutions such as the one around $\Delta \approx 3.3$ for $\chi = 16$ and $\Delta \approx 3.7$ for $\chi = 32$, we can eliminate these spurious solutions based on their stronger dependence on χ and L than the true solutions.

In terms of the accuracies of conformal field theory (CFT) data, we compare the absolute value of relative errors of central charge c and the scaling dimensions upto $\Delta = 3$ at different system size in Fig. 7. One can see that the relative errors of all the quantities approach the exact values and deviate afterwards, which is an expected behavior due to finite-size and finite bond dimension effects [33]. For the CFT data corresponding to the lower indices of the fixed-point tensor, e.g.,

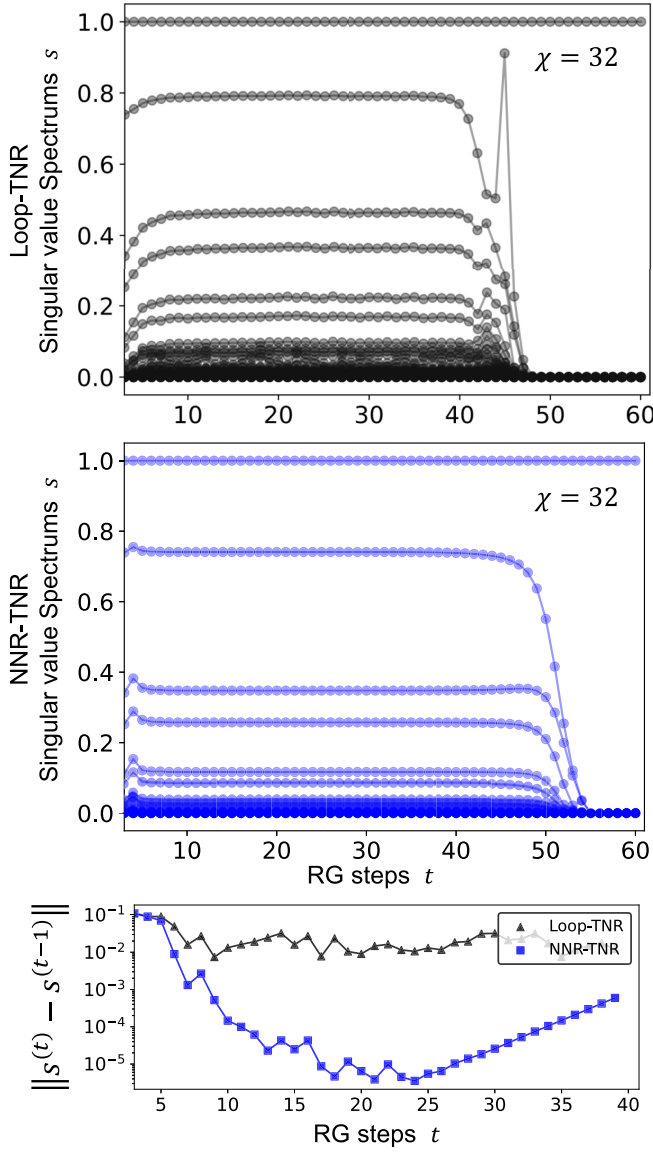


FIG. 5. (Top and middle panels) Singular value spectra $s^{(t)}$ of the coarse-grained tensors $T^{(t)}$, illustrated for Loop-TNR (top) and NNR-TNR (middle), at the exact critical temperature of 2D Ising model for different RG steps t , which is equivalent to linear system size 2^t . At each RG steps, the largest 100 singular values of tensors T are displayed. The largest number of $s^{(t)}$ is normalized to 1 for all t . (Bottom) The distance between singular value spectra $s^{(t)}$ and $s^{(t-1)}$ for both methods. This distance corresponds to the order of the scale-invariance at each RG steps. In this figure, the parameters of NNR-TNR are chosen as $\xi^{\text{init}} = 4 \times 10^{-6}$ and $\rho = 0.85$.

central charge and the scaling dimensions of primary fields, NNR-TNR produces a slightly more accurate and stable result against RG steps than Loop-TNR. In addition, we observe that NNR-TNR tends to significantly improve the accuracies of the higher levels of scaling dimensions compared to Loop-TNR, by an order of magnitude for $\Delta = 3$, for example. It seems to suggest that NNR-TNR can more effectively eliminate short-range correlations than Loop-TNR, thereby reaching a more genuine critical fixed-point tensor.

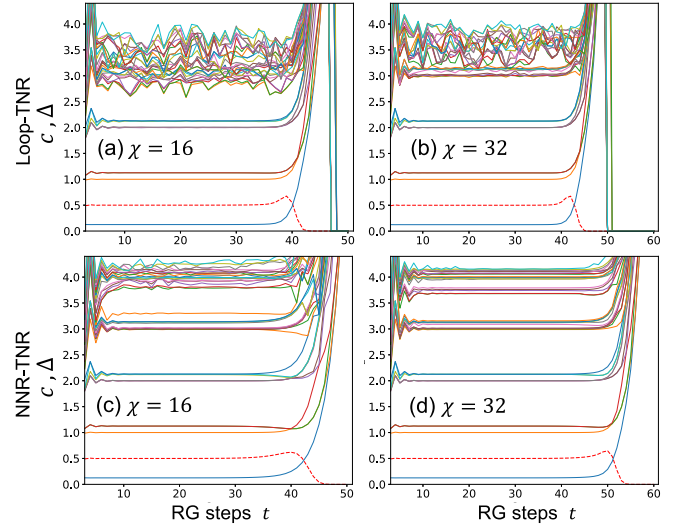


FIG. 6. Comparison of central charges (red dotted line) and scaling dimensions (solid lines) for Loop-TNR and NNR-TNR at different RG steps. Results are computed by Gu and Wens methods [17] using $L = 2$ transfer matrix in Ref. [20]. We chose the parameters of NNR-TNR for (c) and (d) are chosen as $\xi^{\text{init}} = 1 \times 10^{-5}$, $\rho = 0.8$ and $\xi^{\text{init}} = 4 \times 10^{-6}$, $\rho = 0.85$, respectively.

In all cases above, 50 or less iterations were required for NNR loop optimization at each RG step.

IV. CONCLUSION AND DISCUSSIONS

In this paper, we proposed alternative loop optimization for circumventing the local minima related to short-range correlation that lives within tensor space. We pointed out adding nuclear norm regularization to loop optimization resolved the

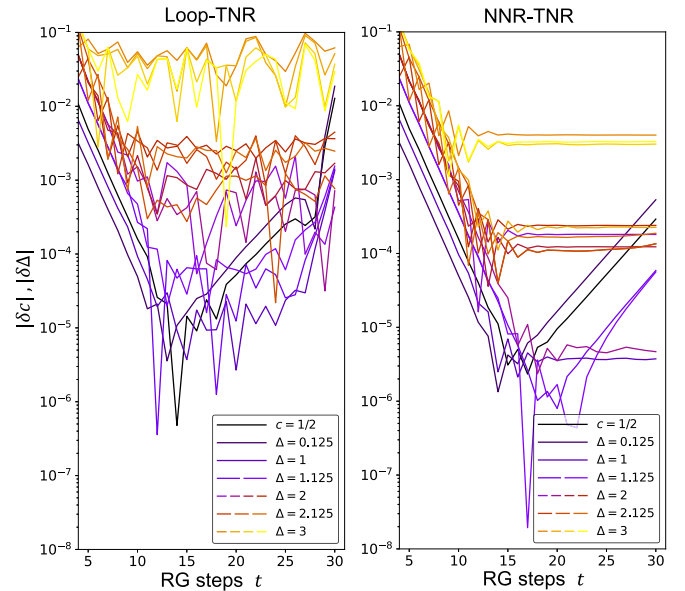


FIG. 7. The RG steps dependence of the absolute value of relative errors of central charge c and scaling dimensions Δ for Loop-TNR and NNR-TNR with bond dimension $\chi = 32$. In this figure, ξ^{init} and ρ were chosen to be 4×10^{-6} and 0.85, respectively.

short-ranged correlation problems and demonstrated that it further improved the RG flow of critical 2D Ising model and higher parts of scaling dimensions became clearly visible by using NNR-TNR.

The major advantage of NNR-TNR over other loop optimization techniques lies in its efficiency in eliminating short-range correlations, allowing us to stably obtain more accurate critical fixed-point tensors through a simple numerical procedure. For example, an iteration of the Loop-TNR consists of several steps, such as entanglement filtering and periodic MPS optimization, to generate the correct fixed-point tensor. In contrast, the present paper shows that the NNR loop optimization alone is sufficient to remove the short-range correlations and further improves the scale invariance of the critical fixed-point tensor. For pursuits of evaluating higher and precise scaling dimensions, there remains several options, such as increasing bond dimensions, length of columns for constructing a transfer matrix, and using gauge symmetries of the model, e.g., Z_2 [34,35], rotational and reflectional symmetries. For those future developments, the NNR loop optimization presented above will serve as the good starting point for even more accurate tensor network algorithms.

From a technical point of view, it should be noted that hyperparameter tuning of NNR loop optimization remain an issue; we chose the parameters ξ and ρ of NNR-TNR, such that the critical fixed-point tensor exhibits scale invariance up to the largest possible system size. However, these parameters are often situation dependent. How to automatically tune these parameters for general cases remain for future research.

An even more challenging and interesting subject for future works is to generalize the NNR loop optimization for higher dimensions. We believe this can be achieved by incorporating the idea of local bond matrix truncation, as introduced in Refs. [13,22], into NNR loop optimization. It would serve as an alternative method of removing redundant correlation efficiently in the 2D quantum or three-dimensional (3D) classical tensor network algorithms such as PEPS [36] and TRG [15,37,38].

ACKNOWLEDGMENTS

We thank Qibin Zhao, Satoshi Morita, Kenji Harada, Xinliang Lyu, Atsushi Ueda, Hyun-Yong Lee, and Katsuya Akamatsu for fruitful discussions. T.O. acknowledges the support from the Endowed Project for Quantum Software Research and Education, The University of Tokyo [39]. This work was supported by JST SPRING, Grant No. JPMJSP2108, JST Grant No. JPMJPF2221, JSPS KAKENHI Grants No. 19H01809 and No. 23H01092, and partially by the joint project of Kyoto University and Toyota Motor Corporation, titled “Advanced Mathematical Science for Mobility Society.” The computation in this work was done using the facilities of the Supercomputer Center, the Institute for Solid State Physics, the University of Tokyo.

APPENDIX A: SOURCE CODE

Our sample code for NNR-TNR is available at Ref. [40]. It can be used to reproduce all the results shown in Sec. III.

APPENDIX B: NUCLEAR NORM REGULARIZED LOOP OPTIMIZATION FOR TENSOR NETWORK

1. Loop optimization based on entropy regularization

In the main text, we choose to penalize the nuclear norm of $S_{[\alpha(n)]}$ to avoid the local minima. As we briefly discussed in the main text, there may be several other choices for the penalty term. Here we consider a term that has the same form as the Shannon entropy. Similar to the nuclear norm, it is a function of the singular values of the matricized tensors. Instead of their simple sum, we consider $-\sum_i \sigma_i \log \sigma_i$ with the normalization $\sum_i \sigma_i = 1$. This can be interpreted as the local mutual information between the extracted leg and the others (after neglecting the environment). We call the resulting regularization the entropy regularization (ER) hereafter. In ER, the cost function Eq. (5) is replaced by

$$\min_S \frac{1}{2} \|\psi(T) - \psi(S)\|_F^2 + \lambda \sum_{n=1}^N E_{[\alpha(n)]}^n, \quad (\text{B1})$$

where $E_{[\alpha(n)]}^n$ is the local mutual information mentioned above. In this setup, the hyperparameter λ is set to be $\lambda = \tilde{\lambda} \|\psi(T) - \psi(S)\|_F^2$ and $\tilde{\lambda}$ is a positive parameter for every RG step.

Intuitively, penalizing the entropy over the diagonal legs in Fig. 1 should also induce the low-rankness and thereby lead to a reduction in redundant correlation, as does the NNR technique. The remaining question is which of the two approaches is numerically preferable for the regularization technique in the context of TNR.

The numerical setup for this demonstration, such as the initial condition of S and normalization of T , is equivalent to the recipes discussed in main text. However, the updating scheme of S and $\tilde{\lambda}$ is slightly different from the main text. We elaborate the updating scheme at each RG step as follows: we perform ten sweeps of the nonlinear conjugate gradient method to optimize Eq. (B1) with $\tilde{\lambda} = 0.01$. Then, we set $\tilde{\lambda} = 0$ and perform 50 sweeps of optimization using the periodic MPS method. Hereafter, we call the TNR based on the ER method as ER-TNR.

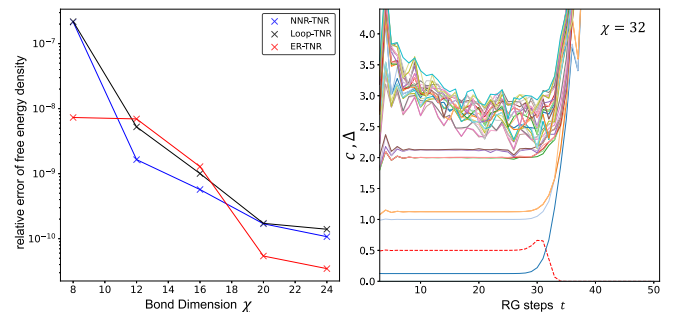


FIG. 8. (Left) Relative error of free energy density of critical 2D classical Ising model as a function of bond dimension using NNR-TNR, Loop-TNR, and ER-TNR with $2^{30} \times 2^{30}$ spins. The blue solid line corresponds to NNR-TNR result with the parameters $\xi^{\text{init}} = 1 \times 10^{-5}$ and $\rho = 0.84$, chosen to be fixed over the full range of bond dimensions for simplicity. (Right) The conformal data extracted by ER-TNR at different RG steps with bond dimension $\chi = 32$.

Figure 8 (left) compares the relative error of the free energy density of the critical 2D classical Ising model with different bond dimensions among ER-TNR, NNR-TNR, and Loop-TNR. It seems that ER-TNR gives comparable results with the other techniques to NNR-TNR or Loop-TNR for $\chi < 20$, and is even more accurate for the higher bond dimension $\chi = 20, 24$. However, the comparison between Figs. 8 (right) and 6(b) and 6(d) shows that the stability of central charge and scaling dimension spectrum as a function of RG step seems to deteriorate around the 30th RG step, and the higher levels of

scaling dimension spectrum could not be obtained for $\Delta > 2$ in Fig. 8 (right). In addition, we also find that the stability of the singular value spectrum and the relative error of the scaling dimension spectrum are rather inferior to the results shown in Figs. 5 and 7.

From the perspectives of the numerical RG algorithm, the NNR seems to be more suitable than the ER, at least for loop optimization in TNR, although the proper norm may depend on the given entanglement structure, e.g., the initial condition.

-
- [1] S. R. White, Density matrix formulation for quantum renormalization groups, *Phys. Rev. Lett.* **69**, 2863 (1992).
 - [2] S. R. White, Density-matrix algorithms for quantum renormalization groups, *Phys. Rev. B* **48**, 10345 (1993).
 - [3] T. Nishino and K. Okunishi, Corner transfer matrix renormalization group method, *J. Phys. Soc. Jpn.* **65**, 891 (1996).
 - [4] T. Nishino and K. Okunishi, Corner transfer matrix algorithm for classical renormalization group, *J. Phys. Soc. Jpn.* **66**, 3040 (1997).
 - [5] G. Vidal, Entanglement renormalization, *Phys. Rev. Lett.* **99**, 220405 (2007).
 - [6] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, Algorithms for finite projected entangled pair states, *Phys. Rev. B* **90**, 064425 (2014).
 - [7] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Classical simulation of infinite-size quantum lattice systems in two spatial dimensions, *Phys. Rev. Lett.* **101**, 250602 (2008).
 - [8] R. Orús and G. Vidal, Simulation of two-dimensional quantum systems on an infinite lattice revisited: Corner transfer matrix for tensor contraction, *Phys. Rev. B* **80**, 094403 (2009).
 - [9] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902 (2003).
 - [10] S. Östlund and S. Rommer, Thermodynamic limit of density matrix renormalization, *Phys. Rev. Lett.* **75**, 3537 (1995).
 - [11] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, Variational optimization algorithms for uniform matrix product states, *Phys. Rev. B* **97**, 045145 (2018).
 - [12] R. Orús and G. Vidal, Infinite time-evolving block decimation algorithm beyond unitary evolution, *Phys. Rev. B* **78**, 155117 (2008).
 - [13] G. Evenbly, Gauge fixing, canonical forms, and optimal truncations in tensor networks with closed loops, *Phys. Rev. B* **98**, 085155 (2018).
 - [14] M. Levin and C. P. Nave, Tensor renormalization group approach to two-dimensional classical lattice models, *Phys. Rev. Lett.* **99**, 120601 (2007).
 - [15] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, Coarse-graining renormalization by higher-order singular value decomposition, *Phys. Rev. B* **86**, 045139 (2012).
 - [16] Z. Y. Xie, H. C. Jiang, Q. N. Chen, Z. Y. Weng, and T. Xiang, Second renormalization of tensor-network states, *Phys. Rev. Lett.* **103**, 160601 (2009).
 - [17] Z.-C. Gu and X.-G. Wen, Tensor-entanglement-filtering renormalization approach and symmetry-protected topological order, *Phys. Rev. B* **80**, 155131 (2009).
 - [18] G. Evenbly and G. Vidal, Tensor network renormalization, *Phys. Rev. Lett.* **115**, 180405 (2015).
 - [19] G. Evenbly, Algorithms for tensor network renormalization, *Phys. Rev. B* **95**, 045117 (2017).
 - [20] S. Yang, Z.-C. Gu, and X.-G. Wen, Loop optimization for tensor network renormalization, *Phys. Rev. Lett.* **118**, 110504 (2017).
 - [21] M. Bal, M. Mariën, J. Haegeman, and F. Verstraete, Renormalization group flows of Hamiltonians using tensor networks, *Phys. Rev. Lett.* **118**, 250602 (2017).
 - [22] M. Hauru, C. Delcamp, and S. Mizera, Renormalization of tensor networks using graph-independent local truncations, *Phys. Rev. B* **97**, 045111 (2018).
 - [23] P. Pippin, S. R. White, and H. G. Evertz, Efficient matrix-product state method for periodic boundary conditions, *Phys. Rev. B* **81**, 081103(R) (2010).
 - [24] F. Verstraete, V. Murg, and J. I. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, *Adv. Phys.* **57**, 143 (2008).
 - [25] Q. Zhao, G. Zhou, S. Xie, L. Zhang, and A. Cichocki, Tensor ring decomposition, [arXiv:1606.05535](https://arxiv.org/abs/1606.05535).
 - [26] L. Yuan, C. Li, D. Mandic, J. Cao, and Q. Zhao, Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion, [arXiv:1809.02288](https://arxiv.org/abs/1809.02288).
 - [27] E. J. Candes and B. Recht, Exact matrix completion via convex optimization, [arXiv:0805.4471](https://arxiv.org/abs/0805.4471).
 - [28] J.-F. Cai, E. J. Candes, and Z. Shen, A singular value thresholding algorithm for matrix completion, [arXiv:0810.3286](https://arxiv.org/abs/0810.3286).
 - [29] B. Recht, M. Fazel, and P. A. Parrilo, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization, *SIAM Rev.* **52**, 471 (2010).
 - [30] M. Jaggi, M. Sulovsk *et al.*, A simple algorithm for nuclear norm regularized problems, in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (ACM, New York, 2010), pp. 471–478.
 - [31] K. Homma, T. Okubo, and N. Kawashima, See Appendix B for nuclear norm regularized loop optimization for tensor network.
 - [32] For our implementation of Loop-TNR, we performed variational periodic MPS method and entanglement filtering, as proposed in the original literature. In addition, we chose the initial value of variational periodic MPS from the method proposed in Ref. [41] to accelerate the convergence. Our

- implementation is supposed to faithfully reproduce the results in Ref. [20].
- [33] A. Ueda and M. Oshikawa, Finite-size and finite bond dimension effects of tensor network renormalization, *Phys. Rev. B* **108**, 024413 (2023).
- [34] S. Singh, R. N. C. Pfeifer, and G. Vidal, Tensor network states and algorithms in the presence of a global U(1) symmetry, *Phys. Rev. B* **83**, 115125 (2011).
- [35] S. Singh, R. N. C. Pfeifer, and G. Vidal, Tensor network decompositions in the presence of a global symmetry, *Phys. Rev. A* **82**, 050301(R) (2010).
- [36] Y. Zheng and S. Yang, Loop update for infinite projected entangled-pair states in two spatial dimensions, *Phys. Rev. B* **102**, 075147 (2020).
- [37] D. Adachi, T. Okubo, and S. Todo, Anisotropic tensor renormalization group, *Phys. Rev. B* **102**, 054432 (2020).
- [38] D. Kadoh and K. Nakayama, Renormalization group on a triad network, *arXiv:1912.02414*.
- [39] <https://qsw.phys.s.u-tokyo.ac.jp/>.
- [40] github.com/KenjiHomma/NNR-TNR.
- [41] L. Wang and F. Verstraete, Cluster update for tensor network states, *arXiv:1110.4362*.