

Scaling dimensions from linearized tensor renormalization group transformations

Xinliang Lyu,^{1,*} RuQing G. Xu,² and Naoki Kawashima^{1,†}

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*

(Dated: February 15, 2021)

We show a way to analyze a renormalization group (RG) fixed point in tensor space: write down the tensor RG equation, linearize it around a fixed-point tensor, and diagonalize the resulting linearized RG equation to obtain scaling dimensions. The tensor RG methods have had a great success in producing accurate free energy compared with the conventional real-space RG schemes. However, the above-mentioned canonical procedure for the fixed-point analysis has not been implemented for general tensor-network-based RG schemes. We extend the success of the tensor methods further to extraction of scaling dimensions through analyzing a fixed-point tensor. This approach is benchmarked in the context of the Ising models in 1D and 2D. The proposed method accomplishes the canonical RG prescription for the tensor RG methods.

I. INTRODUCTION

The renormalization group (RG) is a powerful technique for studying physical systems where fluctuations in all scales of length are important [1]; the most famous example in statistical mechanics is critical phenomena. The main idea behind the RG is to study how a physical system changes as we go from one length scale to another. Conventional RG schemes, such as ϵ -expansion [2] and block-spin methods [3–7], aim at a map from the Hamiltonian of the short length scale to that of the longer one, such that the partition function is unchanged [8]. The map is known as an RG equation. A well-behaved RG equation exhibits fixed points, each corresponding to a conformal field theory (CFT) [9, 10]. A critical system is described by a fixed point. By linearizing the RG equation around the critical fixed point, universal properties like scaling dimensions of the critical system can be extracted. This canonical RG prescription of analyzing a fixed point also provides a theoretical framework to understand universality in critical phenomena. However, for a systematic study with high-precision, the Hamiltonian may not be the most efficient representation of the system.

Recently, ideas from quantum information have stimulated a novel type of RG methods in tensor space. They are versatile numerical RG schemes whose approximations are controlled by an integer, χ , called the *bond dimension*. The RG equation is a map from a tensor encapsulating the Boltzmann weights of local configurations at a short length scale to a new tensor at a longer one. The first realization of this new paradigm is the tensor renormalization group (TRG) [11], followed by many variations [12–17]. These TRG-type techniques have excellent performance in calculations of free energy. For example, the higher-order tensor renormalization group (HOTRG) [14] estimates the free energy of the 2D Ising model with error of order 10^{-7} within a few minutes in a desktop computer. The

estimation error decreases exponentially as χ increases, while the computational cost only grows polynomially.

With all of their success in calculations of free energy, however, the TRG-type techniques encounter obstacles in the fixed-point analysis. Early attempts [18–21] show that if the bond dimension χ of the TRG is larger than 8, the tensor will *never* flow to the critical fixed point of the 2D Ising model; this imposes a very strong restriction on the bond dimension in the fixed-point analysis. For $\chi = 2, 3, 4$, either using the TRG or the HOTRG, the estimated scaling dimension of the energy density operator is similar to the old potential moving tricks [18–20], and that of the spin operator is more than a factor of 2 larger than the exact value [21].

Fortunately, in recent ten years, people have developed many tricks to solve the problem of the unsatisfactory tensor RG flows. In 2009, Gu and Wen [22] was the first to deal with this problem. They followed Levin’s suggestion [11, 23] and focused on a toy model called corner double-line (CDL) tensors, which represent systems with only local correlations. They showed that the CDL tensors are fixed points of the TRG, indicating that the local correlations at the smaller length scales will be carried to the larger ones. A crude algorithm was proposed to filter out the CDL tensors and the problem of the tensor RG flows was partially solved, followed up by an improved algorithm in 2017 [24]. From 2015 to 2017, several similar methods were proposed [25–27]. All of these advanced TRG-type techniques successfully produced critical fixed-point tensors.

With a critical fixed-point tensor in hand, Gu and Wen [22] pointed out that the scaling dimensions can be extracted by diagonalizing a transfer matrix constructed from the fixed-point tensor according to a well-known 2D CFT theorem [28]. Later, Evenbly and Vidal used the tensor network renormalization (TNR) [25, 26] to implement local scale transformation that maps a plane to a cylinder [29]; the spectrum of eigenvalues of a transfer matrix on the cylinder gives scaling dimensions. These methods have been applied to calculate scaling dimensions from then on, while the fixed-point analysis in tensor space

* lyu@issp.u-tokyo.ac.jp

† kawashima@issp.u-tokyo.ac.jp

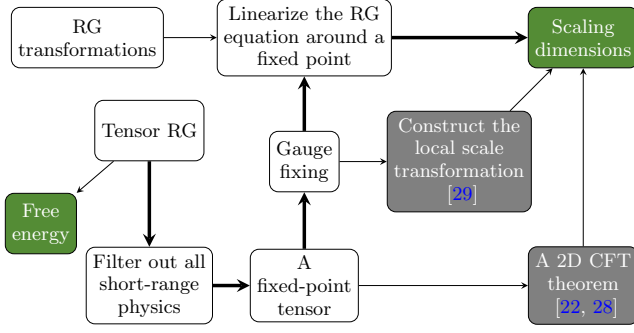


Figure 1. Different ways to extract scaling dimensions using tensor RG methods. The proposed method in this paper corresponds to the path indicated by the thick arrows.

has never followed up.

In this paper, we provide the missing piece of analyzing a fixed point in tensor space at a general bond dimension (see Fig. 1). After laying down the general framework for the canonical RG prescription for the tensor RG methods in Sec. II A, we point out two technical obstacles, local correlations and gauge redundancy, in Sec. II B. The HOTRG is combined with a recently-developed technique, graph-independent local truncation (GILT) [30] in Sec. II C, to generate correct tensor RG flows that will go to a critical fixed point at a general bond dimension. We call this method GILT-HOTRG. In Sec. II D, we show that most gauge redundancy in the tensor description is automatically fixed during GILT-HOTRG, leaving only tractable sign ambiguities. The linearized RG equation for the GILT-HOTRG is easy to implement and has a simple pictorial representation; in practice, it can be generated by automatic differentiation once the GILT-HOTRG is implemented. The scaling dimensions can be extracted from this linearized RG equation. In Sec. III, the canonical RG prescription in tensor space is benchmarked with 1D and 2D classical Ising models. We conclude in Sec. IV.

II. RENORMALIZATION GROUP IN TENSOR NETWORK LANGUAGE

TRG-type methods start with the fact that partition functions of all classical statistical models can be rewritten as tensor network models [11]. Take the square lattice 2D Ising model as a concrete example. The partition function is

$$Z = \sum_{\{\sigma(\mathbf{r})\}} e^{K \sum_{\langle i,j \rangle} \sigma_i \sigma_j}, \quad (1)$$

where σ_i is the shorthand for the spin variable $\sigma(\mathbf{r}_i)$ located at lattice point \mathbf{r}_i and can take values ± 1 , and $K = J/k_B T$. In this paper, we measure temperature in units of J/k_B so it becomes a dimensionless number. The partition function in Eq. (1) can be rewritten as a tensor

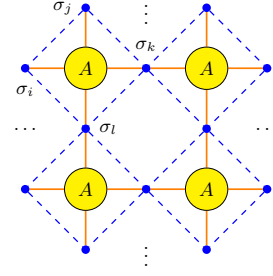


Figure 2. Representation of the local Boltzmann weight in terms of tensors. The dots are where the spin variables locate. They form a square lattice slanted by 45° . The larger circles are tensors A encoding the Boltzmann weight of the configurations of the four surrounding spin variables. The square lattice formed by N copies of A is the tensor network representation of the partition function in Eq. (3).

network by defining a tensor

$$A_{\sigma_i \sigma_j \sigma_k \sigma_l} \equiv e^{K(\sigma_i \sigma_j + \sigma_j \sigma_k + \sigma_k \sigma_l + \sigma_l \sigma_i)} = \text{diagram of tensor A with four legs labeled } \sigma_i, \sigma_j, \sigma_k, \sigma_l. \quad (2)$$

Each index of this tensor can take two values ± 1 and we say the bond dimension of a leg of this tensor is $\chi = 2$. It is now possible to rewrite the partition function of the 2D Ising model in Eq. (1) as the tensor product of N copies of A , with all their indices summed over (Fig. 2)

$$Z = \sum_{\{\sigma(\mathbf{r})\}} \bigotimes_{x=1}^N A_{\sigma_{i(x)} \sigma_{j(x)} \sigma_{k(x)} \sigma_{l(x)}} \equiv \text{tTr} \left(\bigotimes_{x=1}^N A \right). \quad (3)$$

The coarse graining of the tensor network resembles the conventional block-spin methods. We replace a patch of, say, four copies of the original tensor A with one coarse-grained tensor A_c , such that the partition function is approximately described by a coarser tensor network made of $N/4$ copies of A_c

$$Z \approx \text{tTr} \left(\bigotimes_{x=1}^{N/4} A_c \right). \quad (4)$$

The specific procedure for obtaining A_c from A will be discussed later. The map

$$A \xrightarrow{\text{RG}} A_c \text{ or } A_c = \mathcal{T}(A) \quad (5)$$

is the tensor RG equation.

A. General framework

We first define the *canonical* RG prescription in tensor space. To this end, it is helpful to start with a review

of the old approach in Hamiltonian space (we follow the detailed review [21] and textbook [31] closely).

It will be convenient to explain in terms of a specific physical system: classical system spin variables $\sigma \in \{+1, -1\}$ on a lattice, with general short-ranged interactions. The Hamiltonian (or energy) of the system can be parameterized by a set of coupling constants $\mathbf{K} = \{K_j\}$, each of which couples to a possible short-ranged interaction term $s_j(\mathbf{r})$,

$$\mathcal{H} = \sum_{\mathbf{r}} \sum_j K_j s_j(\mathbf{r}). \quad (6)$$

For example, if K_1 is the magnetic field, $s_1(\mathbf{r}) = \sigma(\mathbf{r})$ is the spin variable at lattice point \mathbf{r} ; if K_2 is the nearest neighbor interaction along x direction, $s_2(\mathbf{r}) = \sigma(\mathbf{r})\sigma(\mathbf{r} + a\hat{\mathbf{e}}_x)$, where $\hat{\mathbf{e}}_x$ is the unit vector along x direction and a is the lattice constant. A conventional RG transformation maps the old Hamiltonian \mathcal{H} to a new one \mathcal{H}' with the same form as Eq. (6) but characterized by a set of new coupling constants $\mathbf{K}' = \{K'_j\}$. The map from the old Hamiltonian to the new one $\mathcal{H} \xrightarrow{\text{RG}} \mathcal{H}'$ is then parametrized explicitly as the transformation from the old coupling constants to the new ones,

$$\mathbf{K}' = \mathcal{T}^{\text{old}}(\mathbf{K}). \quad (7)$$

We require that the RG transformation should preserve the partition function of the system and should exhibit a fixed-point Hamiltonian \mathcal{H}^* parameterized by coupling constants \mathbf{K}^* , such that \mathbf{K}^* remains unchanged under the RG transformation,

$$\mathbf{K}^* = \mathcal{T}^{\text{old}}(\mathbf{K}^*). \quad (8)$$

The linearized RG equation around \mathbf{K}^* is defined in the following way. We perturb the coupling constants around the fixed point $\mathbf{K}_p = \mathbf{K}^* + \delta\mathbf{K}$ and perform the RG transformation defined in Eq. (7), $\mathbf{K}'_p = \mathcal{T}^{\text{old}}(\mathbf{K}_p)$. The new coupling constants \mathbf{K}'_p after the RG transformation should be close to \mathbf{K}^* by continuity, so $\mathbf{K}'_p = \mathbf{K}^* + \delta\mathbf{K}'$. The linearized RG equation around \mathbf{K}^* is a matrix \mathcal{R}^{old} telling us how $\delta\mathbf{K}'$ is related to $\delta\mathbf{K}$,

$$\delta K'_i = \sum_j \mathcal{R}_{ij}^{\text{old}} \delta K_j. \quad (9)$$

The matrix \mathcal{R}^{old} has right and left eigenvectors $\{\psi^\alpha\}, \{\phi^\alpha\}$ with the same set of eigenvalues $\{\lambda^\alpha\}$,

$$\sum_j \mathcal{R}_{ij}^{\text{old}} \psi_j^\alpha = \lambda^\alpha \psi_i^\alpha \text{ and } \sum_i \phi_i^\alpha \mathcal{R}_{ij}^{\text{old}} = \lambda^\alpha \phi_j^\alpha. \quad (10)$$

The linear combinations of δK_i according to the components of the left eigenvector ϕ^α are known as scaling fields

$$h^\alpha = \sum_i \phi_i^\alpha \delta K_i, \quad (11)$$

while the linear combinations of interaction terms $s_j(\mathbf{r})$ according to the components of the right eigenvector ψ^α are known as scaling operators

$$o^\alpha(\mathbf{r}) = \sum_j s_j(\mathbf{r}) \psi_j^\alpha. \quad (12)$$

Under the RG transformation with rescaling factor b for a system in dimension d , the scaling fields and the scaling operators transform in a simpler way with

$$(h^\alpha)' = b^{d-x_\alpha} h^\alpha \text{ and } (o^\alpha)' = b^{x_\alpha} o^\alpha, \quad (13)$$

where x_α are the scaling dimensions of the scaling operators $o^\alpha(\mathbf{r})$. Equations (9) to (11) and (13) give the relation between the scaling dimensions $\{x_\alpha\}$ and the eigenvalues $\{\lambda^\alpha\}$ of the linearized RG equation,

$$b^{d-x_\alpha} = \lambda^\alpha. \quad (14)$$

Next, we move on to the tensor approach of the canonical RG prescription. In the tensor RG approach, we skip the Hamiltonian description of the system. Instead, we use a tensor network made of copies of tensor A to represent the partition function Z of the system. The tensor RG equation is a map from the tensor A to the coarser tensor A_c , as is shown in Eq. (5). We claim that the components of the tensor A can be thought of as some proxies of the coupling constants \mathbf{K} (this claim was hinted in Ref. [22]).

To see why this claim is reasonable, note that we can map the partition function of the system with Hamiltonian in Eq. (6) to a tensor network using the method introduced in Ref. [11]. Each component of the initial tensor A is the Boltzmann weight of a given local configuration and depends on the coupling constants,

$$A_{(i)} = f_{(i)}(\mathbf{K}), \quad (15)$$

where we group all legs of A to form a single index, $A_{(i)} \equiv A_{i_1 i_2 i_3 i_4}$. After coarse graining, the components of A_c are still functions of \mathbf{K} but with different functional forms,

$$(A_c)_{(i)} = (f_c)_{(i)}(\mathbf{K}). \quad (16)$$

Now, we require that each component of the coarser tensor A_c should have the same functional form as that of A , but with different coupling constants \mathbf{K}' ,

$$f_{(i)}(\mathbf{K}') = (f_c)_{(i)}(\mathbf{K}), \forall (i). \quad (17)$$

In the old Hamiltonian approach, we need to solve Eq. (17) for \mathbf{K}' in terms of \mathbf{K} , which defines the RG equation from the old \mathbf{K} to the new \mathbf{K}' . However, in the tensor approach, it is enough to know the existence of such \mathbf{K}' . Combine Eq. (16) and Eq. (17), we have

$$(A_c)_{(i)} = f_{(i)}(\mathbf{K}'). \quad (18)$$

At the fixed point, $\mathbf{K} = \mathbf{K}' = \mathbf{K}^*$, equations (15) and (18) give

$$A^* \xrightarrow{\text{RG}} A^* \text{ or } A^* = \mathcal{T}(A^*). \quad (19)$$

Take the total derivative of tensors A and A_c in Eqs. (15) and (18) and set $\mathbf{K} = \mathbf{K}' = \mathbf{K}^*$,

$$\delta A_{(i)} = \sum_n (\partial_{(n)} f_{(i)}) \Big|_{\mathbf{K}=\mathbf{K}^*} \delta K_n, \quad (20)$$

$$(\delta A_c)_{(i)} = \sum_n (\partial_{(n)} f_{(i)}) \Big|_{\mathbf{K}'=\mathbf{K}^*} \delta K'_n. \quad (21)$$

Equations (20) and (21) give the transformation law between the coupling-constant description and tensor description of the canonical RG prescription, with $\partial_{(n)} f_{(i)}$ evaluated at \mathbf{K}^* being the change of basis matrix. Under this transformation, the linearized RG equation in Eq. (9) becomes

$$(\delta A_c)_{(i)} = \sum_j \mathcal{R}_{(i)(j)} \delta A_{(j)}, \quad (22)$$

which defines the linearized RG equation in tensor space. Since Eqs. (9) and (22) are the same linear transformation in two different representations, we can equally well diagonalize the matrix $\mathcal{R}_{(i)(j)}$ and find scaling dimensions according to Eq. (14).

In Sec. III A, we will use the 1D Ising model as a concrete example to demonstrate the general argument above.

B. Technical obstacles

There are two major obstacles for the fixed-point analysis in tensor space. They prevent us from obtaining a fixed-point tensor satisfying Eq. (19). The two obstacles are the problem of local correlations and the gauge redundancy in tensor network language.

Levin and Nave anticipated the first obstacles when looking for fixed points of the RG equation of the TRG [23]. One of the earliest numerical evidence for the peculiar tensor RG flows of the 2D Ising model was provided by Hinczewski and Berker [18]. Their results indicate that the TRG-type techniques have difficulty in integrating out all the local correlations at short distances, so physics at the lattice scale is carried all the way to the physics at larger ones. This shortcoming of the TRG-type techniques makes identification of both non-critical and critical fixed-point tensors very difficult.

To understand how the problem of local correlations at the lattice scale arises in the TRG-type techniques, let us examine the physical picture of the tensor RG transformation. We focus on a concrete example of a tensor network made of $4 \times 4 = 16$ copies of tensor A shown in Fig. 2 with periodic boundary condition. The general picture of a tensor RG transformation is similar to the conventional block-spin methods. For example, we

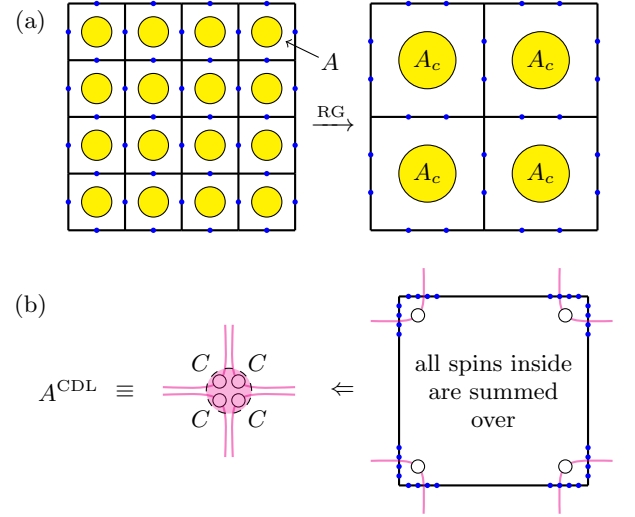


Figure 3. (a) The block-tensor transformation $A \rightarrow A_c$. The spins shared by two tensors A are summed over according to Eq. (24). The squares are larger after the decimation. (b) The origin of the CDL tensors. When the black square becomes large enough, the spins on one edge are far away from those on another, except for the spins around the four corners. The correlations among the corner spins give rise to the CDL tensors, containing physics at the lattice scale.

block a square of four tensors by contracting legs between them and group every two legs in the same side. Call the new tensor A_c ,

$$Z_{4 \times 4} = \text{[Diagram of 4x4 grid of A tensors]} \xrightarrow{\text{RG}} \text{[Diagram of 2x2 grid of A_c tensors]} , \quad (23)$$

where

$$\text{[Diagram of A_c tensor]} \equiv \text{[Diagram of A tensor with summed spins]} . \quad (24)$$

It is enlightening to put the original spin variables back into the tensor network to get a more physical picture of what is happening under such a block-tensor RG transformation. We refrain from drawing legs of A and the dashed lines of the spin lattice in Fig. 2, and surround copies of A with squares on whose sides the spin variables sit. The big picture for the block-tensor transformation in Eq. (23) is shown schematically in Fig. 3(a). The process is similar to the decimation in the conventional approaches. After the spin variables shared by every two A tensors forming the same A_c are summed over, we are left with four bigger squares, with two spin variables sitting on each side of each square. When the squares become large enough as the block-tensor transformation goes on, we expect that, roughly speaking, the spin variables on different edges

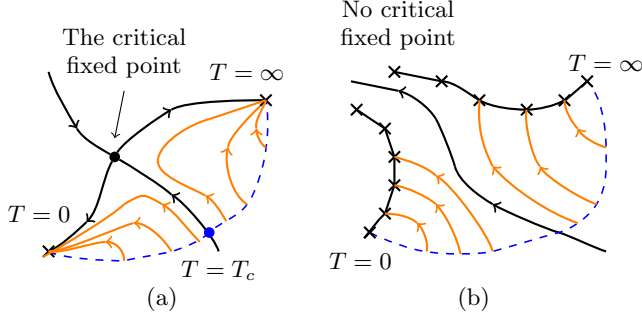


Figure 4. Schematic RG flows of the 2D Ising model. Each point on the dashed line represents the lattice model at a given temperature and is the starting point of an RG transformation. The solid lines with arrows represent different RG flows. (a) The correct RG flow. There are one $T=0$ fixed point, one $T=\infty$ fixed point and one critical fixed point. (b) The RG flows generated by the TRG and the HOTRG. The two trivial fixed points become two fixed lines, and the critical fixed point disappears.

are far away from each other and thus uncorrelated. The only exception is for the spin variables around the four corners. We can use a matrix C in Fig. 3(b) to capture the correlations around the corners; the matrix C must contain physics at the scale of the original lattice constant. Since the spin variables around different corners are far away from each other, the tensor A^{CDL} corresponding to this black square should factorize into the tensor product of four corner matrices C . A tensor with the structure of A^{CDL} is called a CDL tensor.

The CDL tensors are fixed points of the RG equations of the TRG [22, 23, 25, 30] and the HOTRG [32]. This shows that the TRG and the HOTRG have difficulty in integrating out the local interactions among the spin variables around the corners. If we start with two temperatures $T_1 \neq T_2$, both larger than the critical temperature T_c of the 2D Ising model, either of these two methods will generate tensors flowing to two different CDL tensors $A_1^{\text{CDL}} \neq A_2^{\text{CDL}}$, as a natural consequence of the fact that these CDL tensors depend, directly, on the bare interaction constants. At criticality, the previous numerical calculations indicate that we will never reach a critical fixed-point tensor [18, 25]. Their calculations suggest tensor RG flows shown in Fig. 4(b), where the low- and high-temperature fixed points turn into two fixed lines and the critical fixed point disappears. By comparison, the correct RG flow is shown in Fig. 4(a). We will introduce a way to solve the problem of CDL tensors for the HOTRG in Sec. II C.

The second obstacle that prevents us from achieving Eq. (19) is that the tensor network representation of the partition function in Eq. (3) has gauge redundancy. If two tensors \tilde{A} and A are related through the invertible

matrices S_x, S_y by the gauge transformation

$$\tilde{A}_{ijkl} = \sum_{m,n} A_{mnpq} (S_x)_{im} (S_y)_{jn} (S_x^{-1})_{pk} (S_y^{-1})_{ql}, \quad (25a)$$

or pictorially as

$$\tilde{A} = \begin{matrix} & S_y & \\ S_x & A & S_x^{-1} \\ & S_y^{-1} & \end{matrix}, \quad (25b)$$

The two tensor networks formed by A and \tilde{A} represent the same partition function Z . Equation (25) defines an equivalence relation where all the elements in an equivalence class represent a same partition function Z . The gauge redundancy makes the correspondence between a tensor A and the partition function Z no longer one-to-one. In the following, we refer to the equivalence class defined by the gauge transformation in Eq. (25) as $[A]$. In comparison, in the conventional Hamiltonian approach, we fix the form of the new Hamiltonian \mathcal{H}' to be the same as the old \mathcal{H} in Eq. (6), with \mathcal{H}' characterized by a set of new coupling constants \mathbf{K}' . This gives a one-to-one correspondence between a set of coupling constants $\mathbf{K} = \{K_j\}$ and a Hamiltonian \mathcal{H} (or a partition function Z).

The gauge redundancy makes the canonical RG prescription in tensor space less straightforward than that in Hamiltonian space. Even if we have reached a representation tensor A^* of the fixed-point equivalence class $[A^*]$, the tensor RG equation could bring this tensor to another representation of $[A^*]$,

$$A^* \xrightarrow{\text{RG}} \tilde{A}^* \text{ or } \tilde{A}^* = \mathcal{T}(A^*). \quad (26)$$

In general, we must fix the gauge of the tensor during a tensor RG transformation by choosing a preferred set of bases, so that the fixed-point tensor is manifestly fixed, as is shown in Eq. (19). We will show how the gauge is fixed for a HOTRG-like scheme in Sec. II D.

C. GILT-HOTRG

In this subsection, we present a HOTRG-like scheme to solve the first technical obstacle. Compared with the state-of-the-art TRG-type methods [22, 24–27, 30, 33–36] that can generate correct tensor RG flows, our scheme can be most easily generalized to dimensions higher than 2 and is convenient for the subsequent gauge fixing and linearization procedure. The graph-independent local truncation (GILT) [30] is performed to filter out the problematic local correlations before the coarse graining of the HOTRG. We call this scheme GILT-HOTRG.

The key feature of the GILT is that it is a stand-alone procedure to filter out the local correlations and does

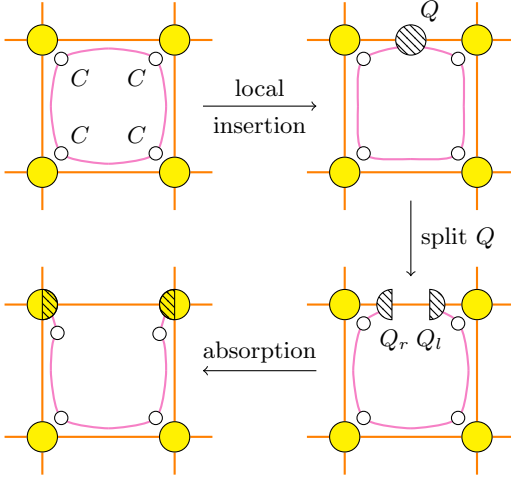


Figure 5. The process of the GILT. Four copies of C matrices are unknown inner structure of the adjacent 4-leg tensors. They are drawn explicitly to make the demonstration clearer. In the first step, a low-rank matrix Q is inserted into a bond. Then we split Q into two pieces using singular value decomposition. The low-rank matrix Q is constructed so that it cuts the legs of the corner matrices C during the splitting. Finally, the pieces of the matrix Q are absorbed into the two neighboring tensors. The original GILT paper [30] presents a nice way to determine the low-rank matrix Q .

not change the geometry of a given tensor network, so it is very flexible. It has been shown that the TRG combined with GILT is able to generate correct tensor RG flows for the 2D Ising model [30] and the 2D ϕ^4 theory [37]. Figure 5 summarizes the basic process of the GILT. The loop containing four matrices C inside the plaquette represents the local correlations (see Fig. 3(b) and imagine putting four CDL tensors together to form a plaquette). The first step, which is the most crucial one, is to insert a low-rank matrix Q into the leg we wish to truncate. The remaining two steps are exact. We split Q into two pieces using singular value decomposition and absorb the two pieces into the adjacent two A tensors. The bond dimension of the leg is smaller and the local correlations on this leg are filtered out.

The low-rank matrix Q is determined by examining the environment E of the bond and performing the singular value decomposition,

$$E \equiv \begin{array}{c} \text{four } C \text{ matrices in a loop} \end{array} \stackrel{\text{svd}}{=} \begin{array}{c} \text{tensor } U \\ \text{diagonal matrix } s \\ \text{tensor } V^\dagger \end{array}, \quad (27)$$

where we refrain from drawing the unknown C matrices in the plaquette. The environment E of the bond should be thought of as a linear map from the vector space of all the legs with ingoing arrows to that of all the legs with outgoing arrows. We can use the tensor U and the diagonal matrix s in Eq. (27) to construct the low-rank matrix Q . To this end, we first define a vector t by

contracting two ingoing legs of the tensor U ,

$$t \equiv \text{contracted } U. \quad (28)$$

Then, we perform a soft truncation of the vector t according to

$$t'_i = t_i \frac{s_i^2}{s_i^2 + \epsilon_{\text{gilt}}^2}, \quad (29)$$

where s_i are the singular values and ϵ_{gilt} is the hyper-parameter of the GILT. Equation (29) says that the components of the vector t will be set to very small values if the corresponding singular values s_i are much smaller than ϵ_{gilt} . The justification for the truncation in Eq. (29) can be found in Ref. [30]. The low-rank matrix Q is constructed from the tensor U^\dagger and the truncated vector t' as

$$Q \equiv \text{contracted } U^\dagger \text{ and } t'. \quad (30)$$

It is proved in Ref. [30] that the matrix Q determined in this way is able to filter out the loop of four C matrices shown in Fig. 5.

Next, we move on to explain the HOTRG. The block-tensor transformation in Eqs. (23) and (24) is exact but not practical, since the bond dimension grows exponentially in the original lattice size. The HOTRG is an approximate tensor RG transformation, which can keep the bond dimension from growing. For the HOTRG in the vertical direction, we aim at the following approximation of a local patch of two copies of A put together vertically,

$$\begin{array}{c} \text{two } A \text{ tensors stacked vertically} \end{array} \approx \begin{array}{c} \text{single } A \text{ tensor} \end{array}, \quad (31)$$

where w is an isometric tensor to be determined and w^\dagger its hermitian conjugate. The isometry w is a linear mapping: $\mathbb{V}_{\tilde{\chi}} \rightarrow \mathbb{V}_{\chi} \otimes \mathbb{V}_{\chi}$, where \mathbb{V}_{χ} denotes a χ -dimensional vector space, and the isometry satisfies $w^\dagger w = \mathbb{1}$. We will later see that the isometric condition of tensor w makes the gauge fixing in the HOTRG easier. It is shown in Ref. [14, 26] that a good approximation can be achieved if the isometry w is a collection of $\tilde{\chi}$ eigenvectors corresponding to the first $\tilde{\chi}$ largest eigenvalues of the χ^2 -by- χ^2 positive semi-definite matrix MM^\dagger , with the matrix M defined as

$$M = \begin{array}{c} \text{contracted } A \text{ and } A \end{array} \equiv \begin{array}{c} \text{single } A \text{ tensor} \end{array}. \quad (32)$$

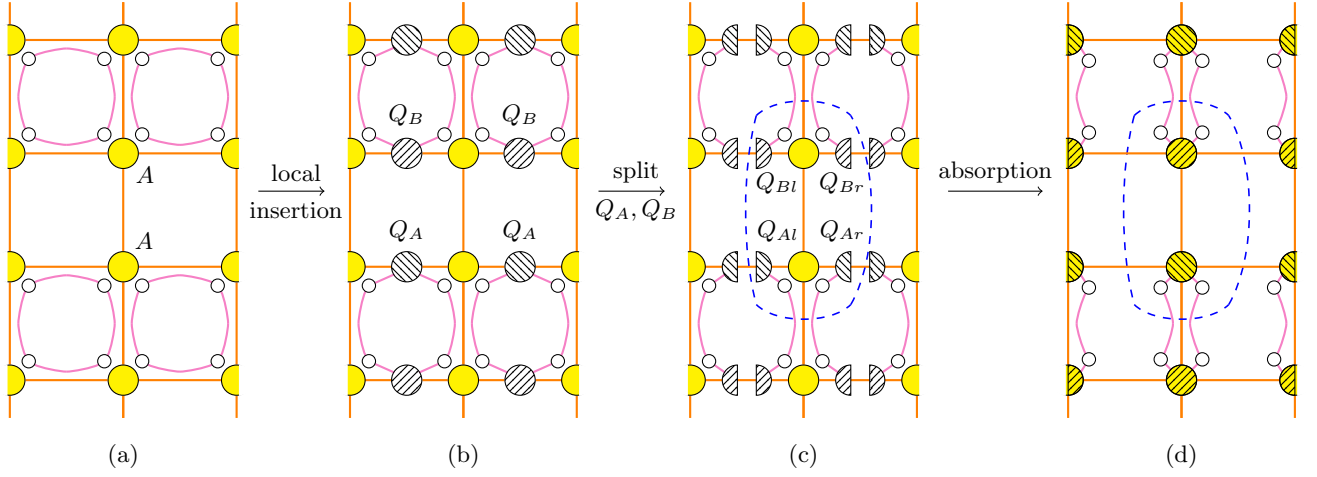


Figure 6. The plaquettes and bonds where the GILT is applied to for the subsequent HOTRG coarse graining. (a) Two copies of A in the center will be coarse grained vertically. The problematic loops of local correlations are drawn explicitly in the plaquettes to make the demonstration clearer. They are unknown inner structure of the main tensor A . (b) Copies of low-rank matrices Q_A, Q_B , determined by the GILT [30], are inserted into the bonds to catch the legs of the loops. (c) Q_A, Q_B are split using singular value decomposition. The GILT ensures the legs of the loops do not leak out. (d) The pieces of Q_A, Q_B matrices are absorbed into the copies of tensor A . The subsequent HOTRG will be applied on the patch of tensors in the dashed circle.

We use the approximation to replace all pairs of A tensors in the tensor network representation of the partition function $Z_{4 \times 4}$ in Eq. (23) to get

$$\begin{aligned}
 Z_{4 \times 4} &\stackrel{(31)}{\approx} \text{[Diagram: A tensor network with two rows of four yellow circles (A) and two rows of four grey circles (C). A dashed blue circle encloses the two central A tensors in the top row.] } \\
 &= \text{[Diagram: A tensor network with two rows of four yellow circles (A') and two rows of four grey circles (C). The A' tensors are connected horizontally, and the C tensors are connected vertically. A dashed blue circle encloses the two central A' tensors in the top row.] }, \quad (33)
 \end{aligned}$$

where in the second step, we contract two A tensors and w, w^\dagger in the dashed circle to get a coarser tensor A' ,

$$\text{[Diagram: A yellow circle labeled A' is shown to be equivalent to a contraction of two yellow circles labeled A with w and w^\dagger legs.] } \quad (34)$$

Notice in the approximation step in Eq. (33), we move the two leftmost w tensors to the right because we have a periodic boundary condition. Equation (34) defines the HOTRG coarse graining in the vertical direction. We

usually choose $\tilde{\chi} \leq \chi_{\max}$ in Eq. (31) to prevent the bond dimension from growing.

It is shown in Ref. [32] that the HOTRG in the vertical direction transforms the A^{CDL} in Fig. 3(b) in the following way,

$$\text{[Diagram: A tensor network with two rows of four yellow circles (A) and two rows of four grey circles (C). A dashed blue circle encloses the two central A tensors in the top row.] } \propto \text{[Diagram: A yellow circle labeled A' with four legs.] }, \quad (35)$$

which means that although the HOTRG can detect and project out four inner C matrices, it can do nothing about the four outer C matrices. Therefore, the GILT should be applied to filter out these four outer C matrices before the HOTRG coarse graining. To this end, we apply the GILT to the plaquettes where the loops of local correlations are drawn explicitly in Fig. 6 and insert two low-rank matrices Q_A, Q_B into the upper and lower bonds for each plaquette. The legs of the unwanted C matrices will be truncated after the splitting of Q_A, Q_B . Finally, we apply the ordinary HOTRG in the vertical direction to the local patch of tensors in the dashed circle in Fig. 6 to get the coarser tensor A'

$$\text{[Diagram: A yellow circle labeled A' is shown to be equivalent to a contraction of two yellow circles labeled A with Q_A and Q_B legs.] } \quad (36)$$

In this way, we can remove all horizontal legs of C matrices: a half of them by the GILT and the other half by

contraction in the HOTRG. We repeat the similar GILT and the HOTRG on A' in the horizontal direction. The coarse graining steps in two directions together define the RG equation of the GILT-HOTRG,

$$(37)$$

The computational cost of the GILT-HOTRG is $O(\chi^7)$, the same as the HOTRG.

The coarse graining defined in Eq. (37) is able to simplify the A^{CDL} tensor in Fig. 3(b) to a single number,

$$(38)$$

Equation (38) shows that the GILT-HOTRG can successfully filter out the local correlations among the spin variables around the corners at the lattice scale (see Fig. 3(b)). Since the CDL tensors are no longer fixed points for the RG equation of the GILT-HOTRG, the peculiar fixed lines in Fig. 4(b) generated by the HOTRG will collapse to fixed points; we expect the RG equation of the GILT-HOTRG is able to exhibit the critical fixed point tensor shown schematically in Fig. 4(a).

D. Gauge fixing and linearization for the GILT-HOTRG

We show how the gauge is fixed and give the explicit expression of the linearized RG equation for the GILT-HOTRG in this subsection.

Part of the gauge can be fixed if the physical model possesses a global internal symmetry. The global symmetry can be incorporated into the tensor network representation of the model [38–40]; it is a generalization of Schur's lemma from matrices to general tensors. For the 2D Ising model, \mathbb{Z}_2 symmetry can be imposed. Each index of the tensor A breaks into even and odd sectors. Half of the gauge is fixed since A is in the bases where the states in the even sector transform trivially and the states in the odd sector is multiplied by -1 under the spin flip operation.

Most of the remaining gauge in the degenerate sectors of A can be fixed by going to the diagonal bases of the tensor. We show how the S_x gauge redundancy in Eq. (25) is fixed. The S_y one can be dealt with in the same way.

Given a tensor A , we first contract its two vertical legs to produce a transfer matrix N_x ,

$$(39)$$

We then find the eigenvalue decomposition of this matrix,

$$(40)$$

where λ is the diagonal matrix encoding eigenvalues. The gauge fixing transformation in the horizontal direction is defined by acting the invertible matrix W_x and its inverse on the horizontal legs of the tensor A ,

$$(41)$$

To see why the gauge fixing procedure in Eqs. (39) to (41) defines a preferred set of bases, let us examine how the tensor \tilde{A} in Eq. (25) transforms under this gauge fixing procedure. The contraction of two vertical legs of \tilde{A} annihilates S_y and S_y^{-1} in the right hand side of Eq. (25b); the resultant \tilde{N}_x is related to N_x through

$$(42)$$

It is straightforward to see that the matrix \tilde{W}_x coming from eigenvalue decomposition of the matrix \tilde{N}_x is related to W_x through

$$(43)$$

where d_x is a diagonal matrix coming from phase ambiguities of eigenvectors, with its diagonal entries to be phases for general complex matrices. For a real symmetric N_x , the diagonal entries of d_x are ± 1 . After the horizontal gauge fixing, the tensor \tilde{A} becomes

$$(44)$$

Compare Eq. (41) with Eq. (44), we see that the gauge redundancies in two horizontal legs are fixed except the phase ambiguities. For 2D classical statistical models with spatial reflection symmetries, for example, the 2D Ising model, the real matrix N_x can be made symmetric, so the phase ambiguities become sign ambiguities.

The gauge fixing procedure described in Eqs. (39) to (41) is general for all TRG-type techniques. However,

this procedure is not necessary for the GILT-HOTRG applied to systems with spatial reflection symmetries like the 2D Ising model, since the RG equation of the GILT-HOTRG has a preferred set of bases. As a result, the gauge redundancy in Eq. (25) collapses into phase ambiguities (or sign ambiguities for real tensors) in the GILT-HOTRG. To make things as simple as possible, we focus on real tensors in the following discussions. The generalization to complex tensors is straightforward.

Write the tensor RG equation in Eq. (37) schematically as $A_c = \mathcal{T}(A)$. For two real tensors A, \tilde{A} that are related by the gauge transformation defined in Eq. (25) where we further restrict S_x, S_y to be orthogonal matrices, the new tensors produced by the GILT-HOTRG according to Eq. (37), $A_c = \mathcal{T}(A), \tilde{A}_c = \mathcal{T}(\tilde{A})$ are equal up to sign ambiguities,

$$(\tilde{A}_c)_{ijkl} = (A_c)_{ijkl} (d_x)_i (d_y)_j (d_x)_k (d_y)_l, \quad (45)$$

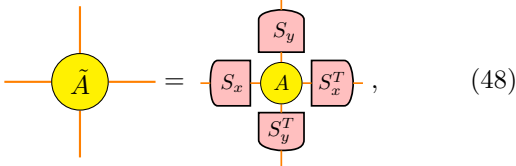
where d_x, d_y are vectors with components ± 1 . Imagine that we manage to fix the sign ambiguities, then the GILT-HOTRG ensures

$$\mathcal{T}(A) = \mathcal{T}(\tilde{A}). \quad (46)$$

Since the orthogonal matrices S_x, S_y are arbitrary, equation (46) says that the whole equivalence class $[A]$ will be mapped into the same tensor A_c . This means that the GILT-HOTRG, after incorporating the sign fixing step, will choose a preferred set of bases. It is worth to mention that the TRG has a similar property [21]. For a fixed-point tensor, equation (46) indicates that we can start with any representation \tilde{A}^* of the equivalence class $[A^*]$, and the GILT-HOTRG will bring \tilde{A}^* to the proper bases; further GILT-HOTRG coarse graining will satisfy Eq. (19),

$$\mathcal{T}(\tilde{A}^*) = \mathcal{T}(\mathcal{T}(\tilde{A}^*)) \equiv A^*. \quad (47)$$

Let us prove the property of the tensor RG equation of the GILT-HOTRG in Eq. (45) and explain how the sign is fixed to have Eq. (46). We focus on the equivalence relation defined as



$$\tilde{A} = S_x A S_y^T, \quad (48)$$

where S_x, S_y are orthogonal matrices. It is sufficient to consider such orthogonal changes of gauge if we restrict to the representations of the equivalence class $[A]$ with spatial reflection symmetries [26],

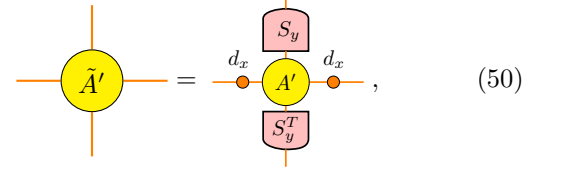
$$A_{kijl} = \sum_{j'l'} (O_y)_{jj'} (O_y)_{ll'} A_{ij'kl'} \quad (49a)$$

and

$$A_{ilkj} = \sum_{i'k'} (O_x)_{ii'} (O_x)_{kk'} A_{i'jk'l}, \quad (49b)$$

where O_x, O_y are orthogonal matrices, also with $O_x^2 = O_y^2 = \mathbb{1}$, and the legs' order convention is as per Eq. (2). It can be shown that, if we start with a tensor with reflection symmetry, the GILT-HOTRG will preserve the reflection symmetry and will rotate the tensor into the set of bases where O_x, O_y become diagonal, with their diagonal entries ± 1 [41].

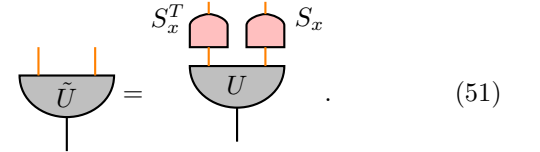
It suffices to discuss the first half of the GILT-HOTRG coarse graining defined in Eq. (36). We want to show that if \tilde{A} is fed into the right hand side of Eq. (36), the \tilde{A}' we obtain in the left hand is related with the original A' by



$$\tilde{A}' = d_x A' d_x, \quad (50)$$

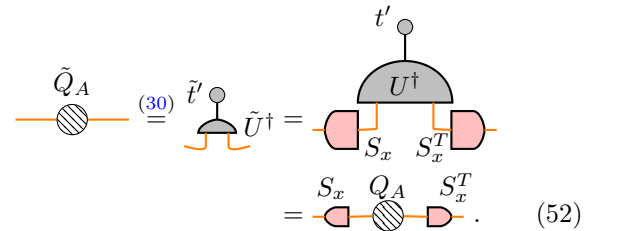
which means that the gauge redundancy in the horizontal legs will be fixed with only sign ambiguities left during the first half of the GILT-HOTRG in the vertical direction. It follows immediately that the full GILT-HOTRG will give Eq. (45).

Let us first figure out the correct \tilde{Q}_A, \tilde{Q}_B matrices in Fig. 6. The environment in Eq. (27) is multiplied by several orthogonal matrices, which will not change the singular values, so $\tilde{s}_i = s_i$. It is easy to check that the tensor U in the singular value decomposition becomes (the sign ambiguities coming from the singular value decomposition does not matter here)



$$\tilde{U} = S_x U S_x^T. \quad (51)$$

The vector \tilde{t} is thus the same as the original t by its definition in Eq. (28), which further gives $\tilde{t}'_i = t'_i$ since the tilde version of the right hand side of Eq. (29) is the same as the original version. Finally, equation (30) gives



$$\tilde{Q}_A \stackrel{(30)}{=} \tilde{t}' U^\dagger = S_x Q_A S_x^T. \quad (52)$$

Equation (52) means that the low rank matrix Q_A transforms in a nice way when we perform a gauge transformation of the tensor A defined in Eq. (48). If the singular values of Q_A do not have degeneracy, after splitting of

Q_A , we have Q_{Ar}, Q_{Al} transform like (the sign ambiguities coming from singular value decomposition of Q_A and \tilde{Q}_A would kick in and contribute to d_x in Eq. (50), but they are not drawn explicitly in the equation below)

$$\tilde{Q}_{Ar} = S_x Q_{Ar} \quad \text{and} \quad \tilde{Q}_{Al} = Q_{Al} S_x^T. \quad (53)$$

The S_x, S_x^T matrices that Q_{Ar}, Q_{Al} pick up will cancel those acting on the A tensor when $\tilde{Q}_{Ar}, \tilde{Q}_{Al}$ are contracted with the \tilde{A} tensor in Eq. (48). The same argument works for Q_B . Equation (53) indicates that all the S_x, S_x^T matrices acting on the four horizontal legs of the local patch in Eq. (36) will be canceled by the low-rank matrices used in the GILT process. The above analysis shows that during the first half of the GILT-HOTRG, the gauge in the horizontal legs will be fixed with only sign ambiguities left, since the GILT favors the bases chosen by the singular value decompositions of Q_A, Q_B .

However, there is one more twist. In practice, we observe that the low-rank matrices are projection operators, which are highly degenerated. As a result, the gauge redundancy in the degenerate subspace will leak out, which will be seen by the subsequent HOTRG process. Luckily, the HOTRG has a similar feature as the GILT process. It favors bases where the positive semi-definite matrix MM^\dagger (see the definition of matrix M in Eq. (32)) is diagonal (the sign ambiguities coming from eigenvalue decomposition of MM^\dagger would similarly kick in here and contribute to d_x in Eq. (50)). It is straightforward to

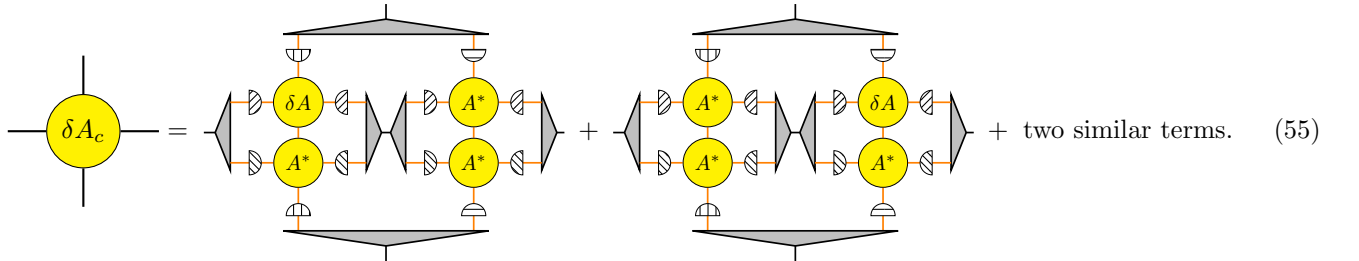
see that the isometry w will pick up the suitable S_x, S_x^T matrices to cancel out the gauge transformation leaking out from the GILT process. There are still concerns about whether degeneracy occurs in eigenvalues of MM^\dagger . Our result in Fig. 9(b) shows, a posteriori, that the potential degeneracy does not cause any problem for the 2D Ising model at criticality.

The sign ambiguities d_x, d_y in Eq. (45) can be determined by comparing the sign of the components of \tilde{A}_c and A_c . For example, upon making sure $(\tilde{A}_c)_{1111}$ and $(A_c)_{1111}$ are both positive, set $j = k = l = 1$ in Eq. (45) to have

$$(\tilde{A}_c)_{i111} = (A_c)_{i111}(d_x)_i. \quad (54)$$

The relative sign of $(\tilde{A}_c)_{i111}$ and $(A_c)_{i111}$ determines $(d_x)_i$. However, this sign fixing method breaks down if both $(\tilde{A}_c)_{i111}$ and $(A_c)_{i111}$ vanish, which occurs as long as there is a symmetry. This is the reason why we first fix part of the gauge by exploiting the global internal symmetry of the physical model. Then, we can apply Eq. (54) in each degenerate sectors of the tensor. The detailed implementation of the sign fixing procedure for \mathbb{Z}_2 symmetric tensors can be found in the source code of this paper (see Appendix A). This completes the arguments for Eqs. (45) and (46).

After reaching the fixed-point tensor A^* in Eq. (47), the next step is to linearize the RG equation of the GILT-HOTRG in Eq. (37). We substitute $A = A^* + \delta A$ into the right hand side of Eq. (37) and collect terms that are first order in δA to get δA_c ,



$$\delta A_c = \text{[Diagrammatic expression]} + \text{two similar terms}. \quad (55)$$

The result resembles the product rule for taking the differentials in calculus. Equation (55) provides a simple pictorial representation of the linearized tensor RG equation \mathcal{R} in Eq. (22) for the GILT-HOTRG. In practice, after the fixed-point tensor A^* , the pieces of low-rank matrices Q_A, Q_B and the isometric tensors w, v in Eq. (37) are determined, automatic differentiation can linearize Eq. (37) around A^* and generate Eq. (55) for us. There are many libraries that support automatic differentiation, including PyTorch [42] and JAX [43].

III. BENCHMARKS

We use the classical Ising model in 1D and 2D to demonstrate how to carry out the canonical RG prescription in tensor space. The Ising model in 1D serves as a concrete example to elucidate the general argument in Sec. II A. The Ising model in 2D provides more nontrivial benchmark results for our method.

A. The Ising Model in 1D

The Ising model in 1D has an exact real-space RG transformation realized via decimation. Even better, the decimation has a natural tensor network representation. This makes the Ising model in 1D a nice example to see the relation between the old and the new approaches of the canonical RG prescription.

The partition function is

$$Z_{1D} = \sum_{\{\sigma_j\}} \exp \left[\sum_{i=1}^N \mathcal{H}(\sigma_i, \sigma_{i+1}) \right], \quad (56)$$

where the local interactions involve the nearest-neighbor term at most

$$\mathcal{H}(\sigma_1, \sigma_2) = g + \frac{h}{2}(\sigma_1 + \sigma_2) + K\sigma_1\sigma_2. \quad (57)$$

The decimation process is shown in Fig. 7. It is realized by summing over all the even-numbered spins and then renumber the remaining odd-numbered spins.

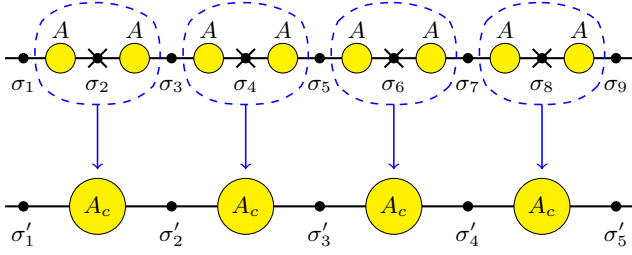


Figure 7. The decimation for the 1D Ising model. The black dots are spin variables. The spins on even sites $\sigma_2, \sigma_4, \dots$ are summed over and the remaining spins $\sigma_1, \sigma_3, \dots$ are renamed $\sigma'_1, \sigma'_2, \dots$ to become new spin variables. In the tensor network language, this decimation is nothing but a matrix multiplication of two transfer matrices to form a coarse-grained matrix $A_c = AA$.

We denote $\sigma'_i = \sigma_{2i-1}$, $s_i = \sigma_{2i}$ and sum over all s -spins in the partition function in Eq. (56) to have

$$Z_{1D} = \sum_{\{\sigma'_j\}} \sum_{\{s_j\}} \exp \left[\sum_i^{N/2} [\mathcal{H}(\sigma'_i, s_i) + \mathcal{H}(s_i, \sigma'_{i+1})] \right], \quad (58)$$

from which we can define the effective local interaction \mathcal{H}' through

$$\exp[\mathcal{H}'(\sigma'_1, \sigma'_2)] = \sum_{s=\pm 1} \exp[\mathcal{H}(\sigma'_1, s) + \mathcal{H}(s, \sigma'_2)], \quad (59)$$

where the effective local interaction has the same form as the old one in Eq. (57) but with new coupling constants g', h', K' ,

$$\mathcal{H}'(\sigma_1, \sigma_2) = g' + \frac{h'}{2}(\sigma_1 + \sigma_2) + K'\sigma_1\sigma_2. \quad (60)$$

The partition function can be fully described by the new σ' -spins,

$$Z_{1D} = \sum_{\{\sigma'_j\}} \exp \left[\sum_{i=1}^{N/2} \mathcal{H}'(\sigma'_i, \sigma'_{i+1}) \right]. \quad (61)$$

Equations (57), (59) and (60) together define the RG equation that maps the old coupling constants (g, h, K) to the new coupling constants (g', h', K') . The explicit expression of the RG equation can be found in Kardar's textbook [44]. The RG equation has two fixed points, one for high-temperature phase and the other for low-temperature phase. Let us focus on the high-temperature fixed point here, where the coupling constants are $g^* = \log(1/2)$, $h^* = 0$, $K^* = 0$. The linearized RG equation around this fixed point gives $\delta g' = 2\delta g$, $\delta h' = \delta h$, $\delta K' = 0 \times \delta K$. The matrix \mathcal{R} is in its diagonal form with eigenvalues 2, 1, 0 for $\delta g, \delta h, \delta K$ respectively.

Next, we translate the above decimation process into tensor network language. We first define the tensor A sitting on the bond connecting two spins shown in Fig. 7 as

$$A_{\sigma_1\sigma_2} = \exp[\mathcal{H}(\sigma_1, \sigma_2)]. \quad (62a)$$

After using the expression for \mathcal{H} in Eq. (57), we have

$$A = \begin{pmatrix} \exp(g+h+K) & \exp(g-K) \\ \exp(g-K) & \exp(g-h+K) \end{pmatrix}, \quad (62b)$$

which is the familiar transfer matrix. Each component of the tensor A is a function of coupling constants g, h, K , as is claimed in Eq. (15). The partition function in Eq. (56) can be rewritten as

$$Z_{1D} = \sum_{\{\sigma_j\}} \bigotimes_{i=1}^N A_{\sigma_i\sigma_{i+1}}. \quad (63)$$

The decimation in the tensor network language is a multiplication of two old A matrices to form a new A_c matrix,

$$A_c = AA. \quad (64)$$

In terms of the new A_c matrix, the partition function is

$$Z_{1D} = \sum_{\{\sigma'_j\}} \bigotimes_{i=1}^{N/2} (A_c)_{\sigma'_i\sigma'_{i+1}}. \quad (65)$$

Equation (64) is the RG equation in the tensor network language. Now, each component of A_c is a function of coupling constants g, h, K but with different functional form, as is claimed in Eq. (16). If we further require that A_c should have the same form as A in Eq. (62) but with \mathcal{H} replaced by \mathcal{H}' , new coupling constants g', h', K' can be solved in terms of the old ones, which is what we do in the conventional approach. The advantage of using

the tensor network language is that the RG equation in Eq. (64) suffices for the canonical RG prescription in tensor space. First, let us set the coupling constants in Eq. (62b) to be the high-temperature fixed point $g^* = \log(1/2)$, $h^* = 0$, $K^* = 0$ to get the fixed-point tensor,

$$A^* = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (66)$$

It can be checked that $A^* A^* = A^*$. The linearized version of Eq. (64) around this fixed-point tensor is

$$\delta A_c = \delta A A^* + A^* \delta A = I \delta A A^* + A^* \delta A I, \quad (67)$$

where in the last equal sign, we add two identity matrices. Write Eq. (67) in its component form, we have $(\delta A_c)_{ab} = \sum_{\alpha, \beta} I_{a\alpha} (\delta A)_{\alpha\beta} (A^*)_{\beta b} + (A^*)_{a\alpha} (\delta A)_{\alpha\beta} I_{\beta b}$. We can read off the matrix of the linearized RG equation as

$$\mathcal{R}_{(ab)(\alpha\beta)} = \frac{(\delta A_c)_{ab}}{(\delta A)_{\alpha\beta}} = I_{a\alpha} (A^*)_{\beta b} + (A^*)_{a\alpha} I_{\beta b}, \quad (68)$$

where we group two indices a, b as a single index (ab) , and α, β as $(\alpha\beta)$. If we put the grouped index into the following order,

$$(11) \rightarrow 1, (12) \rightarrow 2, (21) \rightarrow 3, (22) \rightarrow 4, \quad (69)$$

the matrix takes the following value

$$\mathcal{R} = \begin{pmatrix} 1 & 1/2 & 1/2 & 0 \\ 1/2 & 1 & 0 & 1/2 \\ 1/2 & 0 & 1 & 1/2 \\ 0 & 1/2 & 1/2 & 1 \end{pmatrix}. \quad (70)$$

This matrix \mathcal{R} in Eq. (70) is a symmetric, and we can find its eigenvalues and eigenvectors: $\lambda_1 = 2$, $\mathbf{v}_1 = (1, 1, 1, 1)^T$; $\lambda_2 = 1$, $\mathbf{v}_2 = (1, 0, 0, -1)^T$; $\lambda_3 = 1$, $\mathbf{v}_3 = (0, 1, -1, 0)^T$ and $\lambda_4 = 0$, $\mathbf{v}_4 = (1, -1, -1, 1)^T$. The eigenvalues are the same as what we get in the conventional method.

The relation between the canonical RG prescription in tensor space and the Hamiltonian space can be clarified by noticing that the relation between the coupling constants and the tensor A is given in Eq. (62b). We perturb the coupling constants around the fixed point, $g_p = \log(1/2) + \delta g$, $h_p = \delta h$, $K_p = \delta K$, substitute them into the right hand side of Eq. (62b) and Taylor expand to get the perturbed tensor,

$$A_p = A^* + \frac{1}{2} \delta g \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2} \delta h \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{2} \delta K \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \text{higher-order terms}. \quad (71)$$

We can read off $\delta A = A_p - A^*$ as

$$\delta A = \frac{1}{2} \delta g \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2} \delta h \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{2} \delta K \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad (72)$$

which is Eq. (20) in practice. Recall the order convention in Eq. (69), we see the correspondence $\mathbf{v}_1 \leftrightarrow \delta g$, $\mathbf{v}_2 \leftrightarrow \delta h$ and $\mathbf{v}_4 \leftrightarrow \delta K$.

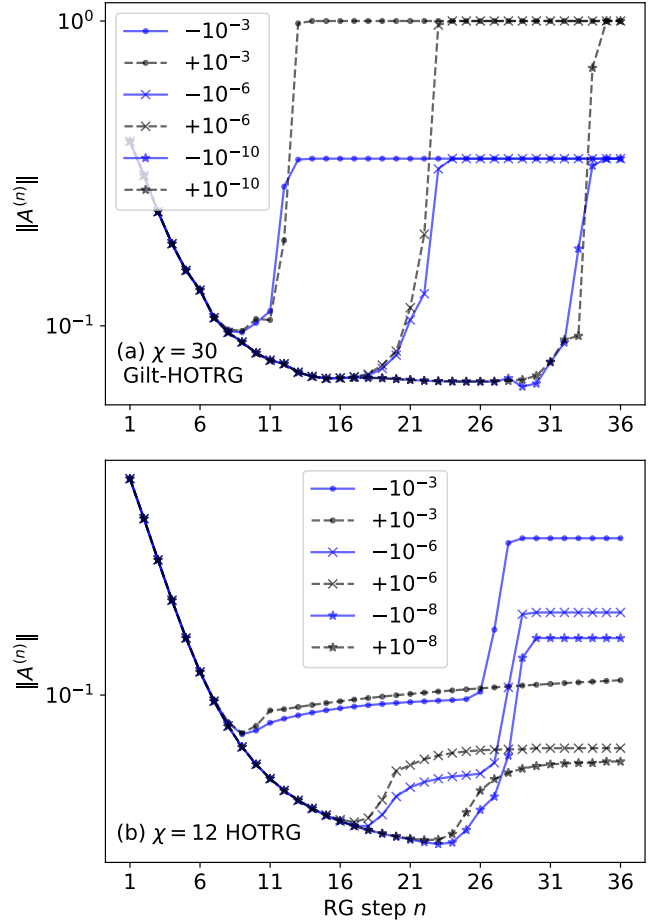


Figure 8. The RG flows of the tensor norms $\|A^{(n)}\|$ at temperatures near the estimated critical temperature $T_c^{[x]}$. Different markers represent different deviations $|\Delta T|$ from $T_c^{[x]}$. Blue solid lines are for $\Delta T < 0$ and black dashed lines for $\Delta T > 0$. (a) For the GILT-HOTRG with $\chi = 30$, $\epsilon_{\text{gilt}} = 6 \times 10^{-6}$, two trivial fixed points are isolated and the critical fixed point can be reached. It corresponds to the schematic RG flows in Fig. 4(a). (b) For the HOTRG with $\chi = 12$, we have fixed lines and there is no exhibition of a critical fixed point. It corresponds to the schematic RG flows in Fig. 4(b).

B. The Ising Model in 2D

There is no exact RG transformation for the Ising model in 2D, so we will use the GILT-HOTRG developed in Sec. II C to generate an RG flow in tensor space. The partition function is given in Eq. (1) and we translate the partition function into a tensor network in Fig. 2. Let us denote the initial tensor in Eq. (2) as $A^{(0)}$. To prevent a rapid grow of the magnitude of the tensor during the RG transformation, we pull out the Frobenius norm of the tensor, $A^{(0)} = \|A^{(0)}\| \mathcal{A}^{(0)}$, to define a normalized tensor $\mathcal{A}^{(0)}$. The normalized tensor $\mathcal{A}^{(0)}$ will be fed into the RG equation of the GILT-HOTRG in Eq. (37) and we denote the output coarse-grained tensor as $A^{(1)}$, from which the norm $\|A^{(1)}\|$ is pulled out and the normalized tensor

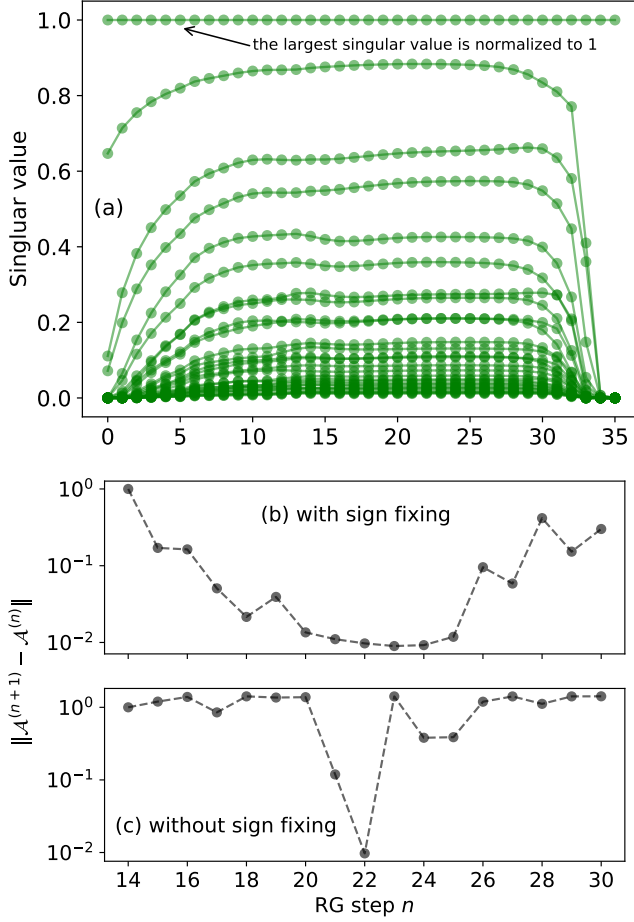


Figure 9. The RG flows of (a) singular values defined in Eq. (73) and (b) the difference between the normalized tensors, $\|\mathcal{A}^{(n+1)} - \mathcal{A}^{(n)}\|$ with sign fixing and (c) without, all at temperature $T_c^{[30]}$ for the GILT-HOTRG with $\chi = 30$, $\epsilon_{\text{gilt}} = 6 \times 10^{-6}$.

$\mathcal{A}^{(1)}$ is defined the same way as the previous step. The process can be repeated so we will have $\mathcal{A}^{(n)} = \|\mathcal{A}^{(n)}\| \mathcal{A}^{(n)}$ at the n -th step. The RG flow in tensor space can be conveniently visualized by examining the evolution of the norms $\|\mathcal{A}^{(n)}\|$ as the RG step n increases.

The RG flows of the norms $\|\mathcal{A}^{(n)}\|$ indicate the GILT-HOTRG is capable of generating a correct RG flow for the 2D Ising model in tensor space shown schematically in Fig. 4(a). For example, for bond dimension $\chi = 30$ and the hyper-parameter of the GILT process $\epsilon_{\text{gilt}} = 6 \times 10^{-6}$, Fig. 8(a) shows several RG flows of the tensor norms $\|\mathcal{A}^{(n)}\|$ at different temperatures. For a given bond dimension χ , there is an estimated critical temperature $T_c^{[\chi]}$ at which the tensor hits the critical surface and will flow to the critical fixed-point tensor $(\mathcal{A}^{[\chi]})_{\text{cr}}^*$. The $T_c^{[\chi]}$ can be determined using the bisection method, and for $\chi = 30$ the difference between the estimated value $T_c^{[30]}$ and the exact T_c , $|T_c^{[30]} - T_c|/T_c$, is of order 10^{-6} . At temperatures off by $\Delta T = \pm 10^{-3}$ from $T_c^{[30]}$, the tensor flows to the high- and low-temperature trivial

Table I. The scaling dimensions for the relevant and marginal operators of the 2D Ising model at criticality from the canonical RG prescription and from the transfer matrix method à la Gu and Wen [22], both using the GILT-HOTRG with $\chi = 30$, $\epsilon_{\text{gilt}} = 6 \times 10^{-6}$ at RG step $n = 22$.

Exact	0.125	1	1.125	1.125	2	2	2	2
RG pres.	0.127	1.009	1.125	1.128	2.002	2.004	2.068	2.073
Trans. mat.	0.125	1.009	1.130	1.148	1.313	1.457	1.558	1.654

fixed-point tensors respectively before it comes near to $(\mathcal{A}^{[30]})_{\text{cr}}^*$. As $|\Delta T|$ becomes smaller to order of 10^{-6} , the tensor will stay in the vicinity of the critical fixed-point tensor $(\mathcal{A}^{[30]})_{\text{cr}}^*$ for a while and then flow away to one of the two trivial fixed-point tensors. If $|\Delta T|$ becomes smaller further to 10^{-10} , the tensor will stay longer near $(\mathcal{A}^{[30]})_{\text{cr}}^*$. By comparison, the RG flow of $\|\mathcal{A}^{(n)}\|$ generated by the HOTRG with bond dimension $\chi = 12$ [45] is displayed in Fig. 8(b). The RG flow shows that the HOTRG has difficulty in exhibiting a critical fixed-point tensor or producing isolated trivial fixed-point tensors. It is interesting to mention that the RG flow generated by the TRG has a similar behavior [18] for bond dimensions $\chi > 8$.

To make sure that the plateau in the RG flow of $\|\mathcal{A}^{(n)}\|$ gives a critical fixed-point tensor $(\mathcal{A}^{[30]})_{\text{cr}}^*$ at the estimated critical temperature $T_c^{[30]}$, we plot the singular values $s^{(n)}$ of tensors $\mathcal{A}^{(n)}$ defined as

$$\mathcal{A}^{(n)} \xrightarrow{\text{svd}} \begin{matrix} \text{---} U \\ \text{---} s^{(n)} \\ \text{---} V^\dagger \end{matrix} \quad (73)$$

The RG flow of the singular values in Fig. 9(a) indicates that we indeed reach a non-trivial fixed-point tensor. The fixed-point tensor is manifestly fixed numerically after adding the sign fixing step in the GILT-HOTRG, which can be confirmed by plotting the Frobenius norm of the difference between the normalized tensors at successive RG steps $\|\mathcal{A}^{(n+1)} - \mathcal{A}^{(n)}\|$, see Fig. 9(b). The norm of the difference starts to decay systematically at RG step $n = 14$, goes all the way down to the order $\sim 10^{-2}$ at $n = 23$ and then increases when the tensor begins to flow away from the critical fixed point. By comparison, we show the RG flow of $\|\mathcal{A}^{(n+1)} - \mathcal{A}^{(n)}\|$ without sign fixing in Fig. 9(c); the sign ambiguities in Eq. (45) prevent us from achieving a manifestly-fixed-point tensor, except at RG step $n = 22$, where the tensor happens to have all signs correct by accident.

We use the automatic differentiation implemented in JAX [43] to generate the linearized tensor RG equation \mathcal{R} in Eq. (55) at RG steps $n = 14, 15, \dots, 28$, when the tensor is very close to the critical fixed-point tensor. The

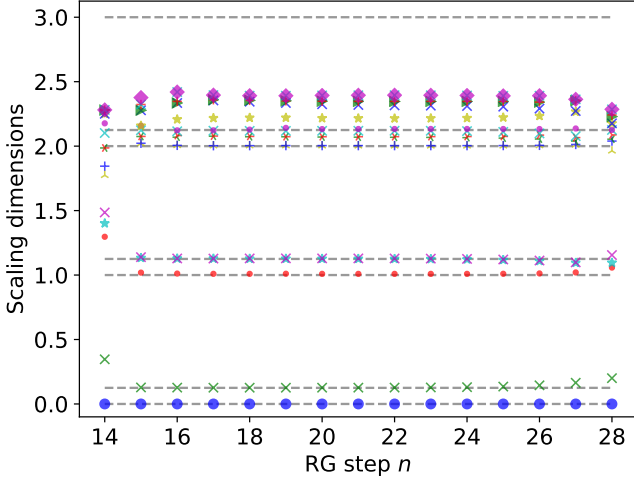


Figure 10. The scaling dimensions of the 2D Ising model from the canonical RG prescription using the GILT-HOTRG with $\chi = 30$, $\epsilon_{\text{gilt}} = 6 \times 10^{-6}$. Dashed lines are the exact values.

scaling dimensions are extracted from the eigenvalues of the matrix \mathcal{R} according to Eq. (14), where $b = 2, d = 2$. In Fig. 10, we show the first few scaling dimensions. The dashed lines are the exact values [46]. For $\chi = 30$, the RG prescription in tensor space gives correct scaling dimensions up to 2.125. The results at RG step $n = 14$ and 28 are unreliable since $\|\mathcal{A}^{(n+1)} - \mathcal{A}^{(n)}\|$ is of order 1 (see Fig. 9(b)). The results for $n = 15, 16, \dots, 27$ indicates that the scaling dimensions from the RG prescription in tensor space are reliable as long as the values of $\|\mathcal{A}^{(n+1)} - \mathcal{A}^{(n)}\|$ have order of or smaller than 10^{-1} .

In Table I, we show the scaling dimensions for all relevant and marginal operators at RG step $n = 22$, compared with the results obtained by Gu and Wen's method [22]. Both methods have similar accuracy for scaling dimensions less or equal to 1.125, but Gu and Wen's method cannot produce higher scaling dimensions. Our method can give two out of total four scaling dimension 2 accurately, but the remaining two are overestimated and more close to 2.125. It is worth mentioning that we only use a single fixed-point tensor to construct the transfer matrix in Gu and Wen's method. If we use two or more copies of the fixed-point tensor to construct a larger transfer matrix, Gu and Wen's method is expected to produce higher scaling dimensions correctly [24].

We end this section with a few remarks on the above calculations. Firstly we impose the \mathbb{Z}_2 symmetry of the tensors [38, 39] when generating the RG flow in tensor space. There are three reasons. Only if the \mathbb{Z}_2 symmetry of the tensor is imposed will the low-temperature fixed-point tensor be stable under the RG. Otherwise, it will flow to the high-temperature fixed point eventually due to numerical errors, which will make the bisection search for the estimated critical temperature $T_c^{[x]}$ less convenient. The second merit of symmetric tensors is that half of the gauge redundancy can be automatically fixed (see

Sec. IID), making the sign fixing procedure in the GILT-HOTRG easier. The third reason is to speed up the computations. However, we roll back to ordinary tensors when performing the RG prescription in tensor space, since the perturbations around the fixed-point tensor do not have to preserve \mathbb{Z}_2 symmetry (for example, the spin operator).

The second remark is about the improvement of the accuracy as the bond dimension χ increases. There are two sources of approximation errors in the above computations. One comes from the truncations of the CDL tensors during the GILT that is necessary for producing the critical fixed point. This error is controlled by the hyper-parameter ϵ_{gilt} . The other source is the leg squeezing step during the HOTRG to prevent the grow of the bond dimension. This error can be reduced by increasing the bond dimension χ . In general, for a given χ , the ϵ_{gilt} should be as small as possible provided that the GILT-HOTRG can exhibit a critical fixed-point tensor. In practice, we tried $\chi = 10, 20, 30$, and ϵ_{gilt} goes down from 6×10^{-4} to 6×10^{-5} and further to 6×10^{-6} . The estimated scaling dimensions converge to the exact results in this process.

The third remark is about the overall multiplication constant in front of the fixed-point tensor. After reaching the critical fixed point, the RG from n -th step to $(n+1)$ -th step is the mapping $\mathcal{A}^* \rightarrow \mathcal{A}^* = \|\mathcal{A}^*\|\mathcal{A}^*$. The shape of \mathcal{A}^* is fixed but its magnitude is still changing under the RG transformation. It has been shown in Ref. [22] that the fixed-point tensor with correct magnitude is simply given by $\mathcal{A}_{\text{inv}}^* = \|\mathcal{A}^*\|^{-1/3}\mathcal{A}^*$, and we will have $\mathcal{A}_{\text{inv}}^* \rightarrow \mathcal{A}_{\text{inv}}^*$ under the RG transformation. Our numerical results have confirmed this statement.

The final remark is that the problem of local correlations could be removed by other methods [22, 24–27, 33–36] other than GILT. For example, the TNR [25, 26] is known to be capable of exhibiting critical fixed-point tensors with its RG equation similar to that of the GILT-HOTRG, and there is a method to fix its gauge [26]. Considering the unprecedented accuracy of the TNR, the estimation of the scaling dimensions might be much better. We develop the canonical RG prescription in tensor space using the GILT-HOTRG in this paper in order to prepare for the further applications to 3D systems.

IV. SUMMARY AND DISCUSSIONS

In this paper, we show how to analyze an RG fixed point in tensor space. The general procedure is summarized as follows: reach a fixed-point tensor using a TRG-type RG equation, fix the gauge redundancy to make the fixed-point tensor manifestly fixed, linearize the RG equation around this fixed-point tensor and finally calculate the scaling dimensions from the eigenvalues of this linearized tensor RG equation. In practice, we propose the GILT-HOTRG that carries out this canonical RG prescription in tensor space. The benchmark results of the 2D classical

Ising model are comparable with the existing method. In our future work, we will generalize the GILT-HOTRG and apply the canonical tensor RG prescription to 3D systems, where there are few practical tensor network methods to extract scaling dimensions efficiently.

ACKNOWLEDGMENTS

We thank Satoshi Morita, Shumpei Iino, Takuhiro Ogino, Yuan Yao and Takeo Kato for fruitful discussions and insightful suggestions, and Glen Evenbly and Guifre Vidal for explanations regarding the TNR and other tensor network methods. We also thank Markus Hauru for clarifying the implementation of the GILT. X.L.

and R.G.X. are grateful to the support of the Global Science Graduate Course (GSGC) program of the University of Tokyo. This work is financially supported by MEXT Grant-in-Aid for Scientific Research (B) (19H01809). The numerical computations were performed on computers at the Supercomputer Center, the Institute for Solid State Physics (ISSP), the University of Tokyo.

Appendix A: Source Code

The source code of this paper can be found at github.com/brucelyu/tensorRGflow. It can be used to reproduce all the results in Sec. III B for the 2D classical Ising model.

-
- [1] K. G. Wilson, The renormalization group and critical phenomena, *Rev. Mod. Phys.* **55**, 583 (1983).
 - [2] K. G. Wilson and M. E. Fisher, Critical exponents in 3.99 dimensions, *Phys. Rev. Lett.* **28**, 240 (1972).
 - [3] L. P. Kadanoff, Scaling laws for ising models near T_c , *Physics Physique Fizika* **2**, 263 (1966).
 - [4] L. P. Kadanoff, Variational principles and approximate renormalization group calculations, *Phys. Rev. Lett.* **34**, 1005 (1975).
 - [5] A. Migdal, Recursion equations in gauge field theories 1975 sov. phys. JETP **42**, 743 (1975).
 - [6] L. P. Kadanoff, Notes on migdal's recursion formulas, *Annals of Physics* **100**, 359 (1976).
 - [7] T. Niemeijer and J. M. J. van Leeuwen, Wilson theory for spin systems on a triangular lattice, *Phys. Rev. Lett.* **31**, 1411 (1973).
 - [8] T. L. Bell and K. G. Wilson, Nonlinear renormalization groups, *Phys. Rev. B* **10**, 3935 (1974).
 - [9] J. Polchinski, Scale and conformal invariance in quantum field theory, *Nuclear Physics B* **303**, 226 (1988).
 - [10] Y. Nakayama, Scale invariance vs conformal invariance, *Physics Reports* **569**, 1 (2015), scale invariance vs conformal invariance.
 - [11] M. Levin and C. P. Nave, Tensor renormalization group approach to two-dimensional classical lattice models, *Phys. Rev. Lett.* **99**, 120601 (2007).
 - [12] 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).
 - [13] H. H. Zhao, Z. Y. Xie, Q. N. Chen, Z. C. Wei, J. W. Cai, and T. Xiang, Renormalization of tensor-network states, *Phys. Rev. B* **81**, 174411 (2010).
 - [14] 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).
 - [15] D. Adachi, T. Okubo, and S. Todo, Anisotropic tensor renormalization group, *Phys. Rev. B* **102**, 054432 (2020).
 - [16] D. Kadoh and K. Nakayama, Renormalization group on a triad network (2019), [arXiv:1912.02414 \[hep-lat\]](https://arxiv.org/abs/1912.02414).
 - [17] S. Morita and N. Kawashima, Global optimization of tensor renormalization group using the corner transfer matrix (2020), [arXiv:2009.01997 \[cond-mat.stat-mech\]](https://arxiv.org/abs/2009.01997).
 - [18] M. Hinczewski and A. N. Berker, High-precision thermodynamic and critical properties from tensor renormalization-group flows, *Phys. Rev. E* **77**, 011104 (2008).
 - [19] K.-I. Aoki, T. Kobayashi, and H. Tomita, Domain wall renormalization group analysis of two-dimensional ising model, *International Journal of Modern Physics B* **23**, 3739 (2009), <https://doi.org/10.1142/S0217979209053357>.
 - [20] Y. Meurice, Accurate exponents from approximate tensor renormalizations, *Phys. Rev. B* **87**, 064422 (2013).
 - [21] E. Efrati, Z. Wang, A. Kolan, and L. P. Kadanoff, Real-space renormalization in statistical mechanics, *Rev. Mod. Phys.* **86**, 647 (2014).
 - [22] Z.-C. Gu and X.-G. Wen, Tensor-entanglement-filtering renormalization approach and symmetry-protected topological order, *Phys. Rev. B* **80**, 155131 (2009).
 - [23] M. Levin, Real space renormalization group and the emergence of topological order (2007), talk at the IPAM workshop "Topological Quantum Computing".
 - [24] S. Yang, Z.-C. Gu, and X.-G. Wen, Loop optimization for tensor network renormalization, *Phys. Rev. Lett.* **118**, 110504 (2017).
 - [25] G. Evenbly and G. Vidal, Tensor network renormalization, *Phys. Rev. Lett.* **115**, 180405 (2015).
 - [26] G. Evenbly, Algorithms for tensor network renormalization, *Phys. Rev. B* **95**, 045117 (2017).
 - [27] M. Bal, M. Mariën, J. Haegeman, and F. Verstraete, Renormalization group flows of hamiltonians using tensor networks, *Phys. Rev. Lett.* **118**, 250602 (2017).
 - [28] J. L. Cardy, Operator content of two-dimensional conformally invariant theories, *Nuclear Physics B* **270**, 186 (1986).
 - [29] G. Evenbly and G. Vidal, Local scale transformations on the lattice with tensor network renormalization, *Phys. Rev. Lett.* **116**, 040401 (2016).
 - [30] M. Hauru, C. Delcamp, and S. Mizera, Renormalization of tensor networks using graph-independent local truncations, *Phys. Rev. B* **97**, 045111 (2018).
 - [31] J. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge Lecture Notes in Physics (Cambridge University Press, 1996) pp. 28–60.
 - [32] H. Ueda, K. Okunishi, and T. Nishino, Doubling of entanglement spectrum in tensor renormalization group, *Phys.*

- Rev. B **89**, 075116 (2014).
- [33] K. Harada, Entanglement branching operator, *Phys. Rev. B* **97**, 045124 (2018).
 - [34] G. Evenbly, Gauge fixing, canonical forms, and optimal truncations in tensor networks with closed loops, *Phys. Rev. B* **98**, 085155 (2018).
 - [35] L. Ying, Tensor network skeletonization, *Multi-scale Modeling & Simulation* **15**, 1423 (2017), <https://doi.org/10.1137/16M1082676>.
 - [36] H.-Y. Lee and N. Kawashima, Tensor-ring decomposition with index-splitting, *Journal of the Physical Society of Japan* **89**, 054003 (2020), <https://doi.org/10.7566/JPSJ.89.054003>.
 - [37] C. Delcamp and A. Tilloy, Computing the renormalization group flow of two-dimensional ϕ^4 theory with tensor networks, *Phys. Rev. Research* **2**, 033278 (2020).
 - [38] S. Singh, R. N. C. Pfeifer, and G. Vidal, Tensor network decompositions in the presence of a global symmetry, *Phys. Rev. A* **82**, 050301 (2010).
 - [39] 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).
 - [40] S. Singh and G. Vidal, Tensor network states and algorithms in the presence of a global su(2) symmetry, *Phys. Rev. B* **86**, 195114 (2012).
 - [41] This is because the pieces of low-rank matrices and the isometric tensors in the RG equation of the GILT-HOTRG in Eq. (37) will inherit the reflection symmetry of the input tensor A .
 - [42] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshine, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019) pp. 8024–8035.
 - [43] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, *JAX: composable transformations of Python+NumPy programs* (2018).
 - [44] M. Kardar, *Statistical Physics of Fields* (Cambridge University Press, 2007).
 - [45] In principle, we can choose $\chi = 30$ here. In practice, however, our calculations show that the problem of local correlations in the HOTRG becomes worse at larger bond dimensions (see Ref. [18] for a similar observation in the context of the TRG), and that $\chi = 12$ is enough to demonstrate this problem.
 - [46] P. Di Francesco, P. Mathieu, and D. Sénéchal, *Minimal models i*, in *Conformal Field Theory* (Springer New York, New York, NY, 1997).