Renormalization group on a triad network

Daisuke Kadoh[1, *] and Katsumasa Nakayama[2, †]

[1]*Physics Division, National Center for Theoretical Sciences,*
National Tsing-Hua University, Hsinchu, 30013, Taiwan
Research and Educational Center for Natural Sciences, Keio University, Yokohama 223-8521, Japan
[2]*NIC, DESY Zeuthen, Platanenallee 6, 15738 Zeuthen, Germany*
(Dated: December 6, 2019)

We propose a new renormalization scheme of tensor networks made only of third order tensors. The isometry used for coarse-graining the network can be prepared at an $O(D^6)$ computational cost in any d dimension ($d \geq 2$), where D is the truncated bond dimension of tensors. Although it is reduced to $O(D^5)$ if a randomized singular value decomposition is employed, the total cost is $O(D^{d+3})$ because the contraction part for creating a renormalized tensor with isometries has D^{d+3} multiplications. We test our method in three dimensional Ising model and find that the numerical results are obtained for large Ds with reasonable errors.

INTRODUCTION

The tensor network is a promising approach to investigate statistical systems to which the Monte Carlo method is not easily accessible. Since this approach is free from the sign problem, it is expected to be an essential tool to study finite density QCD, the theta vacuum, chiral gauge theories and supersymmetric models, and the real-time dynamics of field theories. Since the tensor renormalization group (TRG) was proposed by Levin and Nave [1], it has been improved [2–7], and the TRG and some related methods achieve success in studying two-dimensional lattice field theories [8–27]. However, since it is designed for two-dimensional networks, a new scheme with much less computational cost is needed in order to study theories on three and four dimensions.

The higher order TRG (HOTRG) [28] is a typical example of renormalization schemes in higher dimensions. An isometry is given by the higher order singular value decomposition (HOSVD), and a square lattice network is coarse-grained by taking the contraction between two tensors with the isometries. This method is applicable to any d dimension for $d \geq 2$, and the cost scales as $O(D^{4d-1})$ where D is the truncated bond dimension of tensors at renormalization steps. Recently, another scheme named as an anisotropic TRG (ATRG) was proposed in [6]. Although the cost is reduced to $O(D^{2d+1})$ with a randomized singular value decomposition (RSVD) (See [29, 30]) or other truncation method, it has larger errors than the HOTRG for fixed D. So further studies for making an algorithm with small costs and higher accuracy are needed.

In this paper, we propose a new tensor renormalization scheme by defining it on a tensor network made only of third order tensors. That network and renormalization groups on it, which are referred to as a triad network and Triad RGs in this paper, respectively, are not uniquely determined. We give an example of Triad RGs improving a HOTRG-type renormalization on a triad network. The computational cost is drastically reduced since building blocks of our method are third order tensors. We find that the order of cost for making an isometry does not depend on the dimensionality, but on $O(D^6)$ in any d dimension ($d \geq 2$). This is reduced to $O(D^5)$ with the RSVD. The main cost comes from the contraction of making a renormalized tensor with isometries, which scales as $O(D^{d+3})$ with the RSVD. Then a naive memory usage is proportional to $O(D^{d+2})$ if intermediate tensors of order $d + 2$ are stored on a computer. We test our method in three-dimensional Ising model and find that numerical results are obtained for larger Ds with reasonable errors.

This paper is organized as follows. We firstly present our algorithm in three dimensions. Then we test it in three-dimensional Ising model and compare results to those obtained from the HOTRG and the ATRG. The RSVD used in our method, a review of HOTRG with HOSVD and an extension to d dimension are presented in appendices.

ALGORITHM

We begin with presenting our algorithm in three dimensions starting from a square lattice network made of a sixth order tensor $T_{ijklmn} \in \mathbb{C}$, where all indices $i, j, \cdots$ run from 1 to N. An extension to any dimension is given in the appendix. Without loss of generality, T may be expressed as a canonical polyadic decomposition (CPD):

$$T_{ijklmn} = \sum_{a=1}^{r} W_{ai}^{(1)} W_{aj}^{(2)} W_{ak}^{(3)} W_{al}^{(4)} W_{am}^{(5)} W_{an}^{(6)} \quad (1)$$

where r takes a minimum value in the canonical form, which is called a tensor rank of T. A tensor is derived in

this form for lattice models with nearest neighbor interactions in general.

The tensor network of T is defined on a three dimensional squared lattice $\Gamma = \{(n_1, n_2, n_3)|n_i \in \mathbb{Z}\}$:

$$Z = \text{Tr} \prod_{n \in \Gamma} T_{x_n x'_n y_n y'_n z_n z'_n} \qquad (2)$$

where x_n, x'_n, y_n, y'_n, z_n, z'_n are indices defined on links stemmed from the site n. These indices satisfy $x'_n = x_{n+\hat{1}}$, $y'_n = y_{n+\hat{2}}$, $z'_n = z_{n+\hat{3}}$ where $\hat{\mu}$ stands for the unit vector of μ direction. Tr denotes the summation of all indices. All tensors live on sites and any link shared by two tensors is contracted.

The computational cost of tensor renormalization for d dimensional square lattice network is high when d increases in general because the contraction between two $2d$th order tensors takes a high cost. In order to reduce the cost, we formulate a renormalization group on a network made only of third order tensors.
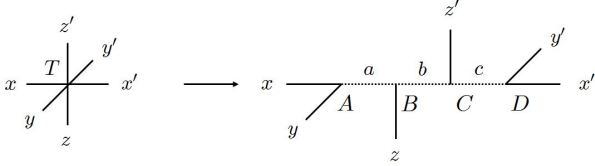

FIG. 1.   Triad representation of T.

It is possible to express T as a product of four 3rd order tensors A, B, C, D:

$$T_{xx'yy'zz'} = \sum_{a,b,c=1}^{r} A_{xya} B_{azb} C_{bz'c} D_{cy'x'}, \qquad (3)$$

where

$$A_{xya} \equiv W^{(1)}_{ax} W^{(2)}_{ay}, \qquad (4)$$
$$B_{azb} \equiv \delta_{ab} W^{(3)}_{az}, \qquad (5)$$
$$C_{bzc} \equiv \delta_{bc} W^{(4)}_{bz}, \qquad (6)$$
$$D_{cyx} \equiv W^{(5)}_{cy} W^{(6)}_{cx}. \qquad (7)$$

Eq.(3) is referred to as a *triad* representation in this paper. Note that it is not unique. Fig.1 shows a triad representation of T shown in eq.(3). Solid lines denote the external indices $x, y, z, \dots$ and dotted lines denote the internal indices a, b, c. The tensor network Z may be regarded as a network made only of third order tensors A, B, C, D by replacing T with a triad unit (3) as shown in Fig.2. We also refer it to as a triad network.

From now on, we assume that all internal and external indices, $a, b, c, \dots$ and $x, x', \dots$, run from 1 to D for $D \geq r, N$. In renormalization steps, D is a truncated bond dimension of triads A, B, C, D. Although a different size may be taken for internal indices to improve the accuracy
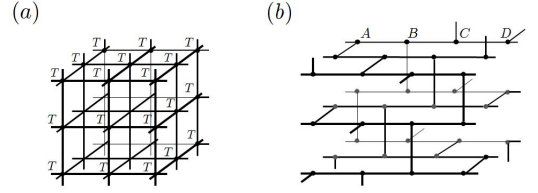

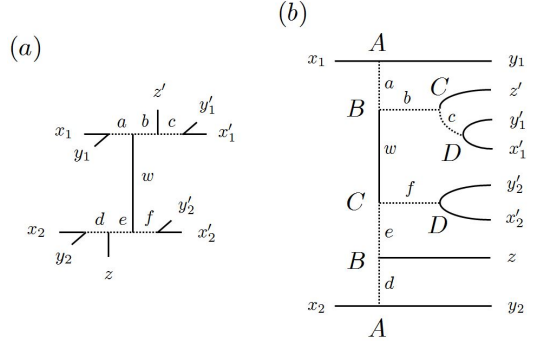FIG. 2.   Square lattice network (a) and triad network (b).

FIG. 3. M in two triads. (b) is another representation of (a).

of results, a common D is employed and $\sum_{i=1}^{D}$ is denoted as $\sum_i$ in the following for simplicity.

We now consider a coarse-graining procedure of a Triad RG. A renormalization along the z direction is carried out by combining two tensors as well as HOTRG:

$$M_{XX'YY'zz'} = \sum_{w} T_{x_1 x'_1 y_1 y'_1 wz} T_{x_2 x'_2 y_2 y'_2 z'w}, \qquad (8)$$

where $X = x_1 \otimes x_2$ and $Y = y_1 \otimes y_2$. Fig.3 shows M in the triad representation. The HOSVD gives

$$M_{XX'YY'zz'} = \sum_{I,J,K,L=1}^{D^2} \sum_{m,n=1}^{D} S_{IJKLmn}$$
$$\times U^L_{XI} U^R_{X'J} V^L_{YK} V^R_{Y'L} W^U_{mz} W^D_{nz'} \qquad (9)$$

where S is the core tensor of M and U, V, W are unitary matrices. See appendix for the detailed definition of HOSVD. U and V are used for coarse-graining the network but W are not needed.

To evaluate U^L associated with the x-axis, we interpret M as a matrix $M'_{X, X'YY'zz'} (\equiv M_{XX'YY'zz'})$ with the row index X and the column index (X', Y, Y', z, z'). Then U^L is a unitary matrix that diagonalizes a hermitian matrix

$$K \equiv M' M'^{\dagger} \qquad (10)$$

as

$$K_{x_1 x_2 x_3 x_4} = \sum_{I=1}^{D^2} U^L_{x_1 x_2 I} \lambda^L_I U^{L\,\dagger}_{Ix_3 x_4}, \qquad (11)$$

where λ_I are eigenvalues sorted in the descending order, $\lambda_1^L \geq \lambda_2^L \geq \cdots \geq \lambda_{D^2}^L \geq 0$. The cost of making K, which is $\mathcal{O}(D^{2d+2})$ in the HOTRG as reviewed in the appendix, is drastically reduced on the triad network.
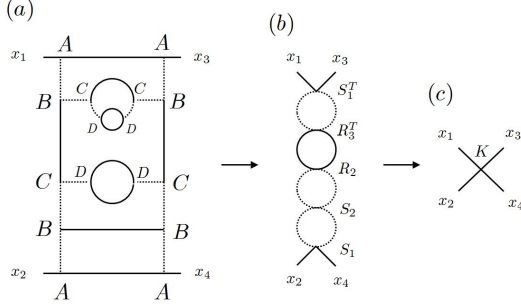

FIG. 4. How to make K in the triad representation.

Fig.4 shows a methodology of making K using the triad representation. In Fig.4 (a), K is represented as a product of M (Fig.3 (b)) and $M^\dagger$ which is a mirror image of M. Fig.4 (b) is obtained by contracting inner lines connecting M and $M^\dagger$ from Fig.4 (a). When K is expressed as a matrix $K'_{x_2 x_4, x_1 x_3} (\equiv K_{x_1 x_2 x_3 x_4})$ with row index x_2, x_4 and column index x_1, x_3, we have

$$K' = S_1 \cdot S_2 \cdot R_2 \cdot R_3^T \cdot S_1^T, \tag{12}$$

with hermitian matrices,

$$(S_1)_{xx', dd'} = \sum_y A_{xyd} A^*_{x'yd'},$$

$$(S_2)_{dd', ee'} = \sum_z B_{dze} B^*_{d'ze'}, \tag{13}$$

$$(R_2)_{ee', ww'} = \sum_{f, f'} C_{ewf} C^*_{e'w'f'} \cdot \sum_{x,y} D_{fyx} D^*_{f'yx},$$

$$(R_3)_{aa', ww'} = \sum_{b, b'} B_{awb} B^*_{a'w'b'} \cdot \sum_z (R_2)_{bb', zz}.$$

Fig.4 (b) can be obtained at an $\mathcal{O}(D^5)$ cost without any approximation since S_i and R_i are computed at this cost. In Fig.4 (c), K is obtained computing matrix products of eq.(12), which takes an $\mathcal{O}(D^6)$ cost.

We thus obtain U^L at an $O(D^6)$ cost by diagonalizing the obtained K. The cost of making K and U^L can be reduced to $O(D^5)$ applying the RSVD to eq.(12) (Fig.4 (b)), as shown in the appendix. The order of cost of making U does not depend on the dimensionality in the triad representation.

The other unitary matrix U^R can also be prepared in the similar manner. We choose each one of U^L and U^R to improve the accuracy of results by comparing the remaining eigenvalues [28],

$$\epsilon_Q = \sum_{i > D} \lambda_i^Q, \quad \text{for} \quad Q = L, R \tag{14}$$

where λ^R are eigenvalues of K' with a different matrix representation $M_{X', XYY'zz'} (\equiv M_{XX'YY'zz'})$. $U = U^L$ for $\epsilon_L < \epsilon_R$ and $U = U^R$ for the others.

A renormalized tensor is defined by

$$T^R_{zz'xx'yy'} \equiv \sum_{X,Y,X',Y'} U^\dagger_{xX} V^\dagger_{yY} M_{XX'YY'zz'} U_{X'x'} V_{Y'y'} \tag{15}$$

where U is U^L (or U^R) and V is V^L (or V^R) which are chosen from a comparison of ϵ_L and ϵ_R. The combined indices x, x', y, y' run from 1 to D by truncating D^2 eigenvalues (functions) to D largest ones.

In the triad representation, we have

$$T^R_{zz'xx'yy'} = \sum_{a,b,e,f} \mathcal{D}_{zxyae} \mathcal{M}_{aebf} \mathcal{U}_{bfy'z'x'}, \tag{16}$$

where

$$\mathcal{U}_{abyzx} = \sum_{c,p,q,p',q'} C_{azc} D_{cqp} U_{pp'x} D_{bq'p'} V_{qq'y}, \tag{17}$$

$$\mathcal{M}_{abcd} = \sum_w B_{awc} C_{bwd}, \tag{18}$$

$$\mathcal{D}_{zxyab} = \sum_{d,p,q,p',q'} B_{dzb} U^*_{pp'x} A_{pqa} V^*_{qq'y} A_{p'q'd}. \tag{19}$$

Note that the cost for making $\mathcal{M}$ is $O(D^5)$ while that for $\mathcal{U}$ and $\mathcal{D}$ is $O(D^6)$, which can be found by taking five contractions in the order of $(p, q') \to (p', q) \to c$ (or d) keeping intermediate fourth order tensors.

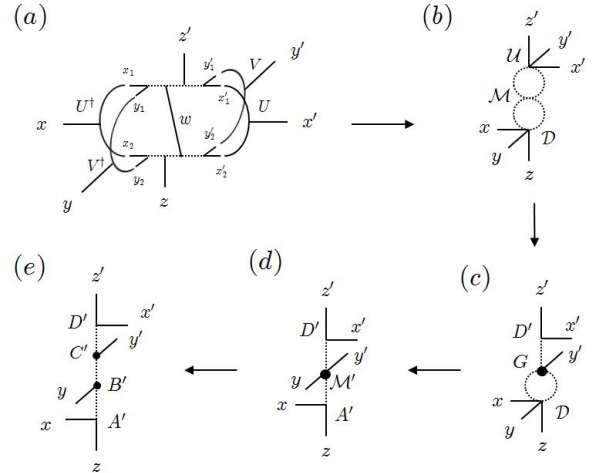Fig.5 shows how to create T^R. Fig.5 (a) and (b) show eq.(15) and eq.(16), respectively. Although a naive con-

FIG. 5. Contraction of two triad units.

traction between $\mathcal{M}$ and $\mathcal{U}$ in eq.(16) takes $O(D^7)$, as presented in appendix, the RSVD provides an approximated decomposition at an $O(D^6)$ cost:

$$\sum_{c,d} \mathcal{M}_{abcd} \mathcal{U}_{cdyxz} \approx \sum_g G_{abyg} D'_{gxz} \tag{20}$$

where D'_{ijk} is a unitary matrix with the row i and the column j, k in terms of the full SVD and G is a 4th order tensor in which the singular values are included. In Fig.5 (c), the black circle implies that G contains the singular values.

Then applying the RSVD to remaining contractions in eq.(16) in the similar manner, we obtain another approximated decomposition at a cost of $O(D^6)$:

$$\sum_{a,b} \mathcal{D}_{zxyab} G_{aby'g} \approx \sum_e A'_{zxe} \mathcal{M}'_{eyy'g} \qquad (21)$$

where A'_{ijk} is a unitary matrix with the row i, j and the column k in terms of the full SVD and $\mathcal{M}'$ is a fourth order tensor in which the singular values are absorbed. Fig.5 (d) represents this decomposition. The SVD finally provides

$$\mathcal{M}'_{eyy'g} \approx \sum_f B'_{eyf} C'_{fy'g} \qquad (22)$$

at an $O(D^6)$ cost, as shown in Fig.5 (e). The RSVD reduces this cost to $O(D^5)$.

Plugging eqs.(20)-(22) into eq.(16), we thus find that T^R is approximately given by a renormalized triad unit:

$$T^R_{zz'xx'yy'} \approx \sum_{a,b,c} A'_{zxa} B'_{ayb} C'_{by'c} D'_{cx'z'}. \qquad (23)$$

An extra axis rotation is not needed since x, y, z of eq.(3) is replaced by z, x, y in (23). In three dimensions, the computational cost of the Triad RG method is $O(D^6)$, and a memory usage is naively $O(D^5)$, which comes from the fifth order tensors $\mathcal{U}, \mathcal{D}$. Repeating this procedure again and again, triad networks can be coarse-grained at this cost keeping the triad representation.

An extension to any d dimension is presented in the appendix. Decompositions such as eq.(20) take $O(D^{d+3})$ as $\mathcal{U}$ and $\mathcal{D}$ are tensors of order $d+2$, which is the main cost of our Triad RG method. The memory usage is naively $O(D^{d+2})$ if $\mathcal{U}$ and $\mathcal{D}$ are stored in the memory.

We make a remark on improvements of the Triad RG method in the rest of this section. As shown in Fig.5 (c)-(e), the singular values denoted as the black circles are inherited to the next decomposition, as done in the ATRG. The best accuracy is achieved with this treatment as long as we tried. Although we have presented a $O(D^{d+3})$ procedure for the contraction part, we can easily create another procedure with an $O(D^5)$ cost neglecting the accuracy of results. For instance, decomposing $\mathcal{M}$ in eq.(16) directly or swapping indices of tensors with the RSVD can reduce a naive cost. We found that these additional SVDs do not work well for the three dimensional Ising model. This could be because such SVDs are optimized only for local tensors and the accuracy for the whole two triad units is lost. However, since there are many variants in the contraction part of our method, further studies are needed to improve the Triad RG method.

A NUMERICAL TEST

We test the Triad RG method in three dimensional Ising model on a periodic lattice with the volume $V = (32768)^3 = (2^{15})^3$ at the critical temperature $T_c = 4.5115$ [6, 28, 31]. The Hamiltonian is given by

$$H = -\sum_{\langle i,j \rangle} \sigma_i \sigma_j, \qquad (24)$$

where $\langle i, j \rangle$ denotes all possible nearest neighbor pairs of lattice sites. The partition function $Z = \text{Tr}(e^{-\beta H})$ with the inverse temperature $\beta = 1/T$ can be expressed as a triad network eq.(3) with $N = r = 2$ and $W^{(\mu)} \equiv W$ given by

$$W = \begin{pmatrix} \sqrt{\cosh(\beta)} & \sqrt{\sinh(\beta)} \\ \sqrt{\cosh(\beta)} & -\sqrt{\sinh(\beta)} \end{pmatrix}. \qquad (25)$$

Note that initial 3rd order tensors A, B, C, D are real and satisfy $A_{xya} = D_{ayx}$ and $B_{azb} = C_{azb}$.

The free energy is evaluated by $F = -\frac{1}{\beta} \log Z$. We take $C = 4$ for the oversampling parameter of RSVD (See the appendix for details of RSVD). The numerical computation is carried out with 2.7 GHz Intel Core i7 and a library *Eigen* for matrix decompositions, and each computation ends in a few hours.

Numerical results are compared to those obtained from the HOTRG and the ATRG. The first version of ATRG is implemented by the RSVD with twice a larger bond dimension $2D$ only for the swapping step. In the Triad RG, an $O(D^5)$ isometry is prepared with the RSVD. The computational cost of HOTRG, ATRG, and Triad RG methods are theoretically $O(D^{11})$, $O(D^7)$, and $O(D^6)$ in three dimensions, respectively.
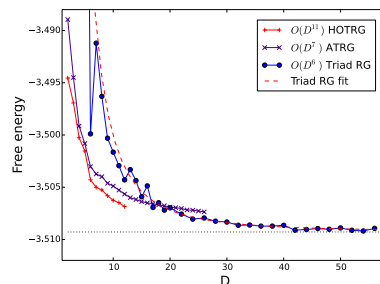

FIG. 6. D-dependence of free energy in 3d Ising model at T_c.

Fig.6 shows the D-dependence of free energy. The accessible D is different among the three methods. Three results decrease toward values around $F = -3.51$ as D increases. The Triad RG has well-controlled errors for larger Ds and shows good convergence.

We extrapolate our result of the Triad RG to the large D limit using a fit function $a + bD^{-c}$ with fitting variables a, b, c. To obtain a precise fit result, we compute the

free energy n times with different random numbers of the RSVD ($n = 100$ for $D \leq 24$ and $n = 4$ for $D > 24$), and use an average value of n trials with error estimated from the standard deviation for the fit. The result for $10 \leq D \leq 56$ is $a = -3.5093(2)$, which is shown as a dotted line in Fig.6. We confirm that this result is stable by changing fit range to $20 \leq D \leq 56$.
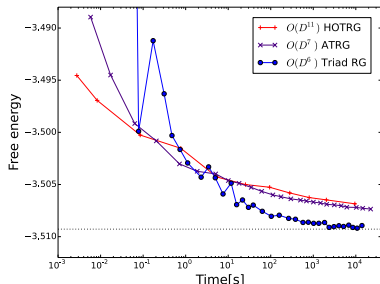

FIG. 7. Free energy as a function of computational time.

Fig.7 shows the free energy against the computational time needed to compute the free energy once. This figure implies that the Triad RG converges faster than the other methods for the same computational time.
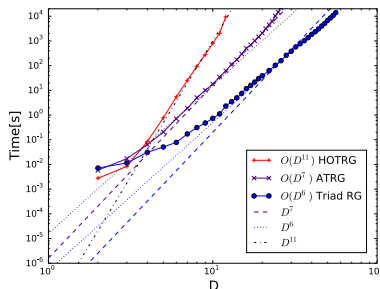

FIG. 8. Computational time against D.

In Fig.8, the computational time is shown as a function of D. Since the theoretical D-dependence is properly reproduced at a practical level, one can consider that the Triad RGs will open a door to studying a wide class of higher dimensional field theory with tensor networks.

ACKNOWLEDGMENTS

* kadoh@keio.jp
† katsumasa.nakayama@desy.de

[1] M. Levin and C. P. Nave, Phys. Rev. Lett. **99**, 120601 (2007), arXiv:cond-mat/0611687 [cond-mat.stat-mech].
[2] Z.-C. Gu and X.-G. Wen, Phys. Rev. B **80**, 155131 (2009), arXiv:0903.1069 [cond-mat.str-el].
[3] G. Evenbly and G. Vidal, Phys. Rev. Lett. **115**, 180405 (2015).
[4] S. Morita, R. Igarashi, H.-H. Zhao, and N. Kawashima, Phys. Rev. **E97**, 033310 (2018), arXiv:1712.01458.
[5] Y. Nakamura, H. Oba, and S. Takeda, Phys. Rev. **B99**, 155101 (2019), arXiv:1809.08030 [cond-mat.stat-mech].
[6] D. Adachi, T. Okubo, and S. Todo, (2019), arXiv:1906.02007 [cond-mat.stat-mech].
[7] W. Lan and G. Evenbly, (2019), arXiv:1906.09283 [quant-ph].
[8] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004), arXiv:cond-mat/0404706 [cond-mat.str-el].
[9] F. Verstraete and J. I. Cirac, (2004), arXiv:cond-mat/0407066 [cond-mat].
[10] Y. Shimizu and Y. Kuramashi, Phys. Rev. **D90**, 014508 (2014), arXiv:1403.0642 [hep-lat].
[11] Y. Shimizu and Y. Kuramashi, Phys. Rev. **D90**, 074503 (2014), arXiv:1408.0897 [hep-lat].
[12] S. Takeda and Y. Yoshimura, PTEP **2015**, 043B01 (2015), arXiv:1412.7855 [hep-lat].
[13] T. Pichler, M. Dalmonte, E. Rico, P. Zoller, and S. Montangero, Phys. Rev. **X6**, 011023 (2016), arXiv:1505.04440 [cond-mat.quant-gas].
[14] H. Kawauchi and S. Takeda, Phys. Rev. **D93**, 114503 (2016), arXiv:1603.09455 [hep-lat].
[15] M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and S. Kühn, Phys. Rev. Lett. **118**, 071601 (2017), arXiv:1611.00705 [hep-lat].
[16] R. Sakai, S. Takeda, and Y. Yoshimura, PTEP **2017**, 063B07 (2017), arXiv:1705.07764 [hep-lat].
[17] M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and S. Kühn, Phys. Rev. **X7**, 041046 (2017), arXiv:1707.06434 [hep-lat].
[18] Y. Shimizu and Y. Kuramashi, Phys. Rev. **D97**, 034502 (2018), arXiv:1712.07808 [hep-lat].
[19] D. Kadoh, Y. Kuramashi, Y. Nakamura, R. Sakai, S. Takeda, and Y. Yoshimura, JHEP **03**, 141 (2018), arXiv:1801.04183 [hep-lat].
[20] Y. Kuramashi and Y. Yoshimura, JHEP **08**, 023 (2019), arXiv:1808.08025 [hep-lat].
[21] D. Kadoh, Y. Kuramashi, Y. Nakamura, R. Sakai, S. Takeda, and Y. Yoshimura, JHEP **05**, 184 (2019), arXiv:1811.12376 [hep-lat].
[22] F. Bruckmann, K. Jansen, and S. Kühn, Phys. Rev. **D99**, 074501 (2019), arXiv:1812.00944 [hep-lat].
[23] Y. Kuramashi and Y. Yoshimura, (2019), arXiv:1911.06480 [hep-lat].
[24] M. C. Banuls, K. Cichy, Y.-J. Kao, C. J. D. Lin, Y.-P. Lin, and D. T. L. Tan, Phys. Rev. **D100**, 094504 (2019),

arXiv:1908.04536 [hep-lat].

[25] Y. Meurice, Phys. Rev. **D100**, 014506 (2019), arXiv:1903.01918 [hep-lat].

[26] A. Bazavov, S. Catterall, R. G. Jha, and J. Unmuth-Yockey, Phys. Rev. **D99**, 114507 (2019), arXiv:1901.11443 [hep-lat].

[27] M. C. Banuls et al., (2019), arXiv:1911.00003 [quant-ph].

[28] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, Physical Review **B86**, 045139 (2012), arXiv:1201.1144.

[29] N. Halko, P.-G. Martinsson, and J. A. Tropp, SIAM Rev **Vol. 53**, pp. 217 (2009), arXiv:0909.4061.

[30] N. B. Erichson, S. Voronin, S. L. Brunton, and J. N. Kutz, (2016), arXiv:1608.02148 [stat.CO].

[31] Y. Deng and H. W. J. Blöte, Phys. Rev. E **68**, 036125 (2003).

Randomized SVD and contraction of tensors

The cost of decomposing (or contracting) tensors is reduced by a randomized SVD (RSVD) [29, 30]. In this appendix, we present technical details of RSVD and how it is used for making projectors U in eq.(11) and for three decompositions eqs.(20)-(22) in the Triad RG method.

Let A be an $m \times n$ matrix with complex entries. The SVD provides a decomposition of A:

$$A = U\Sigma V^\dagger, \tag{26}$$

where U and V are $m \times m$ and $n \times n$ unitary matrices, respectively, and Σ is a non-negative diagonal matrix containing singular values in the descending order, $\Sigma_{ij} = \sigma_i \delta_{ij}$ ($\sigma_1 \geq \sigma_2 \geq \ldots$). The cost of full SVD is $O(mnl)$ for $l = \min\{m, n\}$. However, we often need part of SVD such as k largest singular values for $k \ll m, n$. Then the cost is reduced from $O(mnl)$ to $O(mnk)$ with the RSVD.

For a given $m \times n$ matrix A, k largest singular values and singular vectors are approximately given by the following procedure (RSVD):

1. Generate an $n \times p$ Gaussian random matrix Ω.

2. Construct an $m \times p$ matrix $Y_A = A\Omega$.

3. Compute QR factorization of Y_A as $Y_A = Q_A R_A$.

4. Construct a small $p \times n$ matrix $B_A \equiv Q_A^\dagger A$.

5. Compute an SVD of B_A as $B_A = \tilde{U}\tilde{\Sigma}\tilde{V}^\dagger$.

$p \, (\geq k)$ is a tunable parameter. Note that R_A is an $m \times p$ upper triangular matrix and Q_A is an $m \times m$ unitary matrix (the first $m \times p$ part of Q_A is needed in the fourth step). We thus obtain a low rank approximation of A:

$$A \approx U'\tilde{\Sigma}\tilde{V}^\dagger \tag{27}$$

where $U' \equiv Q_A \tilde{U}$. k largest singular values with singular vectors of A are approximately given by taking k part of eq.(27). In this sense, p provides an oversampling of the index k to improve the accuracy of results.

The costs of the second and fourth steps of RSVD are $O(mnp)$, which are higher than the cost of QR decomposition at the third step $O(p^2 m)$ and the cost of the fifth step $O(p^2 n)$. As long as we take

$$k \leq p \ll m, n \tag{28}$$

the main cost of RSVD, which is $O(pmn)$, is lower than the cost of full SVD. See [29, 30] for a formal discussion of error of RSVD.

The RSVD is also useful to evaluate a matrix product and to give its low-rank approximation. Let E and F be $m \times \ell$ and $\ell \times n$ matrices, respectively. Although the cost of evaluating a matrix product $A = EF$ directly from E and F is $O(mn\ell)$, a low-rank approximation of A can be given at a lower cost using the same procedure of RSVD. At the second step, we can construct Y_A multiplying Ω by E and F in turn instead of A. These multiplications take $O((m+n)p\ell)$, which is smaller than $O(mn\ell)$ under eq.(28). Similarly, the fourth step takes the same cost. Once Q_A and B_A are obtained, without the fifth step, one can give a low rank approximation of $A = EF$ as

$$A_{ij} \approx \sum_{a=1}^{p} (Q_A)_{ia}(B_A)_{aj}. \tag{29}$$

Note that the contraction of rhs takes an $O(mnp)$ cost. Although the matrix product EF takes $O(mn\ell)$, we thus find that it is approximated by rhs of eq.(29) at a lower cost $O((mn + \ell n + \ell m)p)$ for

$$k \leq p \ll \ell, m, n. \tag{30}$$

Moreover, the fifth step provides the SVD of EF at an $O((mn + \ell n + \ell m)p)$ cost.

These tricks are used for tensors in the triad RG method. For the truncated bond dimension D, we basically take $m, n, \ell \geq D^2$ and $k = D$ and $p = Ck(= CD)$ where a fixed integer C is referred to as an oversampling parameter in this paper. Since these parameters satisfy eqs.(28) and (30) for $D \gg 1$, the RSVD effectively works. To estimate the error, we compute a physical value N times with different random numbers. The result with an error is evaluated from the central value and the standard deviation from N trials.

Let us consider a contraction between 4th order tensor $\mathcal{M}$ and 5th order tensor $\mathcal{U}$:

$$\mathcal{A}_{zx,aby} \equiv \sum_{c,d=1}^{D} \mathcal{M}_{abcd}\mathcal{U}_{cdyxz}. \tag{31}$$

Note that $n = D^3, m, \ell = D^2, k = D$ and $p = CD$ in this case and eq.(30) is satisfied. Although the cost of evaluating $\mathcal{A}$ from $\mathcal{M}$ and $\mathcal{U}$ is $O(D^7)$, a low rank approximation of $\mathcal{A}$ such as eq.(29) can be given at a cost of $O(D^6)$. Once $\mathcal{A}$ is approximately given, applying the RSVD to it like eq.(27), we obtain

$$\mathcal{A}_{zx,aby} \approx \sum_{g=1}^{D} G_{abyg}D_{gzx} \tag{32}$$

with

$$G_{abyg} = \tilde{V}^{\dagger}_{g,aby}\tilde{\sigma}_g, \tag{33}$$

$$D_{gxz} = (Q_{\mathcal{A}}\tilde{U})_{zx,g}, \tag{34}$$

where $B_{\mathcal{A}} = \tilde{U}\tilde{\Sigma}\tilde{V}^{\dagger}$, at an $O(D^6)$ cost. The singular values are included in G. This proves eq.(20). Eqs.(21) and (22) are derived in a similar manner.

We apply the RSVD to the $D^2 \times D^2$ matrix K given in eq.(12) in order to obtain D singular vectors $U_{x_1x_2x}$ for $x_1, x_2, x = 1, 2 \cdots, D$. Note that $m = n = \ell = D^2$, $k = D$ and $p = CD$ in this case and eq.(28) is satisfied. We do not construct K directly since it takes $O(D^6)$. Instead, we can evaluate Y_K at an $O(D^5)$ cost multiplying Ω by S_i and R_i alternately where Ω is a $D^2 \times p$ Gaussian random matrix. The fourth step takes the same cost. Once Q_K and B_K are given, K is approximately given at an $O(D^5)$ cost like eq.(29). Applying the RSVD to the constructed matrix $K_{x_1x_2,x_3x_4}$ again, we obtain a low rank approximation of K at an $O(D^5)$ cost like eq.(27):

$$K_{x_1x_2x_3x_4} \approx \sum_{x=1}^{D} U^L_{x_1x_2i}\tilde{\lambda}^L_i\tilde{U}^{L\dagger}_{ix_3x_4}, \tag{35}$$

where $U^L \approx \tilde{U}^L$. This procedure can be easily extended to higher dimensions. We can obtain isometries at an $O(D^5)$ cost with the RSVD in any dimension.

HOTRG

The HOTRG method [28] with the HOSVD is reviewed in this appendix. We will find out that the cost of d dimensional HOTRG is proportional to D^{4d-1}, which comes from final contractions with isometries although the cost of making isometries scales as D^{2d+2}.

To introduce the HOSVD, let us first define the inner product for tensors:

$$\langle \mathcal{T}, \mathcal{T}' \rangle \equiv \sum_{i_1,i_2,\cdots,i_n} \mathcal{T}^*_{i_1i_2\cdots i_n}\mathcal{T}'_{i_1i_2\cdots i_n} \tag{36}$$

where $\mathcal{T}$ and $\mathcal{T}'$ are nth order tensors. The norm of $\mathcal{T}$ is defined as $||\mathcal{T}|| = \sqrt{\langle \mathcal{T}, \mathcal{T} \rangle}$. The HOSVD tells us that an nth order tensor $\mathcal{T}_{i_1i_2,\cdots i_n}$ $(i_k = 1, 2, \cdots, N_k)$ may be expressed as

$$\mathcal{T}_{i_1i_2\cdots i_n} = \sum_{j_1,j_2,\cdots,j_n} \mathcal{S}_{j_1j_2\cdots j_n}U^{(1)}_{i_1j_1}U^{(2)}_{i_2j_2}\cdots U^{(n)}_{i_nj_n} \tag{37}$$

Here $U^{(m)}$ is an $N_m \times N_m$ unitary matrix and $\mathcal{S}_{i_1i_2\cdots i_n}$ is a core tensor of T, which satisfies all-orthogonality:

$$\langle \mathcal{S}_{i_m=\alpha}, \mathcal{S}_{i_m=\beta} \rangle = 0, \quad \text{for} \quad \alpha \neq \beta, \tag{38}$$

and an ordering property:

$$||\mathcal{S}_{i_m=1}|| \geq ||\mathcal{S}_{i_m=2}|| \geq \cdots \geq ||\mathcal{S}_{i_m=N_m}|| \geq 0, \tag{39}$$
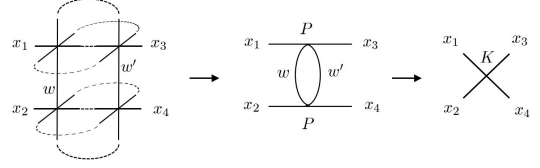

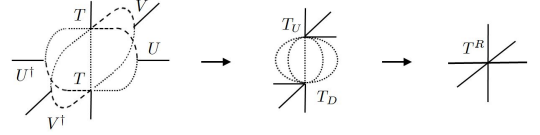FIG. 9. A method of evaluating K for a three dimensional square lattice network.

FIG. 10. Final contractions in the HOTRG.

where $\mathcal{S}_{i_m=\alpha}$ is a sub tensor of order $n-1$ with the m-th index i_m of $\mathcal{S}$ fixed to α.

The renormalization of the HOTRG is carried out for all axes alternately. M in eq.(8) is expressed as eq.(9) in terms of the HOSVD. We evaluate K defined by eq.(12) to give isometries. Fig.9 shows how to create K in three dimensions. We can make a 4th order tensor P by paying an $O(D^{2d+2})$ cost. The final step to make K from two Ps and the diagonalization of K do not need a high cost. We find that the cost of making U^L (or U^R) scales as $O(D^{2d+2})$.

A renormalized tensor is evaluated in a similar way to eq.(15). In d dimension, the number of isometries U is $d-1$.

Figure 10 shows a procedure of making T^R in three dimensions. T_U is a tensor made of upper T and U, V, while T_D is one made of lower T and $U^{\dagger}, V^{\dagger}$. The cost of making T_U (or T_D) is $O(D^{4d-2})$ because their order is $3d-1$ and inner $d-1$ links are contracted. The contraction between T_U and T_D takes $O(D^{4d-1})$ which is the dominant cost of the HOTRG method.

The costs for isometries and final contractions can be reduced with the RSVD although systematic errors could be larger.

Extension to higher dimensions

It is straightforward to extend our method to any d dimension for $d \geq 2$. We may give a d-dimensional triad unit as

$$T_{x_1x'_1x_2x'_2\cdots x_dx'_d} = A^{(1)}_{x_1x_2a_1}A^{(2)}_{a_1x_3a_2}\cdots A^{(d-1)}_{a_{d-2}x_da_{d-1}}$$
$$\times A^{(d)}_{a_{d-1}x'_da_d}\cdots A^{(2d-3)}_{a_{2d-4}x'_3a_{2d-3}}A^{(2d-2)}_{a_{2d-3}x'_2x'_1}. \tag{40}$$

Then the similar calculation as done in three dimensions tells us that $K^{(\mu)}$ in the μ direction ($\mu = 1, 2\cdots, d-1$),

which is defined as eq.(10), is given by

$$K^{(\mu)} = \tilde{S}_{\mu-1} S_\mu \cdots S_{d-1} R_{d-1}$$
$$\times (\tilde{S}_{\mu-1} S_\mu \cdots S_{d-2} R_d)^T, \tag{41}$$

where

$$(S_n)_{ij,kl} = \sum_m A^{(n)}_{imk} (A^{(n)}_{jml})^*, \tag{42}$$

and

$$(\tilde{S}_n)_{ab,cd} = \sum_{i,j,k} A^{(n)}_{iac} (A^{(n)}_{jbd})^* (\tilde{S}_{n-1})_{kk,ij}, \tag{43}$$

$$(R_n)_{ab,cd} = \sum_{i,j,k} A^{(2d-n-1)}_{aci} (A^{(2d-n-1)}_{bdj})^* (R_{n-1})_{ij,kk}, \tag{44}$$

with $(\tilde{S}_0)_{ab,cd} = (R_0)_{ab,cd} = \delta_{ac}\delta_{bd}$. We should note that $\tilde{S}_n$ and R_n are determined recursively at an $O(D^5)$ cost. The cost of evaluating $K^{(\mu)}$ from the matrix products in eq.(41) is $O(D^6)$ and the cost of diagonalizing $K^{(\mu)}$ is also $O(D^6)$. We thus find that the power of cost for making isometries does not depend on the dimensionality

but on $O(D^6)$ in any dimension. This cost is reduced to $O(D^5)$ with the RSVD as well as the three-dimensional case.

The renormalized tensor T^R for d dimension is given in the same manner as eq.(15) with isometries $U^{(\mu)}$. We can show that

$$T^R_{x_d x'_d x_1 x'_1 \ldots x_{d-1} x'_{d-1}} = \sum_{a,b,c,d} \mathcal{D}_{x_d x_1 \cdots x_{d-1} ab}$$
$$\times \mathcal{M}_{abcd} \mathcal{U}_{cd x'_2 \cdots x'_d x'_1}, \tag{45}$$

where $\mathcal{M}_{abcd} = \sum_w A^{(d-1)}_{awc} A^{(d)}_{bwd}$, and $d+2$th order tensor $\mathcal{U}$ is made of $A^{(d)}$, $A^{(d+1)}, \cdots, A^{(2d-2)}$ and $U^{(\mu)}$ such as eq.(17) while $\mathcal{D}$ is made of the other tensors. Since $\mathcal{U}$ and $\mathcal{D}$ are tensors of order $d+2$, the naive memory usage is proportional to $O(D^{d+2})$ if they are stored directly on the memory. It is easy to show that these tensors are created at $O(D^{d+3})$ costs by taking contractions in the appropriate order. Repeating procedures as shown in eqs.(20)-(22), one can obtain a triad representation of d-dimensional T^R at an $O(D^{d+3})$ cost.