4次元のATRGにおける高速化手法について

ON ACCELERATION METHODS FOR THE ATRG IN FOUR DIMENSIONS

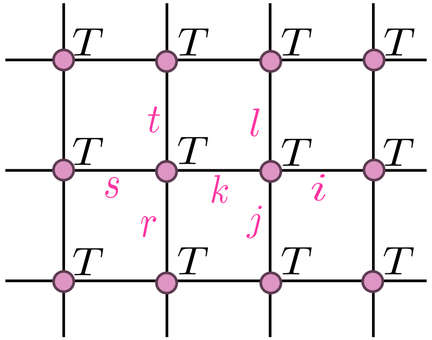
Yuto Sugimoto , Shoichi Sasaki (Nuclear Theory, Tohoku U.)

[M. Levin and C. P. Nave, (2007).] [Xie et al, (2012).]

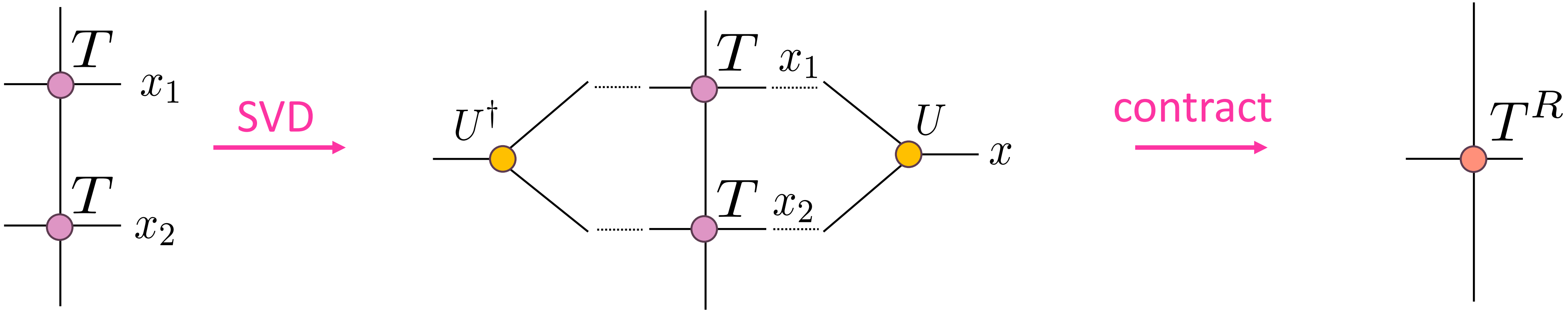
Tensor Renormalization Group = Numerical Real space Renormalization Group

→A candidate for overcoming sign problems in LQCD

$$Z = \sum_{i,j,k,l,m,\dots}^{\chi} T_{ijkl}T_{mnoj}T_{krst}T_{opqr}\cdots = \mathrm{tTr}(\otimes_{x=1}^{N} T) =$$



Approximate by Singular Value Decomposition (SVD)



TRG is applicable for sign problem region, but has large cost at higher dimensions

OTHER APPROACHES TO HIGHER DIMENSIONS

	HOTRG [Xie et al, (2012).]	ATRG [D. Adachi, T. Okubo, and S. Todo, (2020).]	Triad TRG [D. Kadoh and K. Nakayama, (2019).]	MDTRG Triad rep. [K. Nakayama, (2023).]
cost	$O(\chi^{4d-1})$	$O(\chi^{2d+1})$	$O(\chi^{d+3})$	$O(qr^3\chi^{d+3})$
Fundamental tensor	$O(\chi^{2d})$	$O(\chi^{d+1})$	$O(\chi^3)$	$O(\chi^{d+1})$
methods	• Exact Contraction	• Bond-swapping via RSVD • Exact Contraction	• Triad • Contraction via RSVD	• Decomposition of unit-cell tensor • Triad • Internal line oversampling • Contraction via RSVD

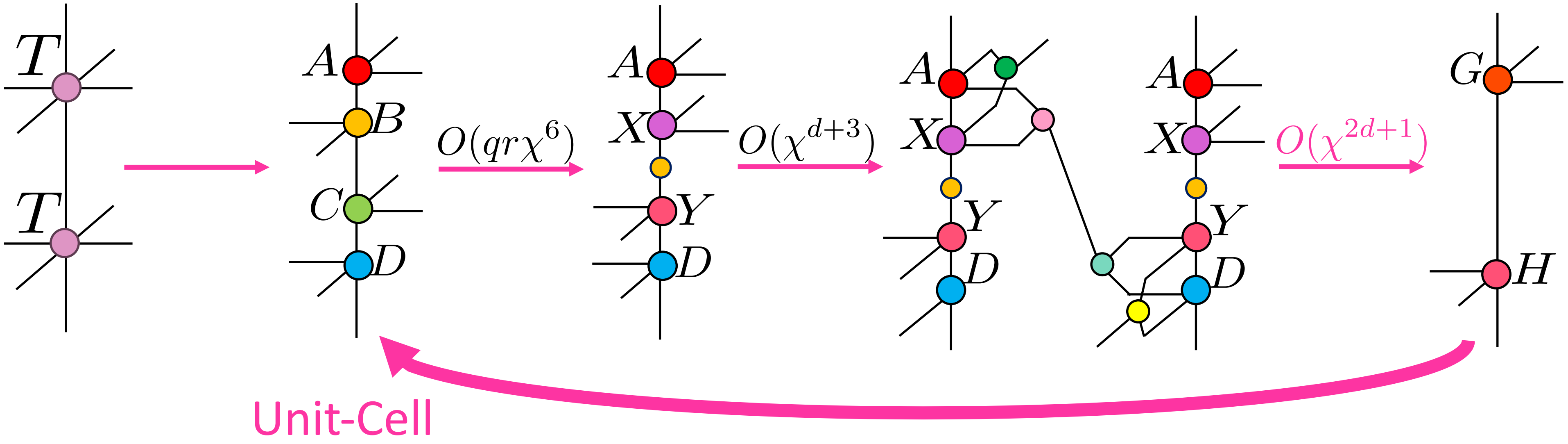
Our motivation is to search for a more efficient algorithm for four-dimensional theories.

[D. Adachi, T. Okubo, and S. Todo, (2020).]

Two reason for the cost reduction

- ✓ The fundamental tensor has d+1 legs.
- ✓ By performing bond swapping, the number of isometries is reduced by 1/2.

The cost of ATRG is χ^{2d+1} ! (HOTRG was χ^{4d-1}) RSVD is used in the bond-swapping step



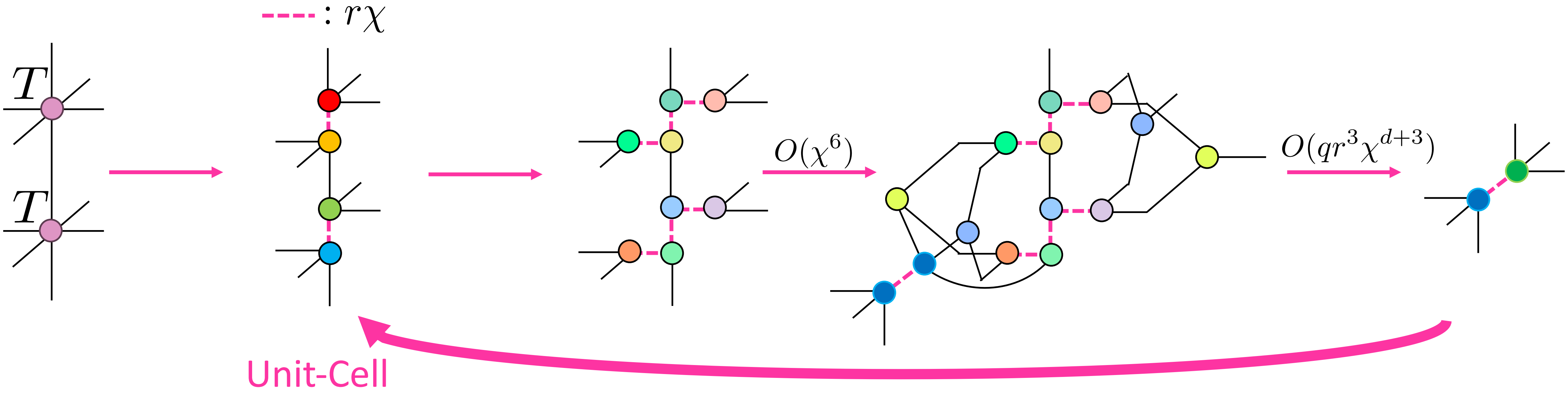
[K. Nakayama, (2023).]

Difference from ATRG is

- ✓ Using RSVD with QR iteration in contraction step (approximated SVD scheme)
- ✓ Oversampling of internal line $\chi \to r\chi$
- ✓ Decomposition of Unit-cell tensor

MDTRG is more accurate than Triad TRG , almost same accuracy of HOTRG

⟶ MDTRG-Triad rep. is improved version of Triad TRG, which cost is $O(qr^3\chi^{d+3})$



- In 4D systems, it is Trade-off between accuracy and computation cost

	HOTRG [Xie et al, (2012).]	ATRG [D. Adachi, T. Okubo, and S. Todo, (2020).]	MDTRG Triad rep. [K. Nakayama, (2023).]
cost	$O(\chi^{15})$ 😰	$O(\chi^{9})$ 🙂	$O(qr^3\chi^7)$?
Accuracy	🙂	😐	?
Problem	• Large cost, difficult to enlarge χ	• Large cost • The convergence of free energy is not as good in the 2D cases.	• Investigation is needed

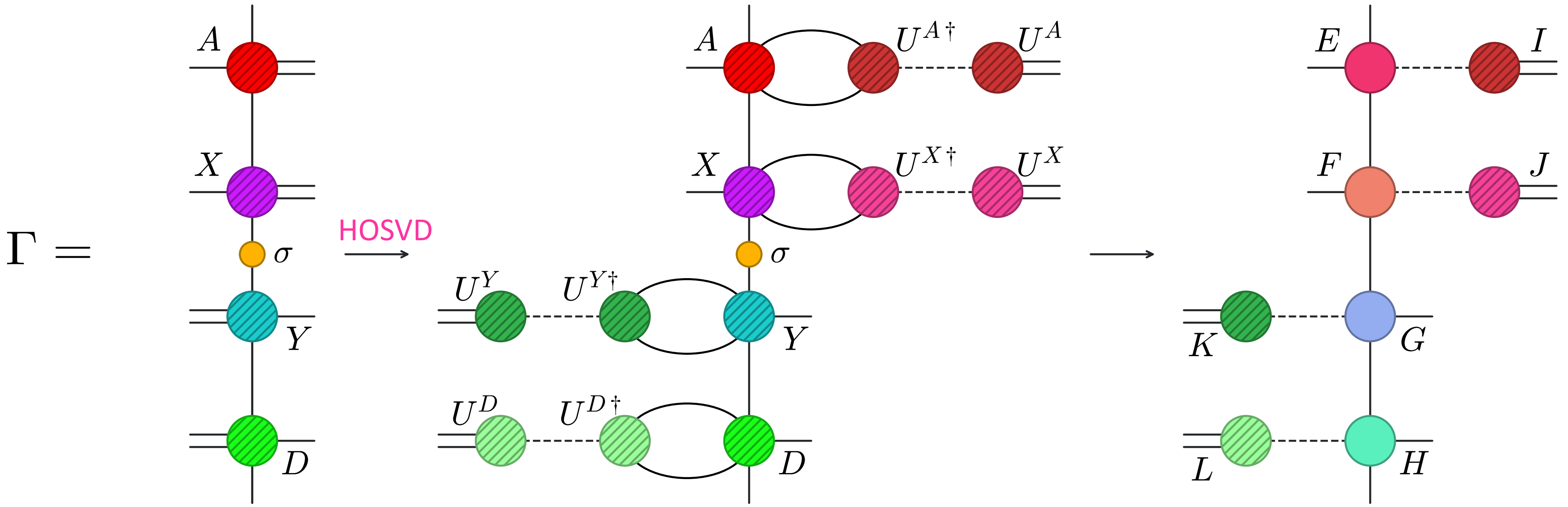
We aim for faster algorithms!

	ATRG [D. Adachi, T. Okubo, and S. Todo, (2020).]	Triad-MDTRG
Cost in 4D	$O(\chi^9)$	$O(qr^3\chi^7)$
methods	• Bond-swapping via RSVD • Exact Contraction • 3 isometry in the contraction step	• Decomp. of Unit-cell tensor • Triad • Internal line oversampling • Contraction via RSVD

Triad-ATRG ?

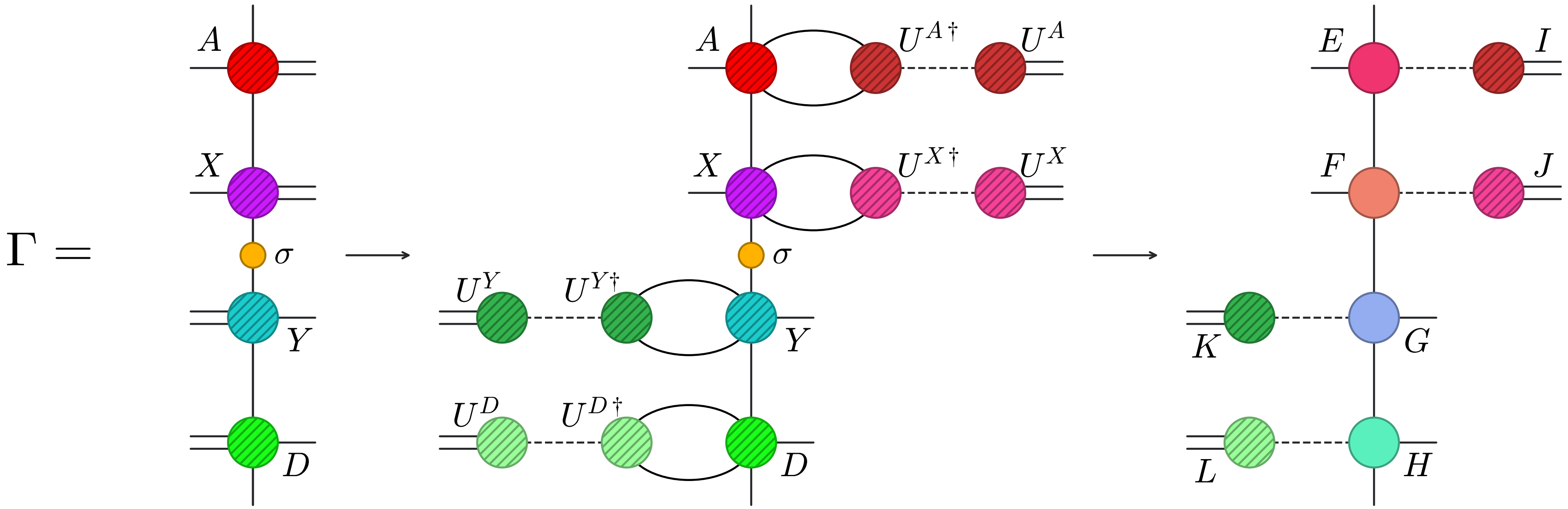
Research

- We consider triad representation of ATRG
- Consider HOSVD of unit cell tensor $\Gamma = AX\sigma YD$ after the Bond swapping
- SVD of. $AX\sigma$ and σYD provides SVD of Γ thanks to canonical form
- Triad legs are oversampled $\chi \to r\chi$
- Computational cost of this procedure is $O(\chi^7)$ (If we use RSVD, $O(qr\chi^6)$)

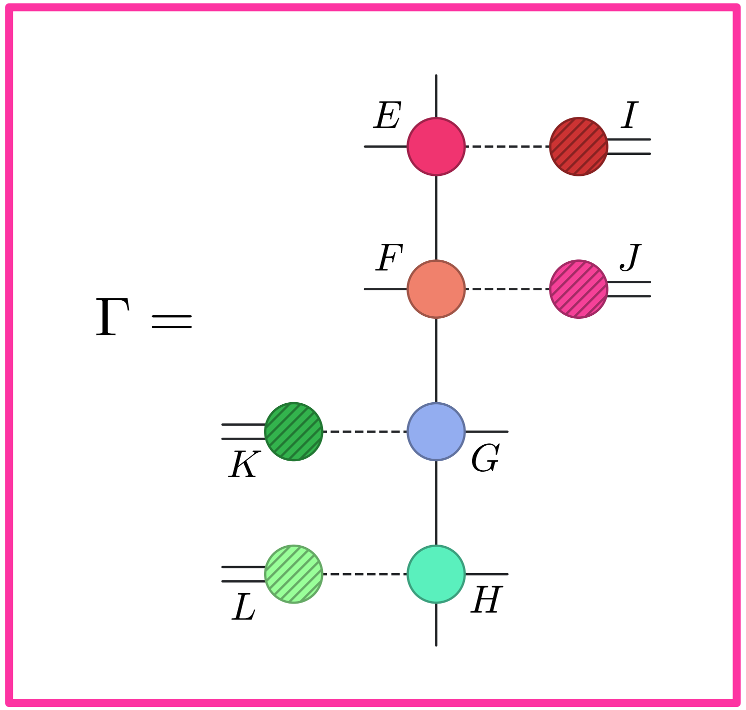


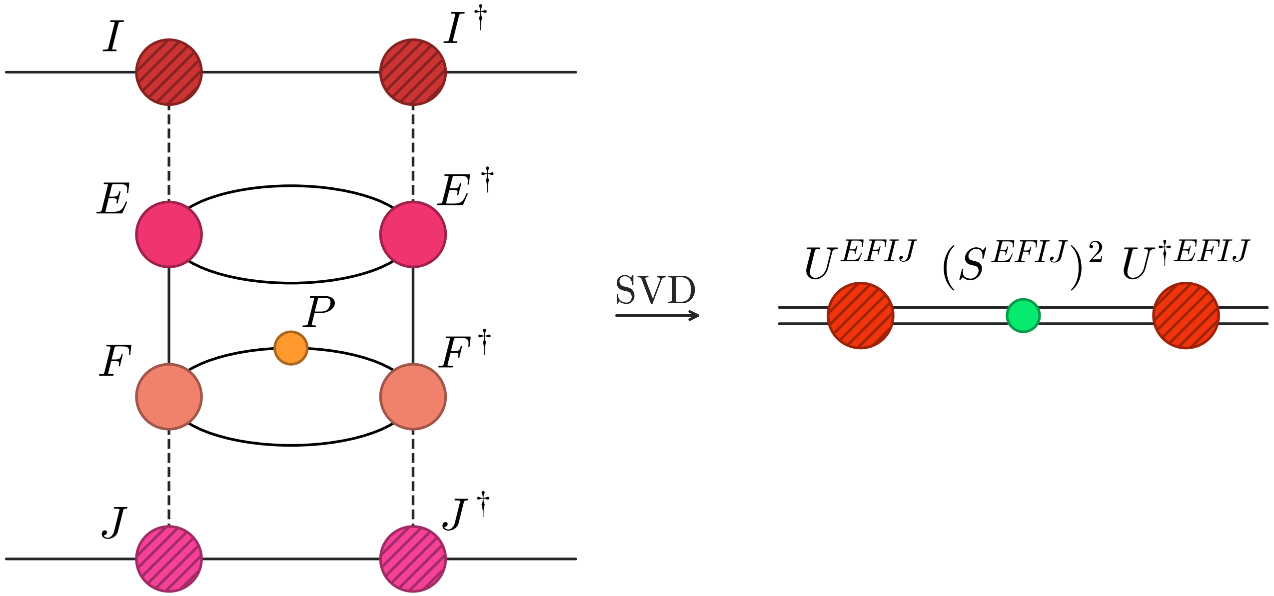
- However, in four dimensions, the order of computational cost does not change even if not all tensors are converted into triad form. Therefore, we use a form with as few decompositions as possible.
- We obtain 4 legs tensors $E, F, G, H \in \mathbb{C}^{\chi\times\chi\times\chi\times r\chi}$ and 3 legs tensors $I, J, K, L \in \mathbb{C}^{\chi\times\chi\times r\chi}$ (we call this form as triad rep.)

oversampled



- We derive squeezers in the same manner of ATRG [S. Akiyama, phd, 2022.]
- since Γ is not canonical form anymore, we must decompose $\Gamma \simeq EFGHIJKL$
- We can calculate separately by introducing the Gramm-matrix of. $EFIJ$ and $GHKL$

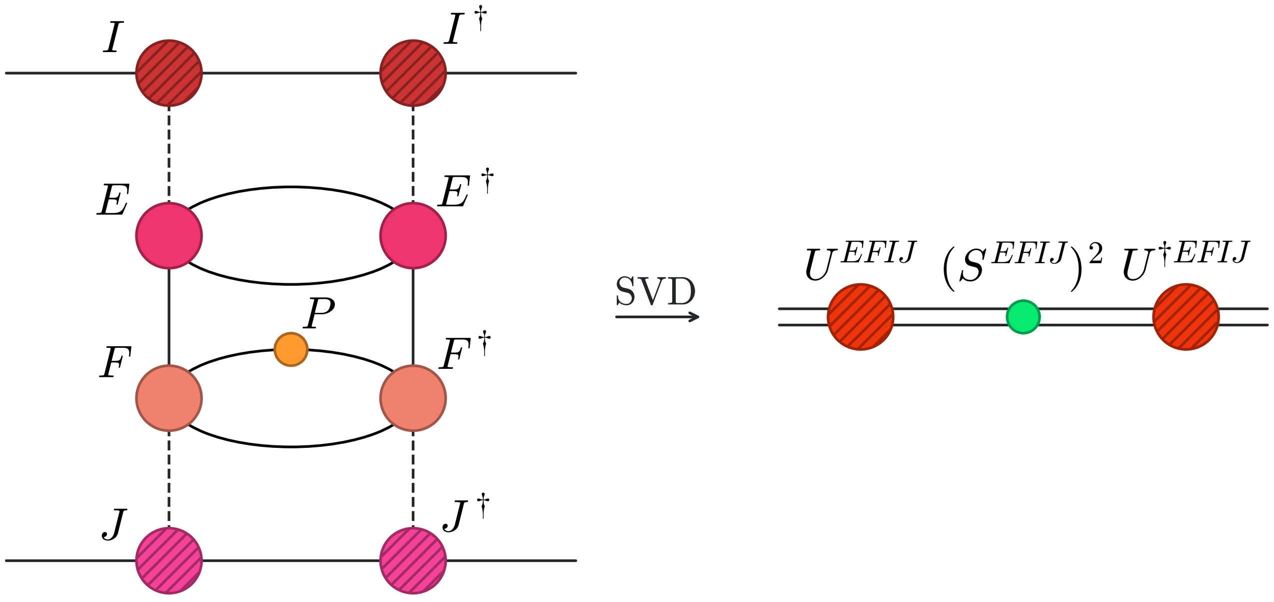




$P = (GHKL)(GHKL)^\dagger$

- Computational cost of this procedure is $\min(O(\chi^7), O(r^2\chi^6))$
- All decomposition in this procedure are SVDs of $\Gamma(n)\Gamma(n+\hat{\mu})$ as in the improved ATRG

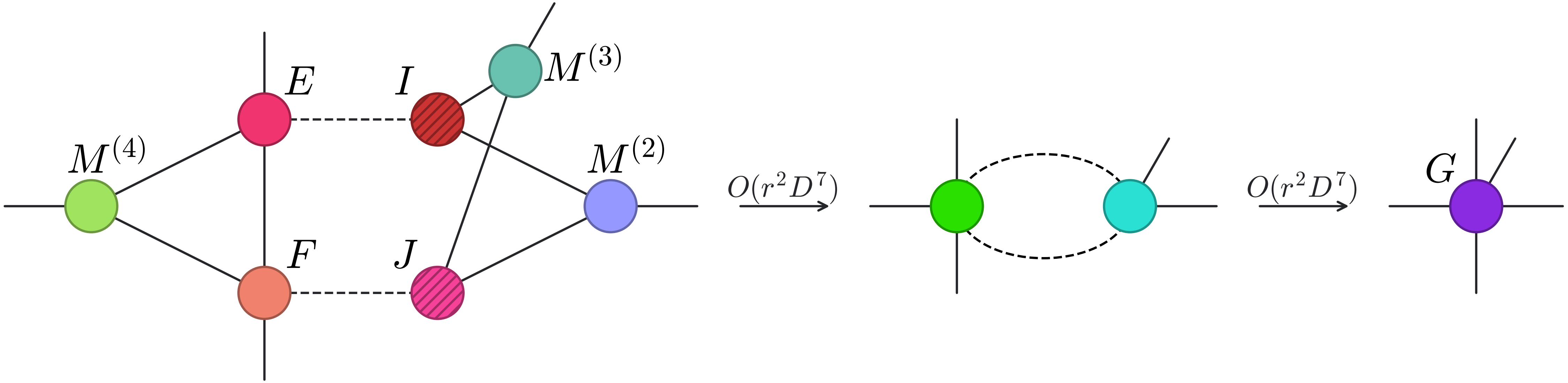
[S. Akiyama, phd, 2022.]
[S. Iino, S. Morita, and N. Kawashima, (2019).]
[D. Adachi, T. Okubo, and S. Todo, (2022).]



$$P = (GHKL)(GHKL)^\dagger$$

CONTRACTION STEP

- Thanks to the Triad form, Computational cost is reduced to $O(r^2\chi^7)$, smaller than ATRG ($O(\chi^9)$)
  Bottleneck
- We do not use RSVD since we already used it once in the bond-swapping step



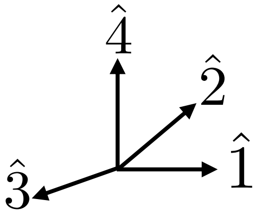
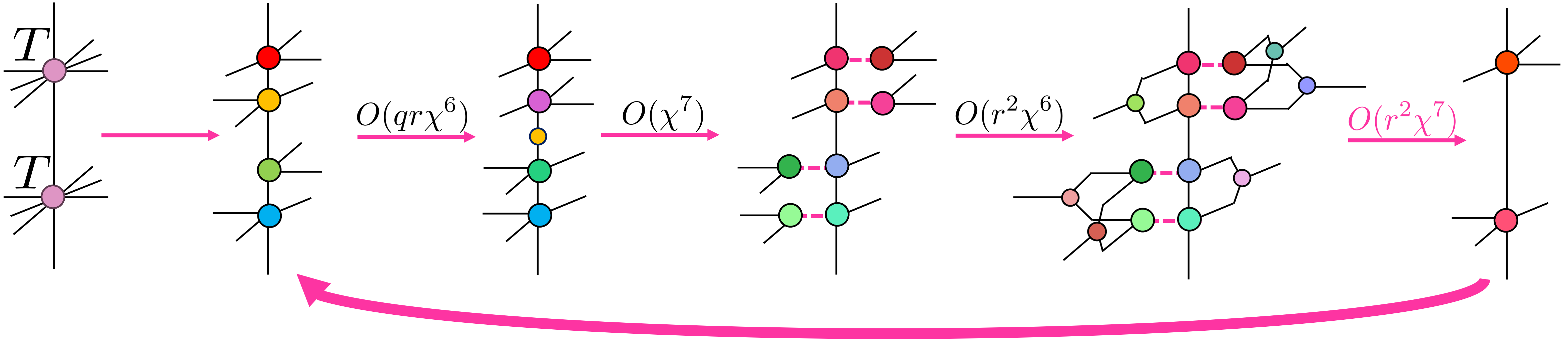


表1 Computational cost

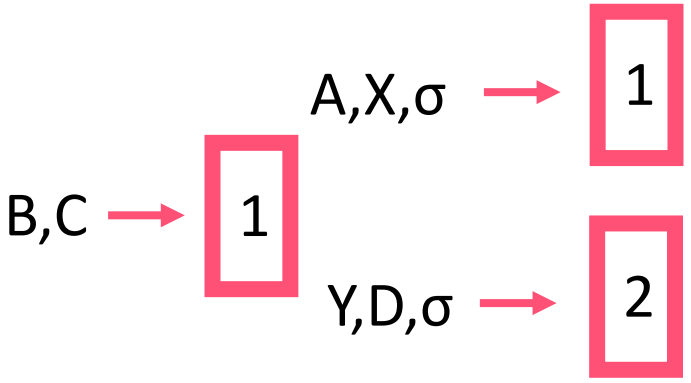
Step	ATRG	Triad ATRG
Bond Swapping	$O(qr\chi^6)$	$O(qr\chi^6)$
Make Triad	None	$O(\chi^7)$
Squeezer	$O(\chi^7)$	$O(\min(\chi^7, r^2\chi^6))$
Contraction	$O(\chi^9)$	$O(r^2\chi^7)$

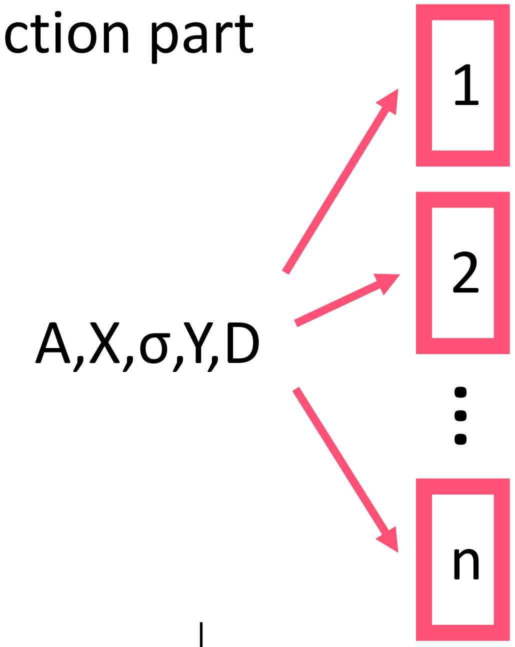
Bottleneck



GPU PARALLELIZATION -ATRG

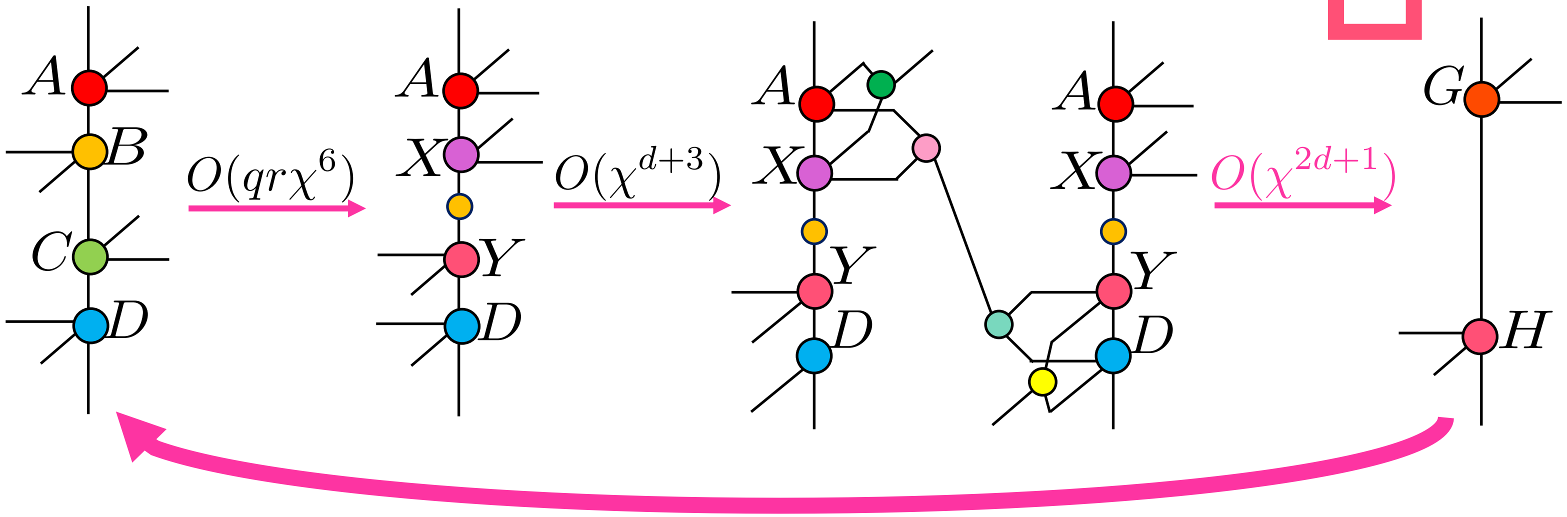
We employ multi-GPU parallelization for contraction part





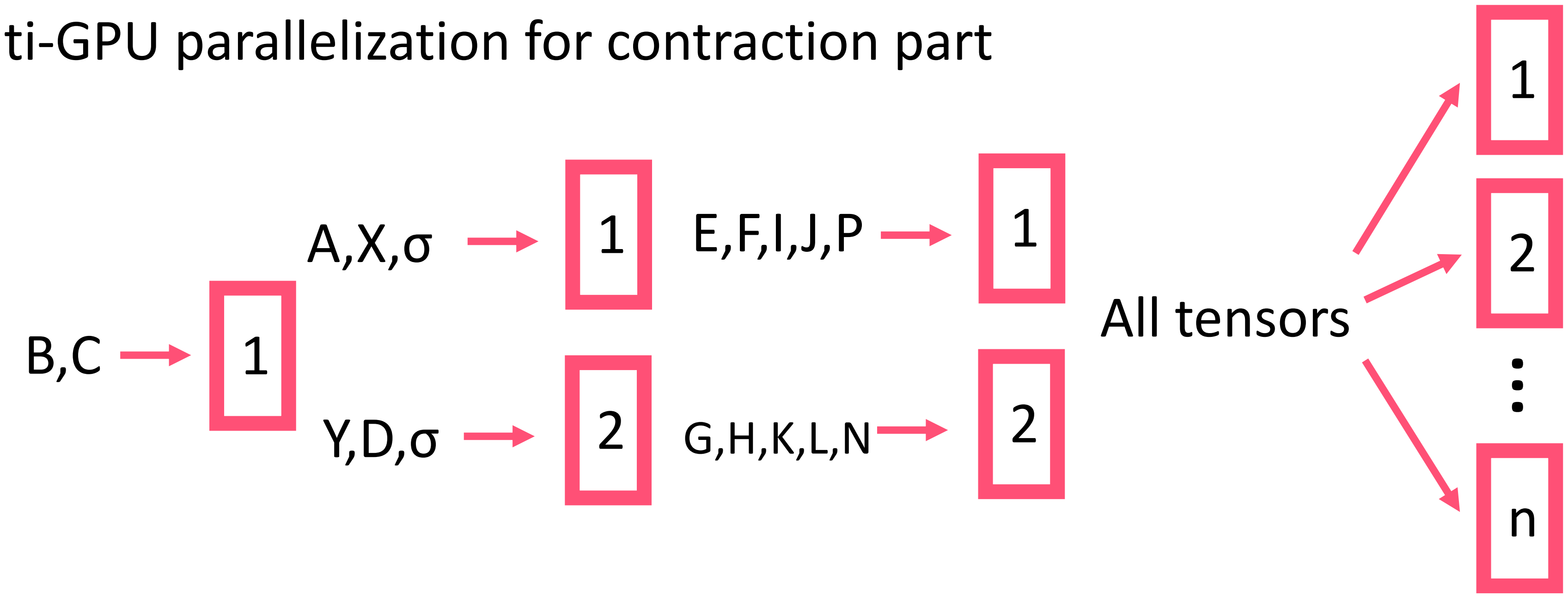
Tensor contraction

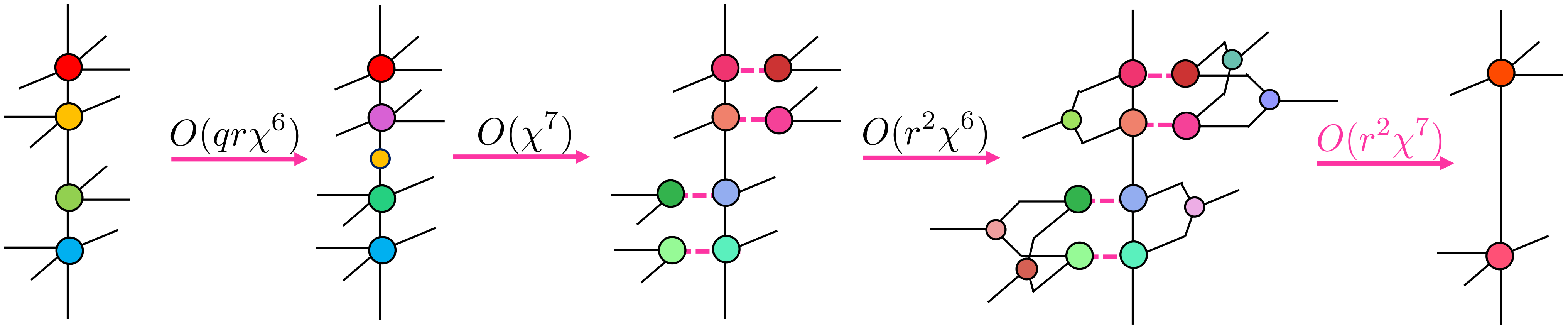
TensorOperations.jl
(A wrapper of cuTENSOR)



GPU PARALLELIZATION –TRIAD-ATRG

We employ multi-GPU parallelization for contraction part



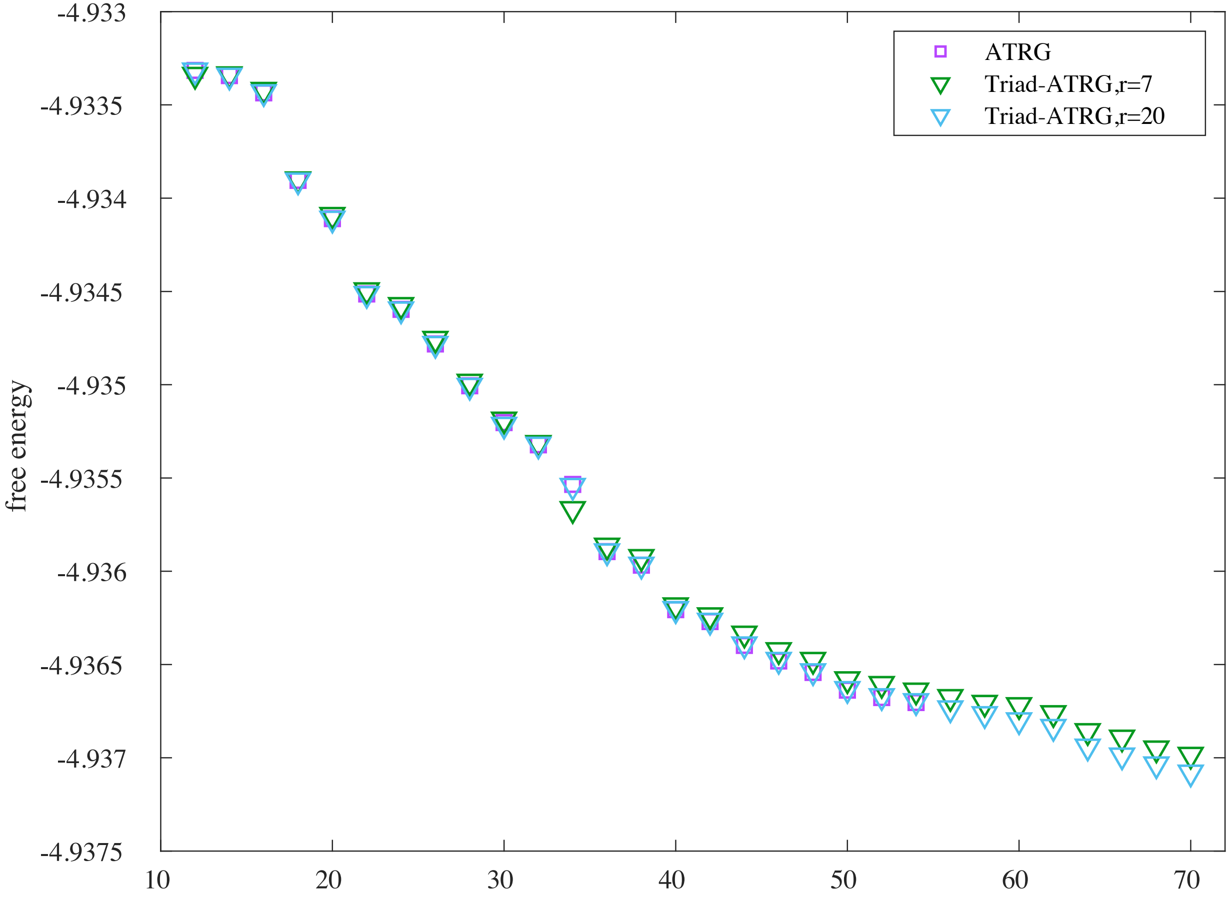


Numerical results on 4D Ising model

FREE ENERGY

- We investigate the convergence of free energy at 4D Ising model in r=7, L=1024,T=6.65035
- The results are in high agreement with ATRG

HOTRG(D=13): $T_c = 6.650365(5)$
[S. Akiyama, Y. Kuramashi, T. Yamashita, and Y. Yoshimura, (2019).]

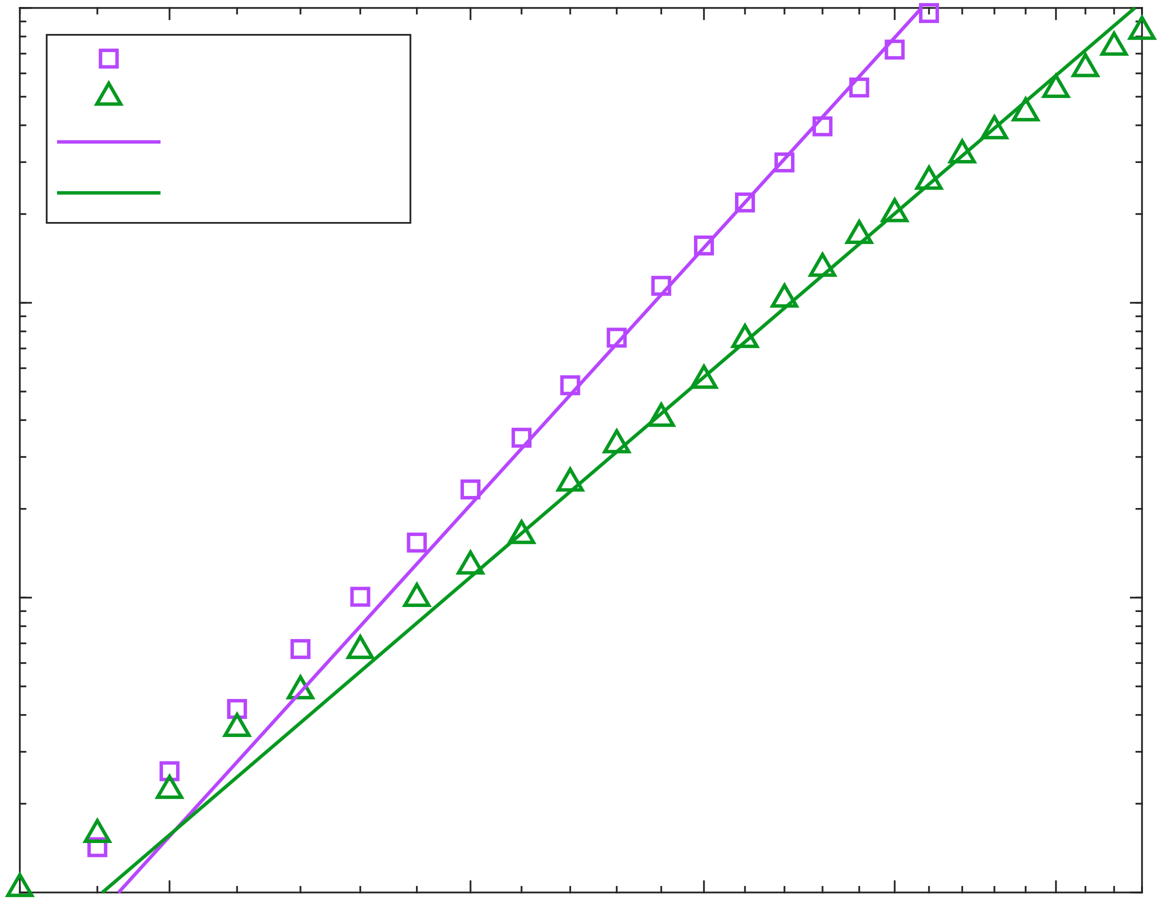


χ	ATRG	r=7	r=20
38	-4.9359675	-4.9359235	-4.9359646
40	-4.9362060	-4.9361825	-4.9362026
42	-4.9362695	-4.9362360	-4.9362648
44	-4.9363974	-4.9363340	-4.9363918
46	-4.9364809	-4.9364227	-4.9364746
48	-4.9365426	-4.9364745	-4.9365357
50	-4.9366373	-4.9365787	-4.9366312
52	-4.9366769	-4.9366039	-4.9366695
54	-4.9367035	-4.9366392	-4.9366959

Difference is only 0.0013%(r=7), 0.00015%(r=20)

COMPUTATIONAL TIME ON A CPU

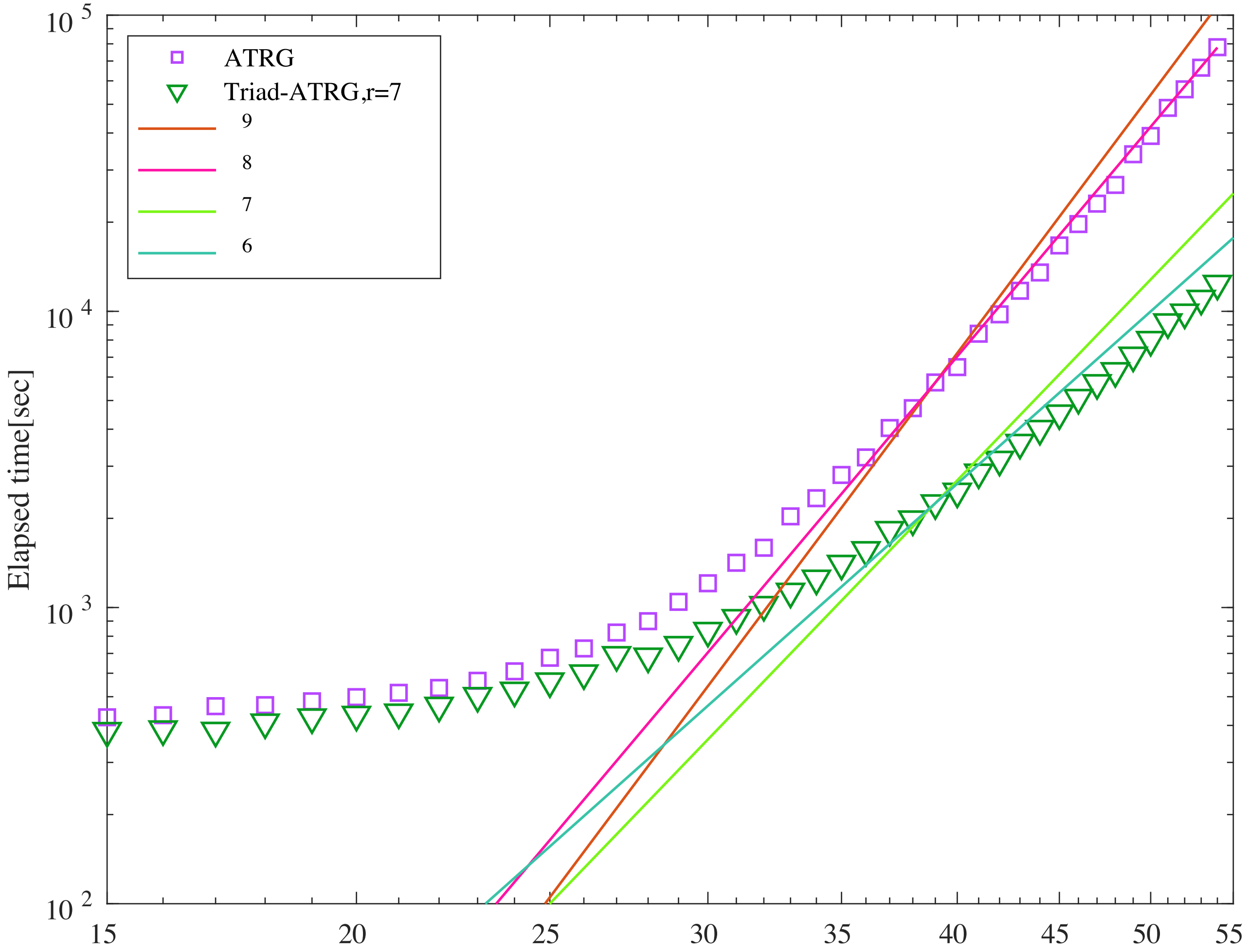
- We investigate the computational time in r=7 using a single CPU calculation
- Scaling of the computational time is $O(\chi^7)$



$O(\chi^7)$

COMPUTATIONAL TIME ON GPUS

- We investigate the computational time in r=7 ,L=1024 by 2 GPU parallelized calculation
- Scaling of the computational time improved significantly



$\simeq O(\chi^8)$

$< O(\chi^6)$

Triad-ATRG could be powerful tool for GPU computations

We have used Tesla V100 16GB PCIE✖2

https://www.elsa-jp.co.jp/wp-content/uploads/2019/03/nvidia_tesla_v100_3qtr.png

PHASE TRANSITION POINT

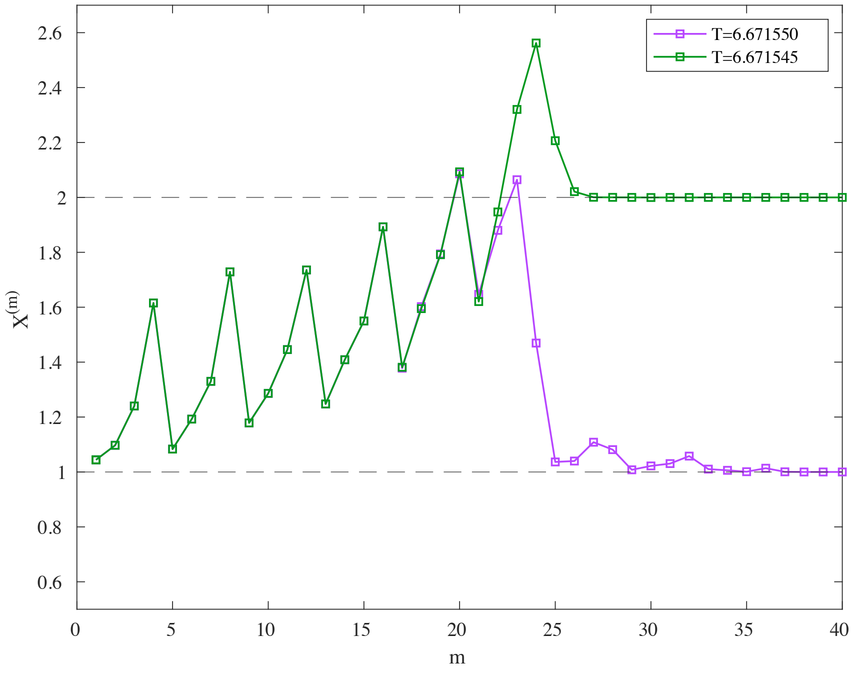
To determine the transition point, we evaluate the following value at each coarse-graining step.

$$X^{(m)} = \frac{(\mathrm{Tr}A^{(m)})^2}{\mathrm{Tr}(A^{(m)})^2}, \ \text{with} \ A^{(m)}_{kl} = \sum T^{(m)}_{i_1 i_2 i_3 k i_1 i_2 i_3 l}$$

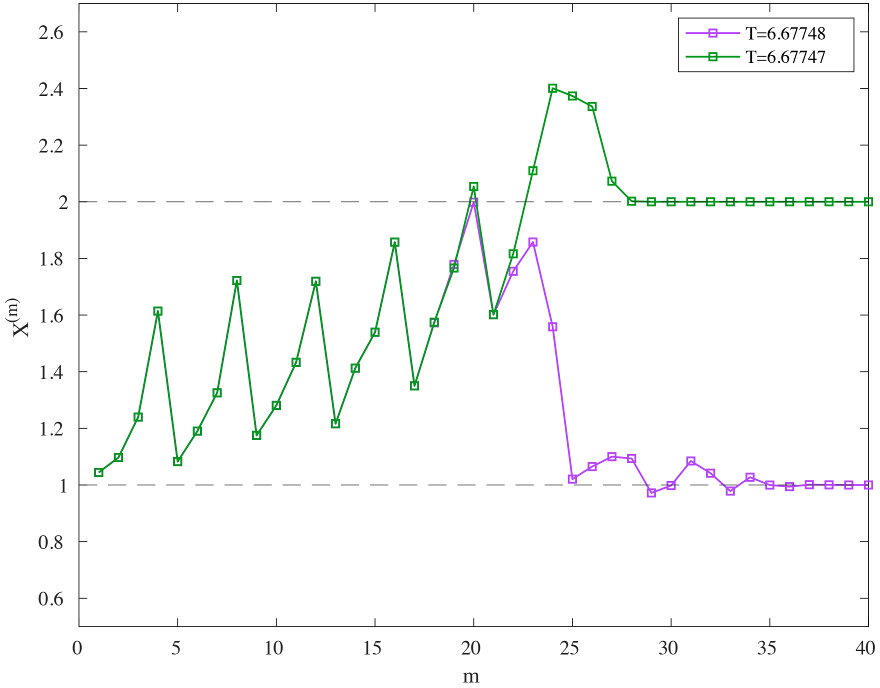
HOTRG(D=13): $T_c = 6.650365(5)$

[S. Akiyama, Y. Kuramashi, T. Yamashita, and Y. Yoshimura, (2019).]

[Z.-C. Gu and X.-G. Wen, (2009).]

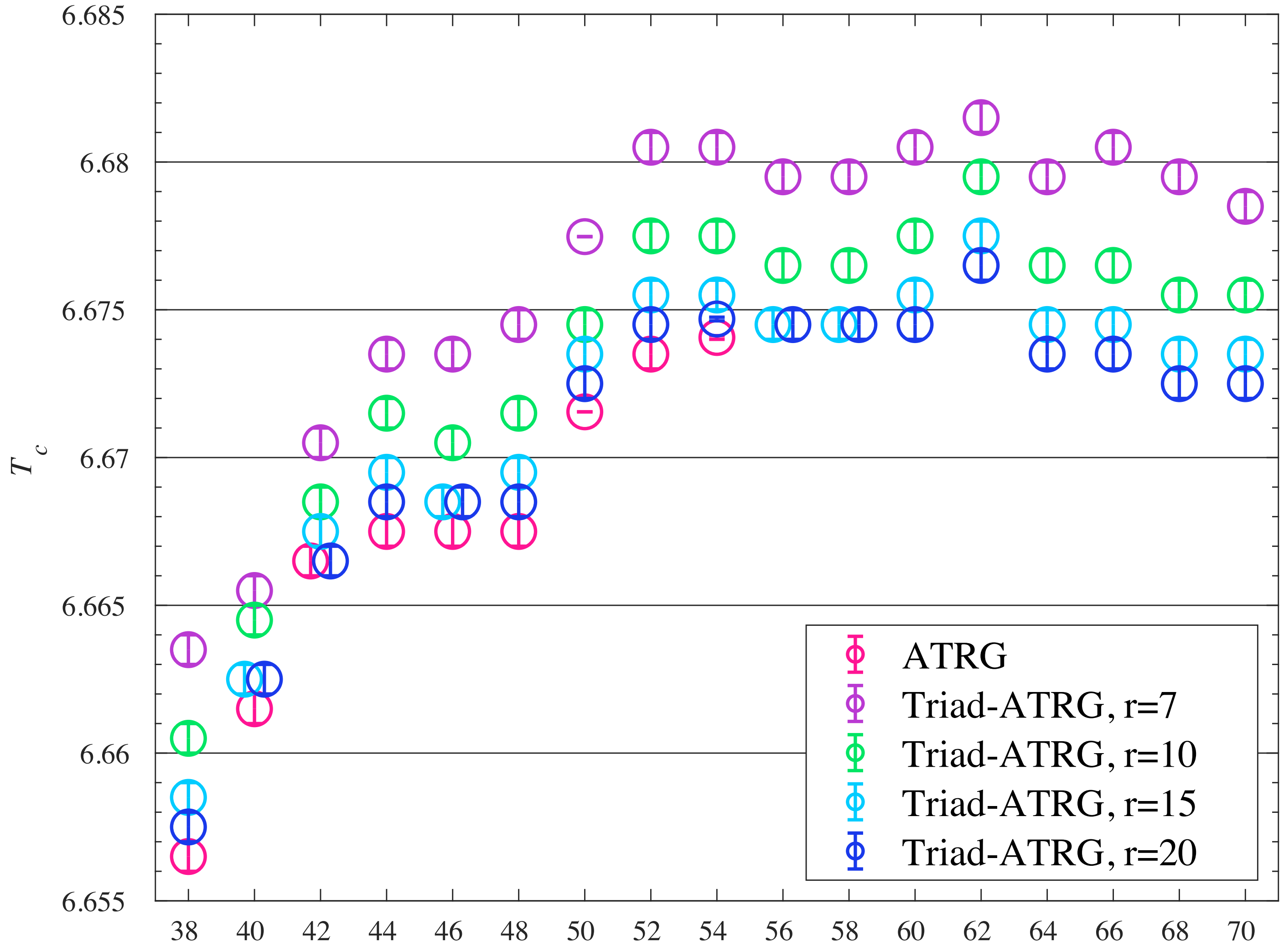


ATRG



Triad ATRG

PHASE TRANSITION POINT



HOTRG(χ=13): $T_c = 6.650365(5)$

[S. Akiyama, Y. Kuramashi, T. Yamashita, and Y. Yoshimura, (2019).]

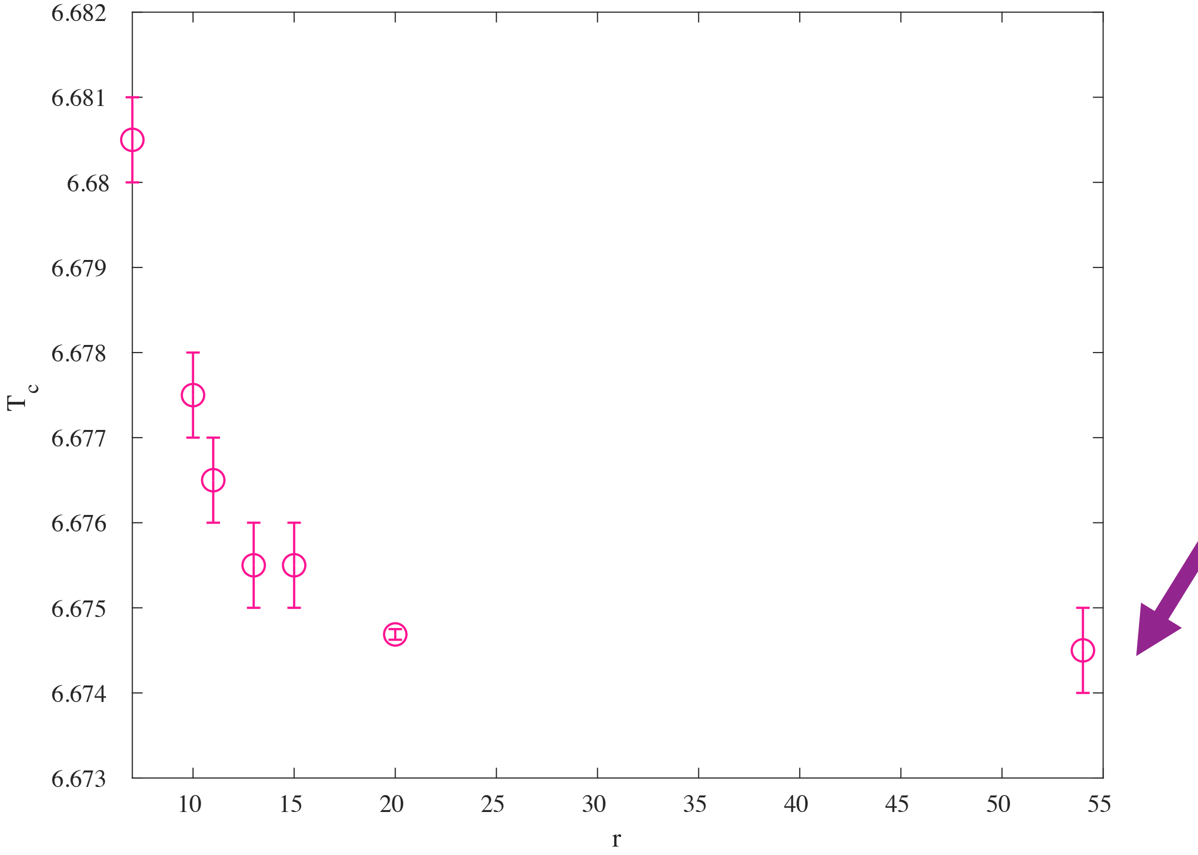
Monte-Carlo: $T_c = 6.6803069(58)$

[P. H. Lundow and K. Markström, (2023).]

Difference from the ATRG results at χ=54 is ~0.1% for r=7, and ~0.04% for r=10,and ~0.09% for r=20

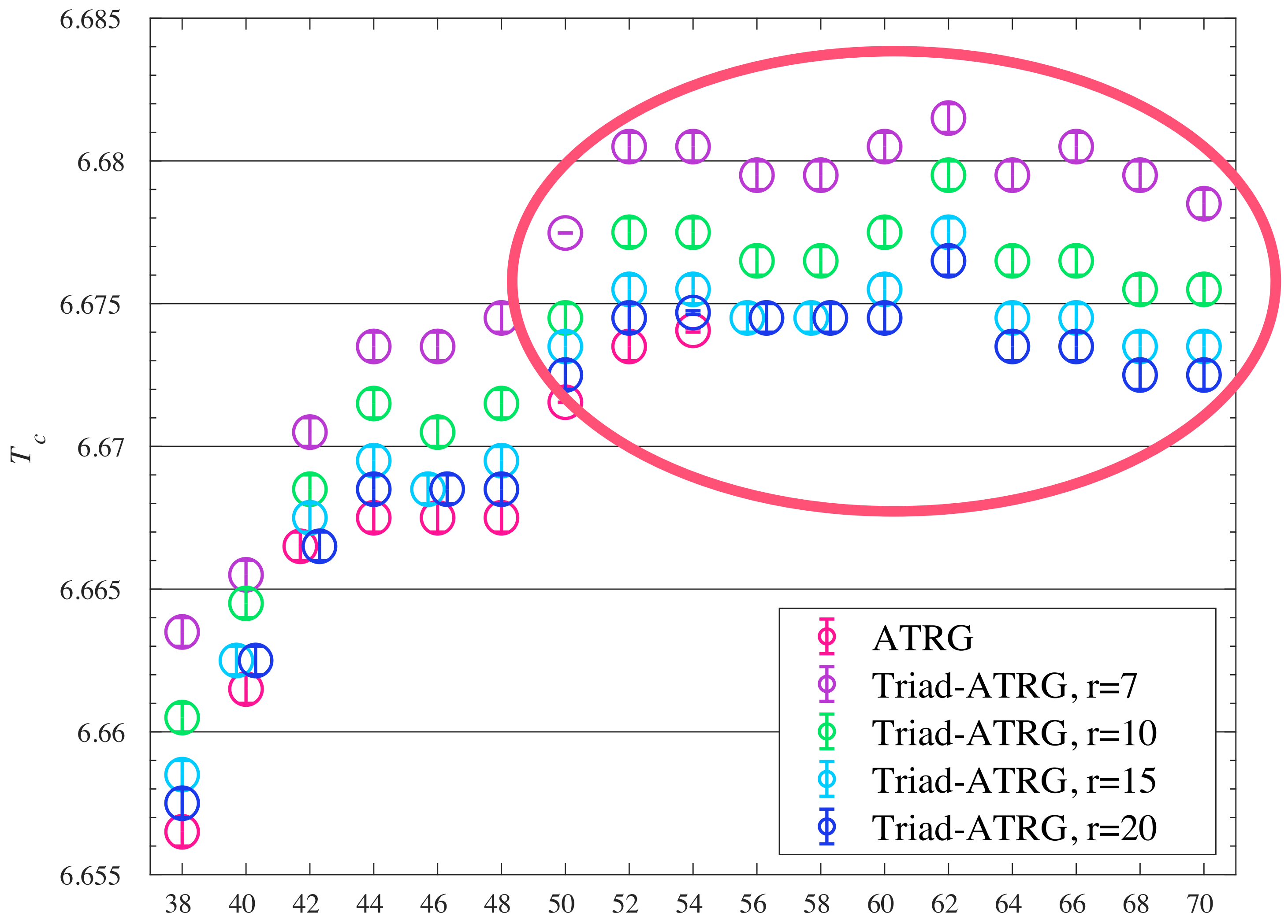
*results of ATRG has not converged well

- Triad-ATRG converges to the ATRG as r increases



ATRG ($\chi = 54,\ r = 54$)

PHASE TRANSITION POINT



HOTRG(χ=13): $T_c = 6.650365(5)$

[S. Akiyama, Y. Kuramashi, T. Yamashita, and Y. Yoshimura, (2019).]

Monte-Carlo: $T_c = 6.6803069(58)$

[P. H. Lundow and K. Markström, (2023).]

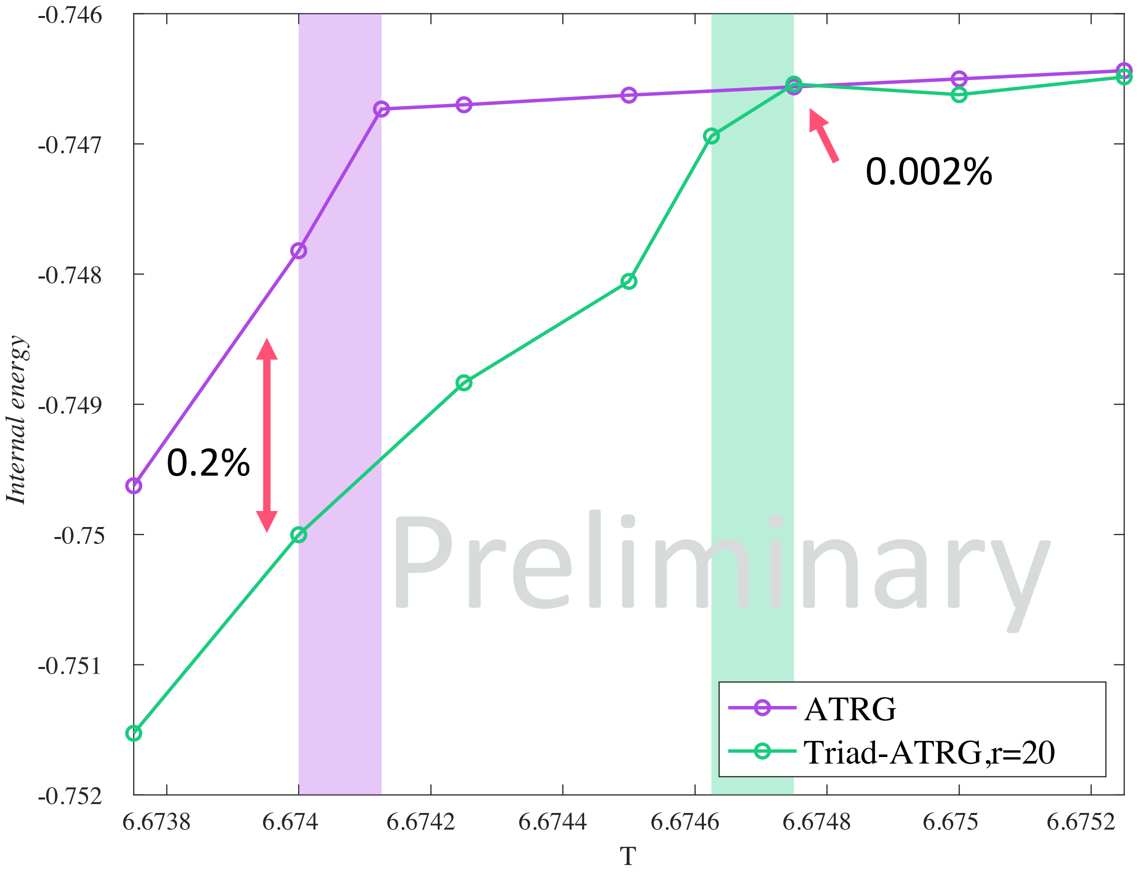
The behavior of Tc seems better for $\chi \geq 56$ but has not converged yet

We have used NVIDIA A100 40GB PCIE ×8

https://www.nvidia.com/ja-jp/data-center/a100/

INTERNAL ENERGY

- We investigate the internal energy with impurity tensor method at $\chi = 54$
- The difference at 0.002% at T=6.67475 and 0.2% at T=6.674



T	ATRG	Triad-ATRG, r=20
6.67375	-0.74962536	-0.75152631
6.674	-0.74782076	-0.75000187
6.67425	-0.74669974	-0.74883493
6.6745	-0.74662574	-0.74805611
6.67475	-0.74656204	-0.74654099
6.675	-0.74650094	-0.74660892
6.67525	-0.74643774	-0.74648570

Colored bands are transition points obtained by the ATRG and Triad-ATRG.

- The results of Triad-ATRG are highly consistent with the ATRG results
- Triad-ATRG significantly improves the computational cost on CPU and GPUs

Triad ATRG would be a powerful tool for 4D systems

Future works

- Calculate in more large χ
- Apply improved HOSVD
- Oversampling at bond-swapping step
- Apply to other 4D systems

END

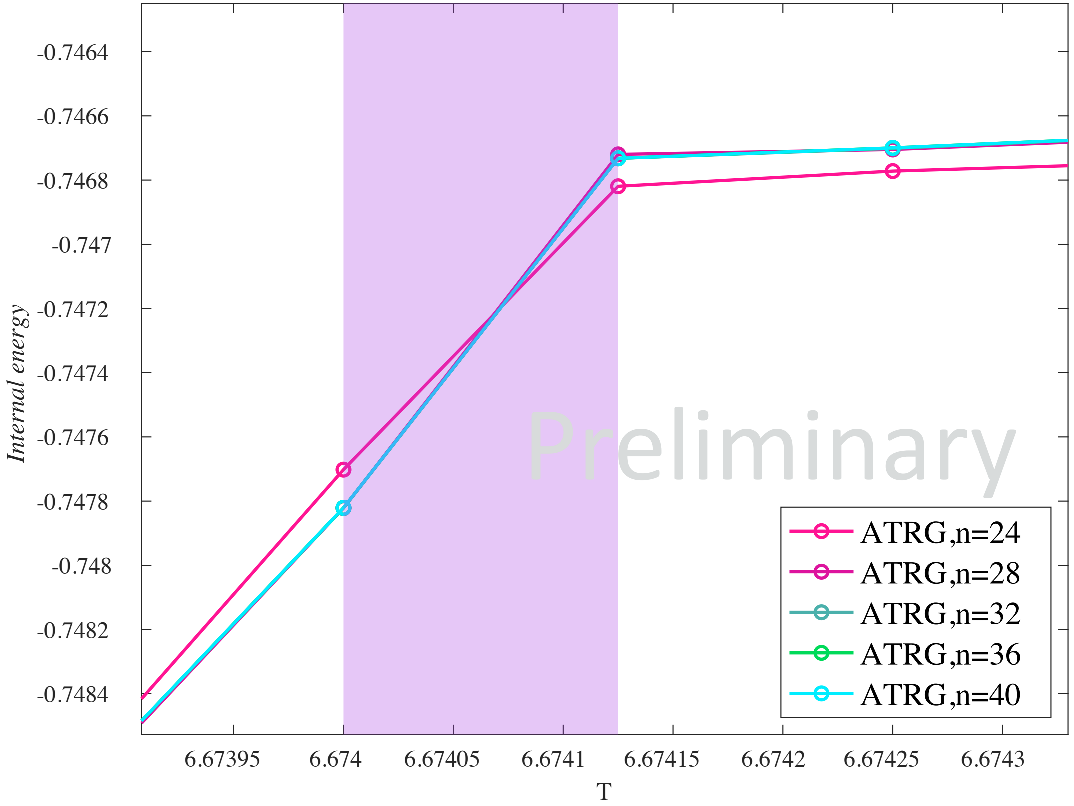
表1　memory cost of D^5

D	size(GB)
45	1.37
50	2.33
55	3.75
60	5.79
65	8.64
70	12.5

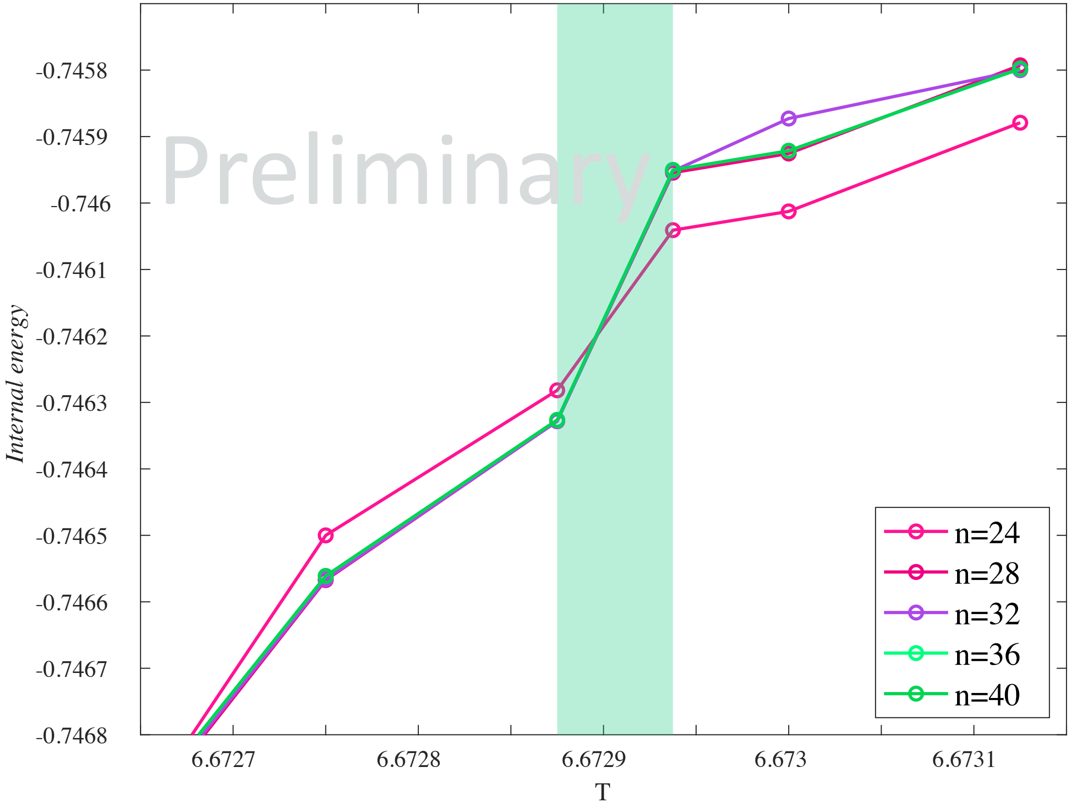
Usually a GPU has 10-80GB memory
→Memory cost is unnegligible

https://www.nvidia.com/ja-jp/data-center/a100/

A mutual crossing still exists→We need more large χ ?



ATRG $\chi = 54$



Triad-ATRG $r = 20, \chi = 70$