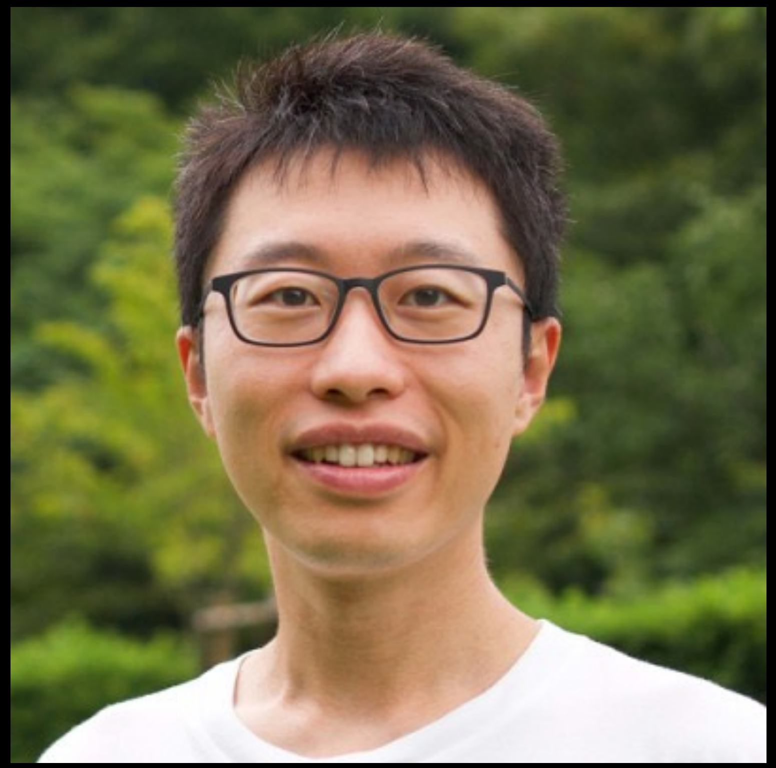
テンソルネットワーク法を用いた量子計算のシミュレーション

白川知功

理化学研究所・計算科学研究センター・量子系物質科学研究チーム（本務）
理化学研究所・量子コンピュータ研究センター・量子計算科学研究チーム
理化学研究所・開拓研究本部・柚木計算物性物理研究室
理化学研究所・数理創造プログラム
理化学研究所・計算科学研究センター・量子HPC連携プラットフォーム部門・量子計算シミュレーション技術開発ユニット

2023年12月25日第4回量子ソフトウェアワークショップ:NISQ計算機時代の量子シミュレーションと量子コンピュータシミュレータ@東京大学本郷キャンパス

Collaborators



Rong-Yang Sun[1,2,3]
(2021~)

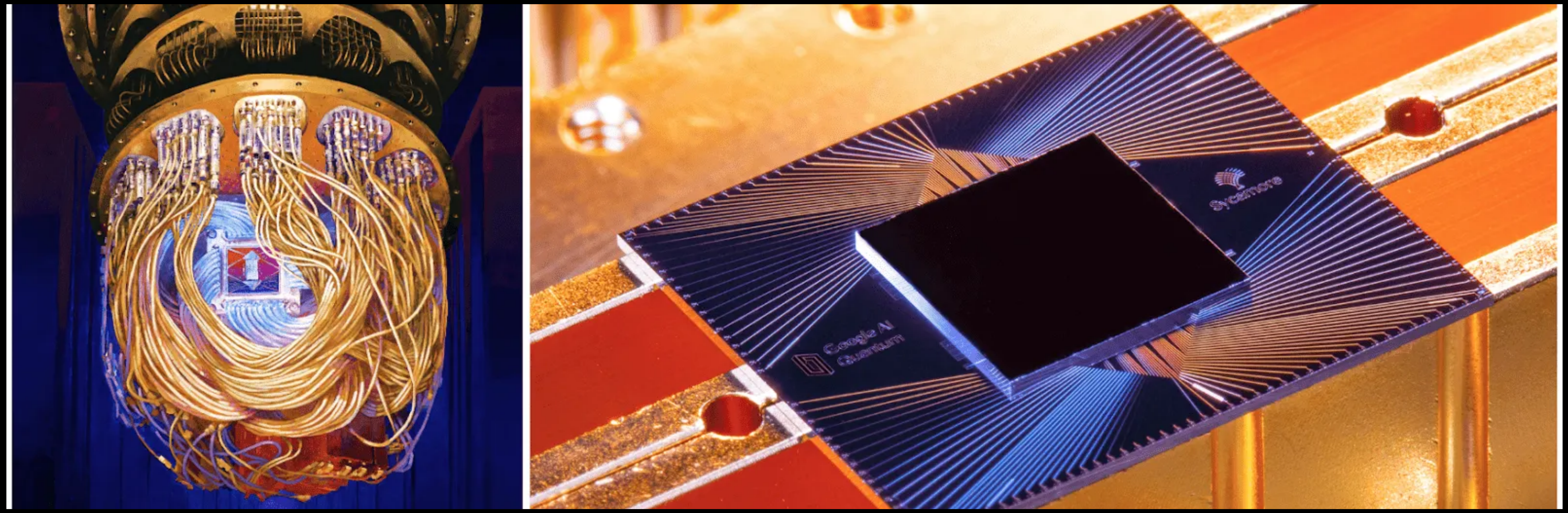
Hidehiko Kohshiro[1]
(2023~,幸城秀彦)

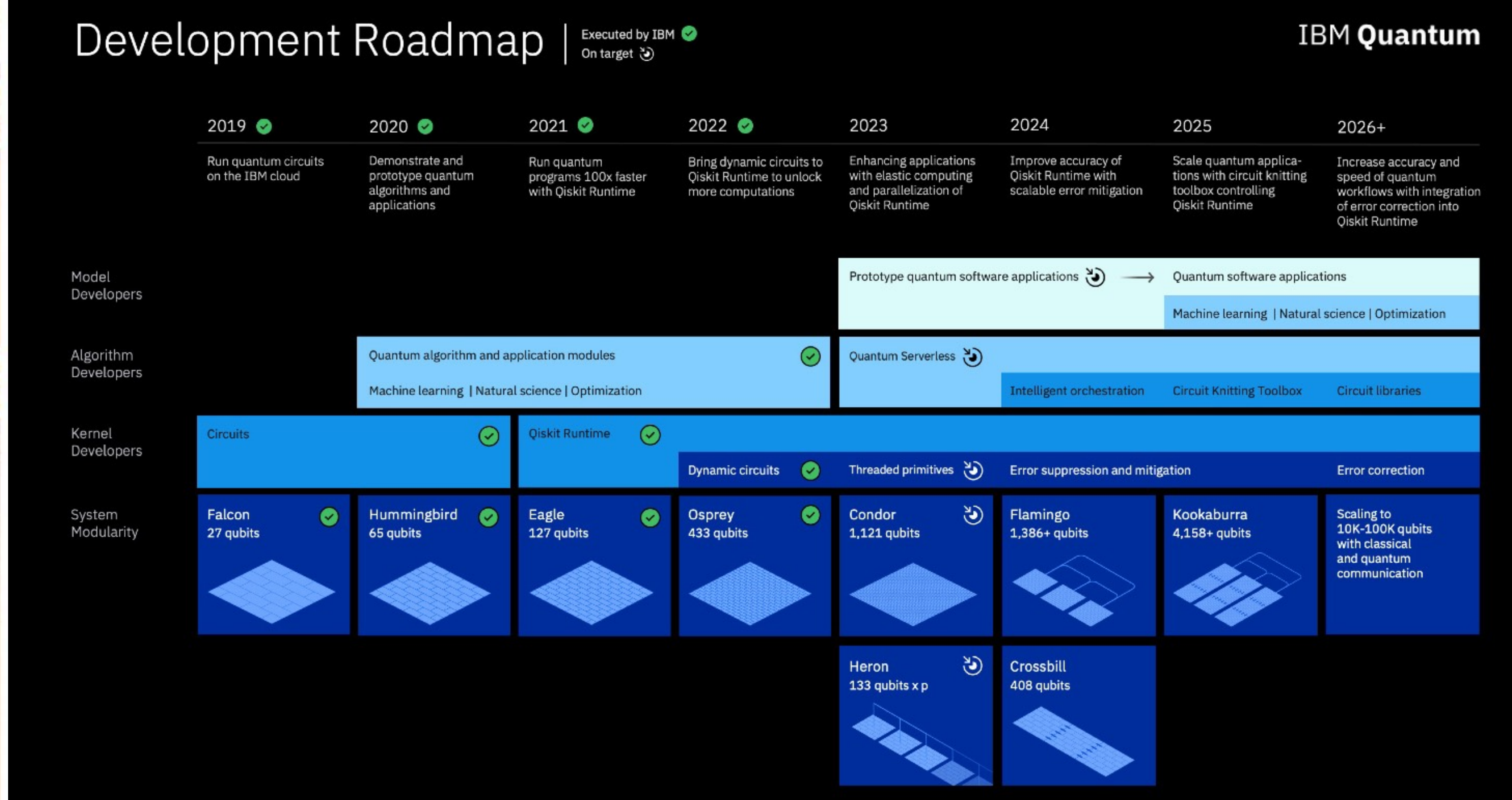
Seiji Yunoki[1,2,3,4,5]
(柚木清司)

[1]Computational Materials Science Research Team, RIKEN R-CCS, Kobe, Hyogo
[2]Quantum Computational Science Research Team, RIKEN RQC, Wako, Saitama

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667]

Near-term quantum devices





- Noisy intermediate-scale quantum (NISQ) era
  - A few $\mathcal{O}(10^2 \sim 10^3)$ qubits without error correction
  - A few $\mathcal{O}(10^1 \sim 10^2)$ depths circuit evolution

Quantum Computing in the NISQ era and beyond

John Preskill

Institute for Quantum Information and Matter and Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena CA 91125, USA

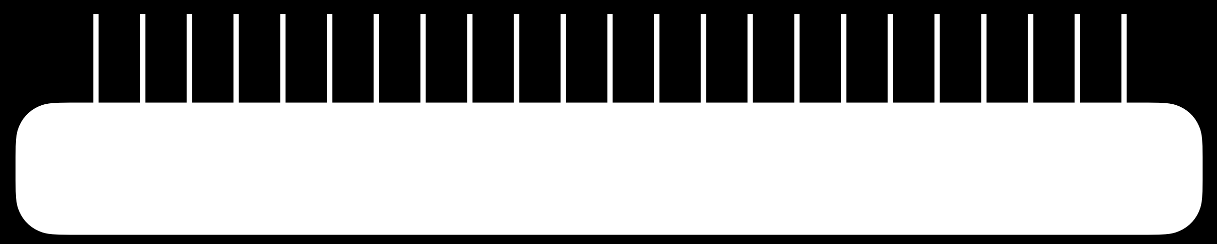
30 July 2018

Near-term aim: achieve useful quantum advantage on NISQ devices

Computational Materials Science

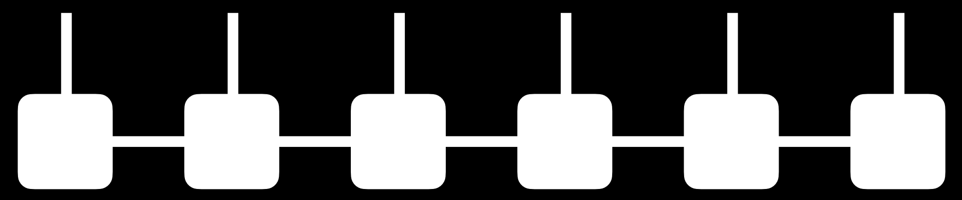
Simulators for quantum computing

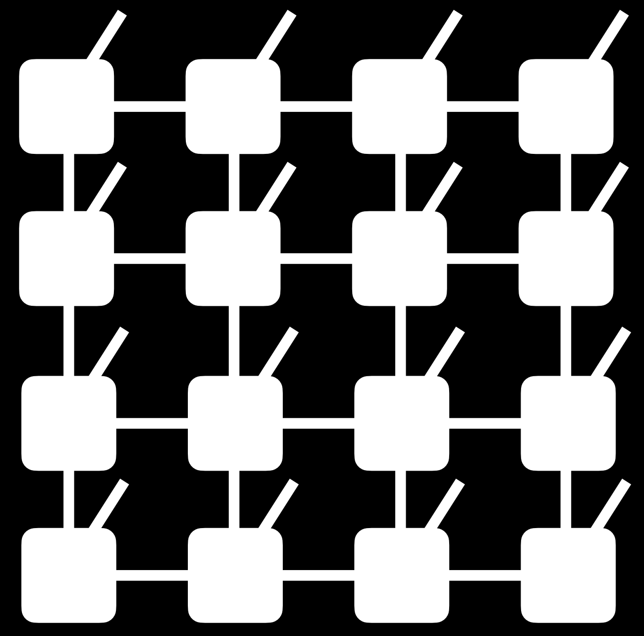
State vector simulator



Can compute any quantum circuits
Limitation on number of qubits

Tensor network simulator





Can compute quantum circuits with large qubits
Limitation on entanglements

Why we need the simulator of quantum computer?

(1) To check the validity of the quantum algorithm assuming that the quantum computer has worked correctly.

In order to explore the useful applications of quantum computers, it is necessary to check the results of quantum computers when they work properly.

(2) To verify that the quantum computer is working properly

Current quantum computing devices are noisy and have no error correction, so they must be evaluated against correct operation.

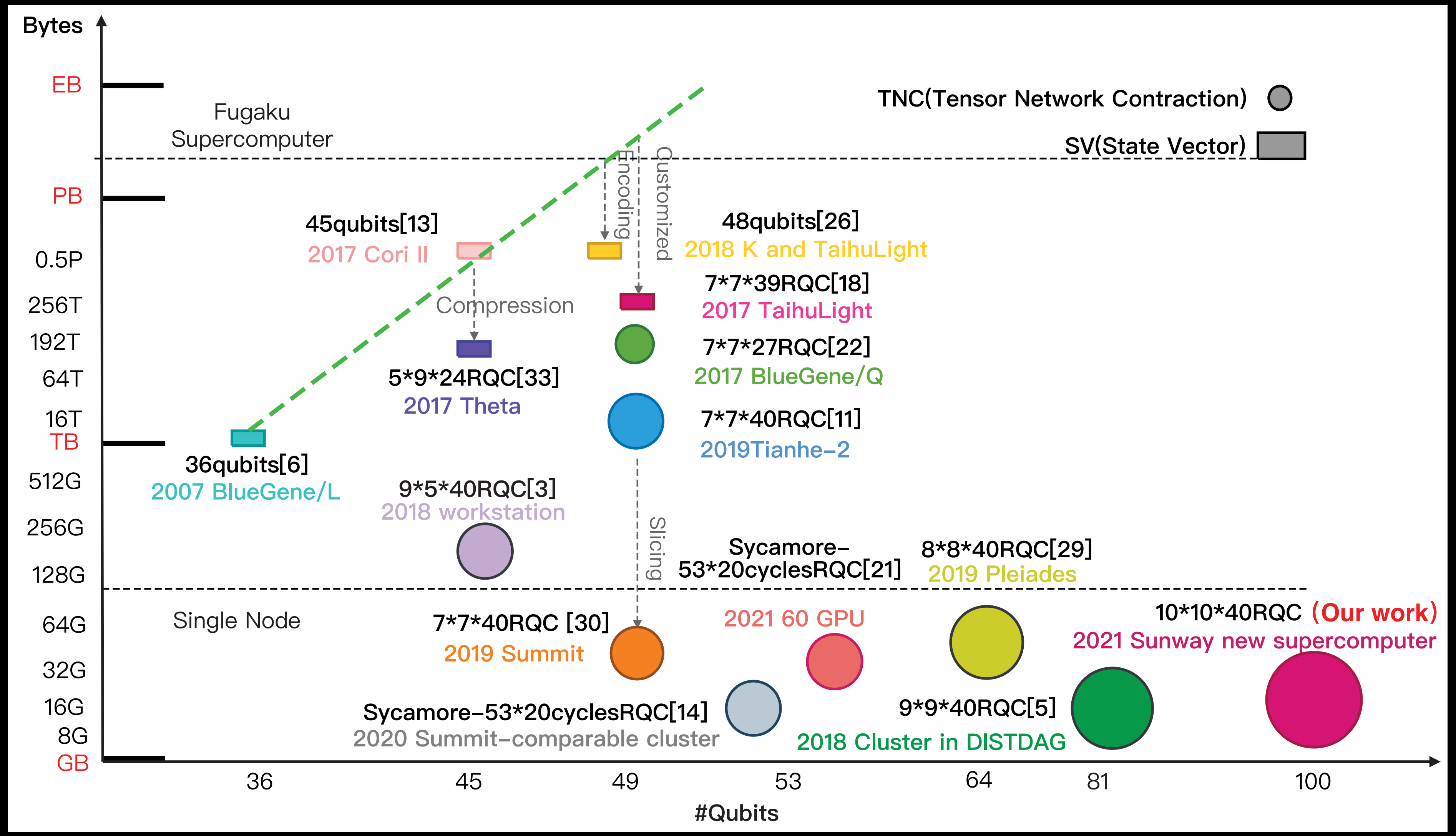
(3) To bridge the classical information and quantum information

For realization of scalable state preparation, hybrid tensor network [Xiao Yuan et al., Phys. Rev. Lett. **127**, 040501 (2021)], etc.

Computational Materials Science

Tensor network simulators

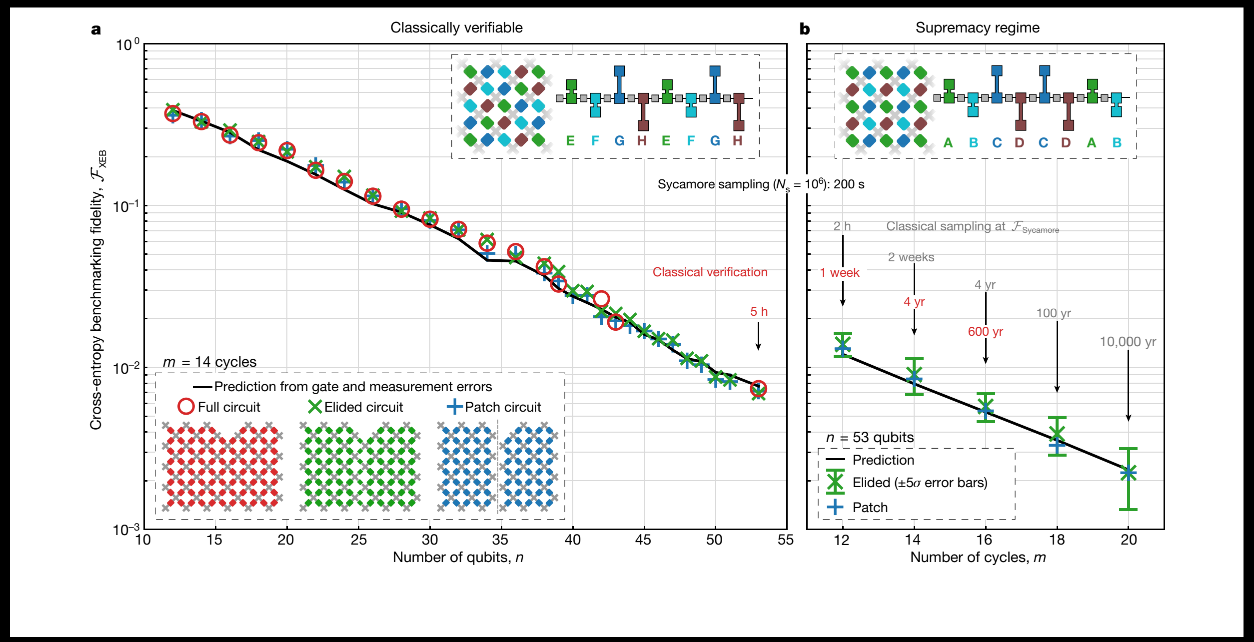
2021 ACM Gordon Bell Prize



- optimal slicing scheme
- three-level parallelization scales to about 42 million cores
- fused permutation and multiplication design for tensor contraction
- mixed-precision scheme

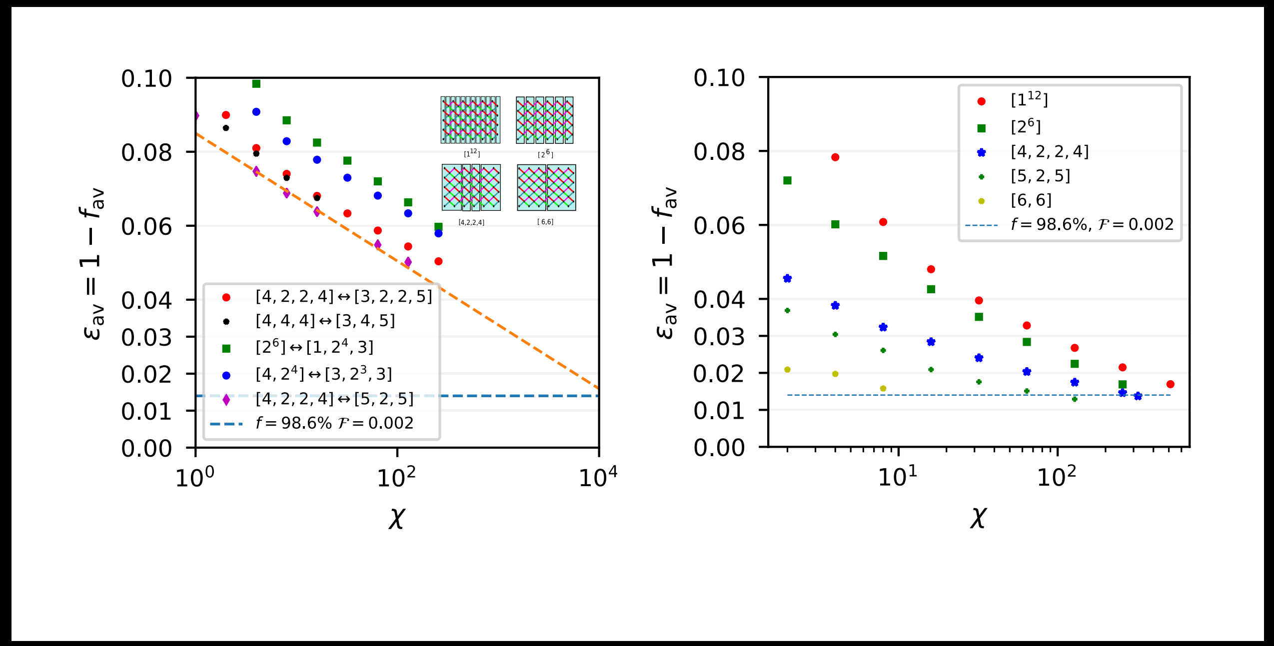
Performance comparison with real quantum devices

[Google, Nature **574**, 505-510 (2019)]



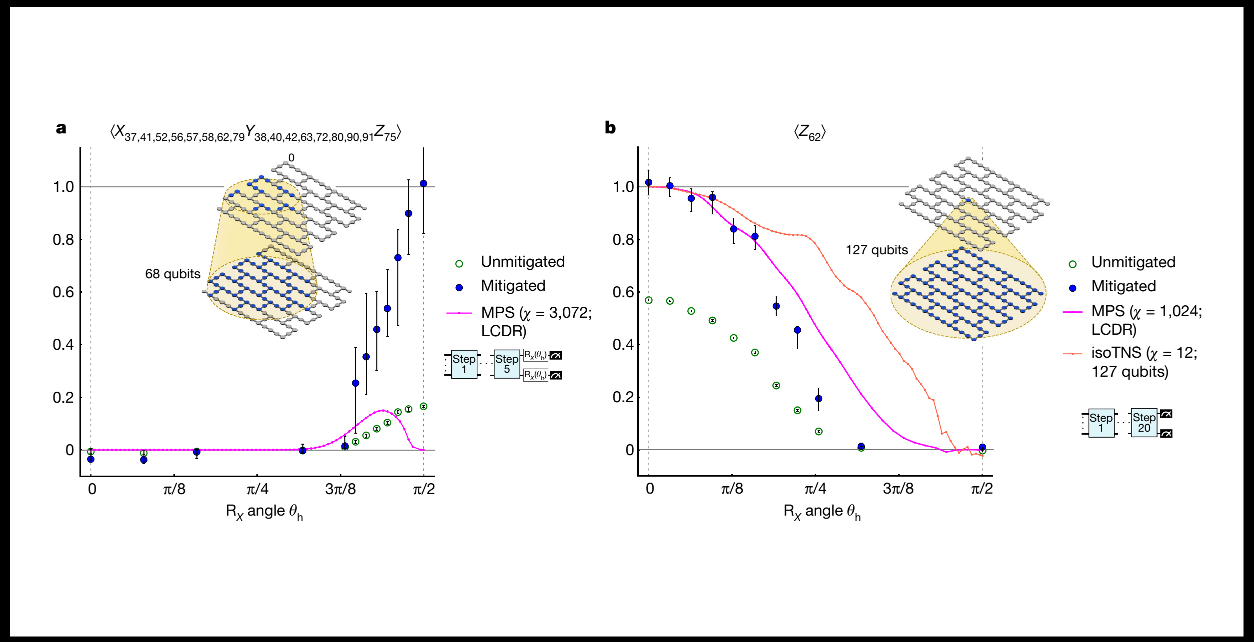
Real-device experiment
for random circuits

[Zhou et al, PRX **10**, 041038 (2020)]



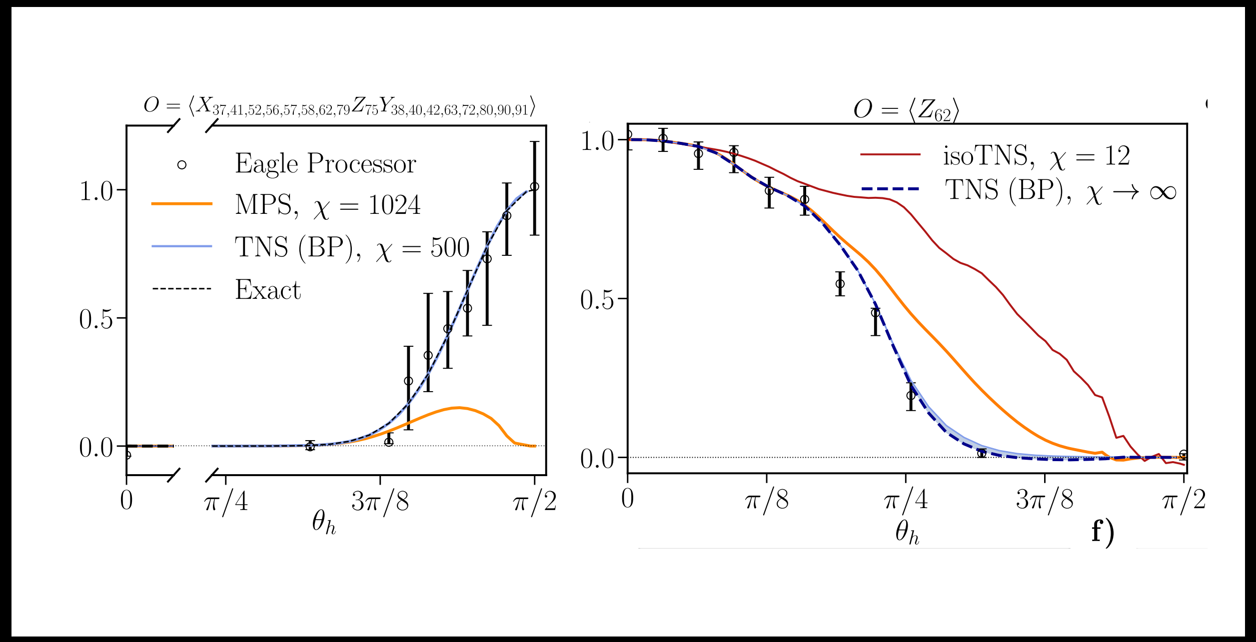
Tensor network simulation

[IBM, Nature **618**, 500 (2023)]



Real-device experiment
for quench dynamics

[Tindall et al., arXiv:2306.14887]



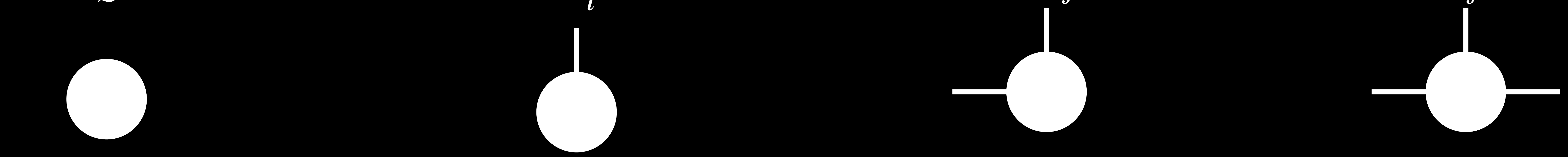
Tensor network simulation
(using Belief propagation technique)

No one knows the limit of performance.

Computational Materials Science

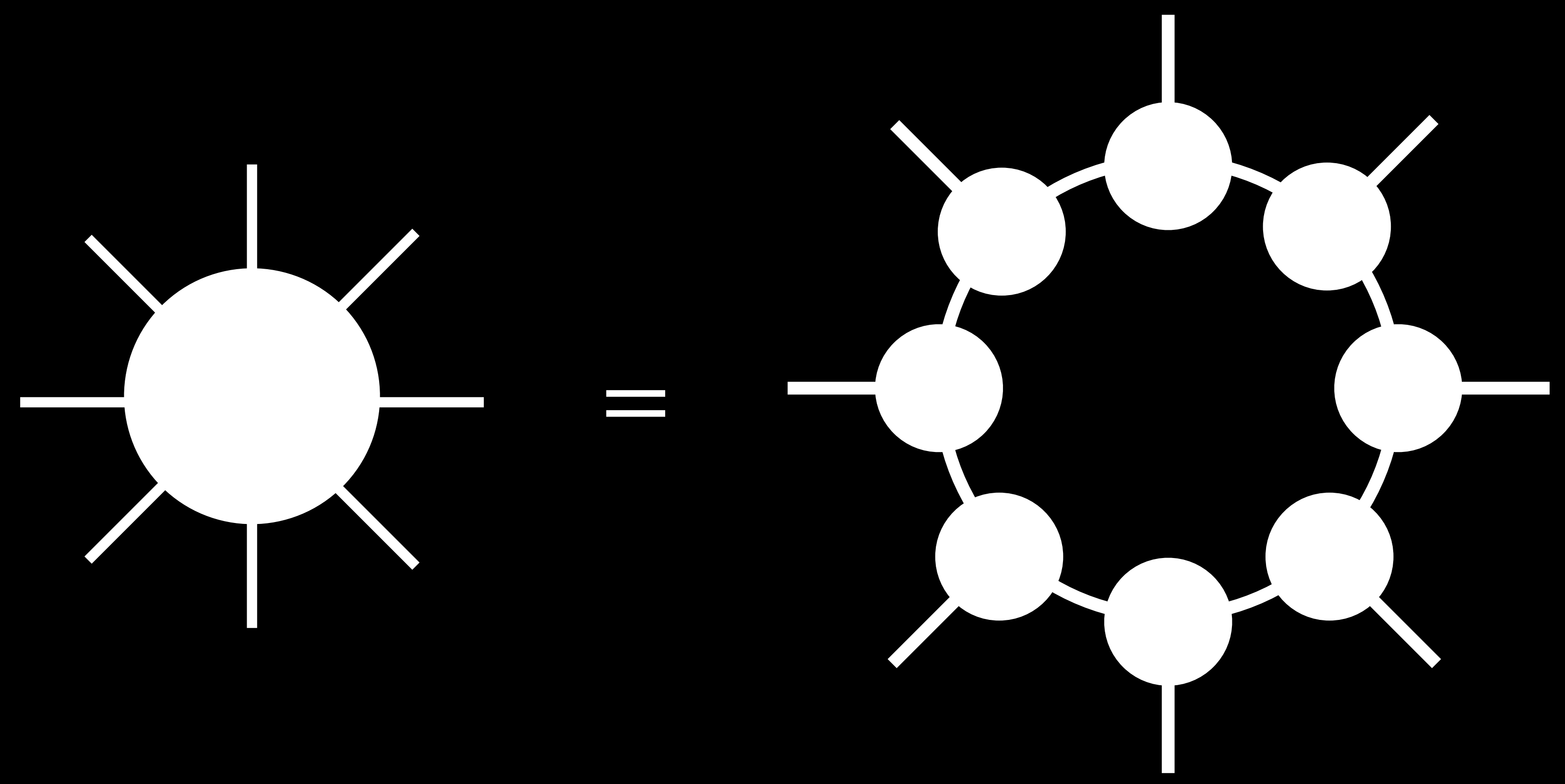
Tensor と Tensor network

rank-n tensor: n-dimensional array

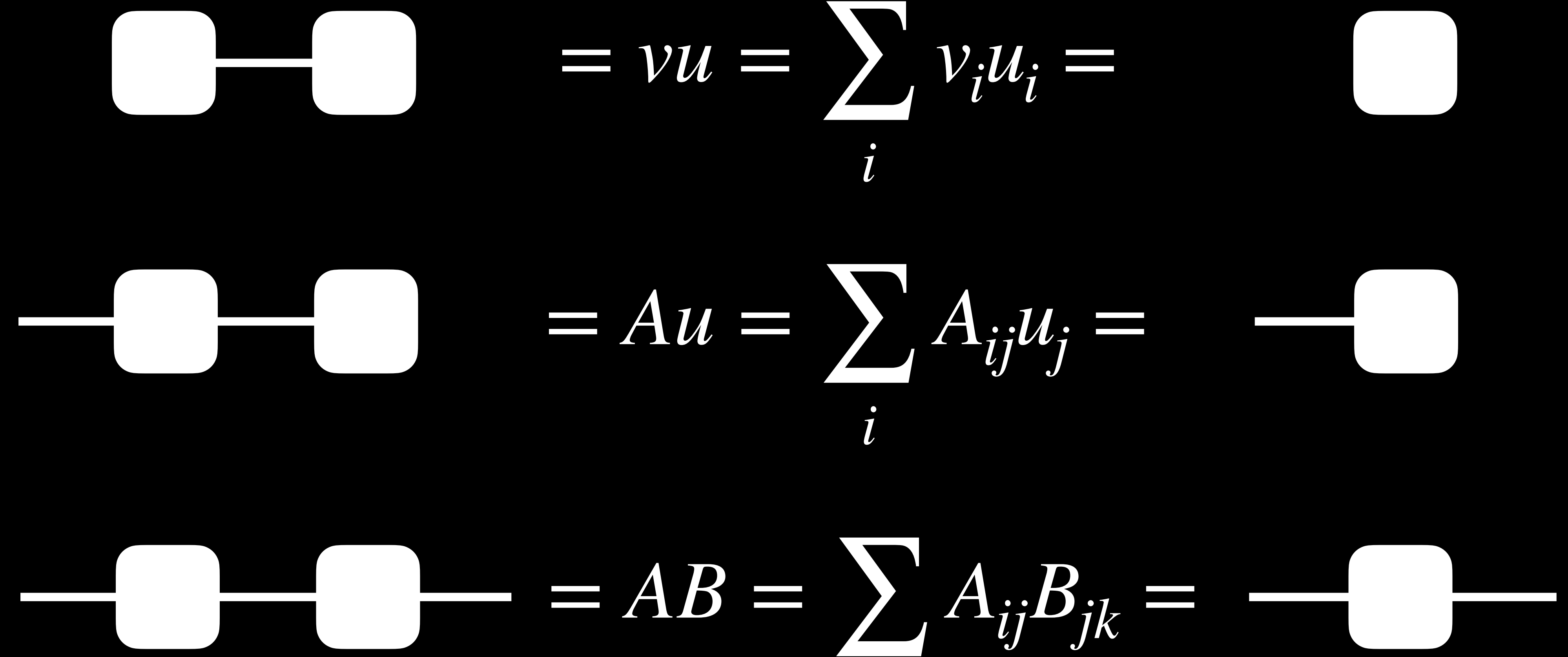


S	V_i	M_{ij}	T_{ijk}
rank-0 tensor = scaler	rank-1 tensor = vector	rank-2 tensor = matrix	rank-3 tensor

tensor network: a decomposition of a **big** tensor



Contraction rule



$$= vu = \sum_i v_i u_i =$$

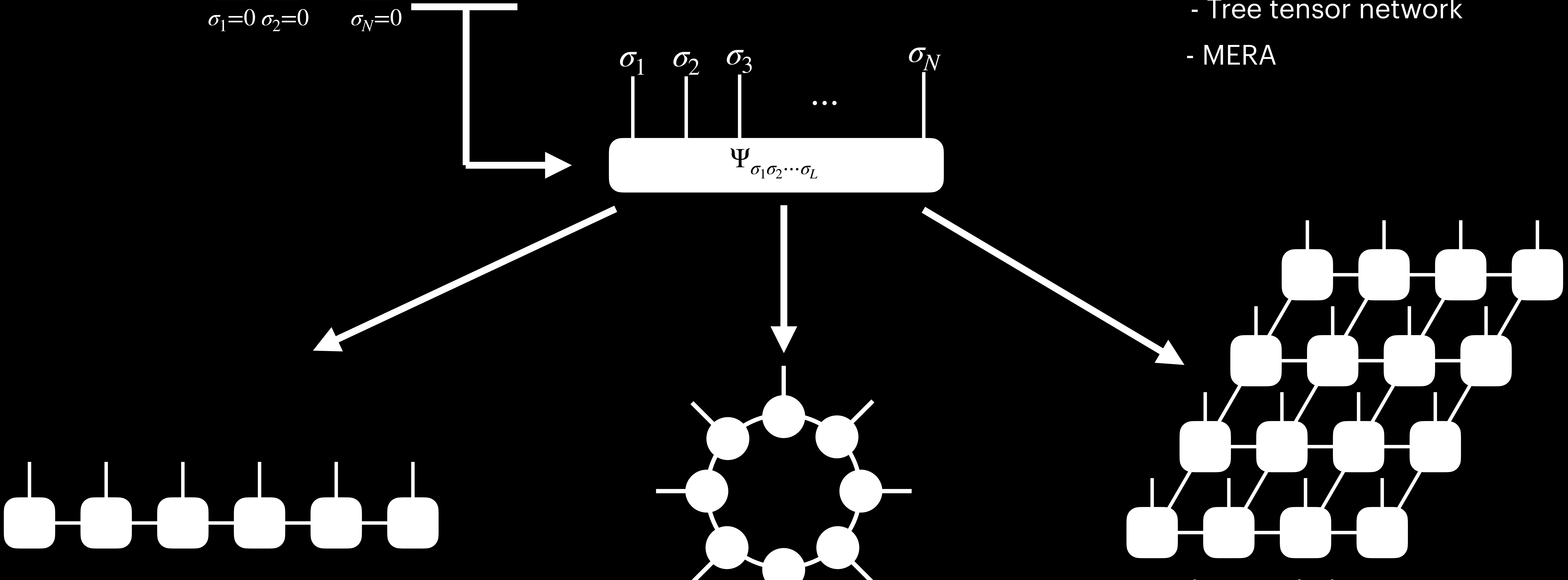
$$= Au = \sum_i A_{ij} u_j =$$

$$= AB = \sum_i A_{ij} B_{jk} =$$

Tensor network state

$$|\Psi\rangle = \sum_{\sigma_1=0}^{d-1} \sum_{\sigma_2=0}^{d-1} \cdots \sum_{\sigma_N=0}^{d-1} \psi_{\sigma_1\sigma_2\cdots\sigma_N} |\sigma_1\sigma_2\cdots\sigma_N\rangle$$

- Tree tensor network
- MERA

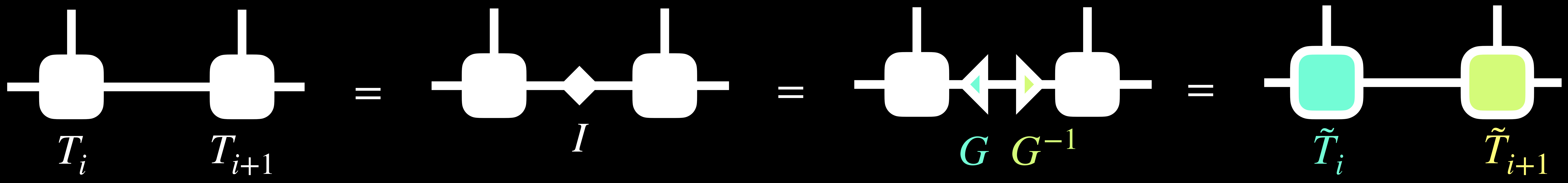


Matrix Product State (**MPS**)

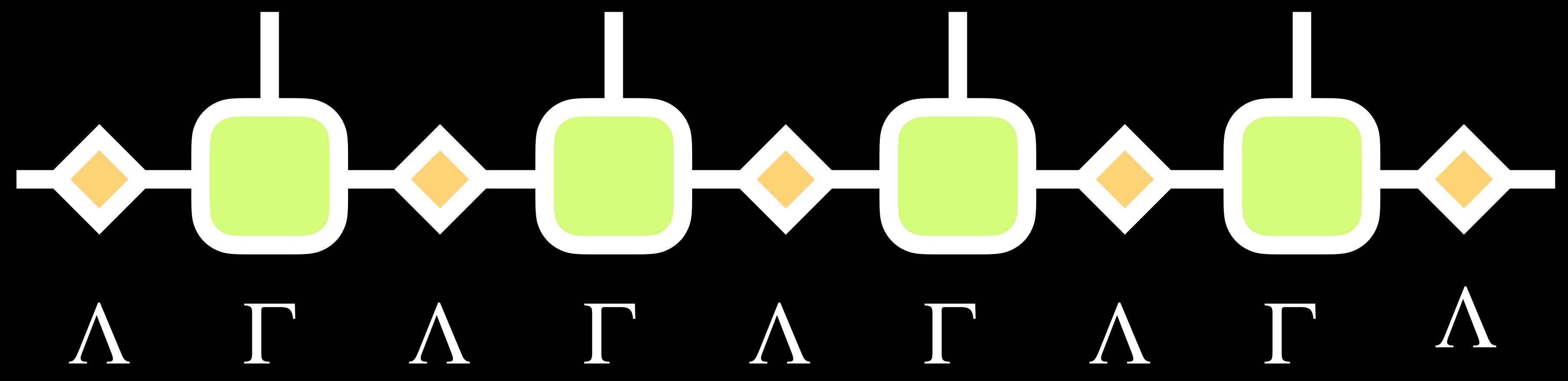
Projected Entangled-Pair State (**PEPS**)

Computational Materials Science

Gauge degree of freedom and canonical form



Canonical form (Vidal form) of MPS [Vidal, PRL 2003, Cirac et al., KMP, 2021]



$A = \Lambda\Gamma \rightarrow A^\dagger A = I$

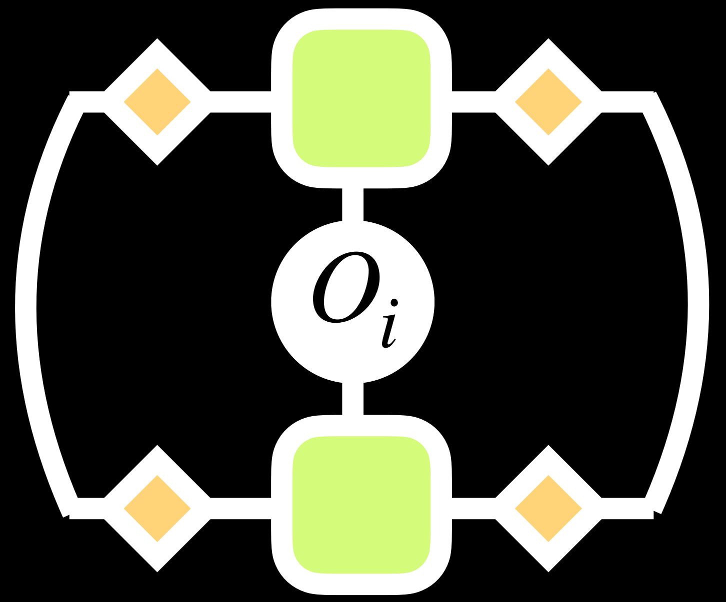
$B = \Gamma\Lambda \rightarrow BB^\dagger = I$

$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_\chi)$

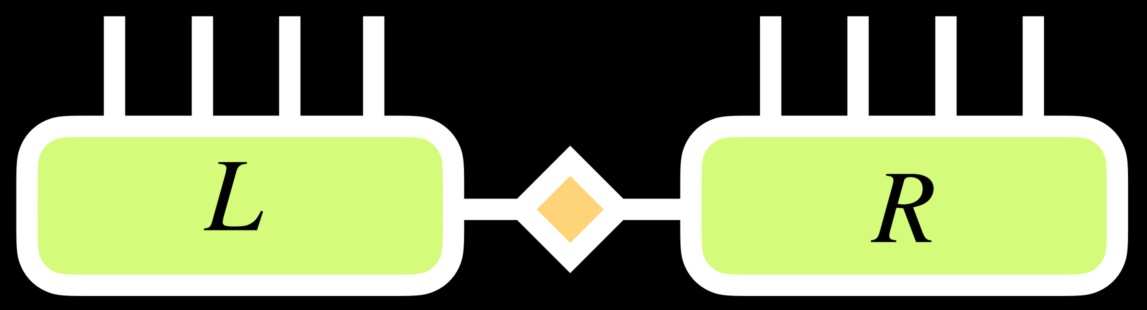
$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\chi$

local observable

$\langle O_i \rangle =$



entanglement



$$S = -\sum_\alpha \lambda_\alpha^2 \ln \lambda_\alpha^2$$

best truncation

$\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_{\tilde{\chi}}, \lambda_{\tilde{\chi}+1}, \cdots, \lambda_\chi)$

↓

$\tilde{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_{\tilde{\chi}}, \cancel{\lambda_{\tilde{\chi}+1}, \cdots, \lambda_\chi})$

MPS algorithms for quantum many-body physics

Ground state method

Density-marix renormalization group method (DMRG) [White 1992]

Real-time dynamics

Time evolving block decimation (TEBD) [Vidal 2003]

Time-dependent variational principle (TDVP) [Haegeman et al., 2011]

Finite temperatur

Minimally entangled typically thermal state (METTS) [White 2009]

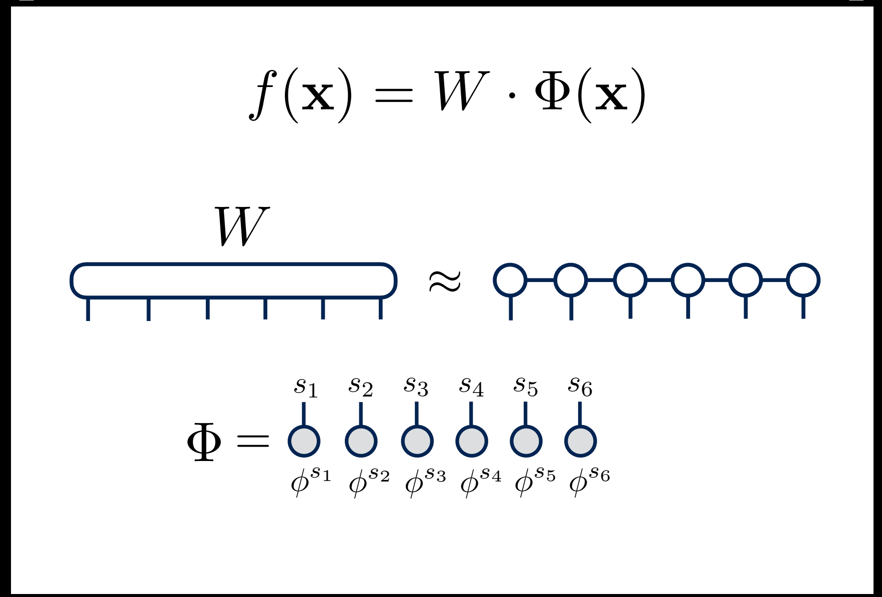
Thermal pure quantum matrix product state (TPQ-MPS) [Iwaki et. al., 2021]

Computational Materials Science

Application field

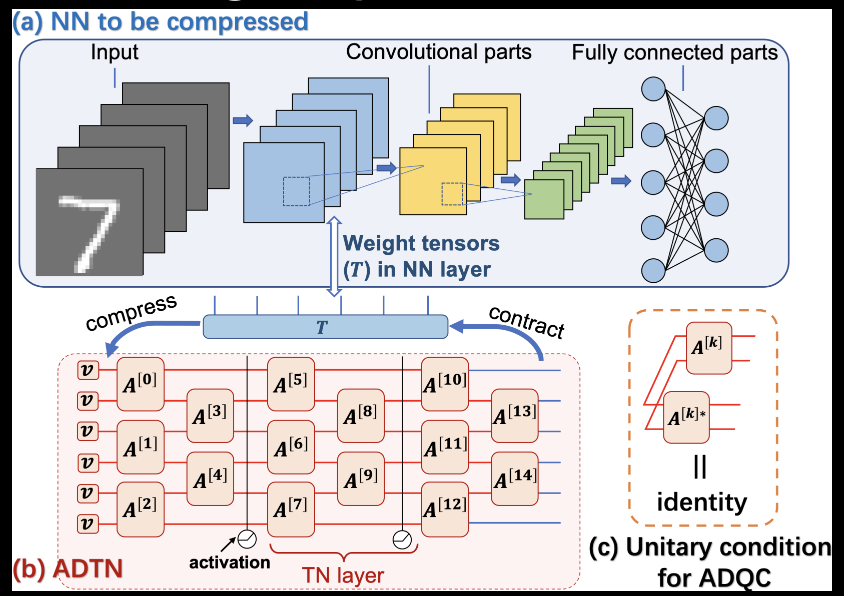
- It has recently attracted attention as an efficient representation of **machine learning** models and as a highly efficient compression method for **big data**.

[Stoundemire&Schwab, 2017]



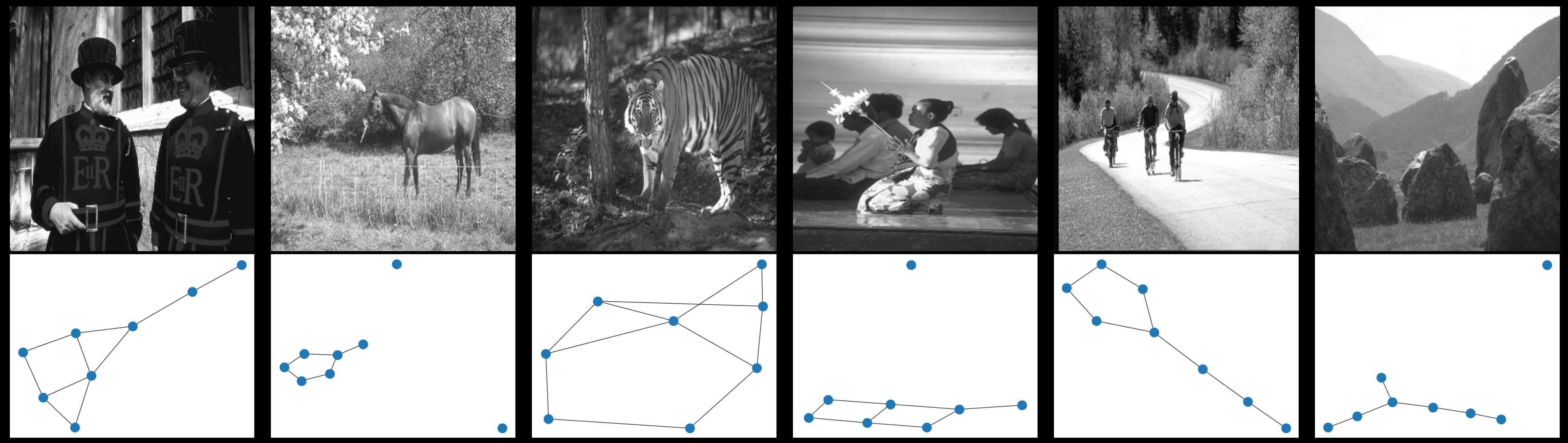
Tensor network as the learning model

[Shi-Ju Ran group, arXiv:2305.06058]



Tensor network as a component of deep neural network

[Chao Li et al (AIP tensor learning team), 2020, 2022, 2023]



Tensor network as a compression of big data / Tensor network structure search problem

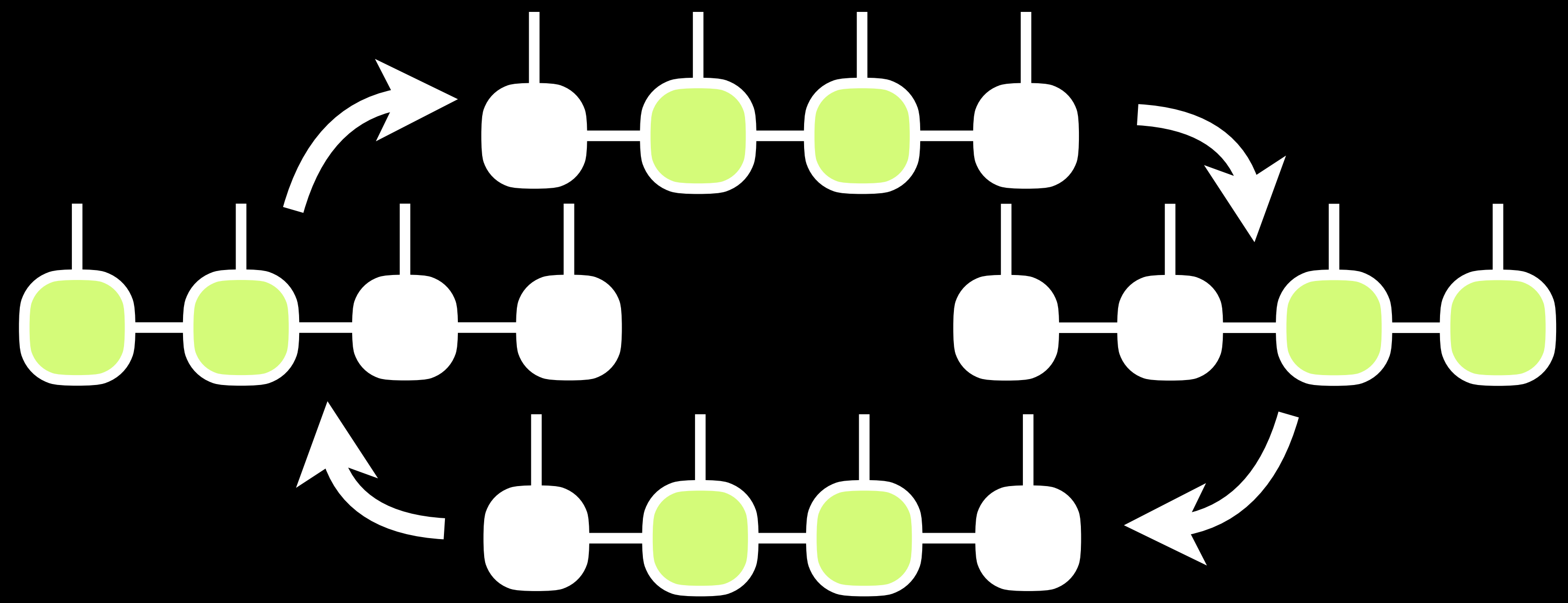
Computational Materials Science

Limitations in current TNS development

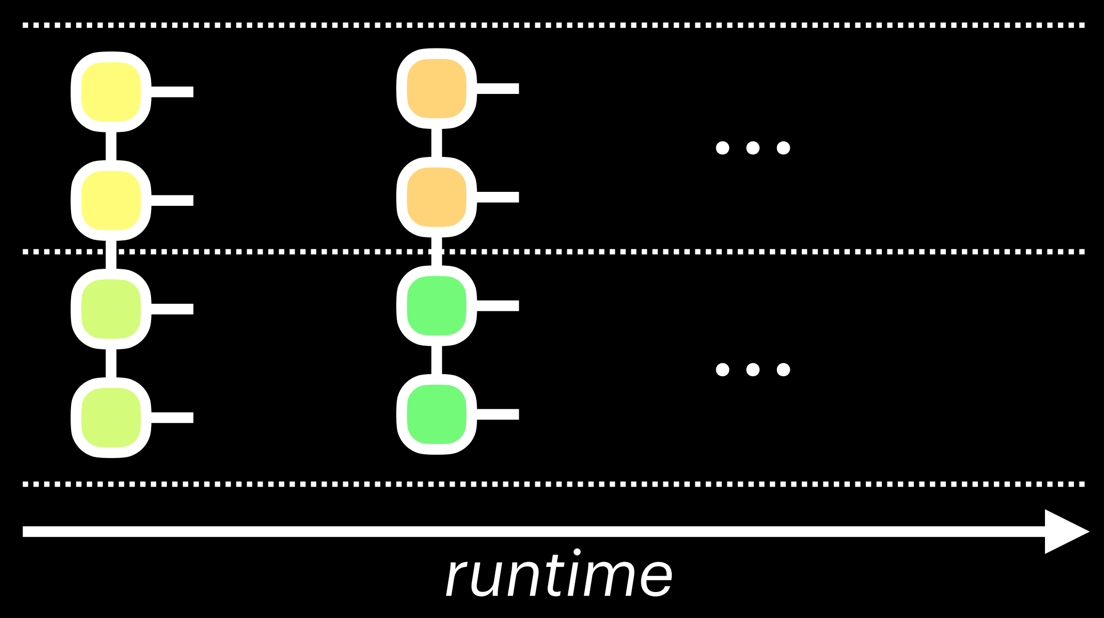
	Limitations	Solution
Algorithms	Sequential nature	Real-space parallelization
Data structure	Shared-memory implementation	Distributed-memory implementation

Algorithms

Limitations: Sequential nature



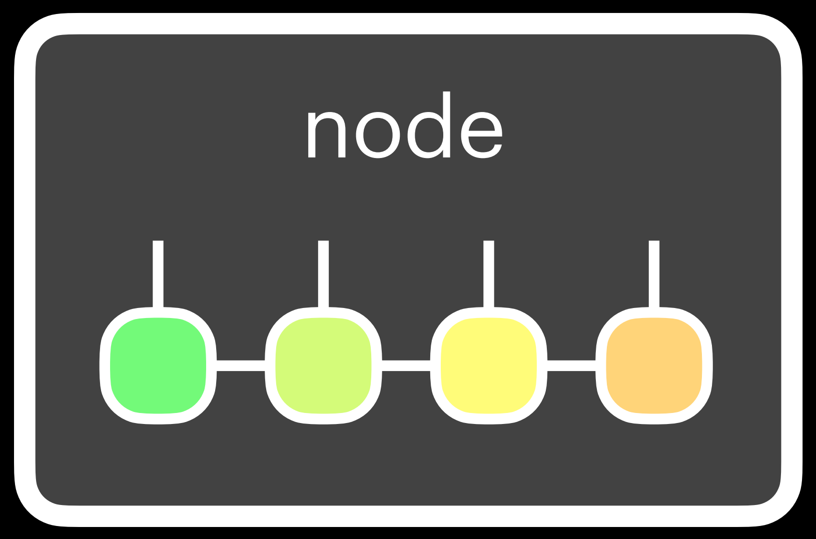
Solution: Real-space parallelization



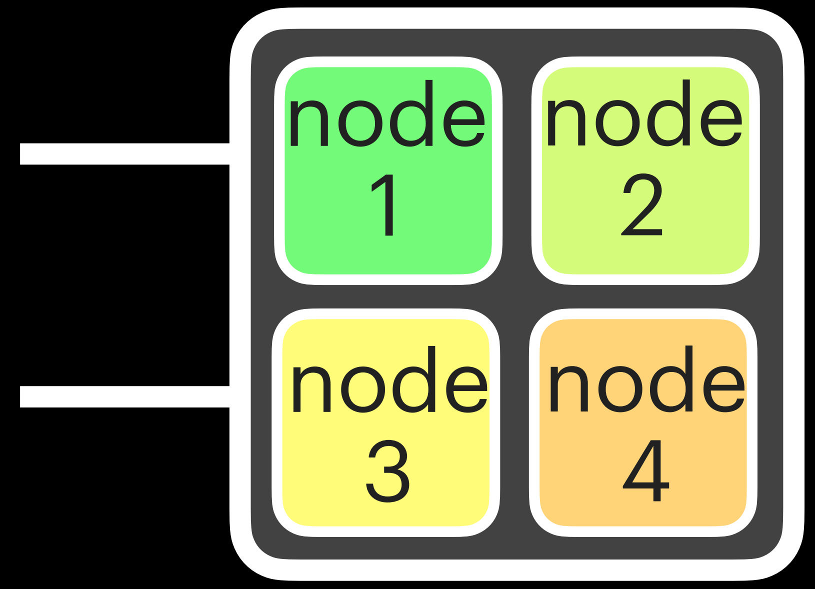
parallel algorithm

Data structure

Limitations: Shared-memory implementation



Solution: Distributed-memory implementation



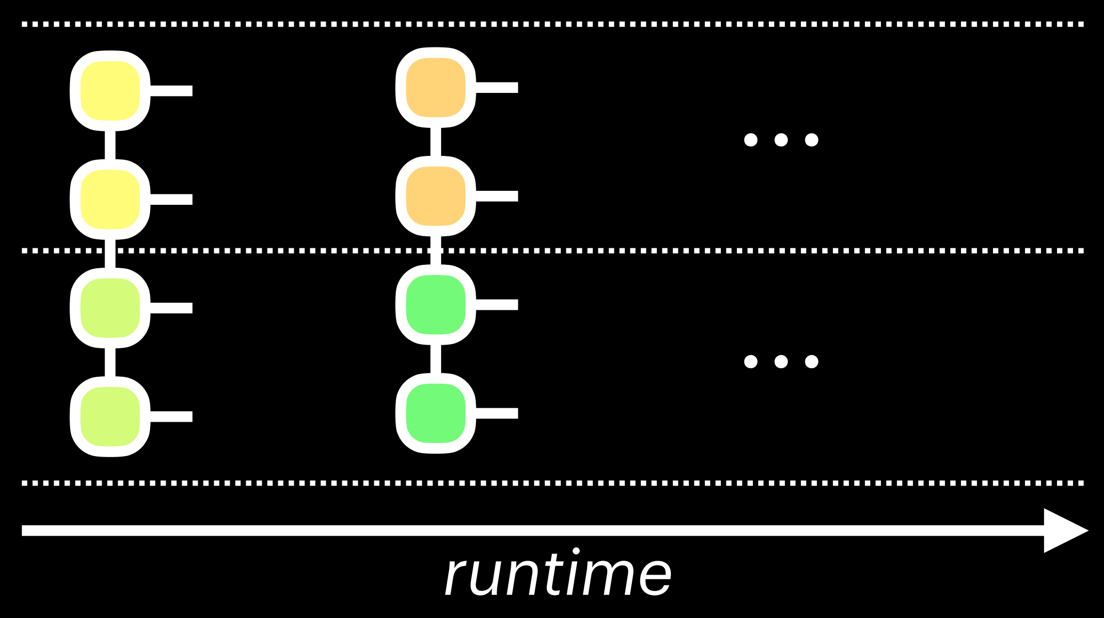
data parallel

Our contributions

Real-space parallelizable MPS algorithm

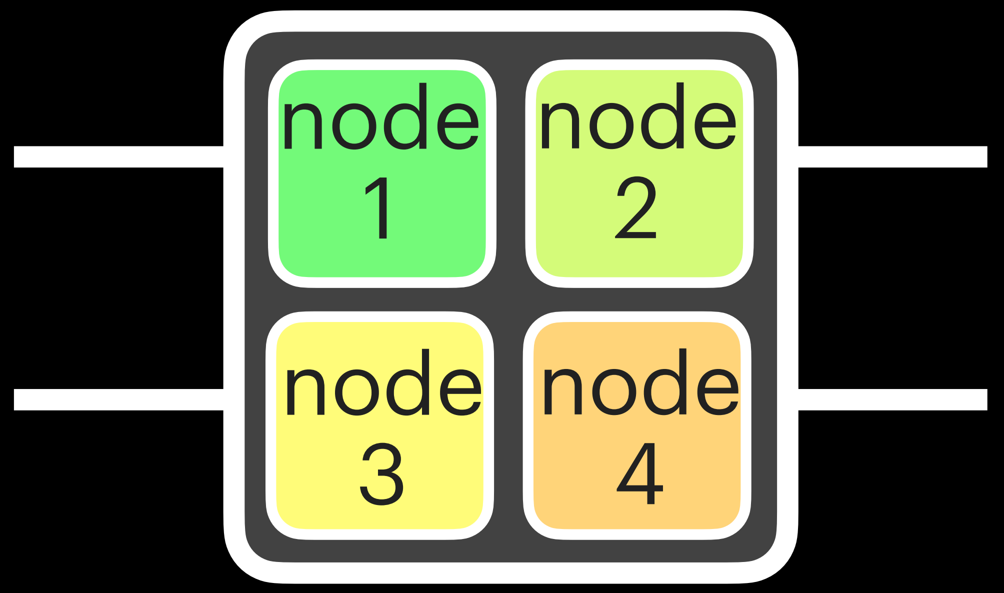
A case study: parallel TEBD (**pTEBD**)

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

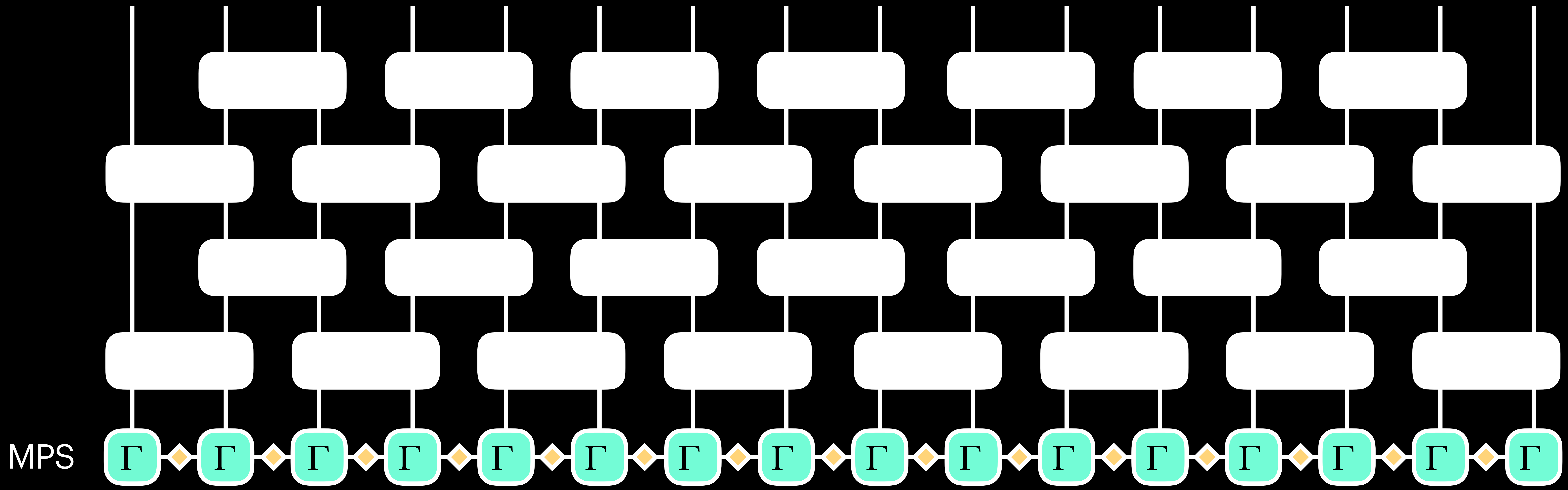


TNS software for Fugaku

https://github.com/gracequantum



Time-evolving block decimation (TEBD)

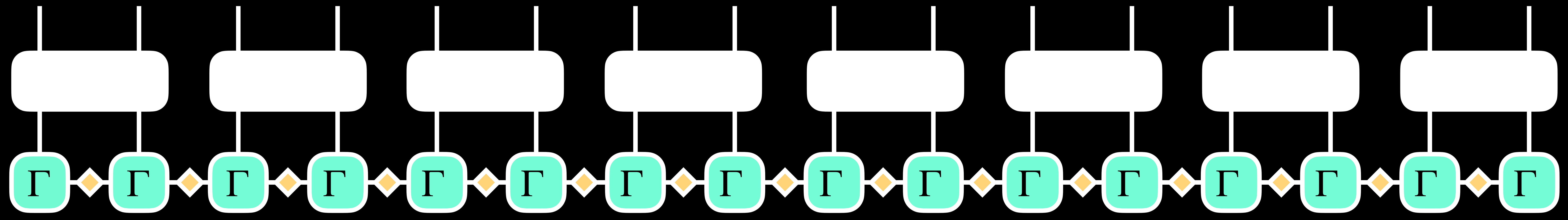


When performing time evolution calculations on MPS, the simplest method is to calculate Trotter slices called **time-evolving block decimation**.

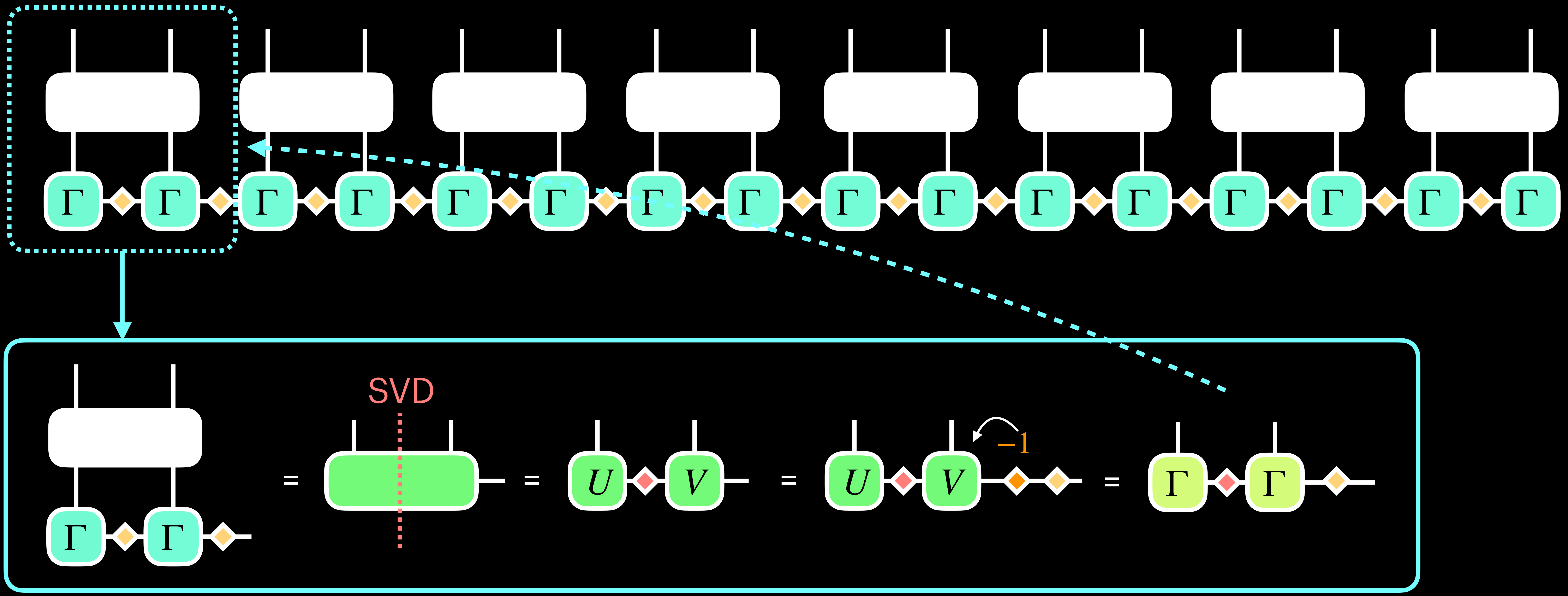
Quantum computing is a time-evolution starting from a trivial initial state (a **direct product state**).

A direct product state is a matrix product state with bond dimension 1.

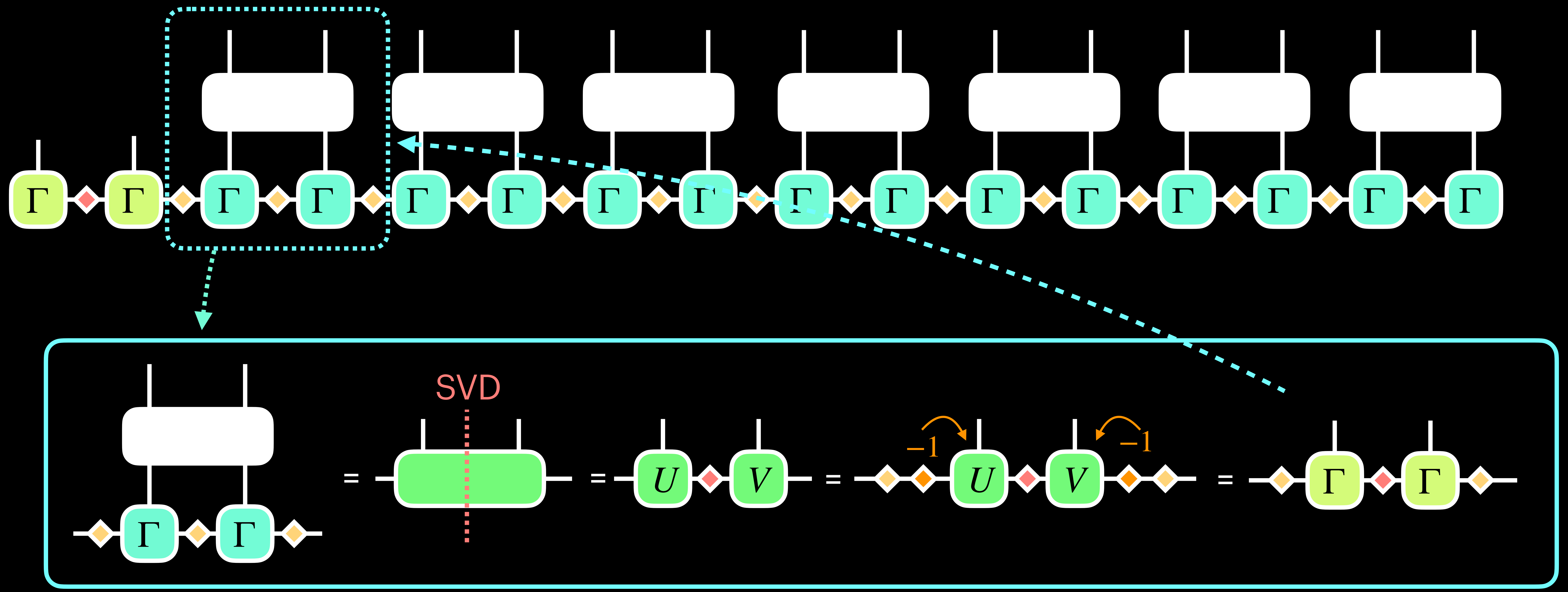
Time-evolving block decimation (TEBD)



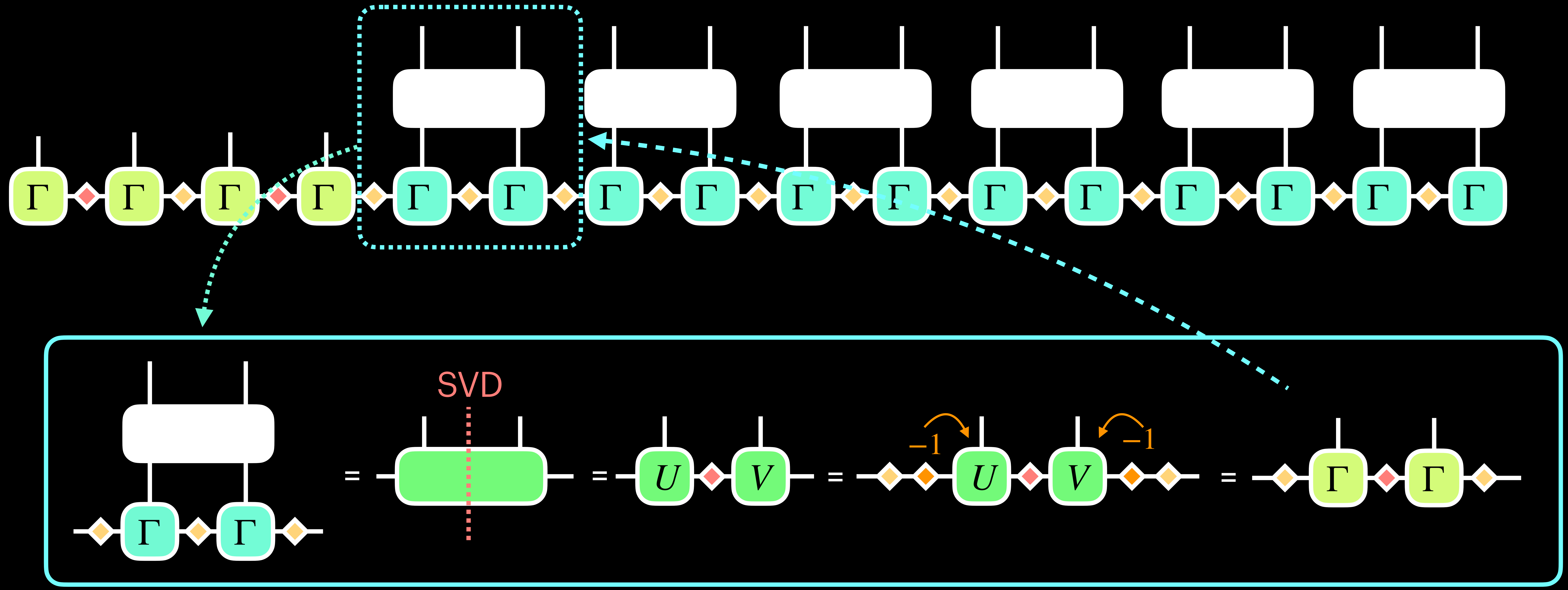
Time-evolving block decimation (TEBD)



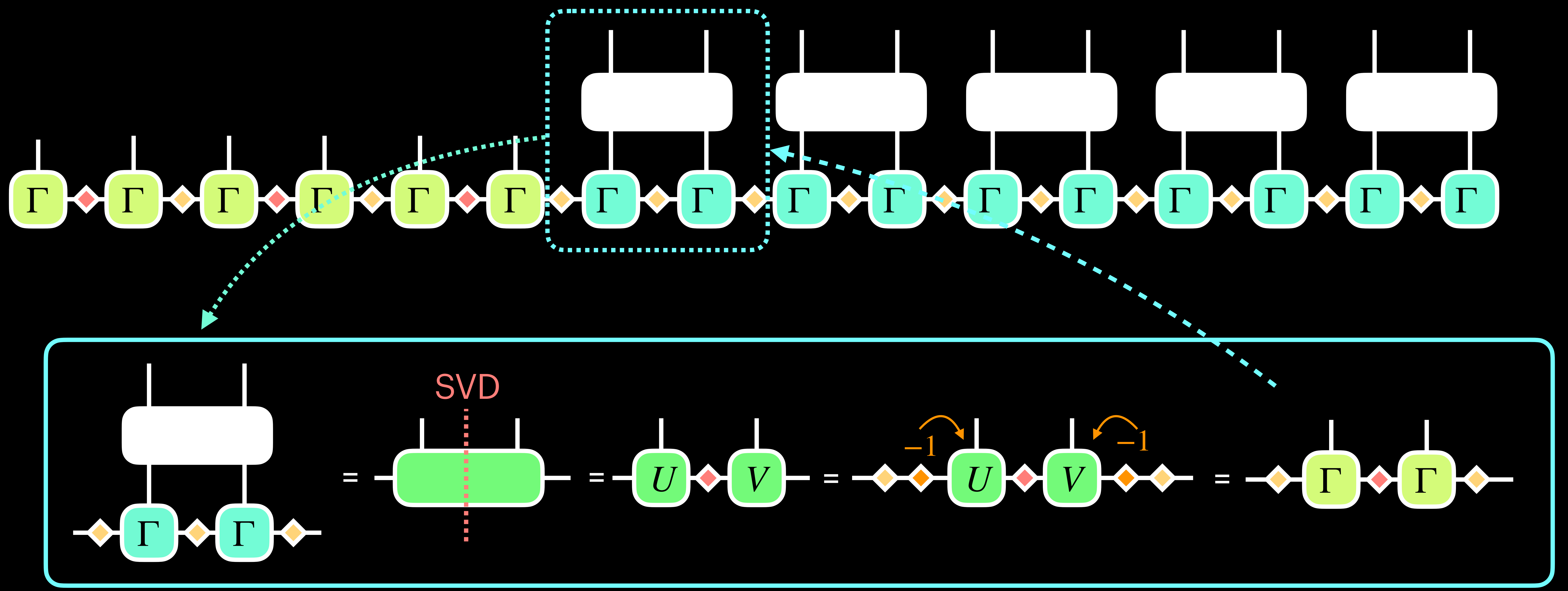
Time-evolving block decimation (TEBD)



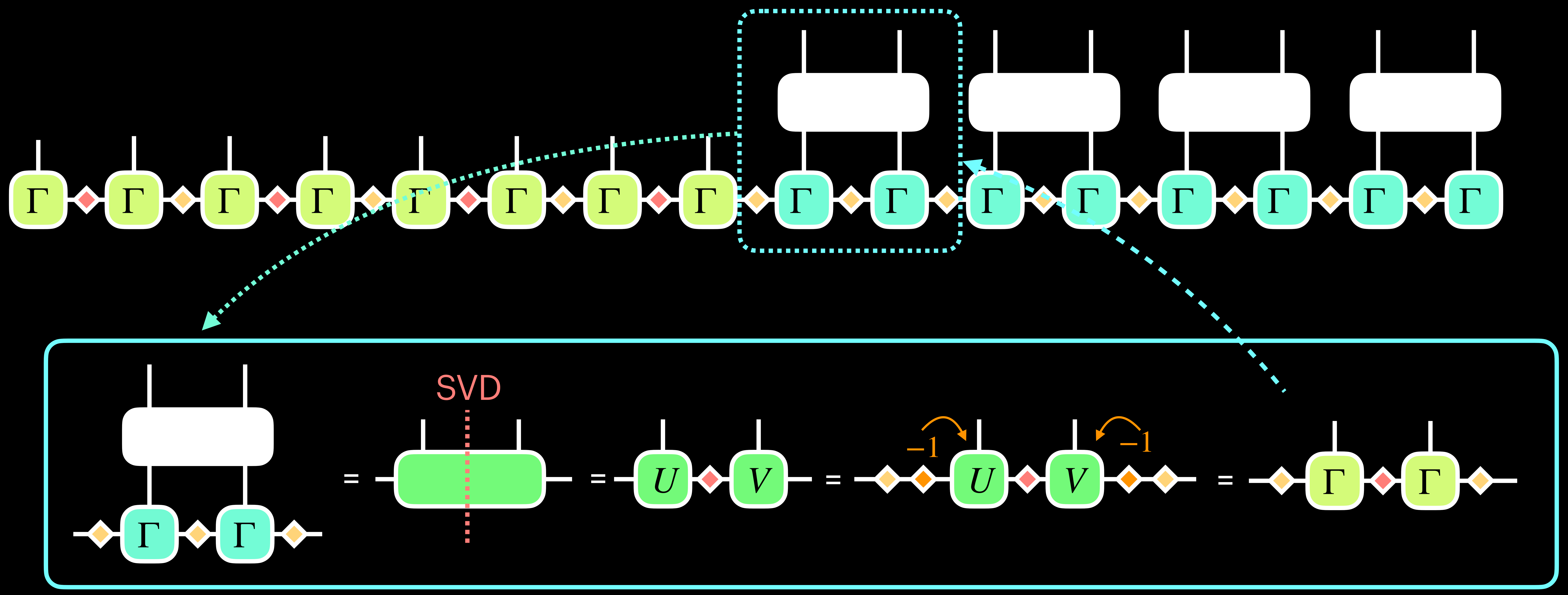
Time-evolving block decimation (TEBD)



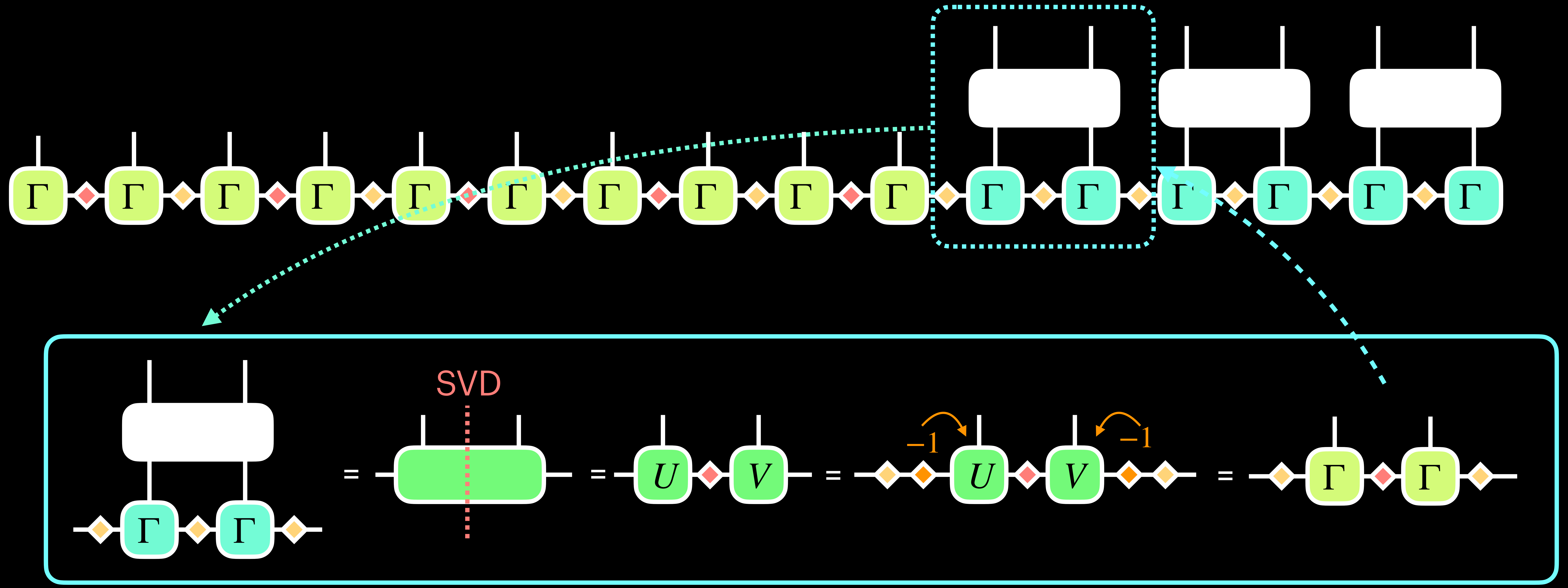
Time-evolving block decimation (TEBD)



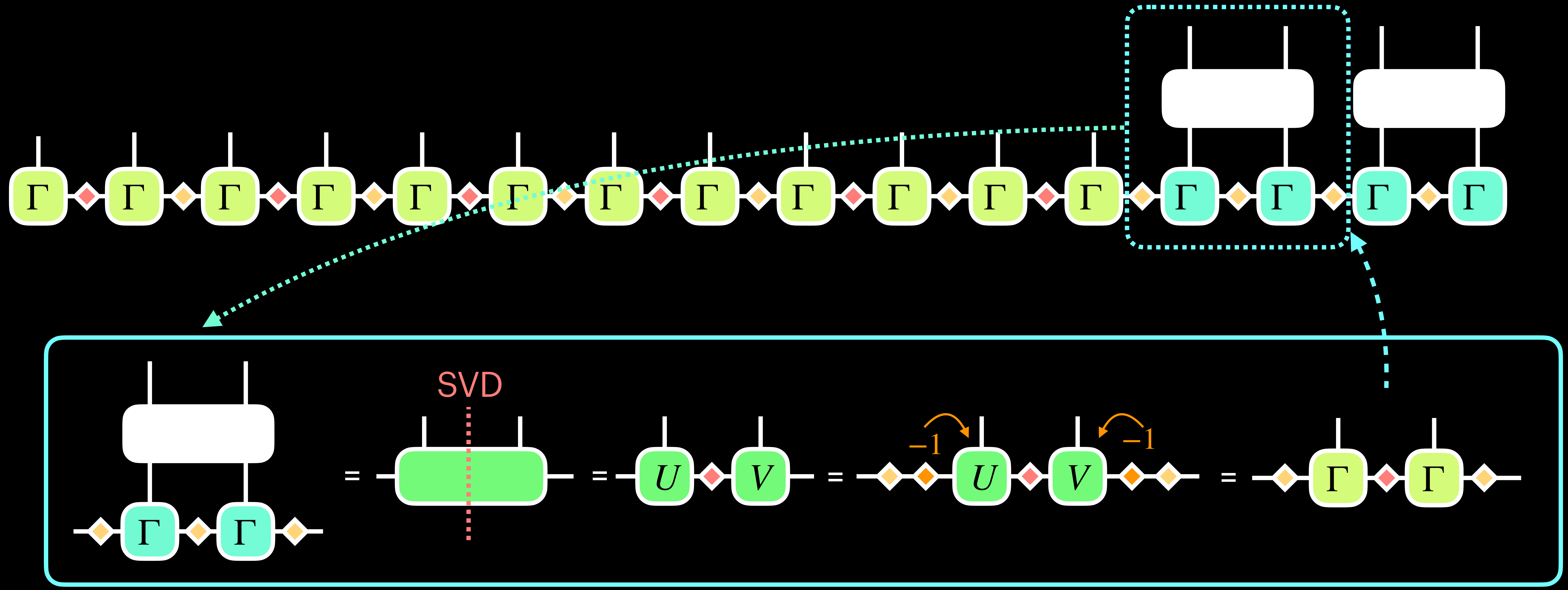
Time-evolving block decimation (TEBD)



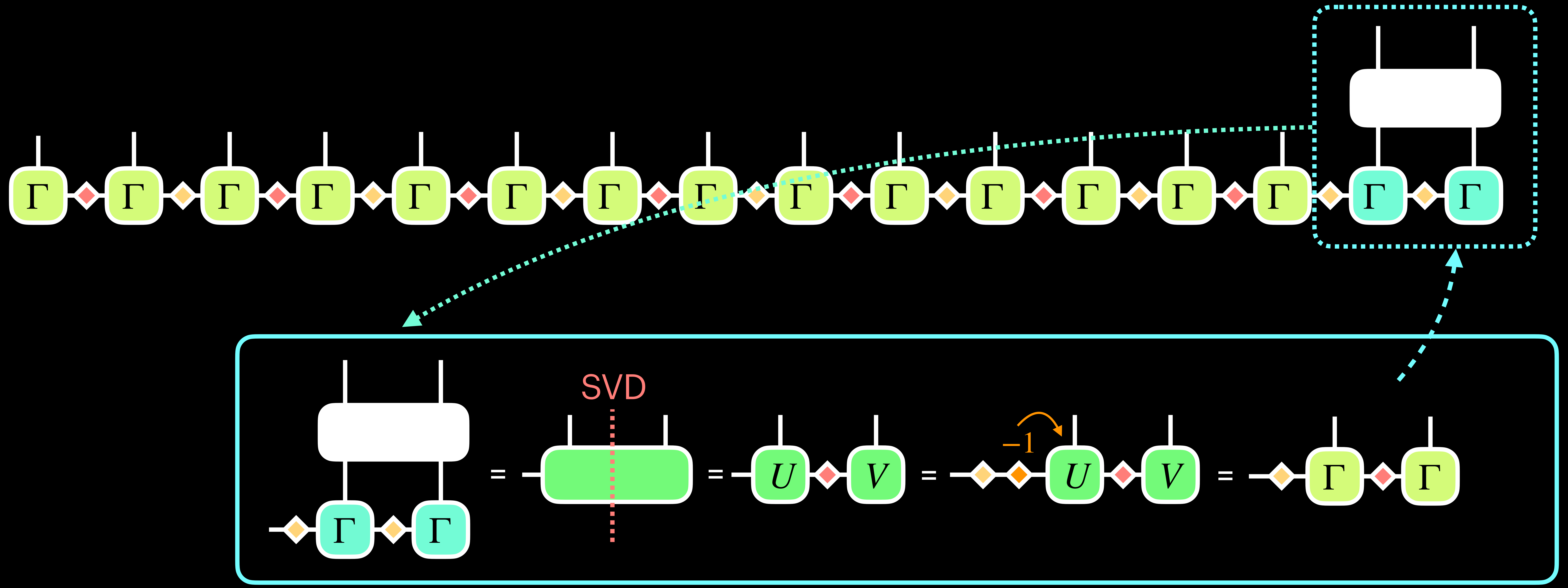
Time-evolving block decimation (TEBD)



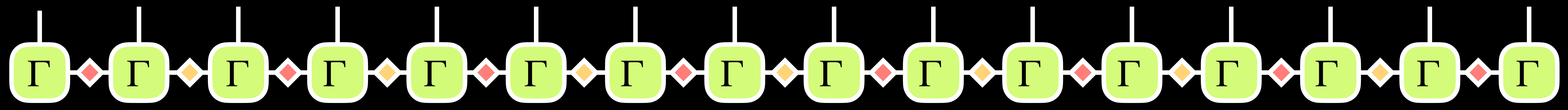
Time-evolving block decimation (TEBD)



Time-evolving block decimation (TEBD)

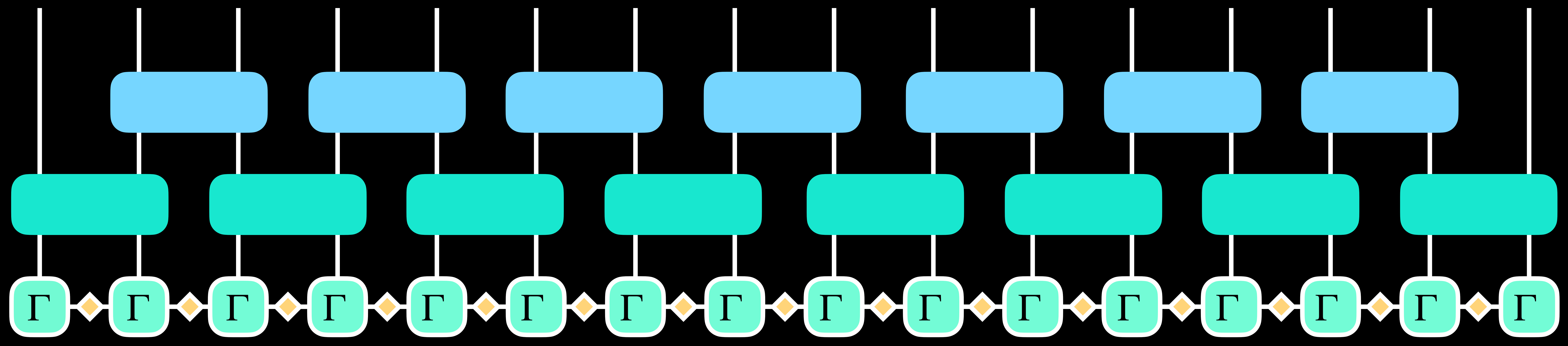


Time-evolving block decimation (TEBD)



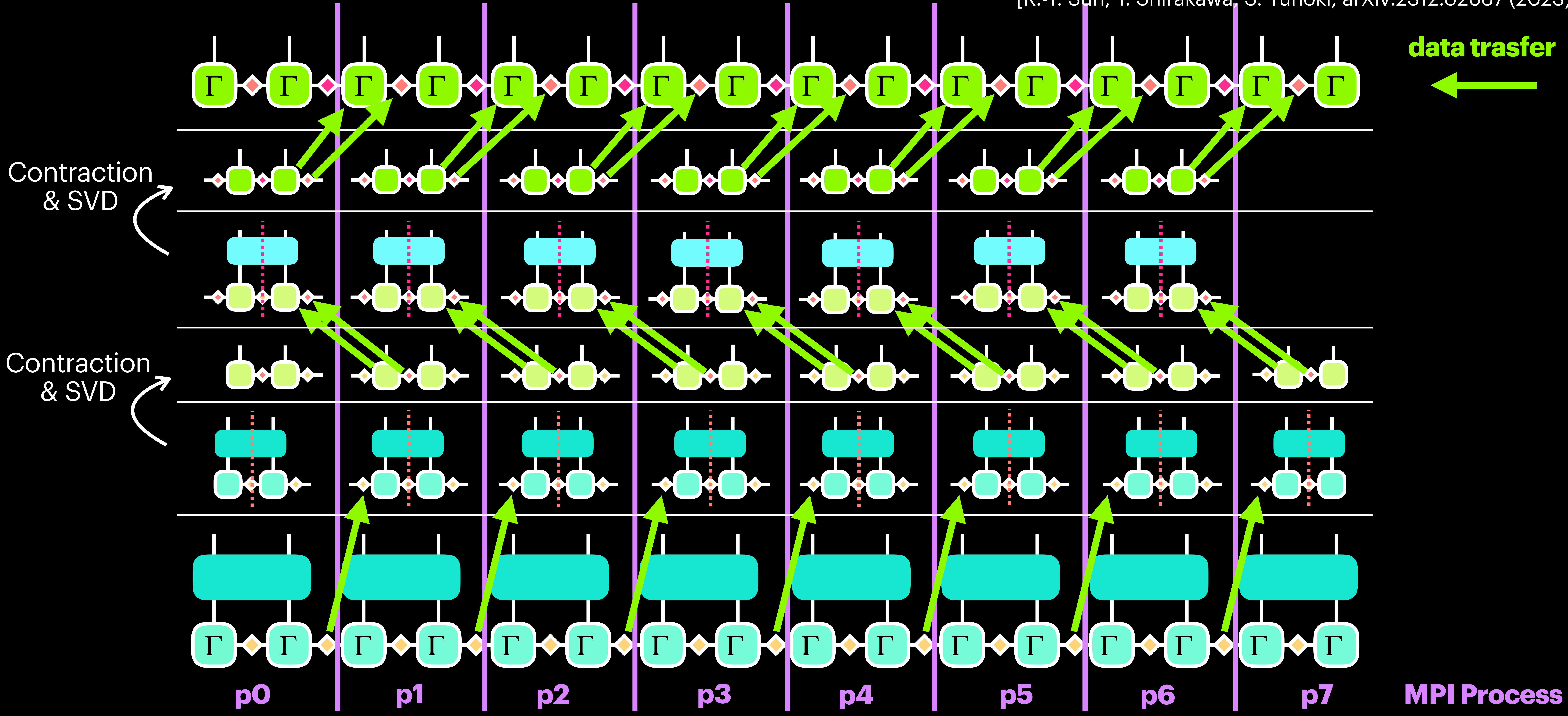
Parallelization of TEBD (pTEBD)

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]



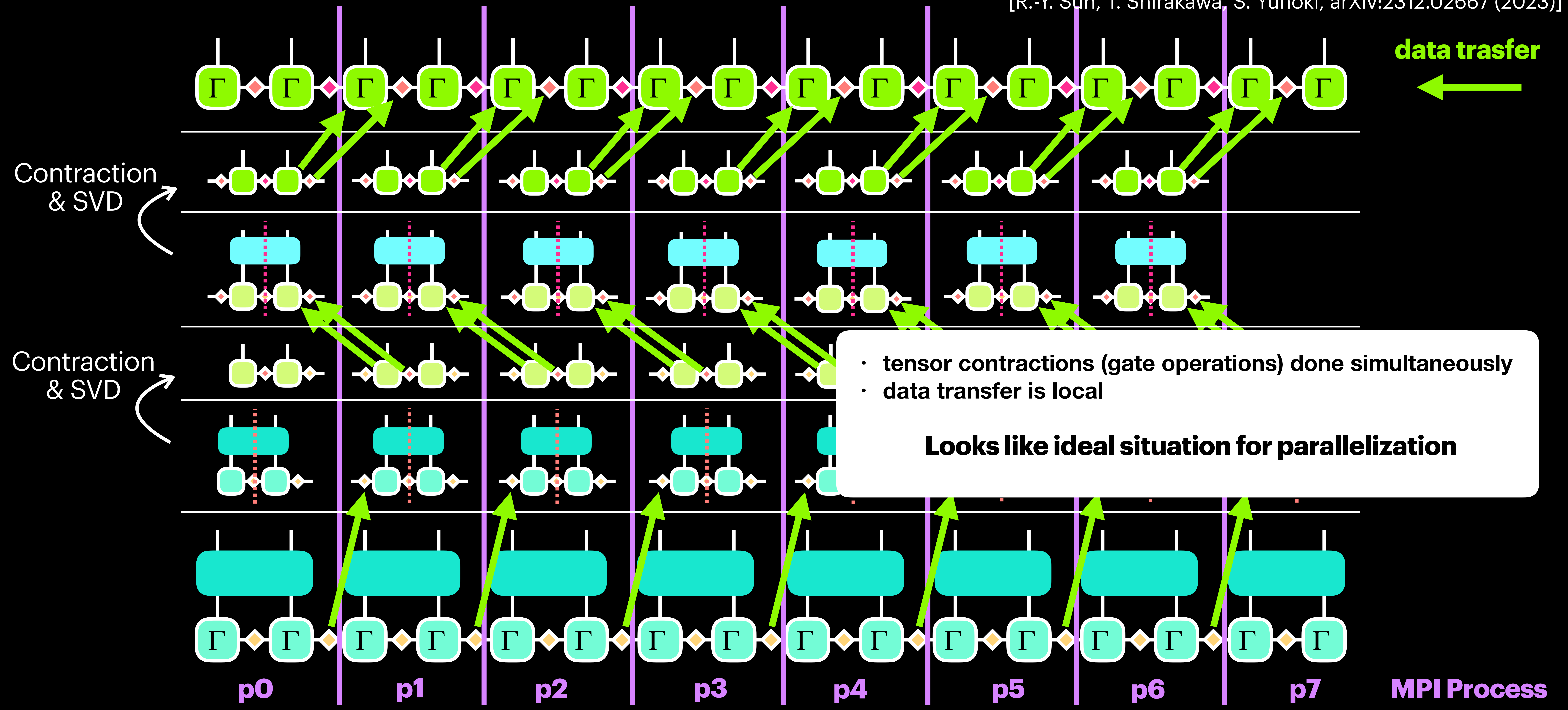
Parallelization of TEBD (pTEBD)

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]



Parallelization of TEBD (pTEBD)

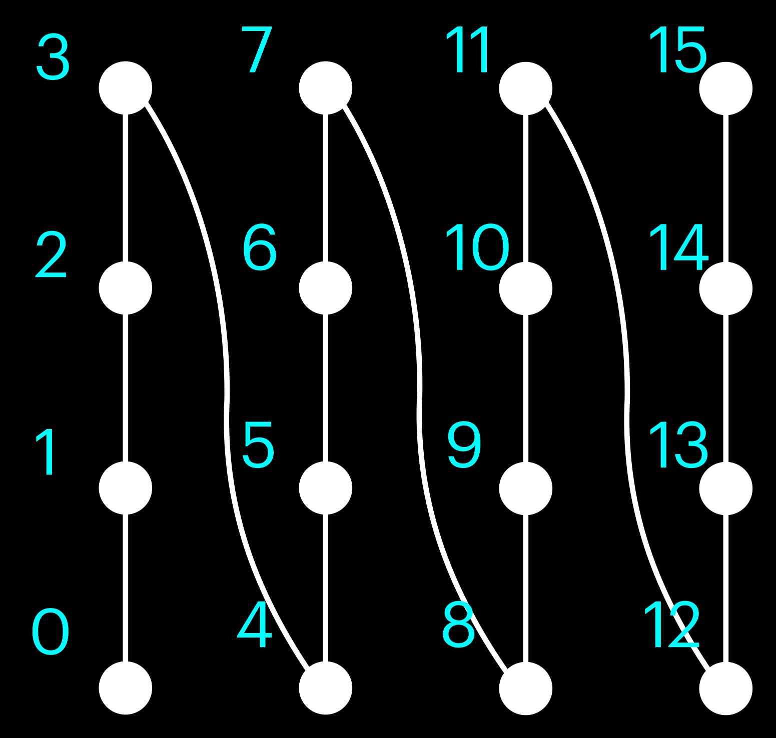
[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

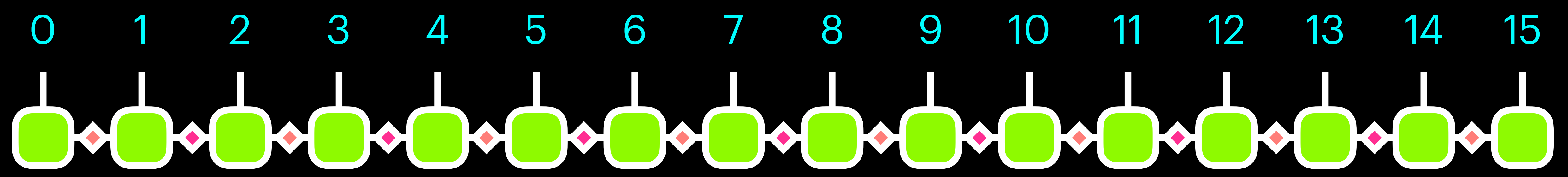


- tensor contractions (gate operations) done simultaneously
- data transfer is local

Looks like ideal situation for parallelization

Simulation for 2D quantum circuit

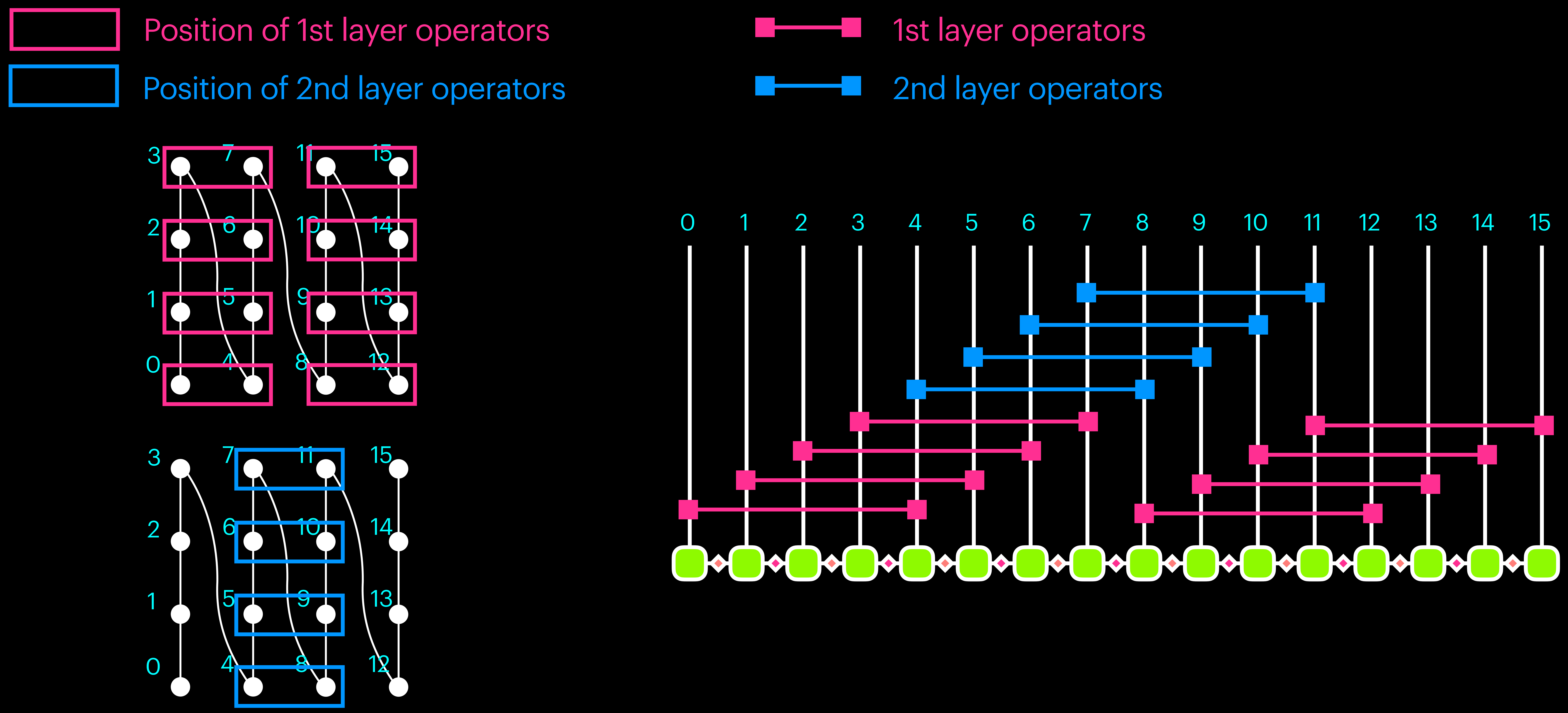




In order to calculate a 2D system using MPS, the 2D system is forcibly regarded as a 1D system.

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Simulation for 2D quantum circuit



Then, the nearest-neighbor operators in the 2D system become distant operators in the virtual 1D system.

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Simulation for 2D quantum circuit

Position of 1st layer operators

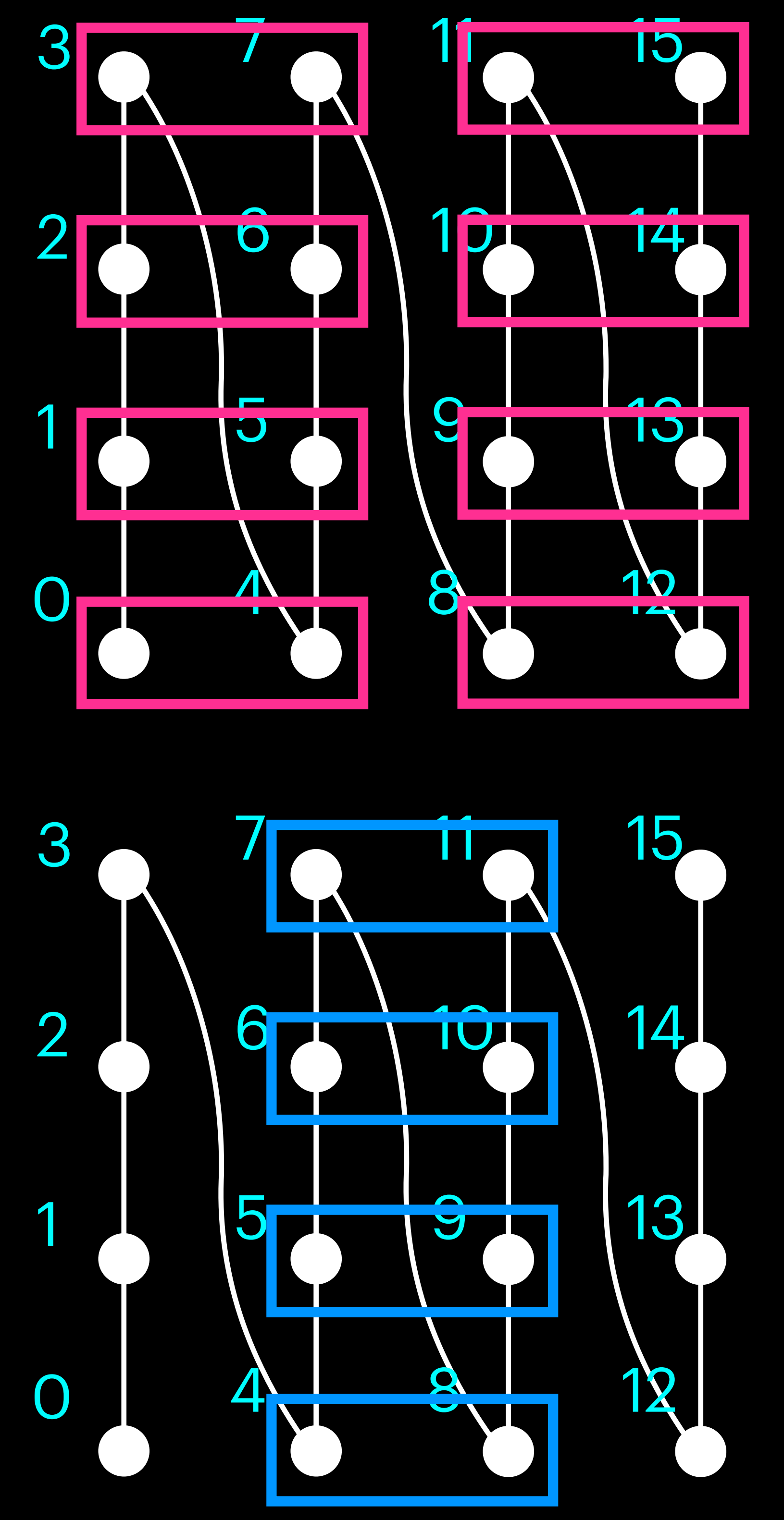
Position of 2nd layer operators

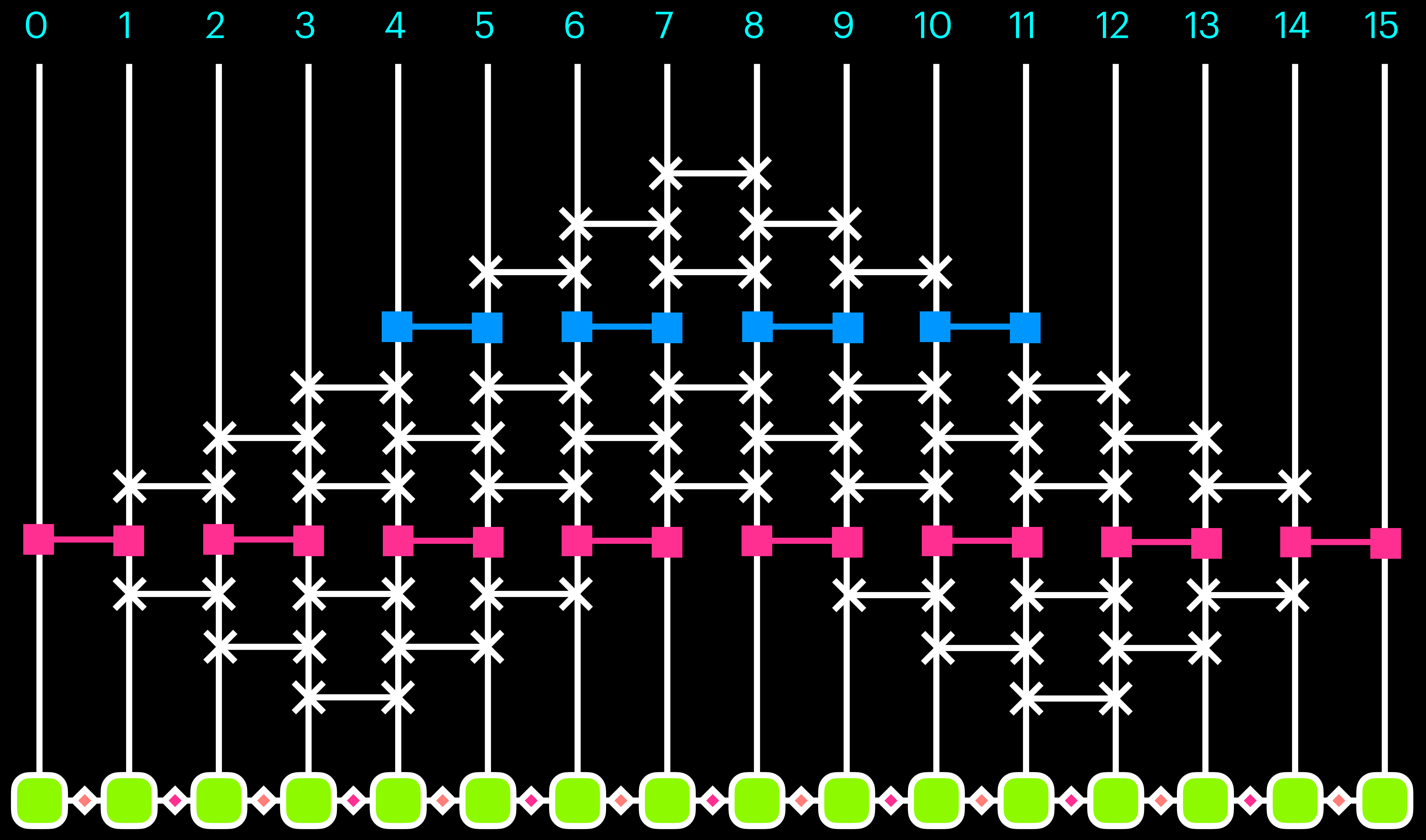
1st layer operators

2nd layer operators

SWAP operator S_{ij}

$S_{ij}\,|\sigma_i\sigma_j\rangle = |\sigma_j\sigma_i\rangle$





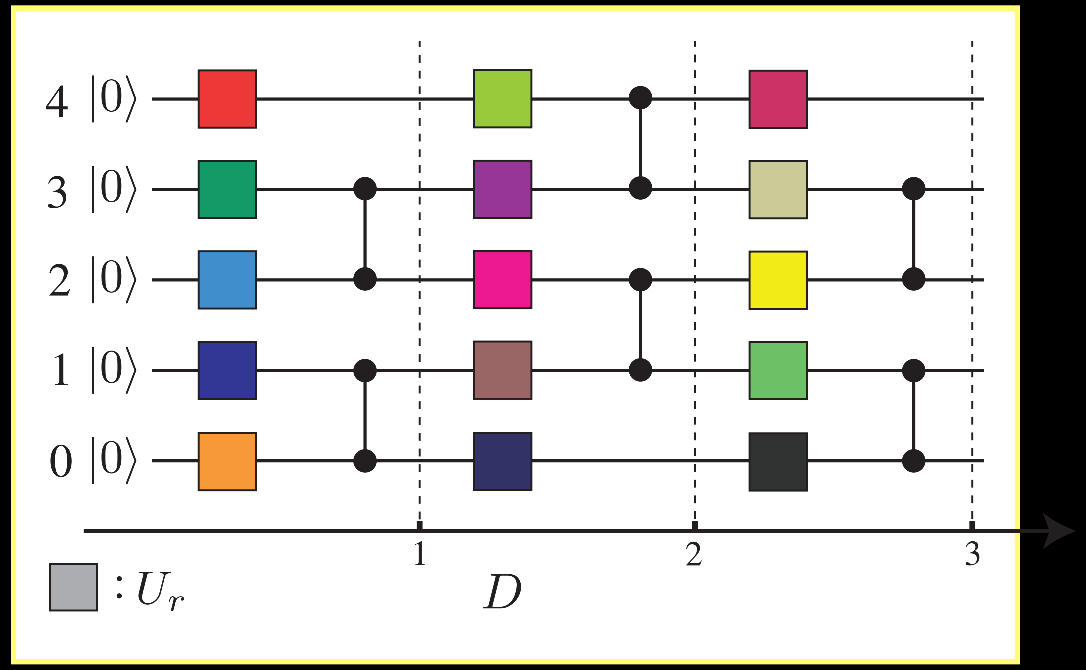
The simplest and most efficient way to handle these bonds in TEBD is by sandwiching the swap operator.

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Benchmark circuits

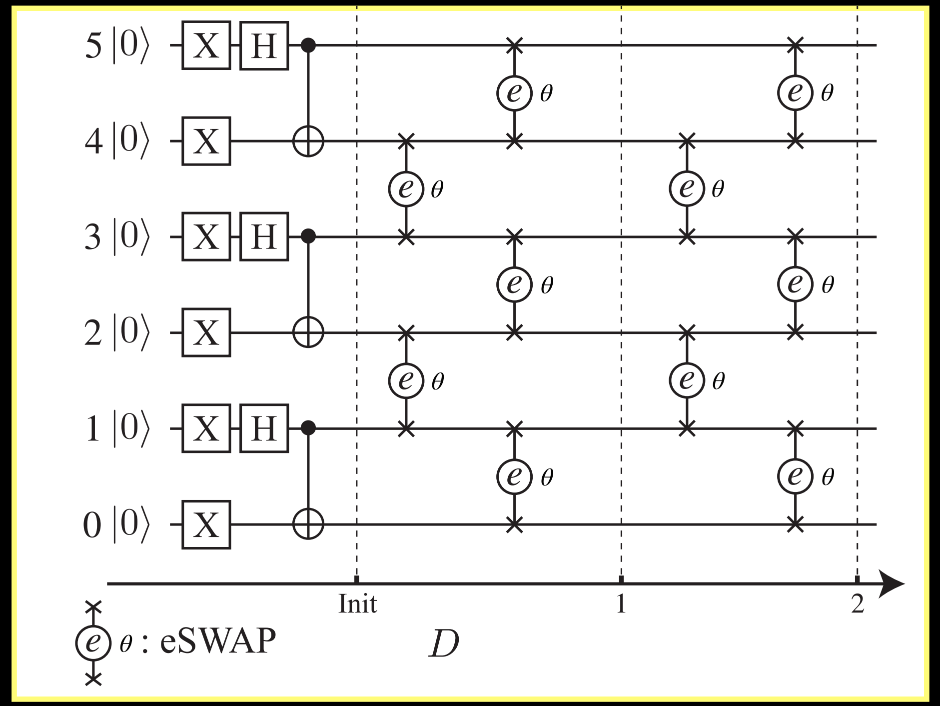
	Random quantum circuit	Parametrized quantum circuit
1D	**RQC-1D**	**PQC-1D**
2D	**RQC-2D(ABCDABCD...)**	**PQC-2D(ABCDABCD...)**

RQC-1D



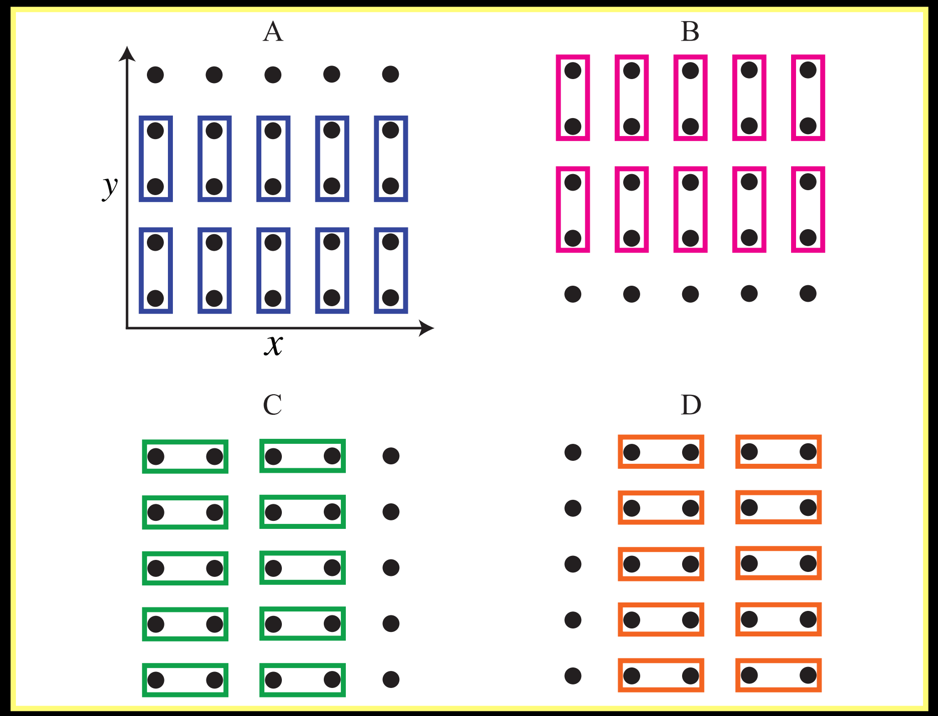
[Y. Zhou et al., PRX **10**, 041038 (2020)]

PQC-1D

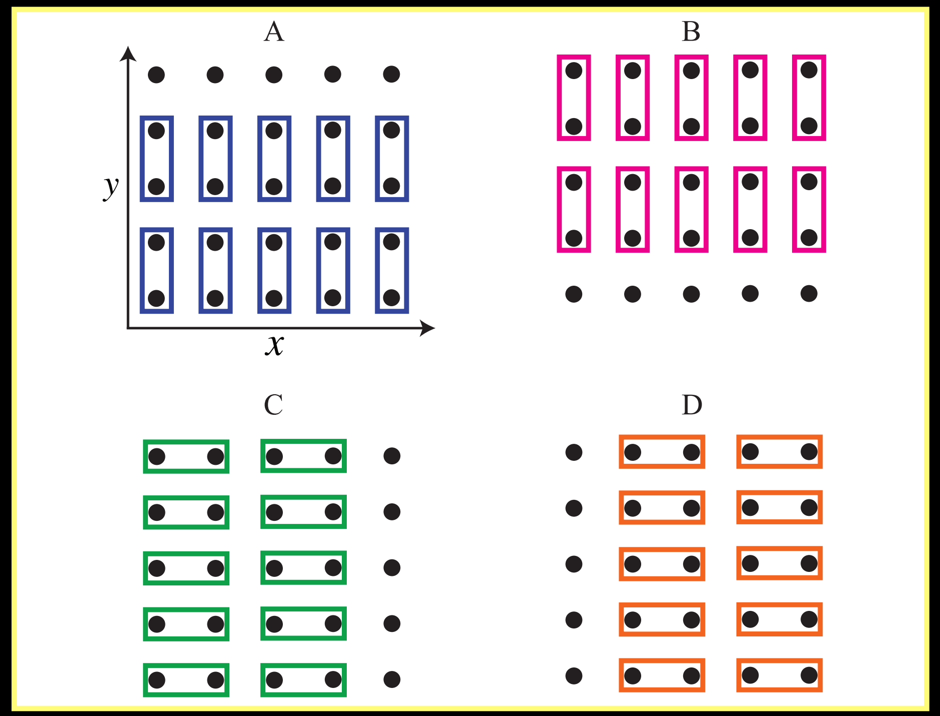


[R.-Y. Sun, T. Shirakawa, S. Yunoki, PRB **108**, 075127 (2023)]

RQC-2D(ABCDABCD...)

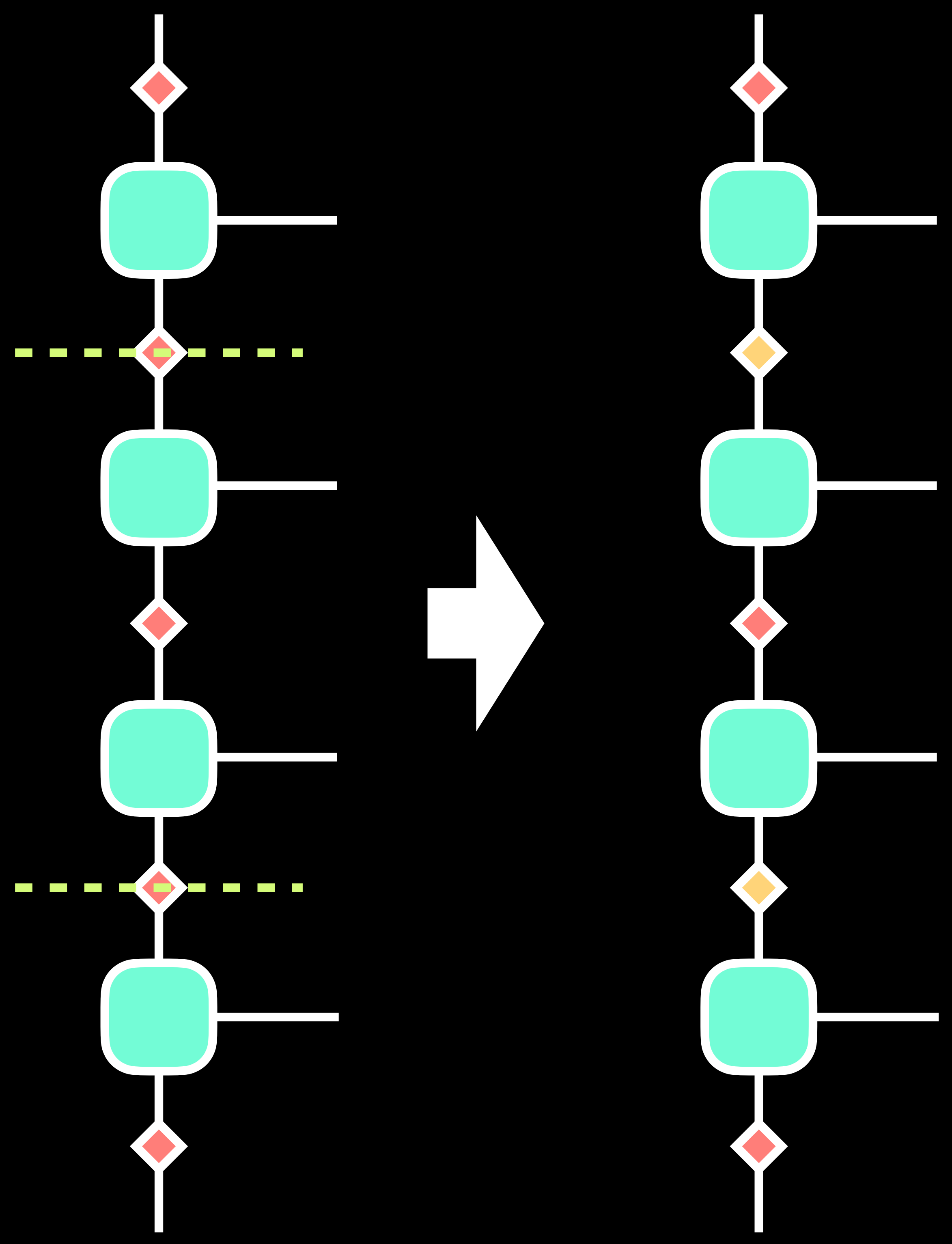


PQC-2D(ABCDABCD...)



[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Parallel MPS compression



$|\Psi(\chi)\rangle$ $\quad$ $|\tilde{\Psi}(\tilde{\chi})\rangle$

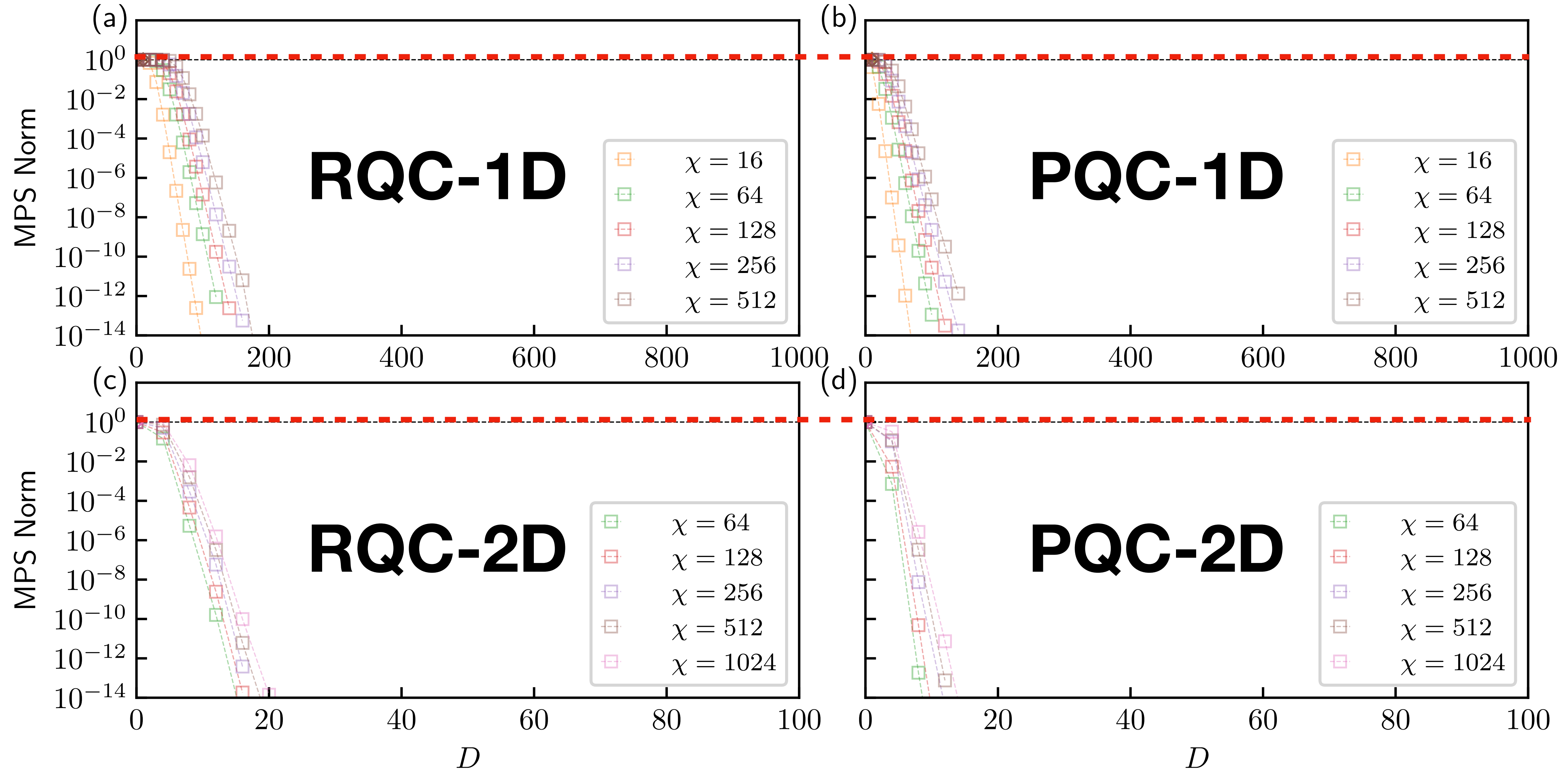
$\tilde{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_{\tilde{\chi}}, \cancel{\lambda_{\tilde{\chi}+1}, \cdots, \lambda_{\chi}})$

Theorem [Verstraete and Cirac, PRB **73**, 094423 (2006)]

$$\Big|\, |\Psi(\chi)\rangle - |\tilde{\Psi}(\tilde{\chi})\rangle \,\Big| \le 2\varepsilon(\tilde{\chi}) \qquad \varepsilon(\tilde{\chi}) = \sum_{i=1}^{N-1} \varepsilon_i(\tilde{\chi})$$

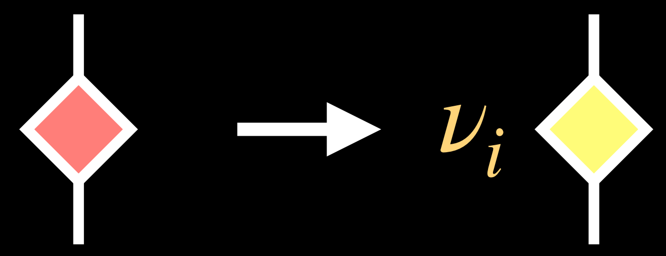
Issue: wavefunction norm decay

$$1 - \sqrt{2\varepsilon(\tilde{\chi})} \le \Big|\, |\tilde{\Psi}(\tilde{\chi})\rangle \,\Big| \le 1 \qquad \varepsilon_i(\tilde{\chi}) = \sum_{n=\tilde{\chi}+1}^{\chi} (\lambda_n^{[i]})^2$$



[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Stabilize the wavefunction norm in parallel

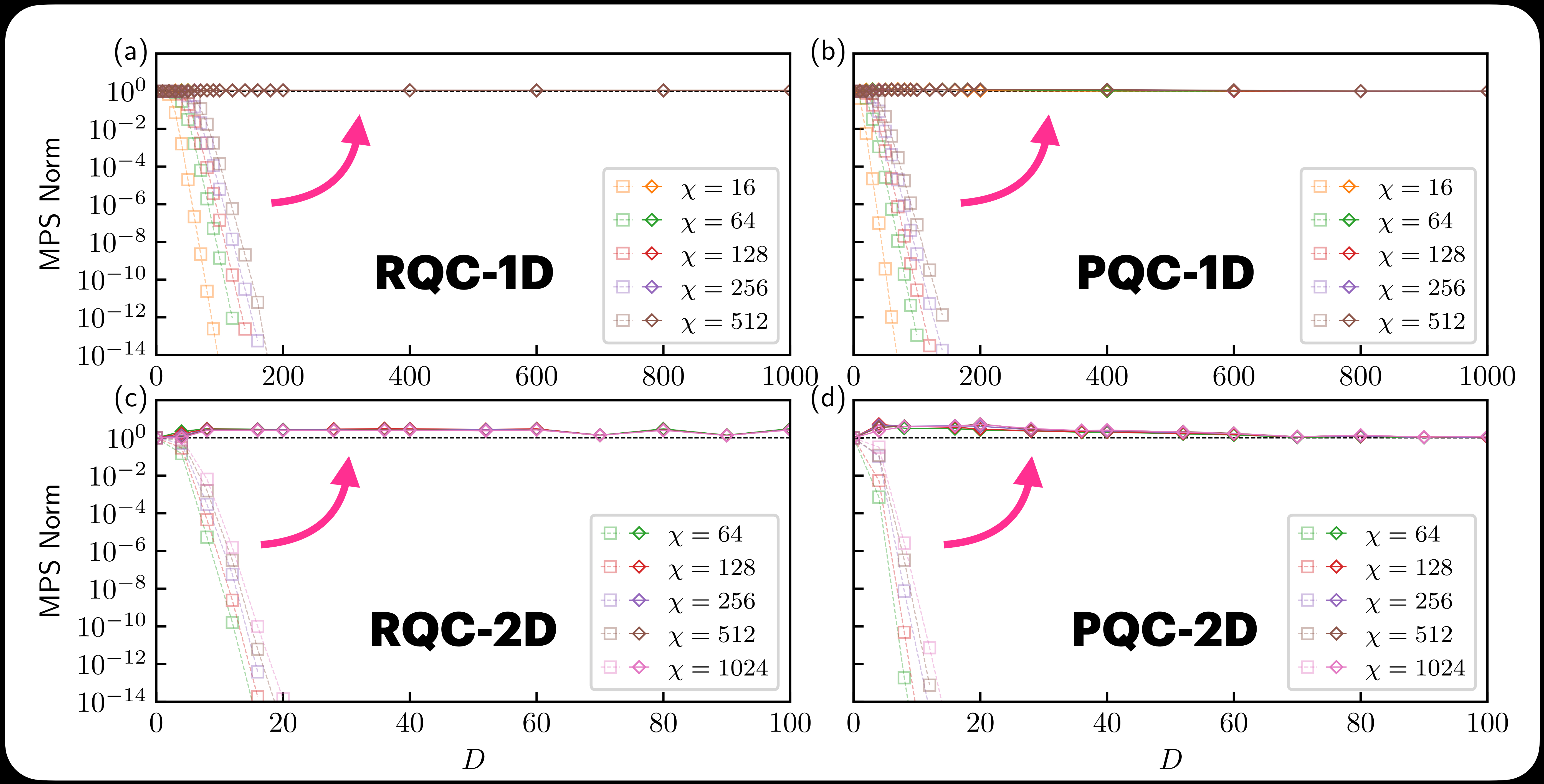


i.e., parallel rescaling.

$$\Lambda^{[i]}(\chi) \rightarrow \nu_i \tilde{\Lambda}^{[i]}(\tilde{\chi}) \qquad \nu_i = [1 - \varepsilon_i(\tilde{\chi})]^{-1/2}$$

$$(1-\sqrt{2\varepsilon(\tilde{\chi})})\prod_{i=1}^{N-1}\nu_i \le \big|\,|\Psi(\tilde{\chi})\rangle\big| \le \prod_{i=1}^{N-1}\nu_i$$

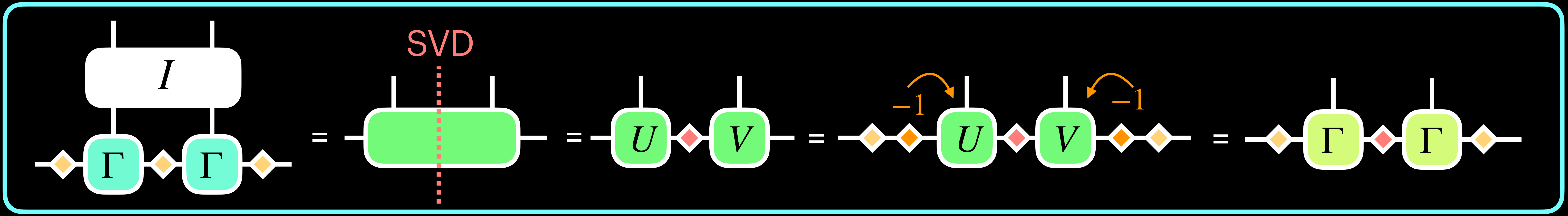
$$(1-\sqrt{2\varepsilon(\tilde{\chi})})\prod_{i=1}^{N-1}\nu_i \le 1,\ 1 \le \prod_{i=1}^{N-1}\nu_i$$



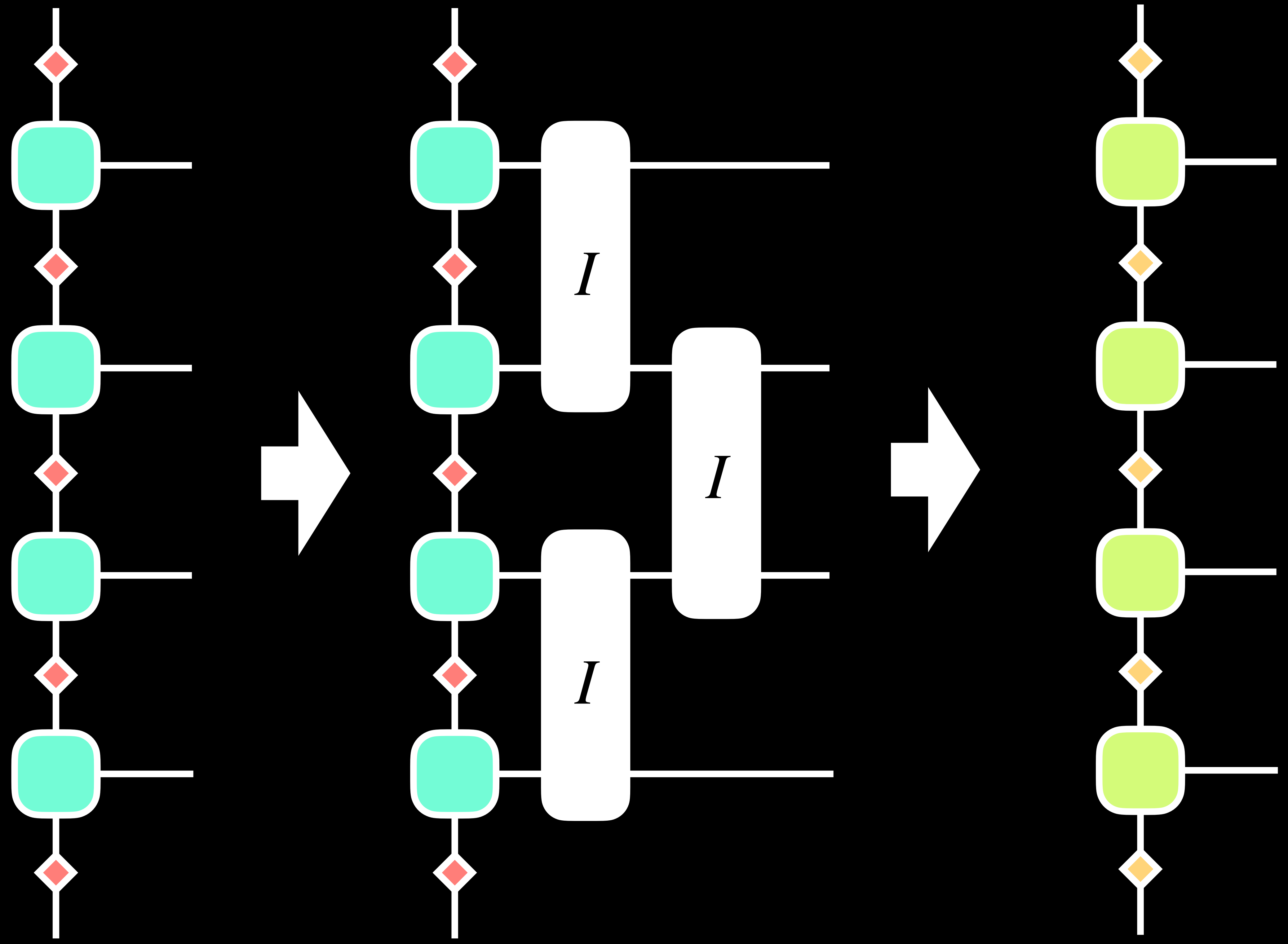
[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Restore the canonical form

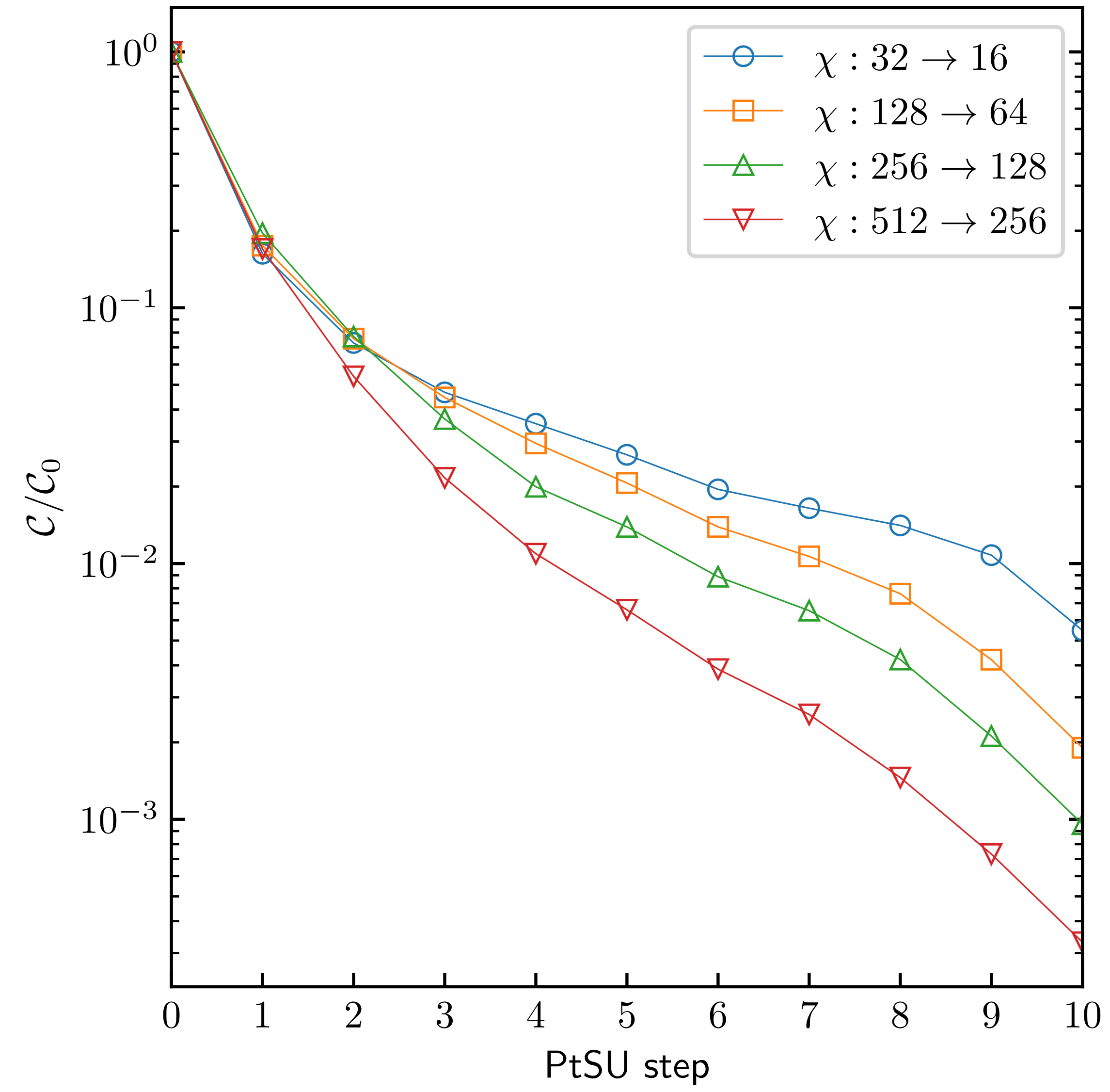
Parallel trivial simple update (PtSU)



Equivalence between belief propagation and trivial simple update [R. Alkabetz and I. Arad, Phys. Rev. Research **3**, 023073 (2021)]



$$C = \frac{1}{2L}\sum_{i=1}^{N}\left(\left|\bar{A}^{[i]}A^{[i]} - I\right|_F + \left|B^{[i]}\bar{B}^{[i]} - I\right|_F\right)$$



[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

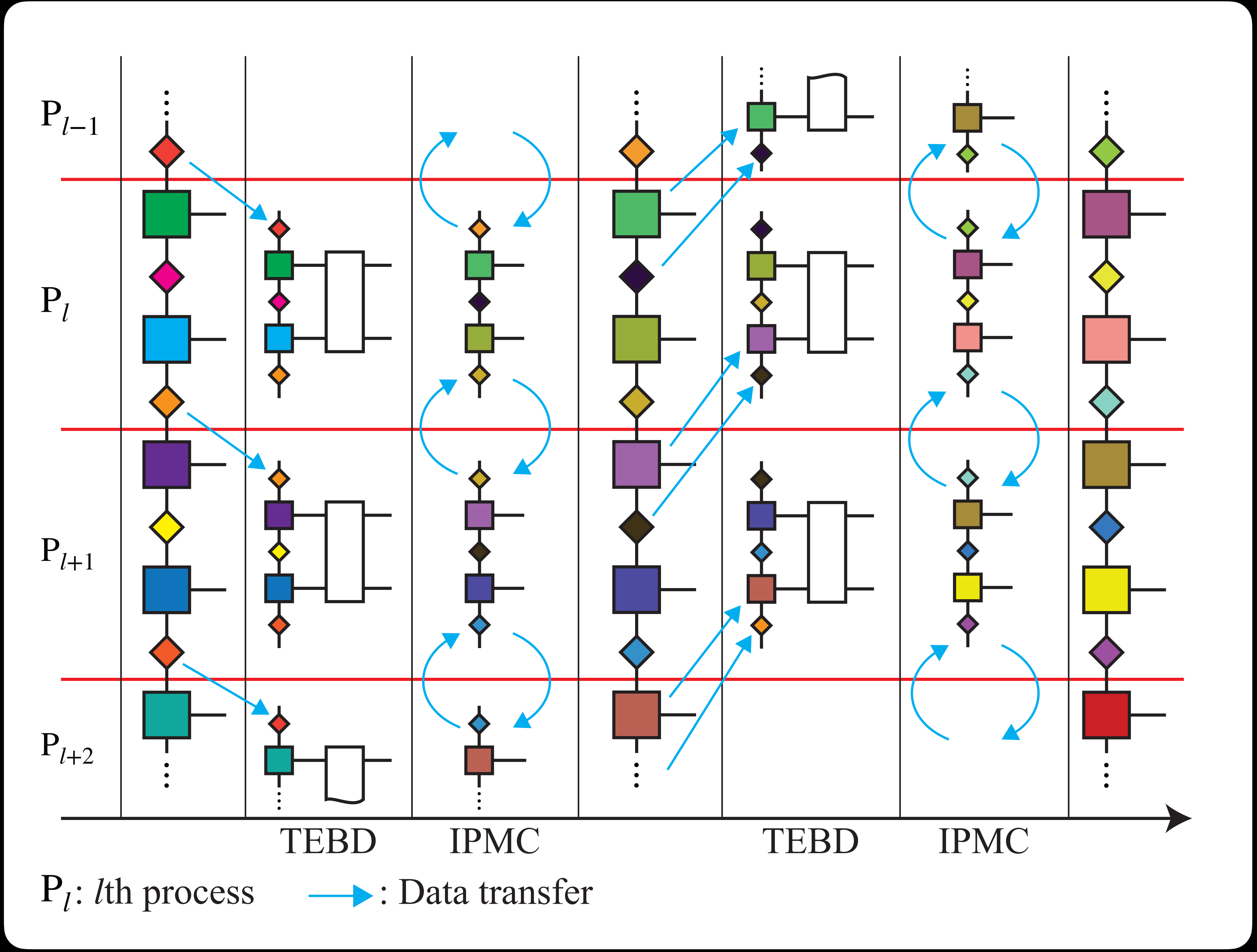
IPMC and pTEBD

IPMC: Improved parallel MPS compression

IPMC = Parallel wavefunction norm stabilization
+ Parallel partial regauging

pTEBD = Parallel gates applications
+ Parallel bond dimension compression
+ IPMC

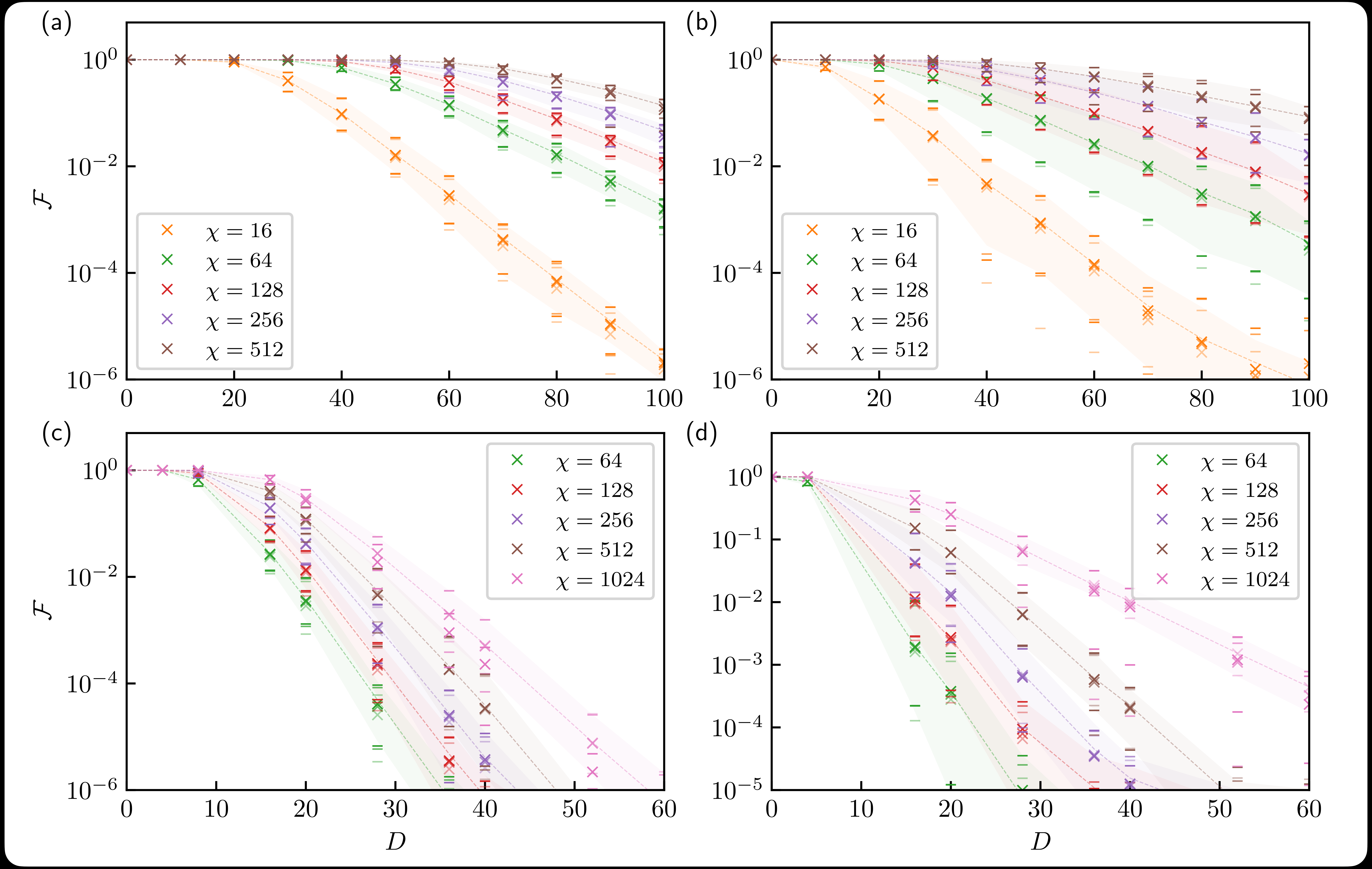
Only neighboring communications !



[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

pTEBD: Accuracy

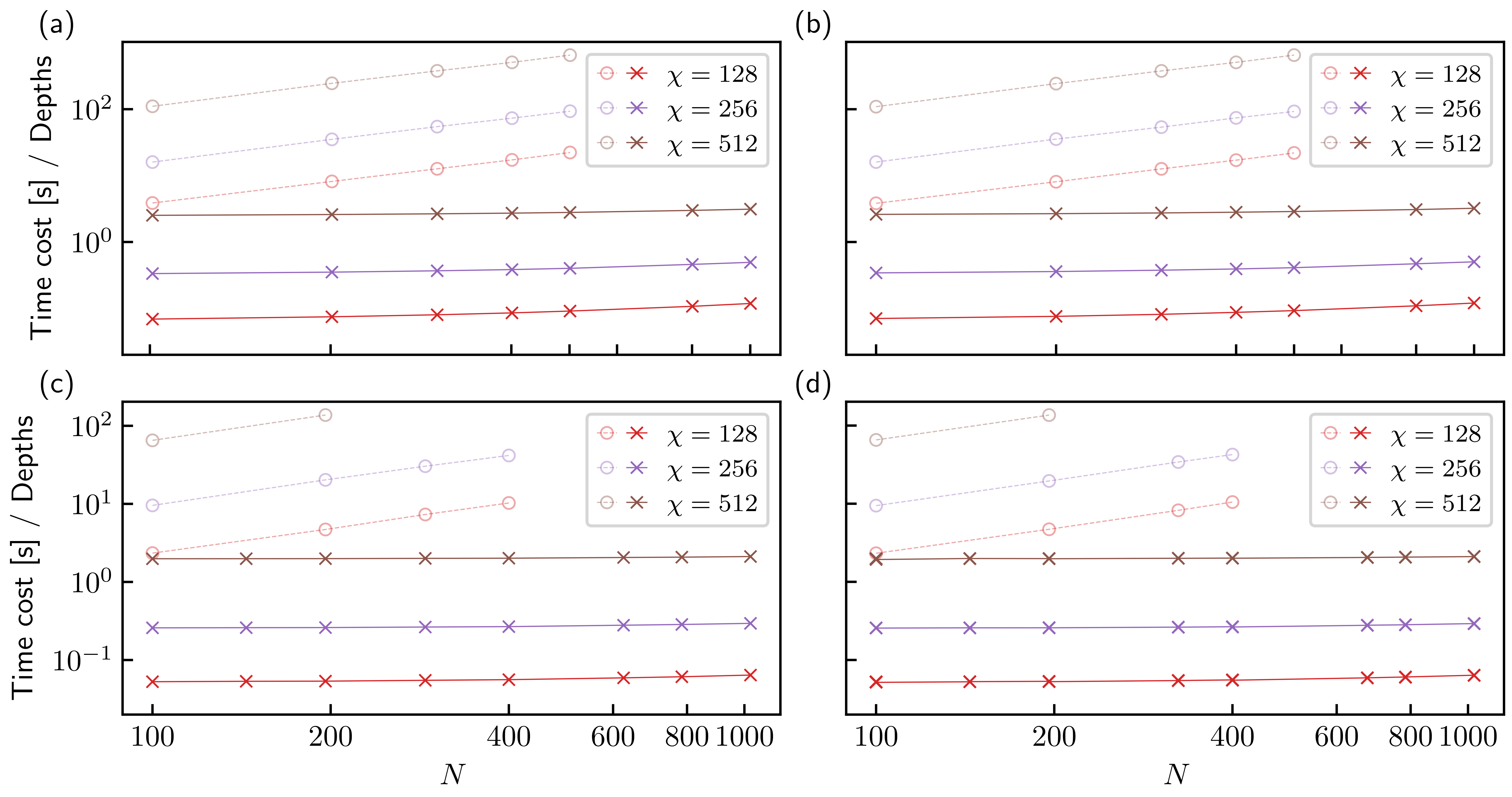
$$\mathscr{F} = \left| \langle \Psi_{\text{exact}} | \Psi(\chi) \rangle \right|$$



$$\mathscr{F}_{\text{pTEBD}} \sim \mathscr{F}_{\text{SeqMPS}}$$

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

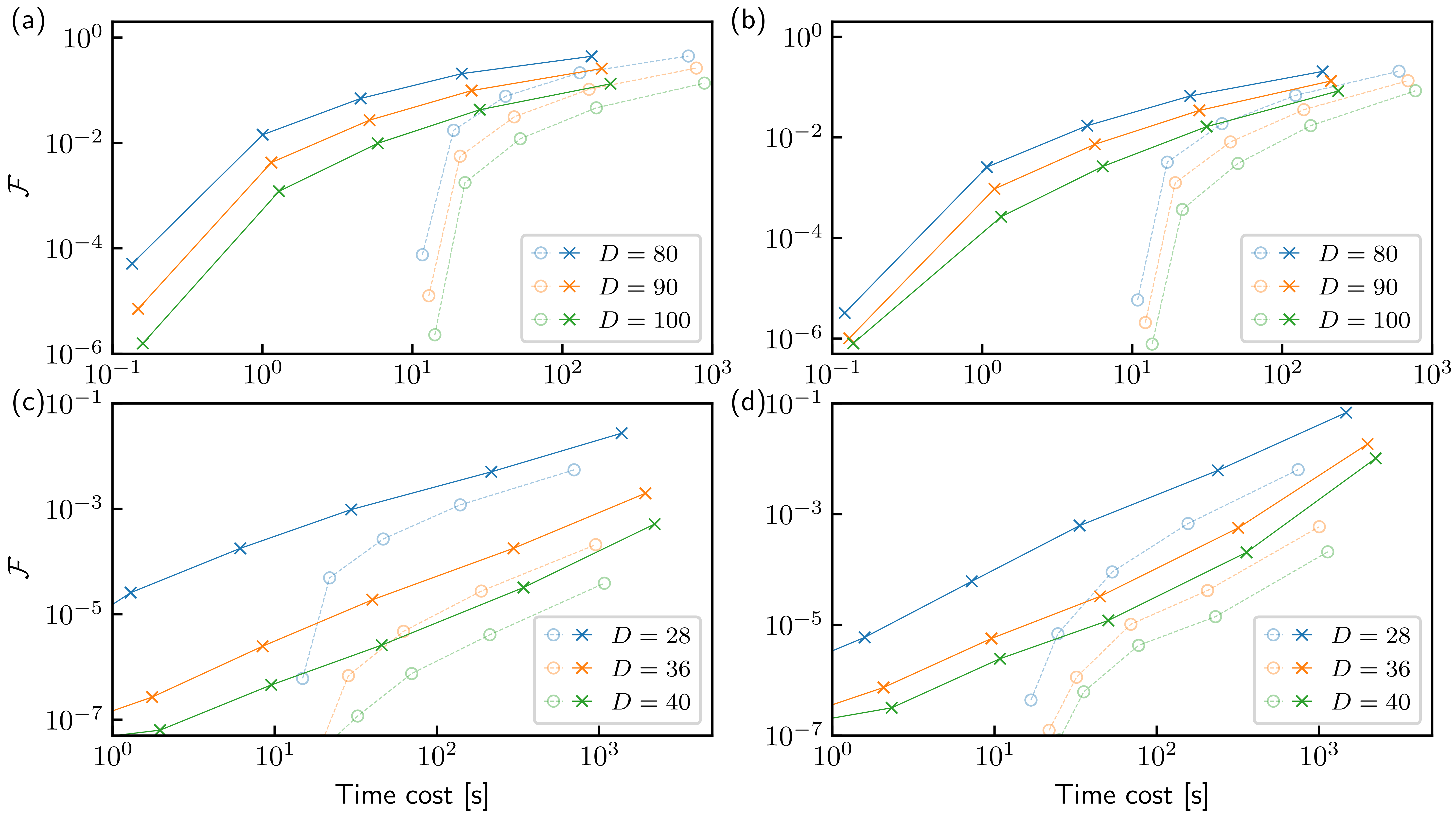
pTEBD: Parallel performance



Perfect weak scaling

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

pTEBD: Cost v.s. Performance



$$T_{\mathrm{pTEBD}} \sim T_{\mathrm{SeqMPS}}/N$$

[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

TNS Software

Goal: Develop high-performance scalable TNS software for Fugaku

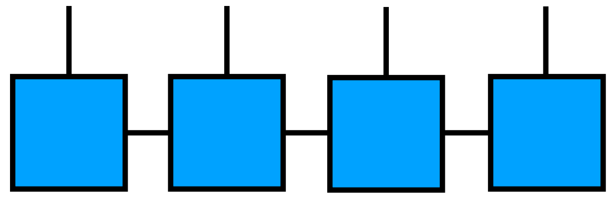
Software stack

TNS Algorithms

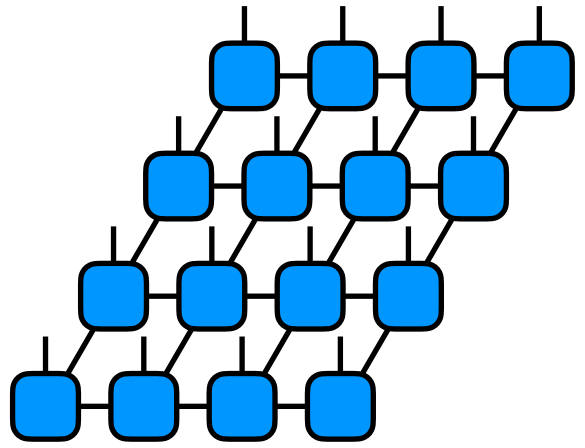
Tensor Library

GraceQ/tnet [TS, H. Kohshiro, R.-Y. Sun]

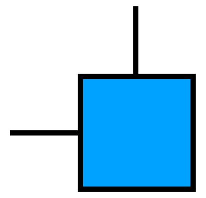
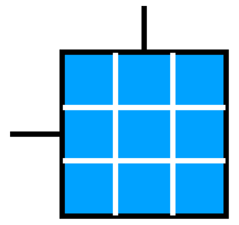
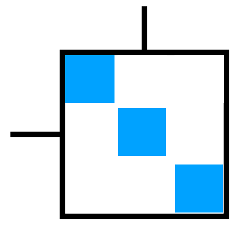
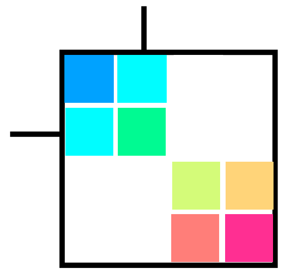
MPS



PEPS



GraceQ/tensor [R.-Y. Sun, TS, H. Kohshiro]

	Shared memory	Distributed memory
Dense	 ✓	 (✓)
Symmetric	 ✓	

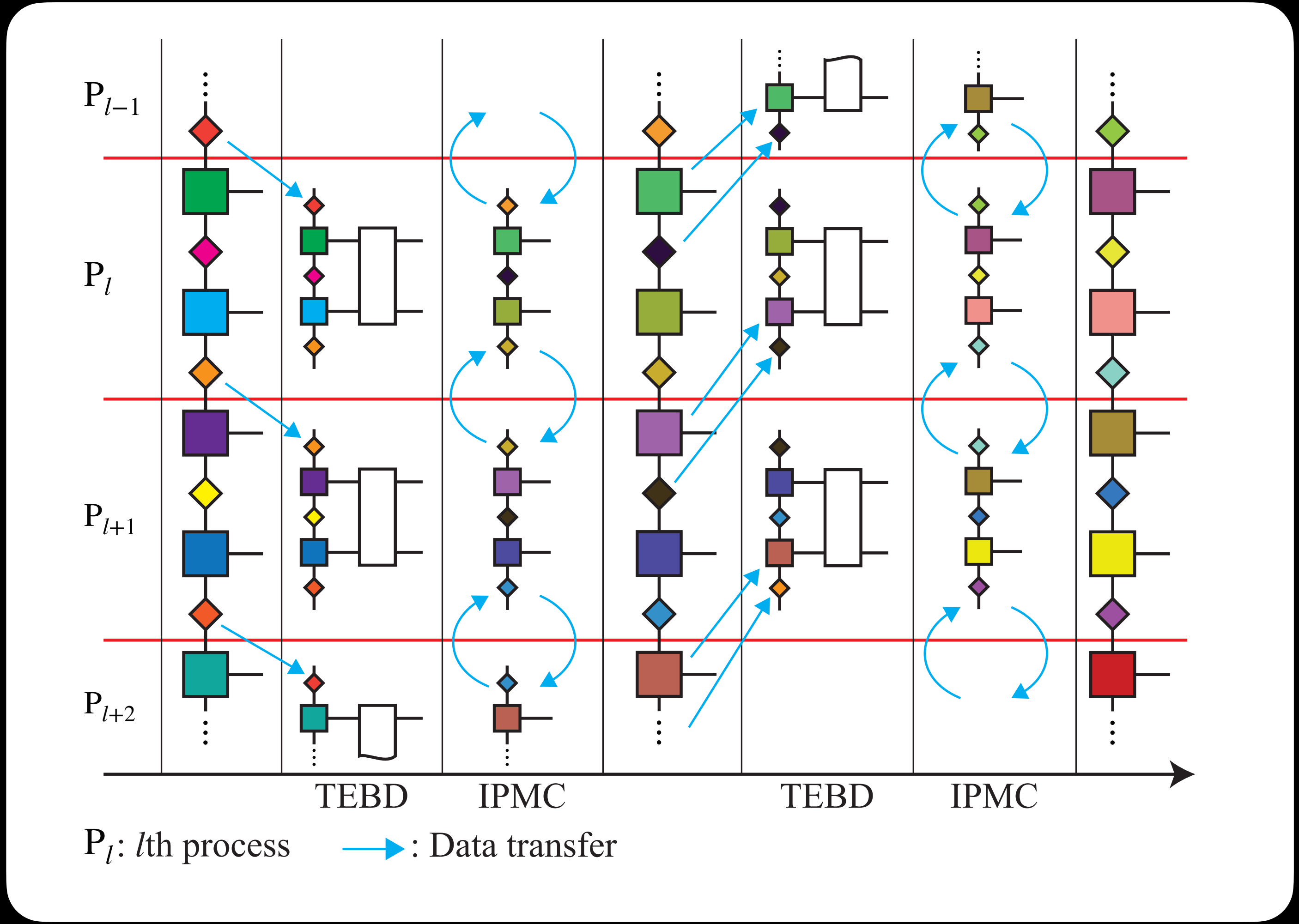
Summary

Development of tensor network method for parallel computer

- Keeping canonical gauge structure is crucial and difficult point in parallelization.

- We introduced a scheme to improve the performance of parallel TEBD: improved parallel MPS compression (**IPMC**).

- Partial gauge fixing by parallel trivial simple update (**PtSU**) may be useful.



[R.-Y. Sun, T. Shirakawa, S. Yunoki, arXiv:2312.02667 (2023)]

Thank you for your attention!