

パイロット講義: 計算科学・量子計算における情報圧縮

Data Compression in Computational Science and Quantum Computing

2021.12.21

#2 Quantum Computer and Simulation

理学系研究科 量子ソフトウェア寄付講座

藤堂眞治 Synge Todo

Lecture materials: <https://github.com/utokyo-qsw/data-compression>

Outline

- Quantum Computer and Quantum Computation
 - quantum computer; quantum supremacy; quantum circuits; typical quantum gates; quantum operations in quantum circuits; quantum measurement
- Simulation of Quantum Circuits
 - Schrödinger simulation; Feynman simulation; Feynman sampling
- Tensor Network Simulation
 - tensor network representation of quantum circuits; memory cost of tensors; computational cost of tensor contraction; optimization of contraction order; slicing
- Contraction Based on Low-Rank Approximation

Quantum systems

Example of quantum system: Array of quantum bits

1 bit

- A quantum bit is represented by two basis vectors.

$$|0\rangle, |1\rangle \text{ or } (|\uparrow\rangle, |\downarrow\rangle)$$

2 bits



The vector space is spanned by four basis vectors.

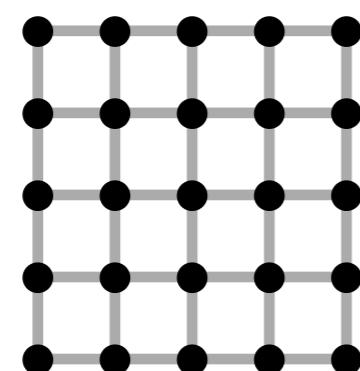
$$|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle$$

$$\text{Simple notation: } |00\rangle, |01\rangle, |10\rangle, |11\rangle$$

$$\rightarrow |\Psi\rangle = \sum_{\alpha, \beta=0,1} C_{\alpha, \beta} |\alpha\beta\rangle$$

$C_{\alpha, \beta}$:complex number

N bits



Dimension of the vector space $= 2^N$

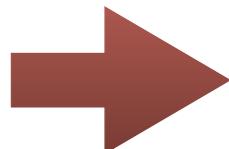
Exponentially large!

$$|\Psi\rangle = \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle$$

Difficulty in quantum many-body problems

Schrödinger equation: $i\hbar \frac{\partial}{\partial t} |\Psi\rangle = \mathcal{H}|\Psi\rangle$

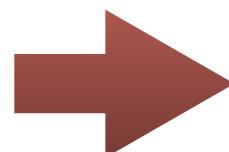
- Dimension of the vector space is **exponentially large**



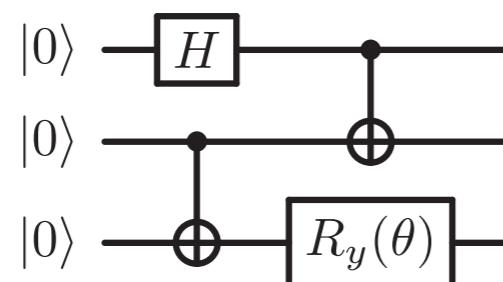
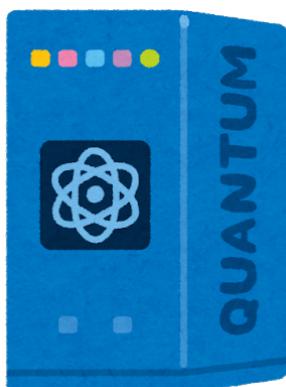
To solve the problem **exactly** by (classical) computer, we need **huge memory** and **huge computation time**.

e.g., We can simulate only ~50 qubits in classical supercomputer.

Quantum computer



It can treat **quantum state directly**, and then (ideally) there is no problems originated form the exponentially large vector space.



Classical computer?

- There are several techniques to treat quantum many body problems.
 - One of them is **a data compression based on tensor networks**.
- We may use them to simulate quantum computers.

Quantum Computer and Quantum Computation

Quantum Computer & Simulation

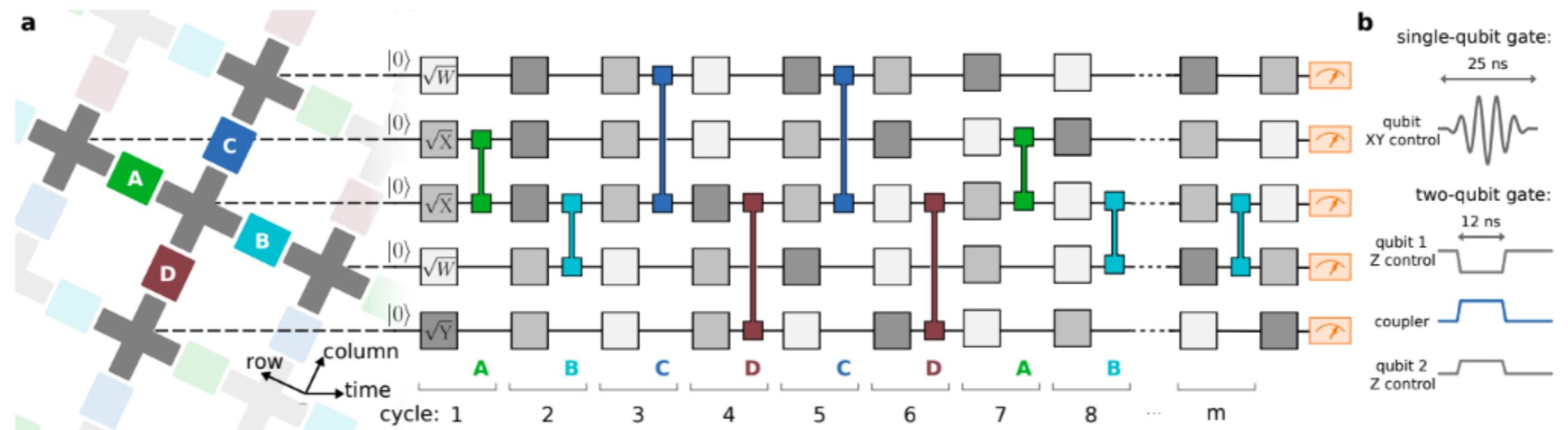
- Quantum simulation
 - execute quantum algorithms with quantum computer
 - Shor's algorithm (prime factorization)
 - factorization of n-digit number: from $\exp n$ to polynomial of n
 - Grover's Algorithm (database search)
 - search problem for unstructured data of size n : from $O(n)$ to $O(\sqrt{n})$
 - simulate quantum many-body systems (e.g., spin systems) with quantum computer
 - proposed by Feynman in 1982 (Simulating physics with computers)
- Quantum computer simulation
 - simulate quantum computers with classical computer
 - today's theme

Quantum Supremacy

- With quantum computers, perform tasks that classical computers can't (in realistic time)
 - Google's experiment with superconducting gates
 - Arute, F., Arya, K., Babbush, R. et al. Quantum supremacy using a programmable superconducting processor. *Nature* 574, 505–510 (2019). <https://doi.org/10.1038/s41586-019-1666-5>
 - random quantum circuit sampling
 - experiments with an optical quantum computer at the University of Science and Technology of China
 - Zhong, H. Sen, Deng, Y. H. et al. Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light. *Physical Review Letters*, 127(18), 180502 (2021). <https://doi.org/10.1103/PhysRevLett.127.180502>
 - Gaussian boson sampling
- How far can we actually go with a classical computer?

Quantum Circuits

- Prepare a set of quantum bits (qubits)
- A number of quantum gates (typically 1-qubit or 2-qubit gates) are applied to qubits in order
 - quantum gates are unitary operations
 - combination of quantum gates are also unitary operation
- Finally, perform measurements to extract information



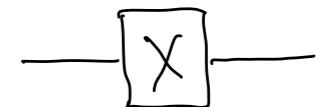
Arute, F., Arya, K., Babbush, R. et al. Quantum supremacy using a programmable superconducting processor. *Nature* 574, 505–510 (2019). <https://doi.org/10.1038/s41586-019-1666-5>

Typical Quantum Gates

- 1-qubit gates (2x2 matrix)

- X gate (NOT)

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



- H gate (Hadamard)

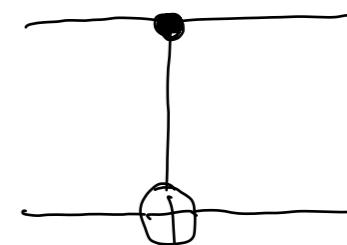
- 2-qubit gates (4x4 matrix)

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$



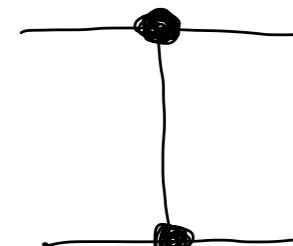
- CX gate (controlled-NOT)

$$CX = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

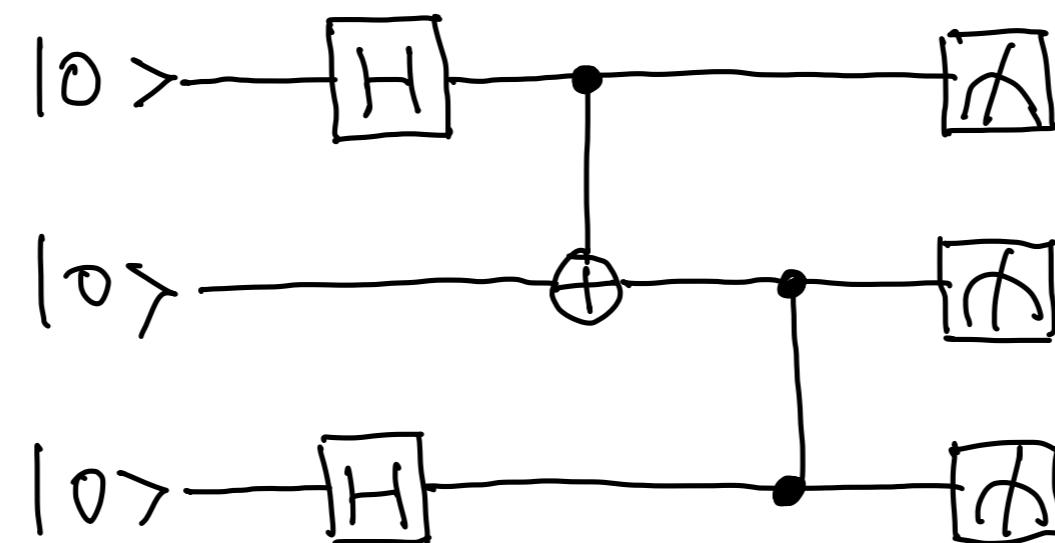


- CZ gate (controlled-Z)

$$CZ = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$



Example of Quantum Circuit



Quantum Operations in Quantum Circuits

- Parallelism
 - inputting a state of superposition produces a superposition of the corresponding outputs.
- Bifurcation
 - states bifurcate (in terms of computational basis) when H-gates, etc. are applied
- Interference
 - superposition coefficients of the states are complex, and they may cancel each other out and vanish
- Collapse
 - collapse to one of the states in the basis used for the measurement

Quantum Measurement

- Pure state
 - In a pure state, entropy is zero
 - density matrix: $\rho = |\Psi\rangle\langle\Psi|$ is a rank-1 matrix
 - the largest eigenvalue is unity, all other eigenvalues are zero
 - Von Neumann entropy: $S(\rho) = - \text{tr} (\rho \log \rho) = 0$
 - Manipulation of states by quantum circuits
 - entropy does not change since operation are unitary transformations
 - if the initial state is a pure state, the entropy stays zero

Quantum Measurement

- Non-selective projective measurement (= measure and don't see the result)
 - entropy always increases
 - projection operators $\{P_i\}$ ($\sum P_i = 1, (P_i)^2 = P_i$)
 - by projective measurement: $\rho \rightarrow \rho' = \sum P_i \rho P_i$
 - $S(\rho') \geq S(\rho)$ (Klein's theorem)
 - a pure state is converted into a mixed state
 - off-diagonal elements vanishes
- Selective projective measurement (measure and see the result)
 - a pure state is sampled according to the measurement probability
 - selection causes the state collapse and the entropy to be zero again

Essence of Quantum Algorithm

- Prepare a superposition of many states, using Hadamard gates, etc.
 - if non-selective projective measurements is performed at this stage, the entropy is extensive (proportional to the number of qubits)
- Manipulate and interfere the states so that desired answer have large amplitude
 - entropy of the quantum state remains zero
 - interference reduces the entropy after non-selective projective measurement
 - a kind of data compression?
- Selective projective measurement (measure and see the result) to get the correct answer with high probability
 - sampling from the state with entropy as small as possible
 - a kind of information extraction?

Simulation of Quantum Circuits

Classical Simulation of Quantum Circuits

- Quantum circuits are one of the quantum many-body systems
 - Is it possible to simulate quantum many-body systems by classical computers?
 - → Yes!
 - Two traditional approaches of classical simulation of quantum systems
 - Schrödinger simulation
 - Feynman simulation
 - Feynman sampling
 - What do we want to obtain in classical simulation?
 - wave function itself
 - distribution of measurement probabilities
 - expectation value of the measurement
 - sample of measurement results, etc

Schrödinger Simulation

- Store coefficients of wave function (superposition coefficients of states) in a vector
 - n qubits \rightarrow complex vector of size 2^n
 - quantum gates $\rightarrow 2^n \times 2^n$ unitary matrices (but generally very sparse)
 - operation with quantum gates \rightarrow matrix-vector product
- Simulation cost
 - n qubits, m gates
 - memory cost: $O(\exp n)$ $n = 50 \rightarrow 4\text{PB}$ (single precision)
 - computational complexity: $O(m \exp n)$ $n = 50, m = 1000 \rightarrow 1\text{EFlop}$

Feynman Simulation

- Summing up the contributions from all "paths"
 - path integral in quantum mechanics
 - branches into different path each time through Hadamard gate
 - contribution (weight) from a path is product of matrix elements
- Simulation cost
 - n qubits, m (branching) gates
 - number of paths: $O(\exp m)$
 - memory cost: $O(n)$ (depending on what do we measure)
 - computational complexity: $O(n \exp m)$

Feynman Sampling

- aka *Quantum Monte Carlo*
 - choose a new state stochastically with probability proportional to the matrix elements at each branching
- Simulation cost
 - n qubits, m (branching) gates
 - computational complexity: $O(m) \times (\text{number of random samples})$
- Negative-sign problem
 - what to do when the matrix element is negative (or complex)
 - Use the absolute value of the matrix element as the "weight".
 - Keep the sign (or phase) separately
 - when the signs are mixed, strong canceling occurs
 - as the circuit gets deeper, it becomes almost random sampling
 - required number of samples: $O(\exp m)$

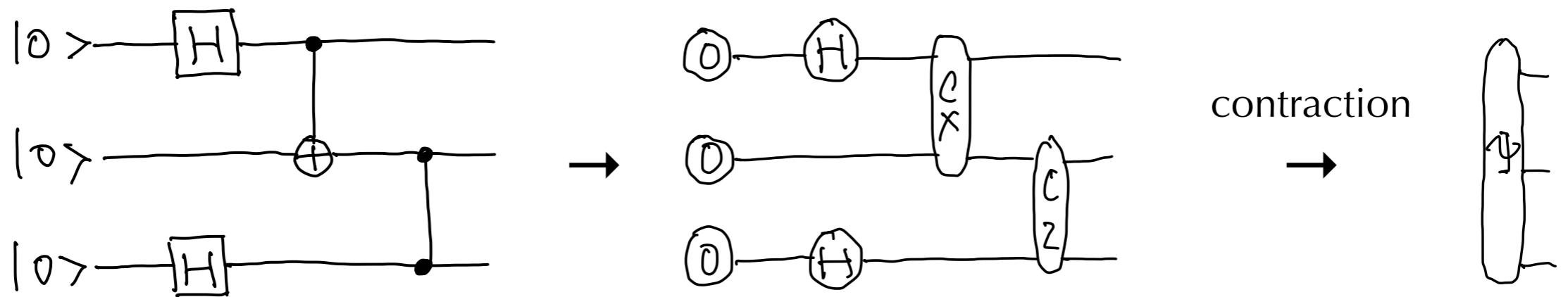
Tensor Network Simulation

Tensor Network Representation of Quantum Circuits

- Tensor representation of quantum gates and states
 - 1-qubit gate:
 - 2×2 matrix \rightarrow 2-leg tensor
 - 2-qubit gate:
 - 4×4 matrix $\rightarrow 2 \times 2 \times 2 \times 2$ tensor \rightarrow 4-leg tensor
 - r -qubit gate:
 - $2^r \times 2^r$ matrix $\rightarrow 2 \times 2 \times \dots \times 2$ tensor $\rightarrow 2r$ -leg tensor
 - initial quantum states
 - product state, e.g., $|0\rangle \otimes |0\rangle \otimes \dots \otimes |0\rangle$
 \rightarrow a set of n 1-leg tensors (vectors)

Tensor Network Representation

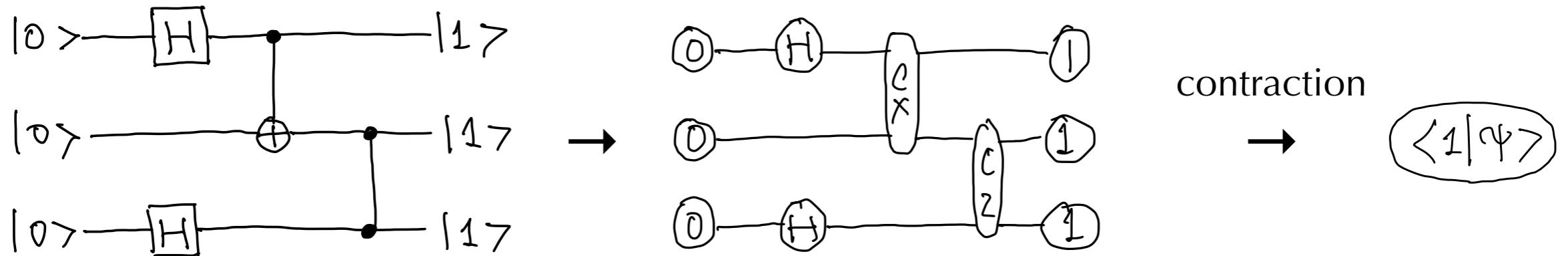
- Tensor network representation of a quantum circuit
 - all bond dimensions are 2



- Taking contraction from the initial state (left to right)
 - equivalent to Schrödinger simulation

Evaluating an Amplitude

- Instead of obtaining wave function itself, let's evaluate an amplitude (a coefficient) in the wave function: e.g., $\langle 111 | \Psi \rangle$

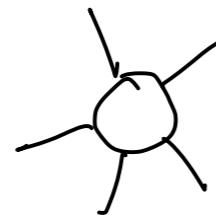


- Order of contraction does not change the final result
- Freedom in contraction order \rightarrow possibility to reduce the cost

Memory Cost of Tensors

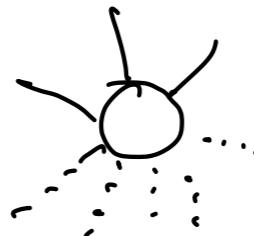
- Assuming all bond dimensions are χ

- 5-leg tensor



$$T_{ijkln} \quad O(\chi^5)$$

- n -leg tensor



$$O(\chi^n)$$

- Memory cost grows exponentially as the number of legs increases

Computational Cost of Tensor Contraction

- Matrix-vector multiplication



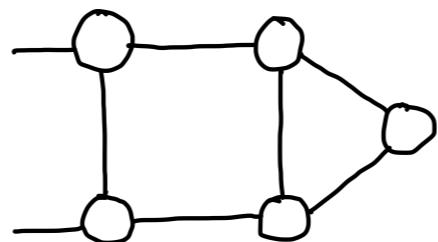
$$w_i = \sum_j T_{ij} v_j \quad O(\chi^2)$$

- Matrix-matrix multiplication



$$A_{ij} = \sum_k T_{ik} R_{kj} \quad O(\chi^3)$$

- Contraction of tensors



$$O(\chi^?)$$

- Cost of tensor contraction strongly depends of contraction order

Optimization of Contraction Order

- General guideline for better contraction order
 - Avoid tensors with a large number of legs in the middle or at the end of computation
- It is known that finding the optimal contraction order is (at least) #P-hard problem
 - can't find the best solution in a realistic time
 - many heuristics have been proposed, c.f.,
 - Schutski, R., Khakhulin, T., Oseledets, I., & Kolmakov, D., Simple heuristics for efficient parallel tensor contraction and quantum circuit simulation. *Physical Review A*, 102(6), 1–11 (2020). <https://doi.org/10.1103/PhysRevA.102.062614>
 - Gray, J., & Kourtis, S., Hyper-optimized tensor network contraction. *Quantum*, 5, 1–22 (2021). <https://doi.org/10.22331/Q-2021-03-15-410>

Slicing Tensor Network

- Slicing
 - for a subset of bonds in the tensor network
 - fix the bonds to some value
 - these bonds disappear from the tensor network
 - for each fixed value
 - perform contraction independently
 - take summation on fixed-value patterns at the outermost
- Advantage
 - memory cost of contraction of sliced tensor network becomes smaller
 - contraction of each sliced network can be performed in parallel

State-of-the-art Tensor-network-based Simulation

- Liu, Y. A., Liu, et al., Closing the “quantum supremacy” gap: Achieving real-Time simulation of a random quantum circuit using a new sunway supercomputer. International Conference for High Performance Computing, Networking, Storage and Analysis, SC (2021). <https://doi.org/10.1145/3458817.3487399>
- Gordon Bell Prize Winner in 2021

The screenshot shows a web browser displaying the ACM awards page. The main content is titled "Real-Time Simulation of Random Quantum Circuit". It describes a team from Chinese institutions winning the 2021 ACM Gordon Bell Prize for their work on simulating a random quantum circuit. The page includes a sidebar with links to view the full list of ACM awards and categories for contributions.

Real-Time Simulation of Random Quantum Circuit

ACM, the Association for Computing Machinery, named a 14-member team, drawn from Chinese institutions, recipients of the 2021 ACM Gordon Bell Prize for their project, [Closing the "Quantum Supremacy" Gap: Achieving Real-Time Simulation of a Random Quantum Circuit Using a New Sunway Supercomputer](#).

The members of the winning team are: [Yong \(Alexander\) Liu](#), [Xin \(Lucy\) Liu](#), [Fang \(Nancy\) Li](#), [Yuling Yang](#), [Jiawei Song](#), [Pengpeng Zhao](#), [Zhen Wang](#), [Dajia Peng](#), and [Huorong Chen](#) of Zhejiang Lab, Hangzhou and the National Supercomputing Center in Wuxi; [Haohuan Fu](#) and [Dexun Chen](#) of Tsinghua University, Beijing, and the National Supercomputing Center in Wuxi; [Wenzhao Wu](#) of the National Supercomputing Center in Wuxi; and [Heliang Huang](#) and [Chu Guo](#) of the Shanghai Research Center for Quantum Sciences.

Quantum supremacy is a term used to denote the point at which a quantum device can solve a problem that no classical computer can solve in a reasonable amount of time. Teams at Google and the University of Science and Technology of China in Hefei both claim to have developed devices that have achieved quantum supremacy.

According to the Gordon Bell Prize recipients, determining whether a device has achieved quantum supremacy for a given task (in a specific scenario) begins with sampling the interactions of the different quantum bits (qubits) in a random quantum circuit (RQC). As the number of possible interactions among qubits in a random quantum circuit is staggeringly large, simulating their interactions is a problem well-suited for a high-performance computer. However, the quantum physics behind the entangled qubits requires that the classical binary bits used in a supercomputer store and compute the information with exponentially-increasing complexity.

In their Gordon Bell Prize-winning work, the Chinese researchers introduced a systematic design process that covers the algorithm, parallelization, and architecture required for the simulation. Using a new Sunway Supercomputer, the Chinese team effectively simulated a $10 \times 10x (1+40+1)$ random quantum circuit (a new milestone for classical simulation of RQC). Their simulation achieved a performance of 1.2 Eflops (one quintillion floating-point operations per second) single-precision, or 4.4 Eflops mixed-precision, using over 41.9 million Sunway cores (processors).

The project far outpaced state-of-the-art approaches to simulating an RQC. For example, the most recent effort, using the Summit supercomputer to simulate a random quantum circuit of the Google Sycamore quantum processor (which has 53 qubits), was estimated to take 10,000 years to perform. By contrast, the Chinese team's approach employing the Sunway supercomputer takes only 304 seconds for a

[View the full list of ACM Awards](#)

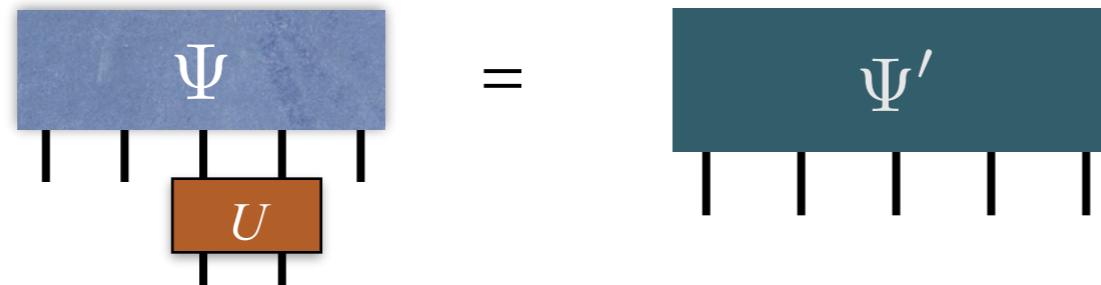
ACM Awards by Category

- + Career-Long Contributions
- + Early-to-Mid-Career Contributions
- + Specific Types of Contributions

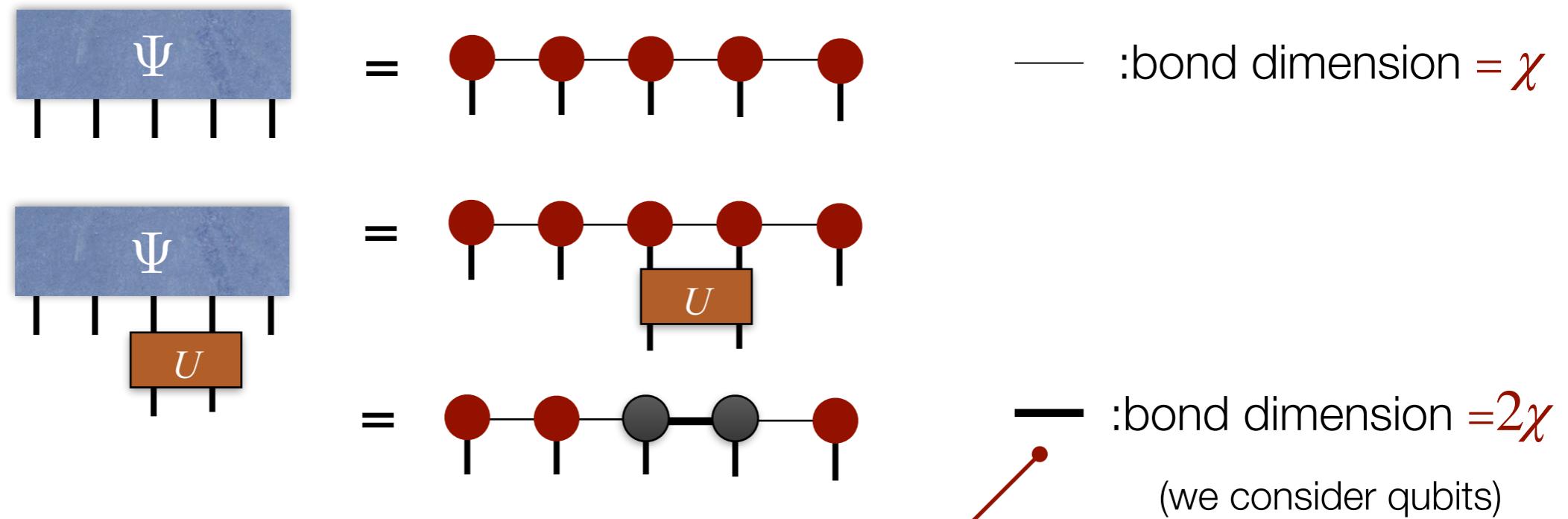
<https://awards.acm.org/bell>

Contraction Based on Low-Rank Approximation

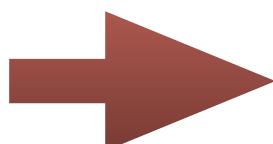
Gate operation in MPS representation



MPS representation:



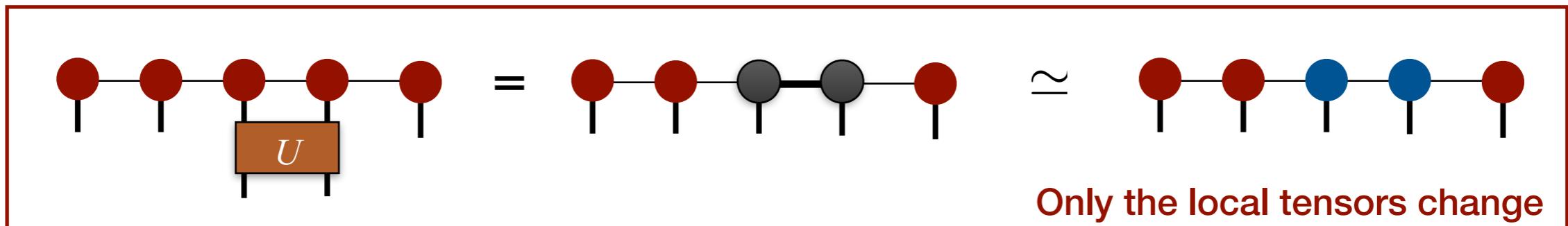
* Even if the initial state is an MPS, the bond dimension of the new state can be increased due to the gate operation.



If we want to repeat gate operations, we need to reduce the bond dimension **by an approximation**.

Approximated gate operation in MPS

Consider an approximation by



- This approximation may not be the best approximation.
 - Generally, local gate operation can affect all tensors in MPS.
 - However, if U is a unitary operator, this approximation is almost optimal.

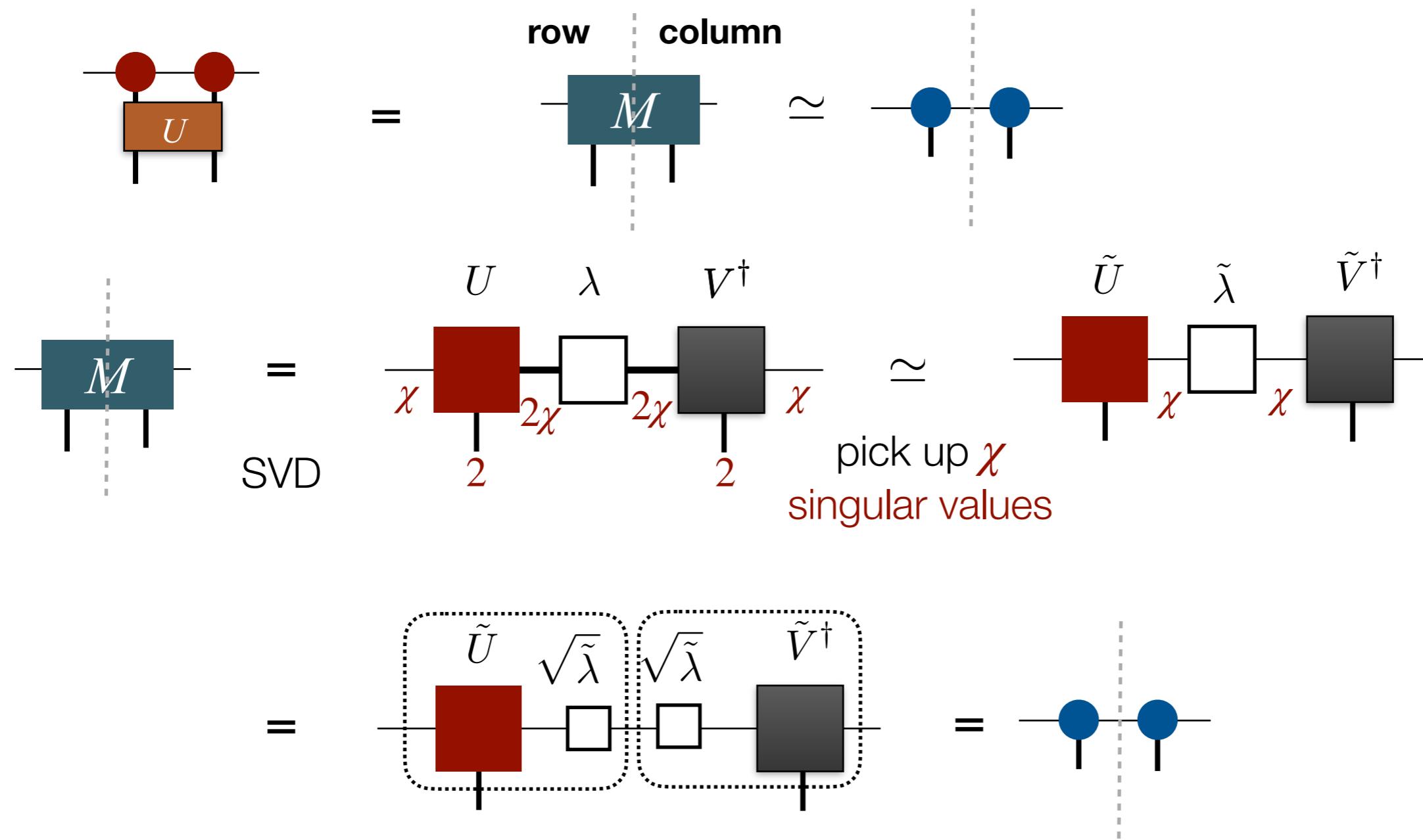
(Rough) approximation neglecting "environment".



Optimal approximation is achieved by SVD.

Approximation by SVD

Approximation by SVD.



Remarks on approximation by MPS

Note that this approximation consider **only** the local structure.

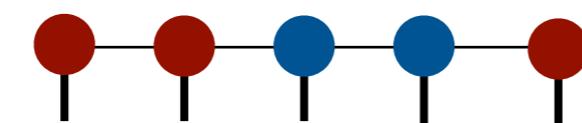
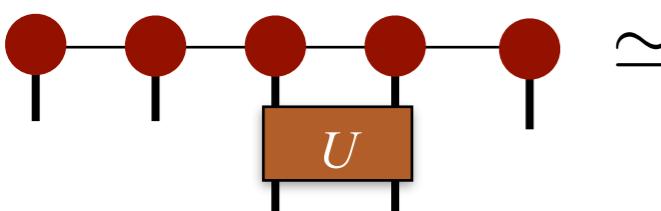
→ In general, the approximation is **not necessarily the best global approximation.**

To include the effect of whole structure, we can use, e.g.,
TEBD algorithm.

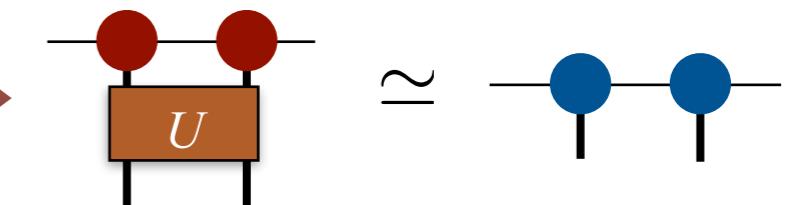
Time evolving block decimation (G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003))

Keywords: Schmidt coefficient, canonical form, entanglement, ...

Global approximation



Local approximation

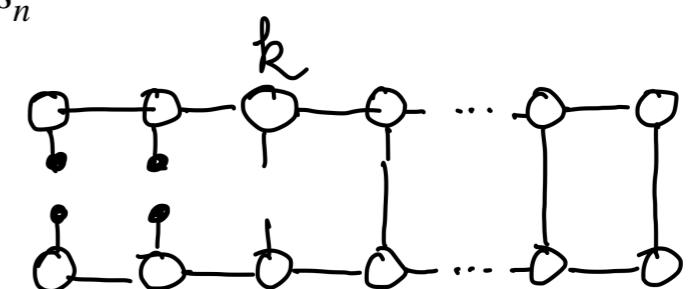
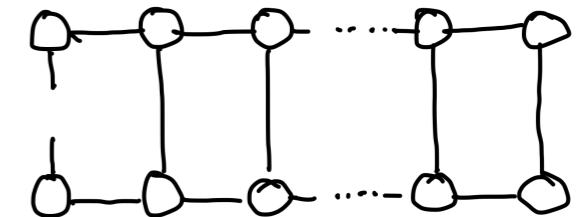


Schrödinger-like Tensor-Network Simulation

- Initial wave function, e.g., $|000\rangle$
 - product state \rightarrow 1 matrix product state (MPS)
- Quantum gate operation
 - 1-qubit operation \rightarrow does not change MPS structure
 - 2-qubit operation \rightarrow contraction of three tensors \rightarrow low-rank approximation by using SVD \rightarrow original MPS structure
- Improving approximation by including environment effects \rightarrow e.g., TEBD
- Generalization to two-dimensional tensor network \rightarrow e.g., PEPS

Sampling States from MPS Wave Function

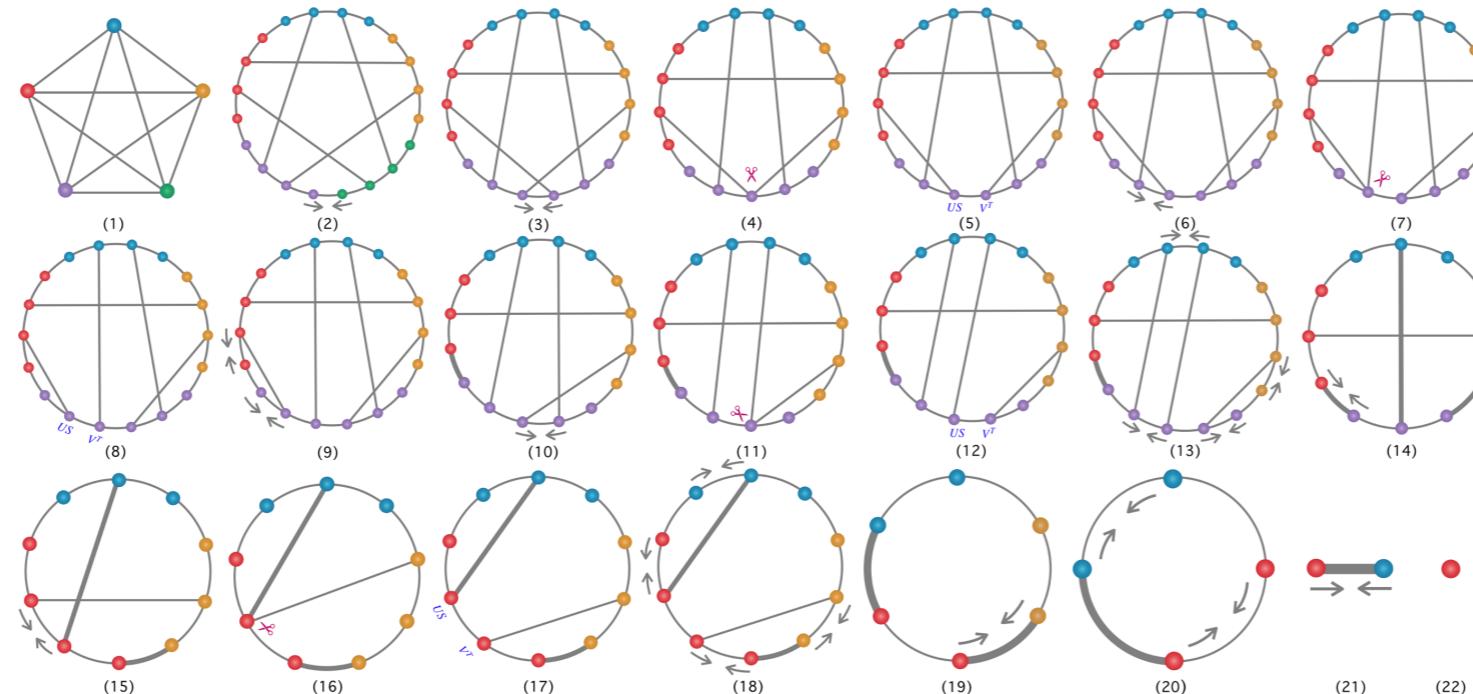
- Sampling in 2^n -dimensional space is difficult
 - $P(s_1, s_2, \dots, s_n) = |\langle s_1 s_2 \cdots s_n | \Psi \rangle|^2 / \langle \Psi | \Psi \rangle$
 - tower sampling: memory cost $\sim O(2^n)$, computation cost $\sim O(2^n)$
- Once (approximate) MPS representation of wave function is obtained
 - it is straightforward to sample a state by using
 - marginal distribution: $P(s_1) = \sum_{s_2, \dots, s_n} P(s_1, s_2, \dots, s_n)$
 - conditional distribution: $P(s_k | s_1, \dots, s_{k-1}) = \sum_{s_{k+1}, \dots, s_n} P(s_1, s_2, \dots, s_n)$



Approximate Arbitrary-order Contraction

- Combination of
 - contraction of tensor network with optimized order
 - low-rank approximation using SVD (if bond dimension exceeds χ)

- e.g.)



- Pan, F., Zhou, P., Li, S., & Zhang, P. Contracting Arbitrary Tensor Networks: General Approximate Algorithm and Applications in Graphical Models and Quantum Circuit Simulations. *Physical Review Letters*, 125(6), 60503 (2020). <https://doi.org/10.1103/PhysRevLett.125.060503>

Schedule of the Pilot Lecture

- 2021/12/14: Tensor network and tensor renormalization group [Okubo]
(テンソルネットワークとテンソル繰り込み群)
- 2021/12/21: Quantum computers and simulations [Todo]
(量子コンピュータ・シミュレーション)
- 2022/1/11: Quantum error corrections and tensor network [Okubo]
(量子誤り訂正とテンソルネットワーク)
- 2022/1/25: Quantum-classical hybrid algorithms and tensor network [Todo]
(量子古典ハイブリッドアルゴリズムとテンソルネットワーク)