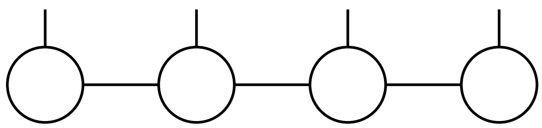
NOTE

YUQING RONG

1. Tensor network & quantum circuit: [1]
2. Parameter shift rule for updating gradients: [2]

1. Isometric tensor network states in two dimensions[3]

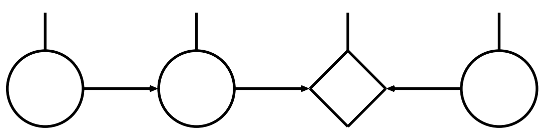
1.1. MPS.



$\psi = T^1 T^2 ... T^N$

not take dimension contraction into account, $\chi_n = \min(d_1 \times ... \times d_n, d_{n+1} \times ... \times d_N)$

isometry condition: $AA^* = I$, $BB^* = I$, then



$\psi = A^1 A^2 ... A^{l-1} \Lambda^l B^{l+1} ... B^N$, Λ is the orthogonal center (also the entanglement spectrum)

$< \psi | O^l | \psi > = < \Lambda | O^l | \Lambda >$ (A...B...are isometry)

1.2. 2D.

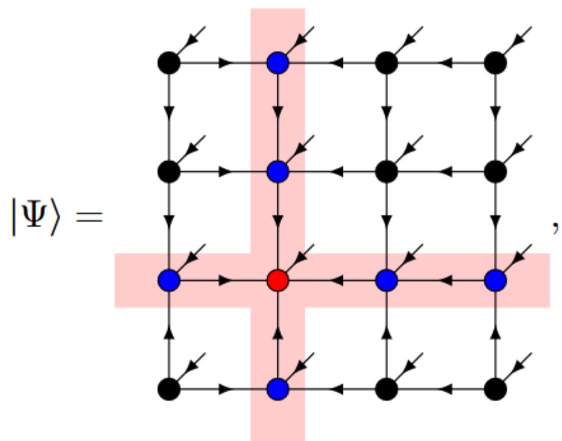


FIGURE 1. isoTNS

the red tensor = orthogonality center (OC)

red and the blue tensors with light red background = the orthogonality hypersurface

1.3. Moses Move.

very close to the optimal variational result; fast.

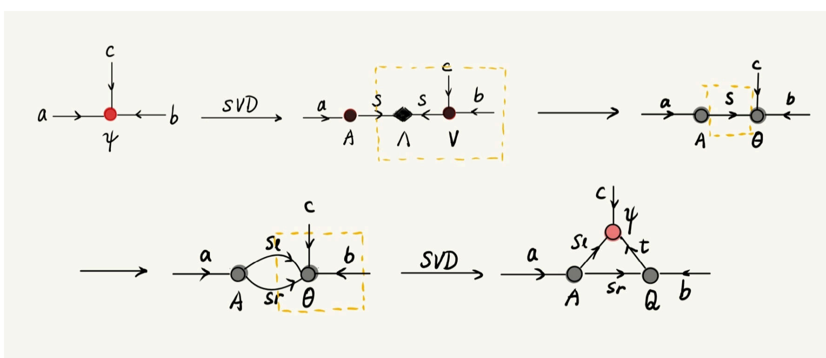


FIGURE 2. Moses Move detail

Repeat the steps above, the orthogonal center moves upward,so we get FIGURE 3:

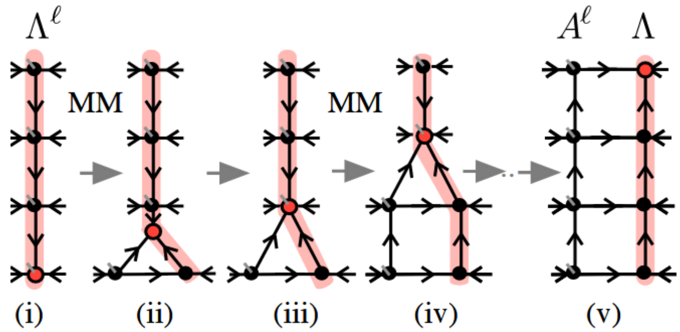


FIGURE 3. Moses Move

Then Λ and B^{l+1} are tied together to form new Λ, $\Lambda^{l+1} = \Lambda B^{l+1}$

1.4. $TEBD^2$ **algorithm.**

(1)Suzuki-Trotter decomposition: $\widehat{\boldsymbol{U}}(\mathrm{dt})=\prod_{r,i} e^{-idtH_i^r} \prod_{c,j} e^{-idtH_j^c}$

(2)Bond truncation local updated at the orthogonality center, just like time evolution operator act on it.

(3)Utilize the SVD and MM to move around the orthogonality center and orthogonality hypersurface.

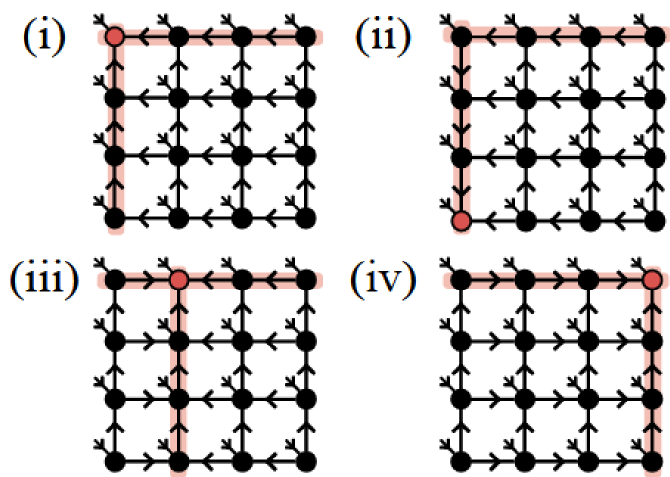


FIGURE 4. TEBD

(i)After 1 round, the isometries rotate 90° counterclockwise, apply $\widehat{\boldsymbol{U}}^{col}(\mathrm{dt})=\prod_{c,j} e^{-idtH_j^c}$ to the states.

(ii)After 2 round, the isometries rotate 180° counterclockwise, apply $\widehat{\boldsymbol{U}}^{row}(\mathrm{dt})=\prod_{r,i} e^{-idtH_i^r}$ to the states.

(iii)After 4 round, the orientation of the isometries back to initialization. Errors are canceled out via symmetrization.

1.5. **DMRG algorithm.**

1.5.1. *1D.*

update each tensor $A^{(n)}$, such that

$E_g = \min_{<\psi|\psi>=1} < \psi|\widehat{\boldsymbol{H}}|\psi > = \min_{<\psi|\psi>=1} \sum_{i,j} < \psi|\widehat{\boldsymbol{H}}_{ij}|\psi >$

eg: update the 2nd tensor(other tensor can be seen as fixed tensors), the effective Hamiltonian H_{eff}:

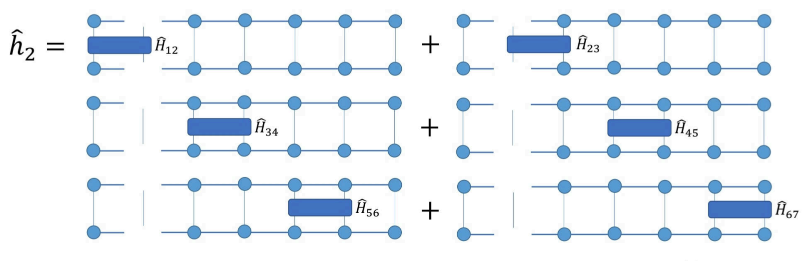


FIGURE 5. effective Hamiltonian

Because all tensors are isometric and canonical transformation, we can move the orthogonality center to the 2nd tensor. Then the problem become

$\min_{<A^{(2)}|A^{(2)^*}>=1} < A^{(2)}|\hat{\boldsymbol{h}}_{\boldsymbol{2}}|A^{(2)^*} >$

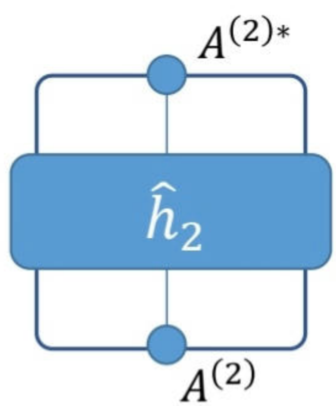


FIGURE 6. contract other tensors

summary:

(1)Initialize all the tensors randomly

(2)SWEEP: Update the tensors from the 1st to Nth in order, then from Nth to 1st as:

(i)Move the orthogonality center to the nth tensor

(ii)Calculate the corresponding effective Hamiltonian h^n

(iii)Update $A^{(n)}$ as the min eigenstate

(3)If MPS converge, done; else if, return to (2)

1.5.2. *2D.*

The operations are similar to 1D, except with the addition of a dimension:

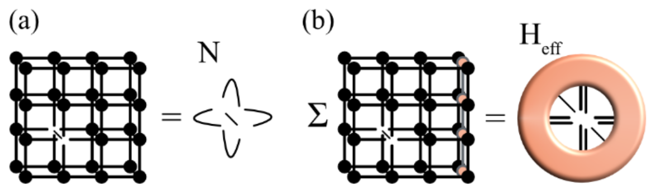


FIGURE 7. 2D DMRG

1.6. **Area Law.**

(1)a pure state's density operator $\rho = |\psi><\psi|$,at zero temperature has vanishing von-Neumann entropy:

S(ρ)=-tr[$\rho \log_2 \rho$]= - $\sum_x \lambda_x \log_2 \lambda_x$

We can use partial trace to find the reduced density matrix: $\rho_A = \mathrm{Tr}_B(\rho)$

(2)area law of entanglement entropy:

divide a D dimension lattice quantum states into two parts, their entropy satisfy:

$S \propto O(l^{D-1})$, l represents the length scale

eg: MPS, D=1, $S \propto O(1)$; PEPS, D=2, $S \propto O(l^1)$

1.7. **truncation error.**

2. SYNERGY BETWEEN QUANTUM CIRCUITS AND TENSOR NETWORKS: SHORTCUTTING THE RACE TO PRACTICAL QUANTUM ADVANTAGE [4]

2.1. **Born rule.**

if a quantum system is described by a wave function $|\psi>$, and an observable (e.g., position, momentum) has eigenstates $|\phi_n\rangle$ with corresponding eigenvalues. the Born rule states that the probability P of measuring the system to be in the state $|\phi_n\rangle$ is given by:

$P = |\langle\phi_n|\psi\rangle|^2$

2.2. **Kullback-Leibler (KL) divergence.**

Given two probability distributions P(true distribution) and Q(approximate distribution), the KL divergence from Q to P is defined as:

(1)discrete case:$D_{KL}(P\|Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$

(2)continuous case: $D_{KL}(P\|Q) = \int P(x) \log \frac{P(X)}{Q(x)} dx$

Properties:

(a)$D_{KL}(P\|Q) \geq 0$;(b)$D_{KL}(P\|Q) \neq D_{KL}(Q\|P)$

2.3. **some gates.**

(1)**U(2)** gate: can represent any rotation on a qubit and can represent any single qubit gates.

$U = \begin{pmatrix} e^{i\alpha}\cos\theta & -e^{i\beta}\sin\theta \\ e^{i\beta}\sin\theta & e^{-i\alpha}\cos\theta \end{pmatrix}$ or $U(\theta,\Phi,\lambda) = \begin{pmatrix} \cos\frac{\theta}{2} & -e^{i\lambda}\sin\frac{\theta}{2} \\ e^{i\Phi}\sin\frac{\theta}{2} & e^{i(\Phi+\lambda)\cos\frac{\theta}{2}} \end{pmatrix}$

(2)**XX,YY,ZZ** gate: represent a rotation around XX, YY, ZZ axis in the space of two qubits; generate entanglement between 2 qubits.

$$XX(\theta) = \exp\left(-i\frac{\theta}{2}X_1X_2\right) = \begin{pmatrix} \cos\frac{\theta}{2} & 0 & 0 & -i\sin\frac{\theta}{2} \\ 0 & \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} & 0 \\ 0 & -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} & 0 \\ -i\sin\frac{\theta}{2} & 0 & 0 & \cos\frac{\theta}{2} \end{pmatrix}$$

X_1 and X_2 are the Pauli-X operators acting on the first and second qubits respectively; θ is a real parameter determines the rotation angle around the XX axis.

(3)**SU(4)** gate: two-qubits quantum gates; satisfy $U^{+}U = I$, determinant=1.

$SU(4)_{ij}(\theta) = U(2)(\theta)_{i(\theta_{1:3})} \times U(2)(\theta)_{j(\theta_{4:6})} \times XX_{ij}(\theta_7) \times YY_{ij}(\theta_8) \times ZZ_{ij}(\theta_9) \times U(2)(\theta)_{i(\theta_{10:12})} \times U(2)(\theta)_{j(\theta_{13:15})}$

(4)**U(4)** gate:

$U(4) = SU(4)(\theta) \times e^{-i\Phi}$, $e^{-i\Phi}$ is a global phase.

3. Encoding of Matrix Product States into Quantum Circuits of One- and Two-Qubit Gates [5]

3.1. **Encoding matrix product state into single-layer quantum circuit.**

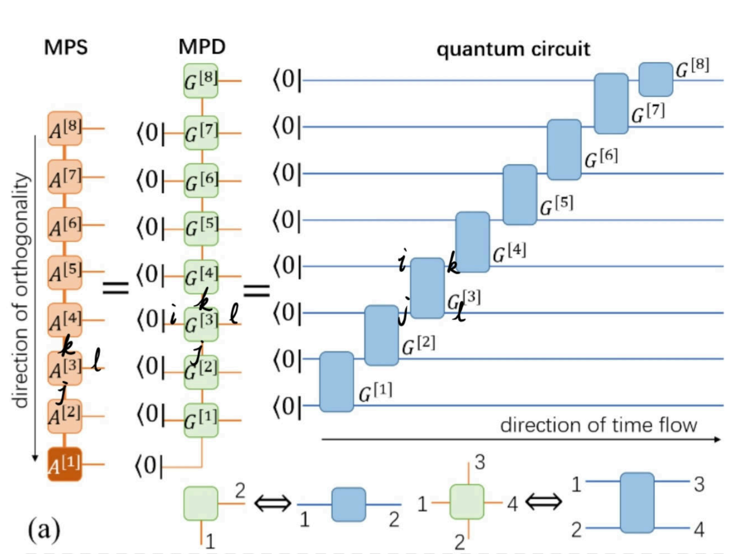


FIGURE 8. MPS → MPD → QC

MPS(N sites):

$|\psi>=\sum_{a_1...a_{N-1}}\sum_{s_1...s_N}A^{[1]}_{s_1,a_1}A^{[2]}_{s_1,a_1,a_2}...A^{[N]}_{s_N,a_{n-1}}\prod_{n=1}^{N}|S_N>$

(1)They satisfy the normalization and left orthogonal conditions.

$\sum_{s_1,a_1}A^{[1]}_{s_1,a_1}A^{[1]*}_{s_1,a_1}=1$

$\sum_{s_n,a_n}A^{[n]}_{s_n,a_{n-1},a_n}A^{[n]*}_{s_n,a'_{n-1},a_n}=I_{a_{n-1},a'_{(n-1)}}$

$\sum_{s_N}A^{[N]}_{s_N,a_{N-1}}A^{[N]*}_{s_N,a'_{N-1}}=I_{a_{N-1},a'_{(N-1)}}$

(2)$\chi=d=2$

MPS: $|\psi>$; MPD: $\widehat{\boldsymbol{U}}$

$|\psi>=\widehat{\boldsymbol{U}}^{+}|0>$

$A_{jkl}\to G_{ijkl}$ (a)i=0, $G^{[n]}_{0jkl}=A^{[n]}_{jkl}$ (b)i $=1,\ \cdot\ \cdot\ \cdot$, d -1,G_{ijkl} are obtained in the kernel of A_{jkl}. $\sum_{kl}G^{n}_{i'j'kl}G^{n}_{ijkl}=I_{i'i}I_{j'j}$

dim(i)=dim(k); dim(j)=dim(l)

(3)negative logarithmic fidelities (NLF) per site:

$F_0=-\frac{\ln|<\psi|\psi^{\sim}_{\chi=1}>|}{N}$

$F_1=-\frac{\ln|<\psi|\widehat{\boldsymbol{U}}^{+}|0>|}{N}$

if $|\psi^{\sim}_{\chi=1}>=|\psi>$, $F_0=0$; The gap between $|\psi^{\sim}_{\chi=1}>$ and $|\psi>$ greater, F_0 greater.

Our goal is minimizing F_1. cost function?

3.2. Deep quantum circuit.

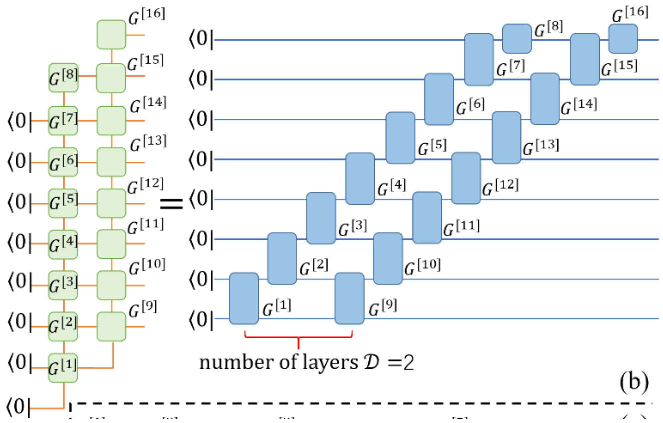


FIGURE 9. D=2 layers quantum circuit

D layers: $|\psi^{\sim}>= \widehat{\boldsymbol{U}}_D^{+}...\widehat{\boldsymbol{U}}_2^{+}\widehat{\boldsymbol{U}}_1^{+}|0>$ $\quad F_D = -\frac{\ln|<\psi|\widehat{\boldsymbol{U}}_D^{+}...\widehat{\boldsymbol{U}}_2^{+}\widehat{\boldsymbol{U}}_1^{+}|0>|}{N}$

4. Simulating Large PEPs Tensor Networks on Small Quantum Devices [6]

4.1. PEPS.

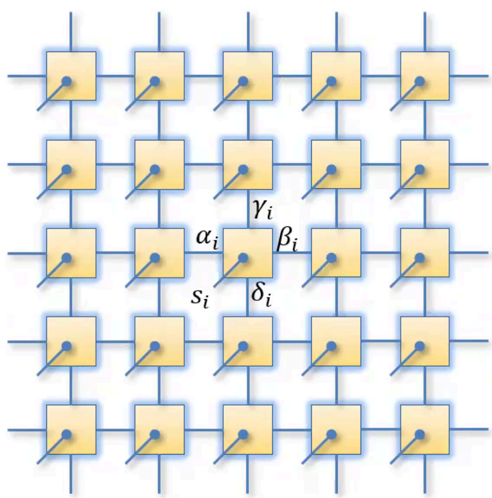


FIGURE 10. PEPS

$|\psi>= \sum_{s=1}^{N}\left(\sum_{\{\alpha,\beta,\gamma,\delta\}=1}^{D} ...A_{\alpha_i,\beta_i,\gamma_i,\delta_i}^{s_i}...\right)|S_n>$

4.2. Relationship of the qubits.

1. virtual qubits: $N_B = \lceil \log \varkappa \rceil$

$\varkappa$ represents the rank of entanglement,that is, the number of possible states that can be shared across the bond. So, $\varkappa = 2^{N_B}$

2. $N \times M$ lattices, total qubits: $N_Q = (N+1) \times N_B + 1$

$N+1$ means open boundary condition, $+1$ means auxiliary qubit uesd to reset or measure.

N_Q is the total number of virtual qubits. Physical qubits are not included because they are not the main resource cost.

3. qubit-efficient:

$N_Q = (N+1) \times N_B + 1 < N \times M$, that is

$M > \lfloor (1 + \frac{1}{N}) N_B + \frac{1}{N} \rfloor$ or $N_B < \lfloor \frac{N \times M - 1}{N+1} \rfloor$

4.3. more output than input, add additional input.

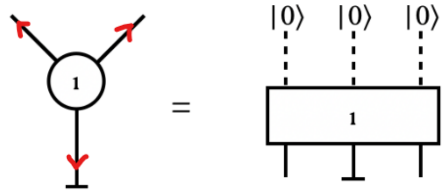


FIGURE 11. PEPS

tensor 1 has 3 output and 0 input, so we add 3 input and fix them |0>

4.4. Example of 4×4.

1. zig-zag pattern

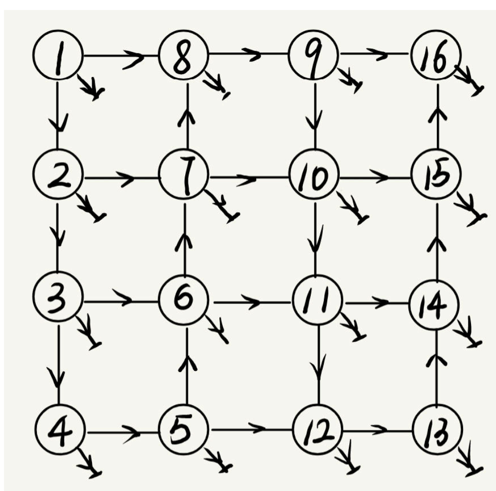


FIGURE 12. zig-zag pattern of 4×5

2. map to quantum circuit

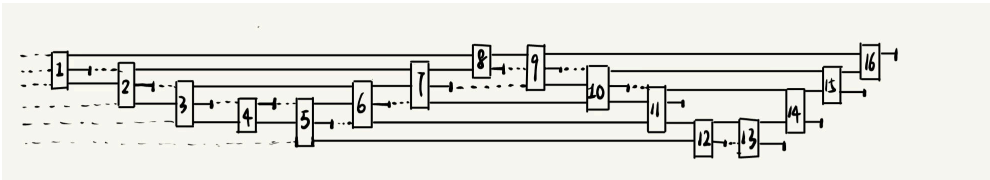


FIGURE 13. 6-qubits circuit of 4×4

the qubits needed only rely on N

5. MPO representation of Hamiltonian

Given an Hamiltonian H, the MPO expresses it as a tensor network:

$H = \sum_{ijk...} M^{[1]} M[2]...M[N]$

M is a tensor encodes local operators.

1. Ising model

$H_{Ising} = -J \sum_i S_i^z S_{i+1}^z - h \sum_i S_i^x$

$M = \begin{pmatrix} I & S^z & S^x \\ 0 & 0 & S^z \\ 0 & 0 & I \end{pmatrix}$

The left boundary: ML=$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$; The right boundary: MR=$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Therefore, only the 1st row 3rd column works.

bond dimension: D=3

2. Heisenberg model

$H_{Heisenberg} = J \sum_i S_i^x S_{i+1}^x + \sum_i S_i^y S_{i+1}^y + \sum_i S_i^z S_{i+1}^z$

M=$\begin{pmatrix} I & S^x & S^y & S^z & 0 \\ 0 & 0 & 0 & 0 & S^x \\ 0 & 0 & 0 & 0 & S^y \\ 0 & 0 & 0 & 0 & S^z \\ 0 & 0 & 0 & 0 & I \end{pmatrix}$

D=5

6. Lanczos method

1. Courant-Fischer Minimax Theorem: If $A \in \mathbb{R}^{n \times n}$ is symmetric, then

$\lambda_k(A) = \max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{y^T A y}{y^T y}$, for k=1:n.

make sense: $y = \sum_{i=1}^k c_i q_i$, $Aq_i = \lambda_i q_i \Rightarrow q_i^T A q_i = \lambda_i \Rightarrow R(A, y) = \frac{y^T A y}{y^T y} = \frac{\left(\sum_{i=1}^k c_i q_i\right)^T A \left(\sum_{i=1}^k c_i q_i\right)}{\left(\sum_{i=1}^k c_i q_i\right)^T \left(\sum_{i=1}^k c_i q_i\right)} = \frac{\sum_{i=1}^k \lambda_i c_i^2}{\sum_{i=1}^k c_i^2}$

$\lambda_1 > \lambda_2 > ... > \lambda_k > ... > \lambda_n$

set of all possible solutions of $\min_{0 \neq y \in S} \frac{y^T A y}{y^T y}$ is $\{\lambda_k, ..., \lambda_n\}$, then, the maximum of them is λ_k

2. Krylov Subspace: $K(A, q_1, k) = [q_1, Aq_1, A^2q_1, ..., A^{k-1}q_1]$

$Q^TAQ = T, QQ^T = I_n$

$Q = [q_1, q_2, ..., q_n]$

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & ... & 0 \\ \beta_1 & \alpha_2 & ... & ... \\ ... & ... & ... & \beta_{n-1} \\ 0 & ... & \beta_{n-1} & \alpha_n \end{pmatrix}$$

since $AQ = QT, Aq_k = \beta_{k-1}q_{k-1} + \alpha_k q_k + \beta_k q_{k+1} \Rightarrow q_k^T A q_k = \alpha_k$,

$r_k = (A - \alpha_k I)q_k - \beta_{k-1}q_{k-1} = \beta_k q_{k+1} \Rightarrow q_{k+1} = \frac{r_k}{\beta_k}$

$\beta_k = \|r_k\| \Rightarrow q_{k+1} = \frac{r_k}{\|r_k\|}$

$q_0 = 0, \beta_0 = 1$, $r_0 = q_1$ is randomly chosen

7. OMEinsum

8. Question

differences between MPS^2 and $PEPS$

Morses Move and zig-zag

References

1. Markov, I. L., Shi, Y.: Simulating quantum computation by contracting tensor networks. SIAM Journal on Computing. 38, 963–981 (2008)
2. Mitarai, K., Negoro, M., Kitagawa, M., Fujii, K.: Quantum circuit learning. Physical Review A. 98, (2018). https://doi.org/10.1103/physreva.98.032309
3. Zaletel, M. P., Pollmann, F.: Isometric Tensor Network States in Two Dimensions. Physical Review Letters. 124, (2020). https://doi.org/10.1103/physrevlett.124.037201
4. Rudolph, M. S., Miller, J., Motlagh, D., Chen, J., Acharya, A., Perdomo-Ortiz, A.: Synergistic pretraining of parametrized quantum circuits via tensor networks. Nature Communications. 14, (2023). https://doi.org/10.1038/s41467-023-43908-6
5. Ran, S.-J.: Encoding of matrix product states into quantum circuits of one- and two-qubit gates. Physical Review A. 101, (2020). https://doi.org/10.1103/physreva.101.032310
6. MacCormack, I., Galda, A., Lyon, A. L.: Simulating Large PEPs Tensor Networks on Small Quantum Devices, https://arxiv.org/abs/2110.00507

, HUSTINGS, Tennessee, TN 59341
Email address: yrong265@connect.hkust-gz.edu.cn
URL: math.ue.edu/~jdoe