



Variational quantum algorithms for state preparation & matrix decomposition

Xin Wang
Baidu Research

PCL Innovation Salon

2020/07/31

Based on arXiv:2005.08797 and 2006.02336.



Overview

- Near-term Quantum Computing
- Quantum Gibbs State Preparation
- Quantum Singular Value Decomposition
- Summary

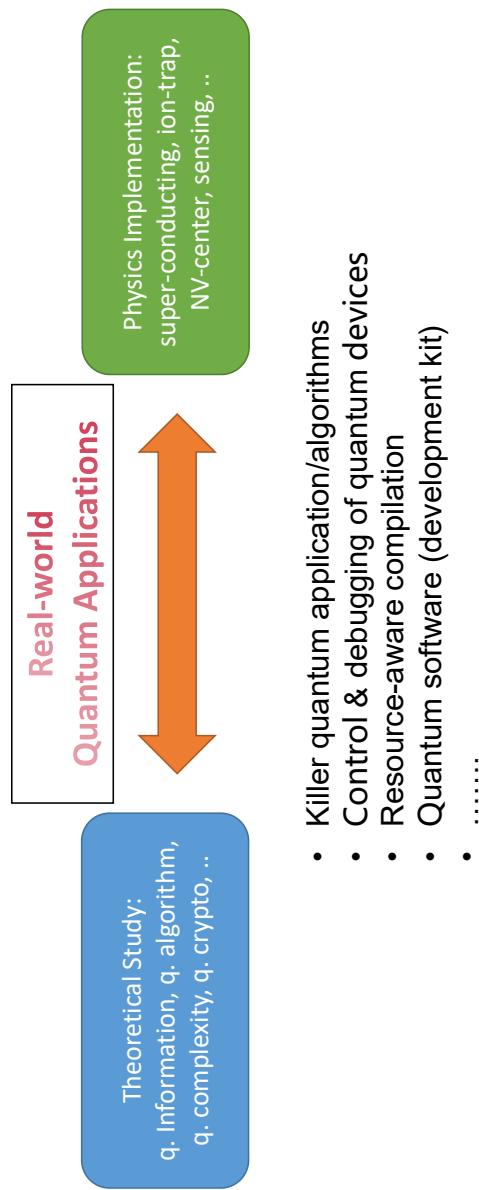


PART 01

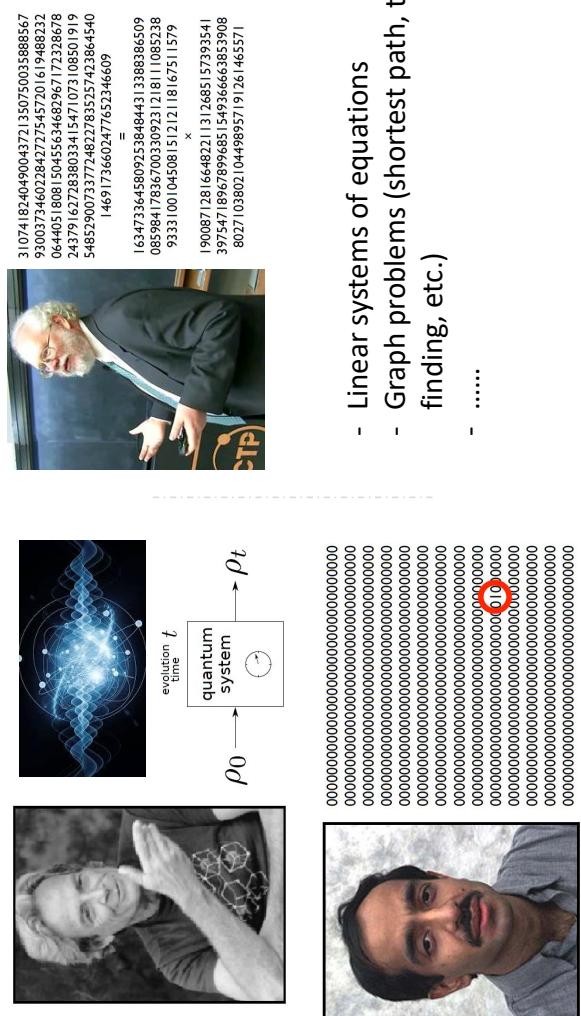
Background

Background

- Major academic and industry efforts are currently in progress to realize scalable quantum hardware and develop powerful quantum software.
- The quantity and quality of physical qubits are continuously increasing!
- This is an *exciting* time for quantum computing!



Existing quantum algorithms for classically hard problems



图片: www.sciencenews.org/, en.wikipedia.org

Towards Near-term Quantum Applications



Requirements	Universal QC	NISQ, 50-200 noisy qubits
Goals	killer apps	executable killer apps
Techs	Algorithm design	ML, optimization, etc.
SDK	Quantum simulator platform	QML platform, etc.  Quantum

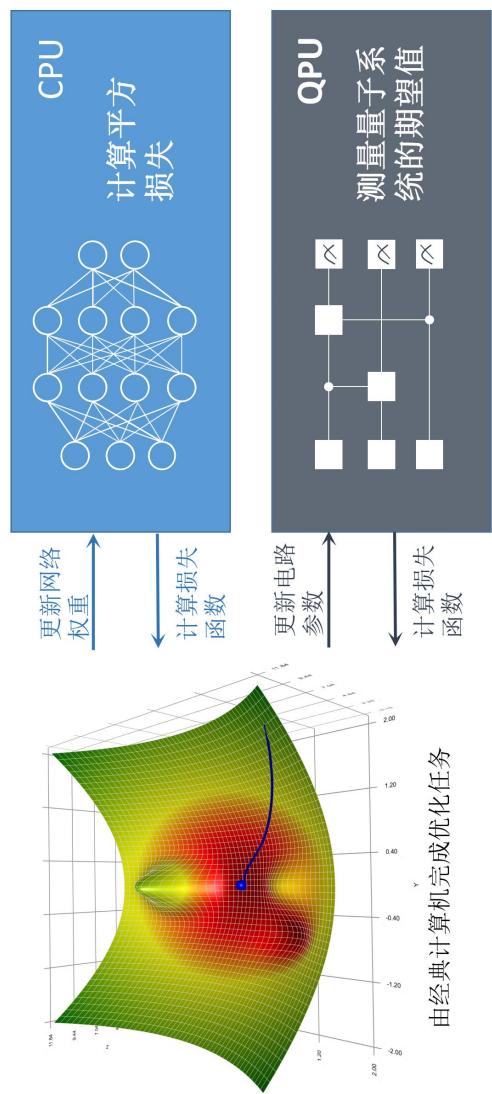
There are still many challenges.

Near-term quantum algorithms

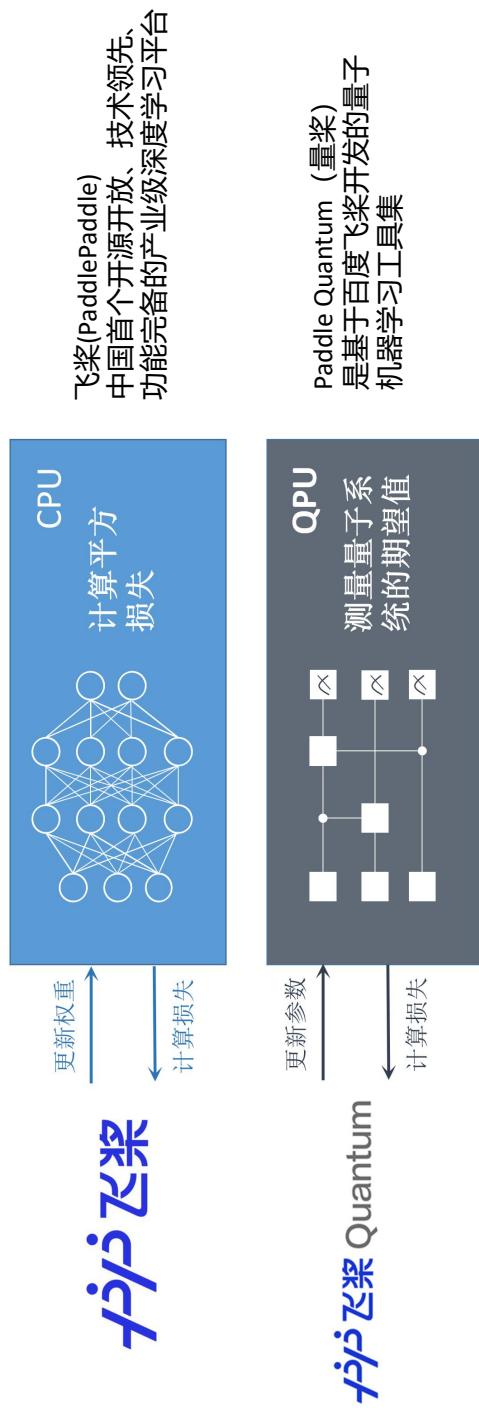
- A trend of near-term quantum algorithms is to employ the promising hybrid quantum-classical algorithms as machine learning models
- Use parameterized circuits to search the Hilbert space and combine classical optimization methods to find optimal parameters.
- Believed to be **best hope for near-term quantum advantage**
- Few rigorous scaling results known for VQAs
- Opportunities and challenges
 - VQAs proposed for:
 - Quantum data compression
 - Quantum eigen-solver
 - Quantum metrology
 - Quantum error correction
 - Quantum state diagonalization
 - Quantum fidelity estimation
 - Quantum simulation
 - Solving linear systems of equations
 - ...
 - Parameterized quantum circuit (PQC) \approx Quantum neural network (QNN)
 - Hybrid quantum-classical algorithm \approx Variational quantum algorithm (VQA)

Review: Benedetti et al. Parameterized quantum circuits as machine learning models.
S. McCarron, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, RMP 2020

Variational quantum algorithms as ML models



Variational quantum algorithms as ML models



- Machine learning may help us better solve problems in quantum computation.
- With QML development platforms, we could focus more on the study of near-term quantum applications.



PART 02

Quantum Gibbs State Preparation

Joint work with Youle Wang and Guangxi Li

arXiv:2005.08797

What is quantum Gibbs state?

- Definition

$$\rho_\beta = \frac{e^{-\beta H}}{\text{tr}(e^{-\beta H})}$$

where H is a Hamiltonian, $\beta = (kT)^{-1}$ is the inverse temperature.

- Why we need Gibbs states?
 - Better understand and study many-body problems
 - Quantum simulation
 - Quantum machine learning
 - Quantum optimization
- In particular, sampling from Gibbs states of Hamiltonians can be applied in solving combinatorial optimization problems, solving SDPs, and training quantum Boltzmann machines.
- However, the preparation of desired initial state is a difficult problem in general (Watrous'09).

Related work and our goal

- Existing methods
 1. Quantum rejection sampling. [PW09, PRL; WKS16, QIC2016]
 2. Quantum walk. [YG12, PNAS]
 3. Dimension reduction. [BB10, PRL]
 4. Dynamic simulation. [KKR17, PRL, RGE12, PRL]
- Require the use of complex quantum subroutines such as quantum phase estimation, which are costly and hard to implement on near term quantum computers.
- How to prepare Gibbs state on NISQ devices?
- A feasible scheme is to employ variational quantum algorithms.

Our Approach

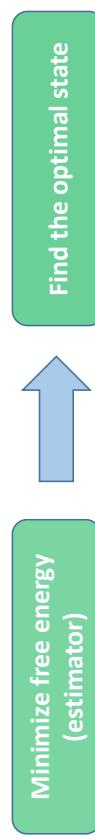
Starting point: A key feature of the Gibbs state is that it minimizes the free energy

$$F(\rho) = \text{tr}(H\rho) - \beta^{-1}S(\rho)$$

- a) F denotes the free energy;
- b) H denotes the Hamiltonian;
- c) ρ denotes the quantum state;
- d) β denotes the inverse temperature;
- e) $S(\rho) = -\text{tr}(\rho \log \rho)$ denotes the von Neumann entropy.

Hence, ρ_G can be approximately prepared by the following variational principle of Gibbs state:

$$\rho_G \approx \underset{\theta}{\operatorname{argmin}} \mathcal{F}(\rho(\theta)),$$



Loss function

- Major obstacle for the free energy evaluation is von Neumann entropy estimation.

- We truncate the von Neumann entropy.

$$S_K(\rho) = \sum_{k=1}^K \frac{(-1)^k}{k} \text{Tr}((\rho - I)^k \rho) = \sum_{j=0}^K C_j \text{Tr}(\rho^{j+1}).$$

- Let H denote the Hamiltonian and $\beta > 0$ be the inverse temperature, we define

$$\mathcal{F}_K(\theta) = \text{Tr}(H\rho(\theta)) - \beta^{-1} S_K(\rho(\theta)).$$

- Can be represented as a linear combination of higher order state overlap
- Given truncation order K , β , let ρ^* denote the output of the algorithm, and let ρ_β denote the true Gibbs state, then

$$F(\rho^*, \rho_\beta) \geq 1 - \sqrt{2\beta \left(\epsilon + \frac{2r}{K+1} (1 - \Delta)^{K+1} \right)}.$$

- r denotes the rank of ρ^* ; ϵ denotes the error tolerance;
- $\Delta \in (0, e^{-1})$ is a constant determined by K .

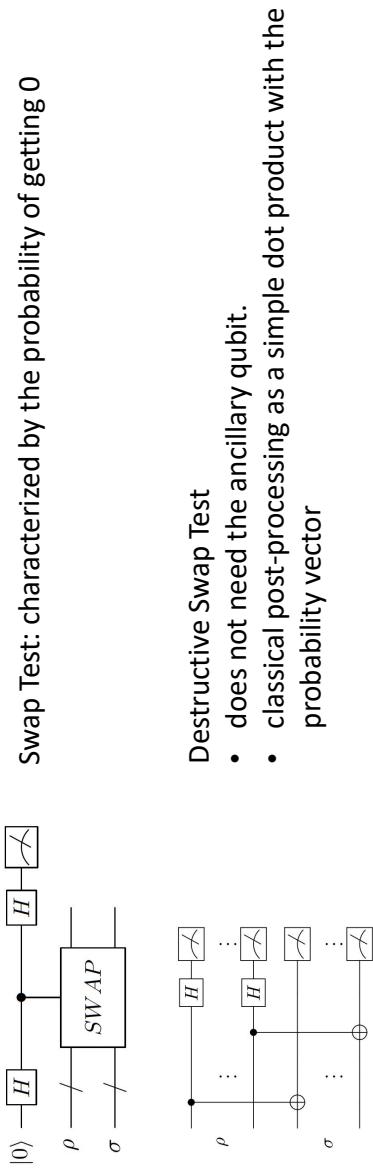
2-truncated free energy

- In particular, we choose the 2-truncated free energy (convex) as the loss function and show that both the loss function and their gradients can be evaluated on NISQ devices.

$$\mathcal{F}_2(\theta) = \text{Tr}(H\rho(\theta)) + \beta^{-1} \left(2 \text{Tr}(\rho(\theta)^2) - \frac{1}{2} \text{Tr}(\rho(\theta)^3) - \frac{3}{2} \right)$$

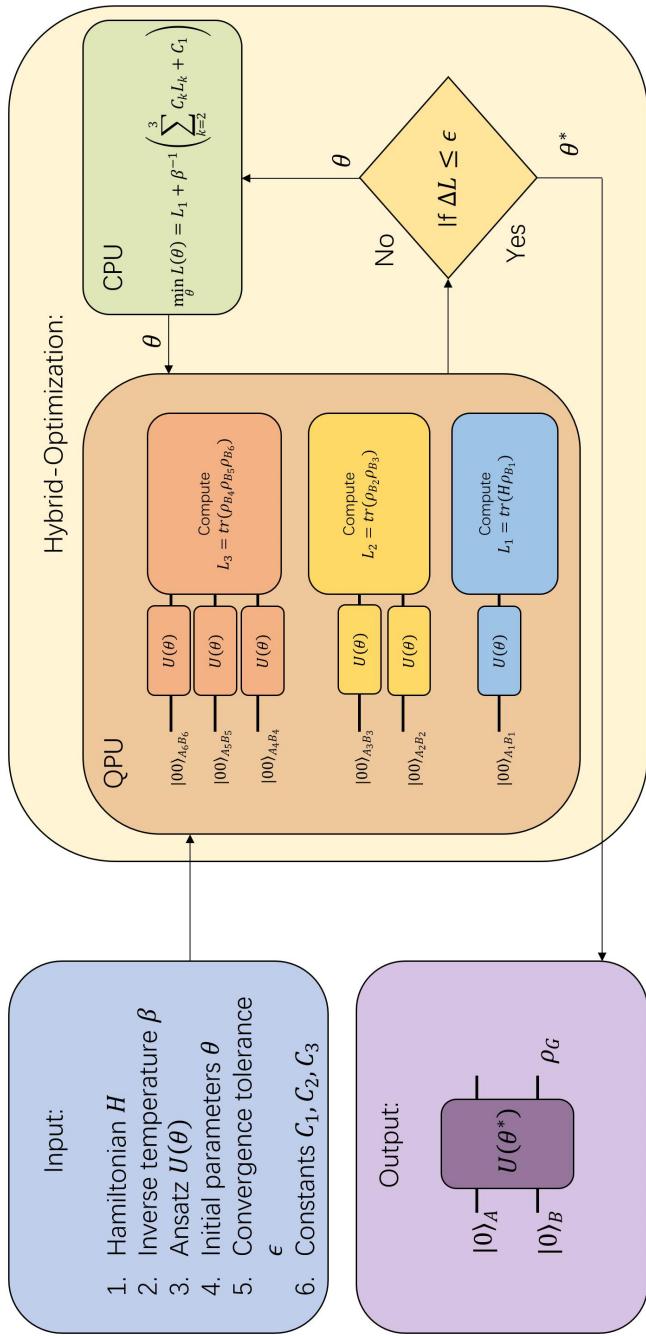
- Analytical gradients $\frac{\partial \mathcal{F}_2(\theta)}{\partial \theta_m} = \frac{1}{2} (\langle K \rangle_{\theta_m + \frac{\pi}{2}} - \langle K \rangle_{\theta_m - \frac{\pi}{2}}) + \beta^{-1} [(2(\langle O \rangle_{\theta_m + \frac{\pi}{2}, \theta_m} - \langle O \rangle_{\theta_m - \frac{\pi}{2}, \theta_m}) - \frac{3}{4} (\langle G \rangle_{\theta_m + \frac{\pi}{2}, \theta_m, \theta_m} - \langle G \rangle_{\theta_m - \frac{\pi}{2}, \theta_m, \theta_m})]$,
- With analytical gradients, one could apply gradient-based methods to minimize the loss function.
 - Either gradient-based or gradient-free optimization methods.
- Gradients for VQA: Mitarai et al. arXiv:1803.00745, Schuld et al. arXiv:1811.11184, Ostaszewski et al. arXiv:1905.09692, Li et al. arXiv:1608.00677

Compute the loss function (gradients) via Swap test

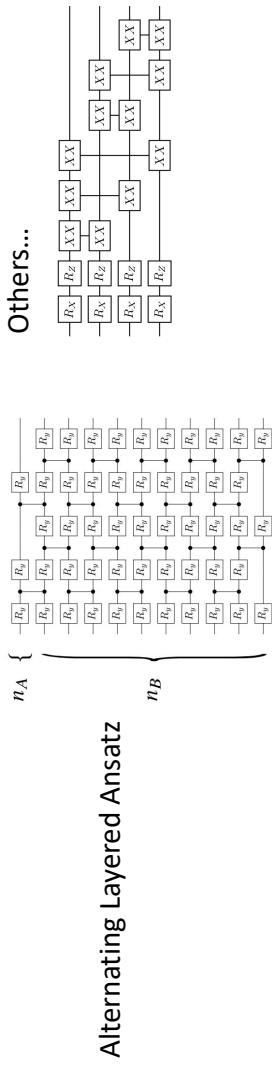
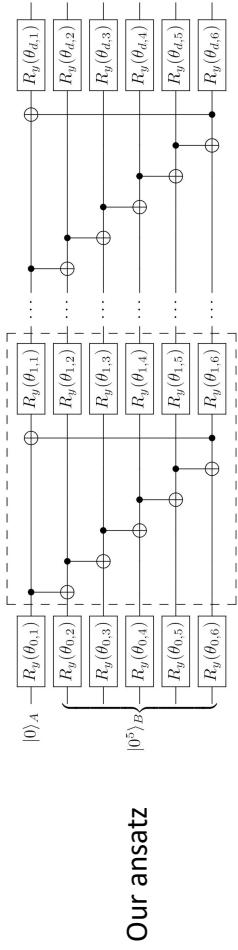


[1] Y. Subasi, L. Cincio, and P.J. Coles, J. Phys. A Math. Theor. 52, 044001 (2019).

Overview of this hybrid quantum-classical algorithm



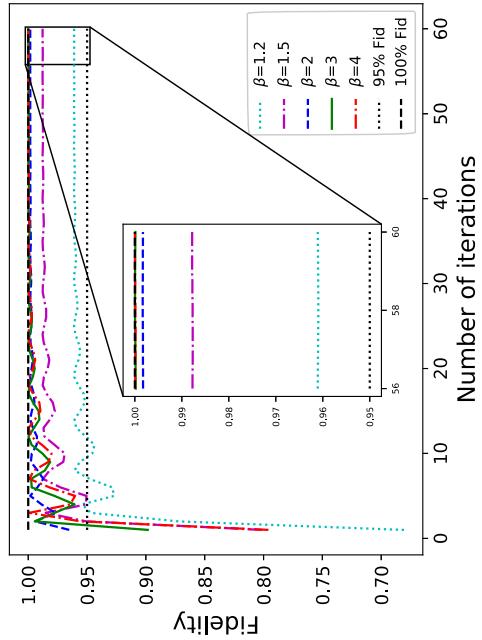
Ansatz for our numerics (Parameterized circuits)



Ising chain model

$$H = - \sum_{l=1}^5 Z^l \otimes Z^{l+1}$$

- Shallow parameterized circuits
- Only one additional qubit in the ansatz
- Prepare the Ising chain Gibbs states with a fidelity higher than 95%.



Findings for Ising chain model

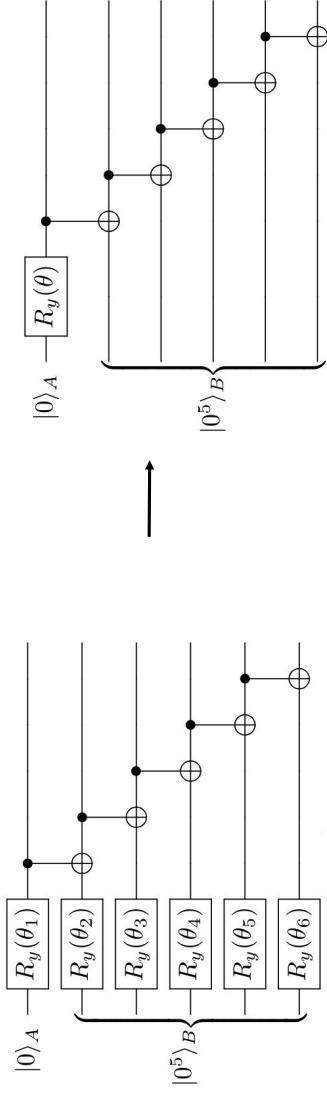
- Both $F(\theta)$ and $F_2(\theta)$ reach their global minimum at $\theta = \frac{\pi}{2}$;
- Fidelity with the target Gibbs state $F\left(\rho\left(\frac{\pi}{2}\right), \rho_\beta\right)_1$ increases fast as β increases

$$F\left(\rho\left(\frac{\pi}{2}\right), \rho_\beta\right) \geq \sqrt{1 + \binom{N}{2} e^{-\beta \Delta}},$$

where N is the dimension;

- This ensures the good performance for large β .

- One **ancilla** suffices to prepare the Gibbs state of Ising model.



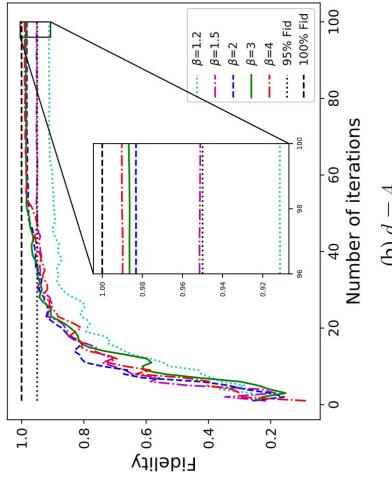
XY spin-1/2 chain model

- Our second instance is the XY spin-1/2 chain of length L=5, with the Hamiltonian

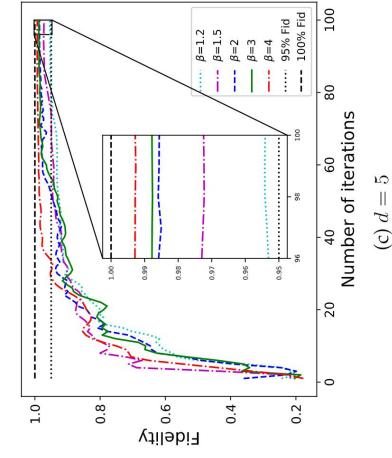
$$H_B = - \sum_{i=1}^L X_{B,i} X_{B,i+1} + Y_{B,i} Y_{B,i+1}$$

and periodic boundary conditions.

- 6-qubit parametrized circuit with one ancillary qubit, where the basic circuit module (which contains a CNOT layer and a layer of single qubit Pauli-Y rotation operators) is repeated d times



(b) $d = 4$



(c) $d = 5$

Summary for Gibbs state preparation

- We propose a variational quantum algorithm for quantum Gibbs state preparation.
- We utilize the truncated free energy to evaluate the free energy.
- We demonstrate our results by providing theoretical evidences and numerical experiments for Ising chain and spin chain Gibbs states.

PART 03

Variational Quantum SVD

Joint work with Zhixin Song and Youle Wang

arXiv:2006.02336

What is Singular Value Decomposition (SVD)?

Given a $m \times n$ dimensional matrix M , there exists the following decomposition:

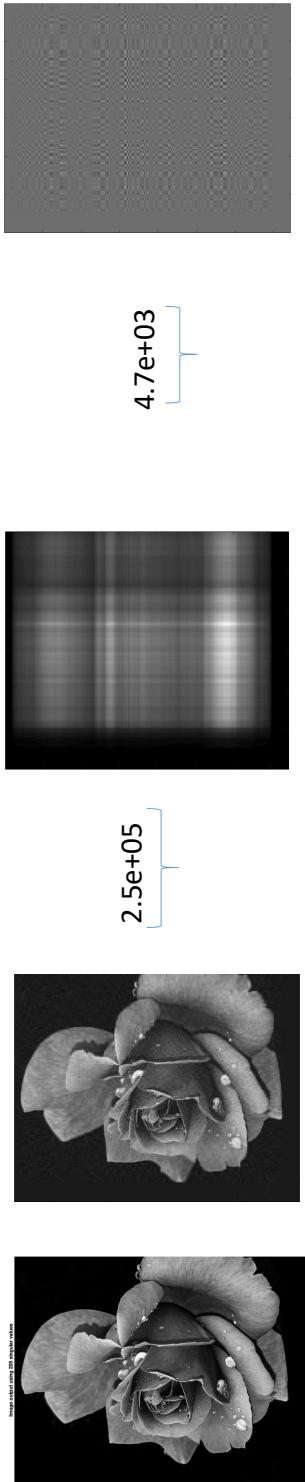
$$M_{m \times n} = U_{m \times m} \times D_{m \times m} \times V^{\dagger}_{n \times n}$$

The diagram illustrates the SVD components:

- $M_{m \times n}$ (Blue bracket) is the input matrix.
- $U_{m \times m}$ (Orange bracket) is a unitary matrix.
- $D_{m \times m}$ (Blue bracket) is a diagonal matrix containing singular values d_1, d_2, d_3, d_4 .
- $V^{\dagger}_{n \times n}$ (Green bracket) is a unitary matrix.
- A central multiplication symbol (\times) indicates the product of $U_{m \times m}$ and $D_{m \times m}$.
- Curly braces group the matrices $U_{m \times m}$ and $D_{m \times m}$ together.
- Curly braces group the matrices $D_{m \times m}$ and $V^{\dagger}_{n \times n}$ together.
- Curly braces group the matrices $U_{m \times m}$, $D_{m \times m}$, and $V^{\dagger}_{n \times n}$ together.

where U and V are two unitaries satisfying $UU^{\dagger} = I_{m \times m}$ and $VV^{\dagger} = I_{n \times n}$

Example (Application in image compression)



Mathematical applications of the SVD

- computing the pseudo inverse
- matrix approximation
- estimating the range and null space of a matrix.
- SVD has also been successfully applied to many areas of science, engineering, and statistics, such as signal processing, image processing, and recommender systems.

Setup and motivation

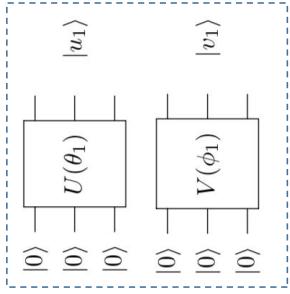
- For a given $n \times n$ matrix M , there exists a decomposition of the form

$$M = UDV^\dagger$$

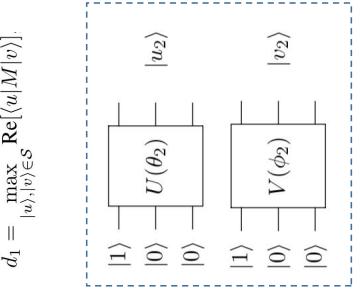
- Assumption on the input matrix as a linear combination of unitaries
- Our goal is to design a quantum algorithm for SVD.
- Motivations
 - Compression of quantum data
 - Analysis of quantum data (e.g., eigenvalues of Hamiltonians/quantum states)
 - Quantum linear system solver
 - Potential speed-up for SVD and many related applications



Starting point: Variational principles of SVD



$$d_1 = \max_{|u\rangle, |v\rangle \in \mathcal{S}} \operatorname{Re}[\langle u | M | v \rangle].$$



$$d_2 = \max_{|u_2\rangle, |v_2\rangle \in \mathcal{S}} \operatorname{Re}[\langle u_2 | M | v_2 \rangle].$$

- A naïve approach is to design QNNs to learn each singular value.
- However, this is not efficient. Can we find a way to learn U and V directly?

Ref on SVD: <https://www.caam.rice.edu/~caam440/pca.pdf>

Our solution

We introduce the following loss function

$$\uparrow M = UDV^\dagger$$

$$L(\alpha, \beta) = \sum_{j=1}^T q_j \times \text{Re}\langle \psi_j | U(\alpha)^\dagger M V(\beta) | \psi_j \rangle$$

where the weights $q_1 > q_2 > \dots > q_T > 0$, and $\{|\psi_j\rangle\} = \{|000\rangle, |010\rangle, \dots, |111\rangle\}$

This loss function has several nice properties

- Could find **all** the singular values and singular vectors via training
- Theoretical guarantee of the ideally optimized solution
- Could be computed on near-term quantum devices (**Hadamard Test**)

Theoretical reason for choosing $L(\alpha, \beta)$

- Let's assume that $\langle \psi_j | U(\alpha)^\dagger M V(\beta) | \psi_j \rangle = m_j$ are real numbers for simplicity.
- We have

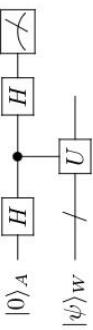
$$\begin{aligned}
 L(\alpha, \beta) &= \sum_{j=1}^T q_j \times \operatorname{Re} \langle \psi_j | U(\alpha)^\dagger M V(\beta) | \psi_j \rangle \\
 &\leq \sum_{j=1}^T q_j \times m_j \\
 &= \sum_{j=1}^T (q_j - q_{j+1}) \sum_{t=1}^j m_t^\downarrow \quad \xrightarrow{\text{Theorem 2 Ky Fan Theorem states that}} \quad \sum_{j=1}^T d_j = \max_{\text{orthonormal } \{v_j\}, \{u_j\}} \sum_{j=1}^T \langle u_j | M | v_j \rangle. \\
 &\leq \sum_{j=1}^T (q_j - q_{j+1}) \sum_{t=1}^j d_t. \quad \xleftarrow{\text{---}}
 \end{aligned}$$

Main message: the QNNs could learn the SVD after training via $L(\alpha, \beta)$

- The loss function is maximized if and only if $\langle \psi_j | U(\alpha)^\dagger M V(\beta) | \psi_j \rangle$ extracts the singular value d_j for each j .
- The QNNs U and V learn the singular vectors in the sense that $\langle \psi_j | U(\alpha)^\dagger M V(\beta) | \psi_j \rangle$ extract the left and right singular vectors, respectively.

Compute the loss function via Hadamard Test

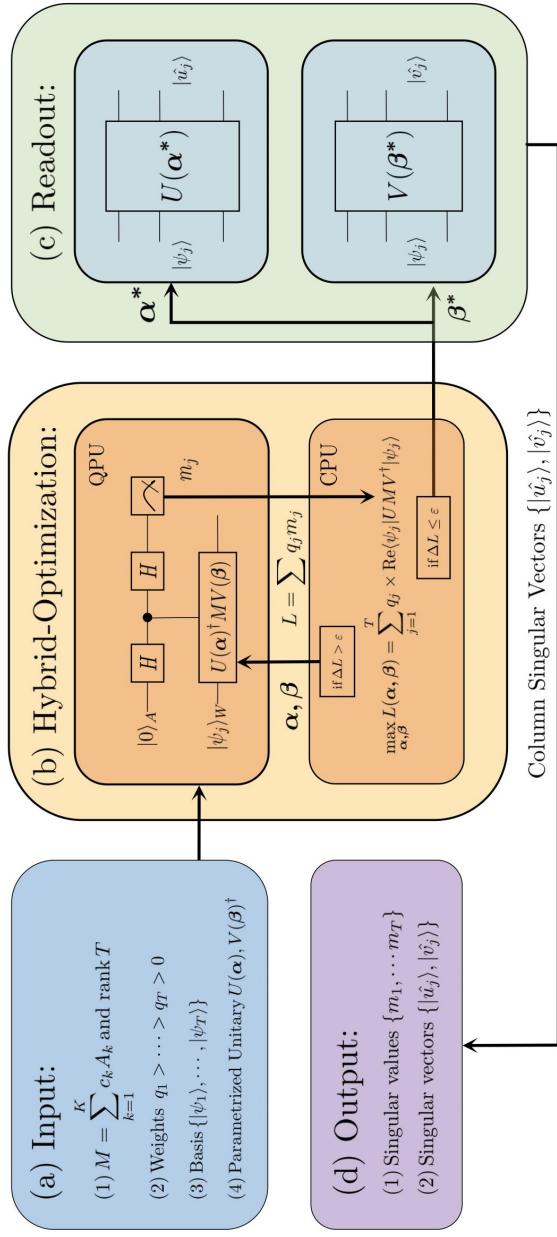
- The loss function only requires estimating the quantity $\text{Re}\langle\psi|U|\psi\rangle$ for a general unitary U
- Apply a controlled unitary U , conditioned on the qubit in register A and prepare state $\frac{1}{\sqrt{2}}(|0\rangle_A|\psi\rangle_W + |1\rangle_A U|\psi\rangle_W)$
- Then apply a Hadamard gate on the ancillary qubit A , and measure it. If the measurement outcome is 0, set the output to be 1; otherwise, set it to be -1. Thus the expectation of output is $\text{Re}\langle\psi|U|\psi\rangle$
- We could also estimate $\text{Im}\langle\psi|U|\psi\rangle$ via the state $\frac{1}{\sqrt{2}}(|0\rangle_A|\psi\rangle_W - i|1\rangle_A U|\psi\rangle_W)$



Quantum circuit for implementing Hadamard test

[1] D. Aharonov, V. Jones, and Z. Landau, *Algorithmica* 55, 395(2009), arXiv:0511096 [quant-ph].

Schematic diagram of vQSVD algorithm



Optimization

- Both gradient-based and gradient-free methods could be used to do the optimization.
- We show that **analytical gradients** in our VQSVD could be estimated easily on near-term devices by a “*parameter shift rule*” [1].

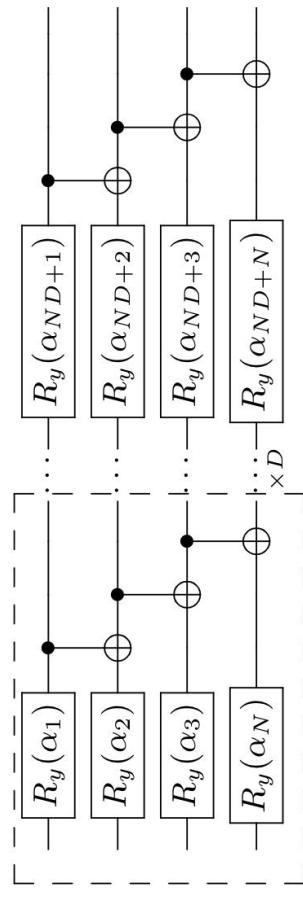
$$\begin{aligned}\frac{\partial L(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_\ell} &= \frac{1}{2} \sum_{j=1}^T q_j \times \text{Re} \langle \psi_j | U^\dagger (\alpha_\ell + \pi) M V(\boldsymbol{\beta}) | \psi_j \rangle & \frac{\partial L(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_k} &= \frac{1}{2} \sum_{j=1}^T q_j \times \text{Re} \langle \psi_j | U^\dagger (\boldsymbol{\alpha}) M V(\beta_k - \pi) | \psi_j \rangle \\ &= \frac{1}{2} L(\alpha_\ell + \pi, \boldsymbol{\beta}) & &= \frac{1}{2} L(\boldsymbol{\alpha}, \beta_k - \pi).\end{aligned}$$

- Compare to the finite difference method (FDM), we only need to rotate the angle **once not twice**.
- Additionally, Harrow & Napp [2] find positive evidence that circuit learning using the analytical gradient outperforms any FDM.

[1] <https://arxiv.org/pdf/1811.11184.pdf>
[2] <https://arxiv.org/pdf/1901.05374.pdf>

Numerical experiments

We use the following Hardware-efficient Ansatz [1] as our circuit model:

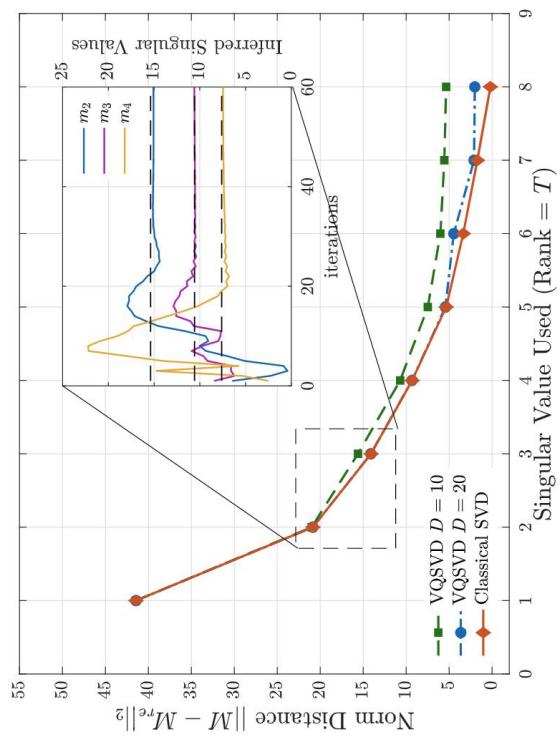


The above ansatz works well for problems with real numbers.

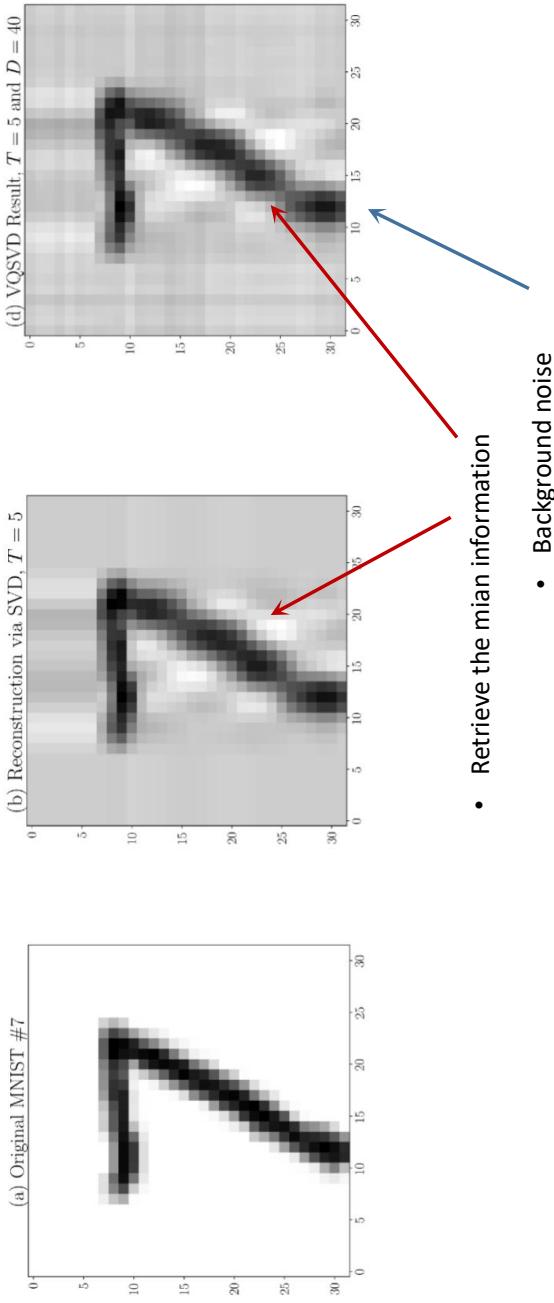
SVD for random matrices

$$M_{re} = \sum_{i=1}^T m_i |u_i\rangle \langle v_i|$$

$$\|A_{n \times n}\|_2 = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$$



Toy example in image compression



Summary of vQSVD and future directions

- ◆ A novel loss function to train the QNNs to learn the left and right singular vectors and output the target singular values.
- ◆ Positive numerics for SVD of random matrices and image compression
- ◆ Extensive applications in solving linear systems of equations.

-
- ◆ How to load classical data into quantum devices efficiently?
 - ◆ How would quantum noise affect the performance of QML algorithms?
 - ◆ How to better train the QNNs and avoid barren plateaus issues?
 - ◆ More applications?
 - ◆ New QNN architectures?



主要功能

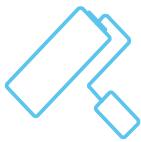
Quantum Features of Paddle Quantum



Easy to use



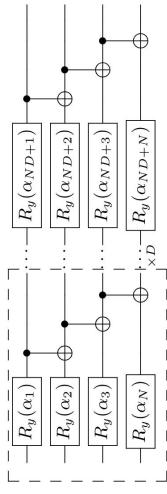
Extensibility



Featured toolkits

- Easy-to-build QNN
- Fruitful tutorials
- Support general circuit model
- Hybrid quantum-classical algorithms
- Provide toolkits for quantum chemistry, QAOA
- Self-innovate QML applications

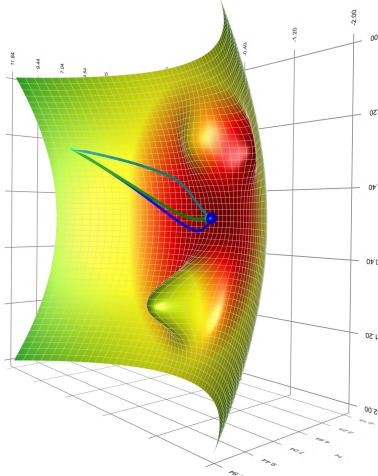
Make QML Development Easier!



```

def U_theta(theta, input_state, N, D):
    cir = UAnsatz(N, input_state=input_state)
    for repeat in range(D):
        for i in range(1, N):
            cir.cnot.control=[i, i + 1]
        for i in range(N):
            cir.ry(theta=theta[repeat][0][i], which_qubit=i + 1)
    return cir.state

```



```

# 初始化paddle动态图机制
with fluid.dygraph.guard():

    # net:
    net = vqe_demo(shape=[D + 1, 3, N])

    # 一般来说 我们利用Adam优化器来获得好的收敛,当然你也可以使用SGD或者RMS prop.
    optimizer = fluid.optimizer.adagradOptimizer(learning_rate = LR,
                                                parameter_list = net.parameters())

    # 优化循环
    for itr in range(ITER):

        # 前向传播计算损失函数
        loss = net(initial_state, N, D)

        # 在动态图模式下, 反向传播最小化损失函数
        loss.backward()
        optimizer.minimize(loss)
        net.clear_gradients()

    # 记录学习曲线
    loss_list.append(loss.numpy()[0])
    print('iter:', itr, ' loss: %.4f' % loss.numpy())

```

Welcome submissions to Quantum!



the open journal for quantum science

HOME PUBLICATIONS CALL FOR EDITORS INSTRUCTIONS PEOPLE ABOUT

Doi, title, author, arXiv id, ...

Quantum has published 300 Papers in 4 Volumes, as well as 42 Views in Quantum Views.

PERSPECTIVE

Causality meets resource theory

Rafael Chaves,

Quantum Views 4, 42 (2020).

When presenting Bell's theorem [1] to audiences seeing it for the first time, be they students or surprisingly physicists from areas with little regard to the Quantum foundations, often...

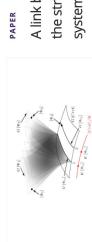


PAPER

Law without law: from observer states to physics via algorithmic information theory

Markus A. Müller,
Quantum 4, 30 (2020).

According to our current conception of physics, any valid physical theory is supposed to describe the objective evolution of a unique external world. However, this condition is challenged by...



PAPER

A link between symmetries of critical states and the structure of SLOCC classes in multipartite systems

Oskar Slavik, Martin Hebenstreit, Barbara Kraus, and Adam Smolek,
Quantum 4, 30 (2020).
Central entanglement theory is characterized by local transformations among pure multipartite states. As a first step

SEARCH

SUBMIT

Quantum is an open-access peer-reviewed journal for quantum science and related fields.

Quantum is non-profit and community-run; an effort by researchers and for researchers to make science more open and publishing more transparent and efficient.

Sign up for our monthly digest of papers and other news.

PEOPLE

Steering Board

Anne Brodbeck
Harry Buhrman
Jens Eisert
Debbie Leung
Chi-Yang Lu
Ana Mariá Ray
Anna Sepera
Umesh Sinha
Robert W. Spekkens
Reinhard Werner
Birgitta Whaley
Andreas Winter

Editors

Ahsan Nazir

Antonio Achúcarro

Carlo Bennink

<https://quantum-journal.org/>

Thanks!

Recruitment

- ✉ Mathematics/ Computer Science / Quantum Physics
- ✉ Passion, persistence, and patience
- ✉ Opening positions:
 - Researchers
 - Interns
 - Visiting Scholars, "Polaris Program" (> 2 months).
- ✉ Contact quantum@baidu.com