SCIENCE CHINA
Information Sciences

• LETTER •

February 2023, Vol. 66 129502:1–129502:2
https://doi.org/10.1007/s11432-021-3382-2

A hybrid quantum-classical Hamiltonian learning algorithm

Youle WANG[1,2], Guangxi LI[1,2] & Xin WANG[2*]

[1]*Center for Quantum Software and Information, University of Technology Sydney, Sydney* NSW 2007*, Australia;*
[2]*Institute for Quantum Computing, Baidu Research, Beijing* 100193*, China*

Received 1 August 2021/Revised 28 September 2021/Accepted 10 December 2021/Published online 14 November 2022

Citation Wang Y L, Li G X, Wang X. A hybrid quantum-classical Hamiltonian learning algorithm. Sci China Inf Sci, 2023, 66(2): 129502, https://doi.org/10.1007/s11432-021-3382-2

Dear editor,

Hamiltonian learning is a central task in studying quantum physics systems and the experimental realization of quantum computers. It can be applied to predict the quantum system's locality and describe the effective interactions between particles, which plays a crucial role in quantum technology, such as quantum simulation and adiabatic quantum computation. In many-body physics, the system's Hamiltonian is composed of a sequence of Pauli strings and associated interaction coefficients. The goal of Hamiltonian learning is to recover the interaction coefficients from the measurement data of the system. Despite the number of these coefficients in general scales polynomially in the system's size, it is pretty challenging to do this. The most straightforward method is tomography, which gives the classical description of the Hamiltonian at the expense of exponential costs. Recently, it shows that the problem can be solved by repeatedly preparing and measuring the associated Gibbs state [1]. This method could cost polynomially many copies of the Gibbs state, but the measurement data processing can be demanding. The rapid development of quantum computers has brought hope since they promise computational capacity over their classical counterparts. However, utilizing near-term available quantum devices for Hamiltonian learning remains unknown.

In this study, we propose a hybrid quantum-classical Hamiltonian learning algorithm (HQHL) that is friendly to near-term quantum computers. The main idea is to utilize the measurement data processing of [1], which involves optimizing a strongly convex function. In particular, we exploit a gradient-descent method to find the desired interaction coefficients, with the help of quantum computers to estimate the gradient.

Problem statement. We assume that the target Hamiltonian H is composed of n-qubit Pauli strings $\{E_\ell\}_{\ell=1}^m$ and interaction coefficients, i.e., $H=\sum_{\ell=1}^m \mu_\ell E_\ell$. The measurements corresponding to $\{E_\ell\}_{\ell=1}^m$ are performed on the Gibbs state $\rho_\beta = \mathrm{e}^{-\beta H}/\mathrm{Tr}(\mathrm{e}^{-\beta H})$ at an inverse temperature β, and the measurement results are denoted by $e_\ell = \mathrm{Tr}(\rho_\beta E_\ell)$, $\forall \ell = 1,\ldots,m$. In our method, we take $\{e_\ell\}_{\ell=1}^m$ as inputs since many methods have been proposed to efficiently obtain the measurement results of various measurement operators.

To recover the interaction coefficients, we first adopt a strategy that was proposed in [1], which transforms the Hamiltonian learning problem into an optimization problem by using Jaynes' principle (or maximal entropy principle) [2]. This strategy is to find a quantum state with the maximal entropy from all states whose measurement results under $\{E_\ell\}_{\ell=1}^m$ match $\{e_\ell\}_{\ell=1}^m$.

$$\begin{aligned} &\max_{\rho}\ S(\rho) \\ &\text{s.t.}\ \ \mathrm{Tr}(\rho E_\ell) = e_\ell,\ \forall \ell = 1,\ldots,m, \\ &\qquad \rho > 0,\ \mathrm{Tr}(\rho) = 1. \end{aligned} \tag{1}$$

It has been shown in [2] that the optimal state is of the following form:

$$\sigma = \frac{\exp(-\beta\sum_{\ell=1}^m \mu_\ell^* E_\ell)}{\mathrm{Tr}(\exp(-\beta\sum_{\ell=1}^m \mu_\ell^* E_\ell))}. \tag{2}$$

Here, the state σ is a quantum Gibbs state of a Hamiltonian with interaction coefficients $\boldsymbol{\mu}^* = (\mu_1^*,\ldots,\mu_m^*)$. As a result, Ref. [1] shows that coefficients of σ are the target interaction coefficients, i.e., $\boldsymbol{\mu}^* = \boldsymbol{\mu}$. Moreover, Ref. [1] also points out an approach for obtaining $\boldsymbol{\mu}^*$ that is to solve the dual optimization problem, which is derived by applying the Lagrange multiplier method to the problem in (1).

$$\boldsymbol{\mu} = \mathrm{argmin}_{\boldsymbol{\nu}} \log Z_\beta(\boldsymbol{\nu}) + \beta\sum_{\ell=1}^m \nu_\ell e_\ell. \tag{3}$$

Here, $Z_\beta(\boldsymbol{\nu}) = \mathrm{Tr}(\mathrm{e}^{-\beta\sum_{\ell=1}^m \nu_\ell E_\ell})$ denotes the partition function, parameterized by $\boldsymbol{\nu} = (\nu_1,\ldots,\nu_m) \in [-1,1]^m$.

Our method. Note that the objective function is strongly convex since the log-partition function $\log Z_\beta(\nu)$ is strongly convex [1]. We, therefore, develop a gradient descent method to steadily find the desired solution and recover the unknown Hamiltonian. The main obstacle is to compute the gradients of the objective function $L(\boldsymbol{\nu}) = \log Z_\beta(\boldsymbol{\nu}) + \beta\sum_{\ell=1}^m \nu_\ell e_\ell$, which involves computing the corresponding

* Corresponding author (email: wangxin73@baidu.com)

partition function. Particularly, the component of the gradient $\nabla L(\boldsymbol{\nu})$ is given by

$$\frac{\partial L(\boldsymbol{\nu})}{\partial \nu_l} = -\beta \mathrm{Tr}(\rho_\beta(\boldsymbol{\nu}) E_l) + \beta e_l, \tag{4}$$

where $\rho_\beta(\boldsymbol{\nu}) = \mathrm{e}^{-\beta H(\boldsymbol{\nu})}/Z_\beta(\boldsymbol{\nu})$ and $H(\boldsymbol{\nu}) = \sum_{\ell=1}^m \nu_\ell E_\ell$.

Our method employs two principal subroutines to overcome this issue: one utilizes the parameterized quantum circuits (PQCs) to extract the information about the spectrum of the Hamiltonian, which we call stochastic variational quantum eigensolver (SVQE). The other takes convex optimization tools to estimate the log-partition function $\log Z_\beta(\boldsymbol{\nu})$.

Inspired by the subspace-search variational quantum eigensolver [3], our method combines the PQC with the eigenvalues' property to learn the Hamiltonian spectrum. To be specific, the PQC, denoted by $U(\boldsymbol{\theta})$, consists of single-qubit rotations and the two-qubit CNOT/CZ gate (An example can be found in the Supplementary file). On the other hand, we design the loss function of the variational quantum algorithm using the fact that eigenvalues majorize the diagonal elements, and the dot function with an increasingly ordered vector is Schur concave [4]. Suppose the current Hamiltonian is $H(\boldsymbol{\nu})$, and let N be the dimension and $\{|\psi_j\rangle\}_{j=1}^N$ denote the computational basis. Then the loss function is defined as

$$M(\boldsymbol{\theta}) = \sum_{j=1}^N q_j \cdot \langle \psi_j | U^\dagger(\boldsymbol{\theta}) H(\boldsymbol{\nu}) U(\boldsymbol{\theta}) | \psi_j \rangle, \tag{5}$$

where $\boldsymbol{q} = (q_1, \ldots, q_N)$ is a probability distribution such that $0 < q_1 < q_2 < \cdots < q_N$.

When training $\boldsymbol{\theta}$ to minimize $M(\boldsymbol{\theta})$, a gradient descent method is used to do the optimization. Particularly, the gradient is computed according to the shift rule. That is, we only need to shift the phase of θ_k by $\pm\pi/2$ and evaluate the corresponding loss function if we wish to compute $\partial_{\theta_k} M(\boldsymbol{\theta})$. After optimization, the final parameters will enable PQC to learn the eigenvectors of $H(\boldsymbol{\nu})$ and output eigenvalues.

Computing $M(\boldsymbol{\theta})$ can be daunting for large Hamiltonians, since there are exponentially many values $\langle \psi_j | U^\dagger(\boldsymbol{\theta}) H(\boldsymbol{\nu}) U(\boldsymbol{\theta}) | \psi_j \rangle$ that need to evaluate. Thus, we use importance sampling to estimate $M(\boldsymbol{\theta})$ by regarding it as the expectation of probability $\boldsymbol{q}$. The usage of Chebyshev's inequality implies the required number of measurements scales polynomially with precision and qubit counts.

Our second subroutine is to compute the log-partition function $\log Z_\beta(\boldsymbol{\nu})$. The key idea is to use the relationship between log-partition function and free energy defined as $F(\rho) = \mathrm{Tr}(H(\boldsymbol{\nu})\rho) - \beta S(\rho)$, i.e., $\log Z_\beta(\boldsymbol{\nu}) = -\beta \min_\rho F(\rho)$. From this equation, we need to find the global minimum of the free energy. For this purpose, we choose an alternative version of this equation.

$$\log Z_\beta(\boldsymbol{\nu}) = -\beta \min_{\boldsymbol{p}} \sum_{j=1}^N p_j \cdot \lambda_j + \beta^{-1} \sum_{j=1}^N p_j \log p_j, \tag{6}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$ is the vector of eigenvalues of $H(\boldsymbol{\nu})$, and $\boldsymbol{p} = (p_1, \ldots, p_N)$ represents a probability distribution.

The optimization in (6) is a standard convex optimization. We can use existing convex optimization tools to solve it if we know the eigenvalues λ_j and efficiently evaluate the objective function. On the one hand, we can use the PQC $U(\boldsymbol{\theta})$ to learn the eigenvalues of the Hamiltonian as discussed earlier. On the other hand, we efficiently compute the $\sum_{j=1}^N p_j \cdot \lambda_j$ by combining $U(\boldsymbol{\theta})$ and the sampling technique. As for the quantity $\sum_{j=1}^N p_j \log p_j$, it can be calculated easily since $\boldsymbol{p}$ is stored on classical computers.

With the aforementioned subroutines, we are able to estimate the gradient. In particular, we show that the gradient can be computed in the sense that

$$\frac{\partial L(\boldsymbol{\nu})}{\partial \nu_\ell} \approx -\beta \sum_{j=1}^N \widehat{p}_j^* \cdot \langle \psi_j | U^\dagger(\boldsymbol{\theta}) E_\ell U(\boldsymbol{\theta}) | \psi_j \rangle + \beta e_\ell, \tag{7}$$

where $\widehat{\boldsymbol{p}}^*$ denotes the global optimum of (6). The intuition is that $\sum_{j=1}^N \widehat{p}_j^* U(\boldsymbol{\theta}) |\psi_j\rangle\langle\psi_j| U^\dagger(\boldsymbol{\theta})$ is a well-approximation of $\rho_\beta(\boldsymbol{\nu})$. To be more specific, $U(\boldsymbol{\theta})$ can prepare the eigenvectors of $\rho_\beta(\boldsymbol{\nu})$ since $H(\boldsymbol{\nu})$ commutes with $\rho_\beta(\boldsymbol{\nu})$, and the optimum $\widehat{\boldsymbol{p}}^*$ of (6) approximates well the eigenvalues of $\rho_\beta(\boldsymbol{\nu})$. In addition, we also use the sampling technique to suppress the computational resources in gradient estimation.

Ultimately, we synthesize these subroutines with the gradient-descent method to give the Hamiltonian learning algorithm in Algorithm 1.

Algorithm 1 Hybrid quantum-classical Hamiltonian learning algorithm (HQHL)

Input: Pauli operators $\{E_\ell\}_{\ell=1}^m$, coefficients $\{\nu_\ell\}_{\ell=1}^m$, constants $\{e_\ell\}_{\ell=1}^m$, constant β.

```
Initialize coefficients {ν_ℓ}_{ℓ=1}^m;
Set number of iterations I and l = 1;
Set parameterized quantum circuit U(θ);
Set learning rate r;
while l ⩽ I do
    Set Hamiltonian H(ν) = Σ_{ℓ=1}^m ν_ℓ E_ℓ;
    Train U(θ) by SVQE with H(ν);
    Derive a probability p̂* by performing a log-partition
    function estimation with U(θ) and β;
    Compute gradient ∇L(ν) by a gradient estimation with
    U(θ), p̂*, and β;
    Update coefficients ν ← ν − r∇L(ν);
    Set l ← l + 1;
end while
return The final coefficients ν.
```

Conclusion. This study develops a concrete near-term quantum algorithm for Hamiltonian learning and demonstrates its effectiveness. In particular, we show that learning the spectrum of Hamiltonians during the learning process could produce high-precision estimates of the target interaction coefficients. Our work may have applications in quantum device certification, quantum simulation, and quantum machine learning.

Acknowledgements Youle WANG and Guangxi LI acknowledge support from Australian Research Council (Grant No. DP180100691) and Baidu-UTS AI Meets Quantum project. Guangxi LI acknowledges the financial support from China Scholarship Council (Grant No. 201806070139).

Supporting information The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Anshu A, Arunachalam S, Kuwahara T, et al. Sample-efficient learning of interacting quantum systems. Nat Phys, 2021, 17: 931–935
2 Jaynes E T. Information theory and statistical mechanics. Phys Rev, 1957, 106: 620–630
3 Nakanishi K M, Mitarai K, Fujii K. Subspace-search variational quantum eigensolver for excited states. Phys Rev Res, 2019, 1: 033062
4 Roberts A W, Varberg D E. Convex functions-coll. pure and applied mathematics, 1973. https://archive.org/details/convexfunctions0000robe