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, NSW 2007, Australia
[2]Institute for Quantum Computing, Baidu Research, Beijing 100193, China

1. Introduction

Hamiltonian learning is crucial to the certification of quantum devices and quantum simulators. This work tackles the problem of Hamiltonian learning and proposes a hybrid quantum-classical algorithm to find the coefficients of the Pauli operator components of the Hamiltonian. The algorithm builds on minimizing the system's free energy, and its primary goal is to compute the log-partition function. Specifically, we devise a stochastic variational quantum eigensolver to diagonalize the Hamiltonians and then exploit the obtained eigenvalues to compute the free energy's global minimum using convex optimization. Our approach not only avoids the challenge of estimating von Neumann entropy in free energy minimization but also reduces the quantum resources via importance sampling in Hamiltonian diagonalization, facilitating the implementation of our method on near-term quantum devices. As the proof of principle, we demonstrate our approach's validity by conducting numerical experiments with Hamiltonians of interest in quantum many-body physics.

2. Problem Setting

1. Suppose we are given a many-body Hamiltonian H that is decomposed to the form of

$$H(\mu) = \sum_{l=1}^{m} \mu_l E_l, \tag{1}$$

where m_l denote the real coefficients, and E_l denote the n-qubit tensor products of Pauli matrices. Especially, the number of term is $m = O(poly(n))$.
2. We also suppose we are given measurement results $\{e_l\}_{l=1}^m$ of a Gibbs state

$$\rho = \frac{e^{-\beta H}}{\mathrm{Tr}(e^{-\beta H})}. \tag{2}$$

Here, measurement result is given by

$$e_l = \mathrm{Tr}(\rho E_l). \tag{3}$$

3. Given the measurements $\{e_l\}_{l=1}^m$, the task is to find a vector $\widehat{\mu} \in R^m$ such that

$$|\widehat{\mu}_l - \mu_l| \leq \epsilon, \forall 1 \leq l \leq m. \tag{4}$$

5. Log-partition function estimation

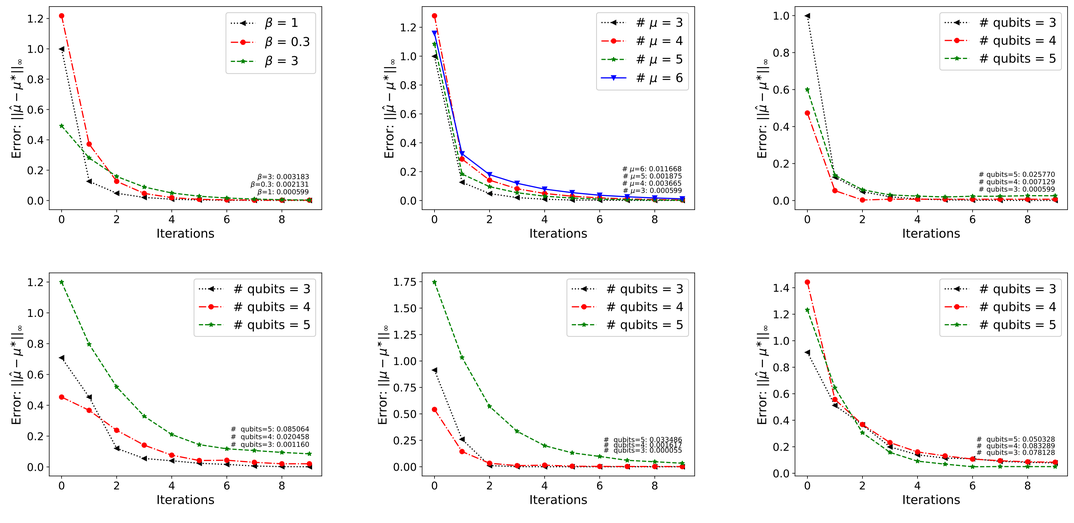
Our main motivation is the relationship between log-partition function and free energy defined as $F(\rho) = \mathrm{Tr}(H(\nu)\rho) - \beta S(\rho)$, i.e., $\log Z_\beta(\nu) = -\beta \min_\rho F(\rho)$. From this relationship, we compute the log-partition function by solving the associative optimization program. However, it is still infeasible to directly minimize the free energy on NISQ devices since estimating entropy is quite challenging. To overcome this challenge, we choose an alternative version, which is

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

where $\lambda = (\lambda_1, ..., \lambda_N)$ is the vector of eigenvalues of $H(\nu)$, and $\vec{p} = (p_1, ..., p_N)$ represents a probability distribution stored on classical computers. On the one hand, $\beta \sum_{j=1}^N p_j \log p_j$ can be computed immediately since $\vec{p}$ is classical. On the other hand, value $\sum_{j=1}^N p_j \cdot \lambda_j$ can be easily estimated via the sampling mean, assuming we have access to the eigenvalues λ.

6. Numerical simulation

We verify our method mainly from three aspects: different β, different numbers of μ (# μ) and a different number of qubits (# qubits). Meanwhile, we demonstrate the performance of our algorithm for quantum many-body models. Particularly, we consider learning the one-dimensional nearest-neighbour Ising model, XY model, and Heisenberg model.



The curves in (a), (b), (c) represent the infinity norm of the error of μ with different β, different number of μ, and different number of qubits, respectively. In (d), (e), (f), the curves represent the infinity norm of the error of μ for different many-body Hamiltonians with the number of qubits varies from 3 to 5. The numbers on the line represent the values of the last iteration. These numbers close to 0 indicate that our algorithm is effective.

3. Main idea

We adopt a strategy based on Gibbs states, proposed recently by Anshu et al., 2020. That is, solving the optimization problem below

$$\mu = \mathrm{argmin}_\nu \log Z_\beta(\nu) + \beta \sum_{\ell=1}^{m} \nu_\ell e_\ell. \tag{5}$$

Here, $Z_\beta(\nu) = \mathrm{Tr}(e^{-\beta \sum_{\ell=1}^m \nu_\ell E_\ell})$ is just the partition function with $N = 2^n$ dimensional input vector $\nu = (\nu_1, ..., \nu_N)$.

To tackle this problem, we use the gradient-descent method to find the global minimum. Particularly, **we utilize variational quantum algorithms and classical optimization to devise a subroutine to diagonalize the Hamiltonian and compute the log-partition function. In addition, the gradients are computed using these tools accordingly.**

4. Hamiltonian diagonalization

We take a parameterized quantum circuit (PQC) $U(\theta)$ and computational basis, denoted by $|\psi_j\rangle$, to generate parameterized states $U(\theta)|\psi_j\rangle$, for all $j = 1, ..., N$. Specifically, we train the PQC $U(\theta)$ to minimize the objective function given as follows:

$$M(\theta) = \sum_{j=1}^{N} q_j \cdot \langle \psi_j | U^\dagger(\theta) H(\nu) U(\theta) | \psi_j \rangle. \tag{7}$$

where $\vec{q} = (q_1, ..., q_N)$ is a probability distribution such that $q_1 < q_2 < ... < q_N$. As a result, we will lead to a PQC that has capability to learn the eigenvectors of Hamiltonian $H(\nu)$ and output eigenvalues.

Notice that terms in Eq. (7) may be exponentially many for matrices with particular real-world applications, endowing a potential obstacle to the loss evaluation. To overcome this challenge, we use importance sampling to estimate $M(\theta)$ by regarding it as the expectation of probability $\vec{q}$. We show the required number of measurements scales polynomially with the precision and qubit counts from Chebyshev's inequality. Viewed from this perspective, we call our method stochastic variational quantum eigensolver.