

Learning Unitary Transformation by Quantum Machine Learning Model

Yi-Ming Huang¹, Xiao-Yu Li^{1,*}, Yi-Xuan Zhu¹, Hang Lei¹, Qing-Sheng Zhu² and Shan Yang³

¹School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, 610054, China

²School of Physics, University of Electronic Science and Technology of China, Chengdu, 610054, China

³Department of Chemistry, Physics and Atmospheric Sciences, Jackson State University, Jackson, MS, 39217, USA

*Corresponding Author: Xiao-Yu Li. Email: xiaoyuuestc@uestc.edu.cn

Received: 08 January 2021; Accepted: 09 February 2021

Abstract: Quantum machine learning (QML) is a rapidly rising research field that incorporates ideas from quantum computing and machine learning to develop emerging tools for scientific research and improving data processing. How to efficiently control or manipulate the quantum system is a fundamental and vexing problem in quantum computing. It can be described as learning or approximating a unitary operator. Since the success of the hybrid-based quantum machine learning model proposed in recent years, we investigate to apply the techniques from QML to tackle this problem. Based on the Choi–Jamiołkowski isomorphism in quantum computing, we transfer the original problem of learning a unitary operator to a min–max optimization problem which can also be viewed as a quantum generative adversarial network. Besides, we select the spectral norm between the target and generated unitary operators as the regularization term in the loss function. Inspired by the hybrid quantum-classical framework widely used in quantum machine learning, we employ the variational quantum circuit and gradient descent based optimizers to solve the min-max optimization problem. In our numerical experiments, the results imply that our proposed method can successfully approximate the desired unitary operator and dramatically reduce the number of quantum gates of the traditional approach. The average fidelity between the states that are produced by applying target and generated unitary on random input states is around 0.997.

Keywords: Machine learning; quantum computing; unitary transformation

1 Introduction

With the enormous variety of applications, machine learning has already impacted modern life [1,2]. It not only provides powerful tools for mining the potential information from massive data but also helps us explore the laws of nature from a new perspective. In recent years, machine learning has caught the attention of researchers from various fields such as physics and chemistry [3,4]. They intend to develop new tools based on machine learning techniques



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

for scientific research. Quantum machine learning is one such rising and promising branch that combines the approaches of data processing from machine learning with the outperformance of quantum computing. As well known, there is computational difficulty in simulating a quantum system by classical methods. As the size of a quantum system increases, the resources required for simulating will increase exponentially.

Based on this fact, quite a lot of quantum machine learning models that harness the quantum advantages have been proposed in recent years. Quantum support vector machine (QSVM) and its physical implementation that runs on a small-scale quantum device result in an exponential speed-up over the corresponding classical one [5]. Quantum principal component analysis (QPCA) is more than a quantum extension of the classical algorithm. It also provides a technique to construct unitary transformation efficiently [6]. Quantum generative adversarial networks (QGAN) inherit the framework of classical models, consisting of a generator and a discriminator. However, the components of QGAN can be constructed by variational quantum circuits such that generating quantum data is more efficient like quantum state preparation [7–10]. These quantum machine learning models are classified into two categories. One category is full quantum that the models only rely on full quantum algorithms including quantum Fourier transformation, quantum phase estimation and HHL algorithm etc. [11,12]. The representative works are QSVM, QPCA, and so on. Another category is hybrid that the models consist of the classical part and quantum part. That is the reason why it is called the hybrid quantum-classical model. These models are formalizing the problem of interest as a variational optimization problem and finding an approximate solution by parametrized quantum circuit and classical optimizer [13]. Compared to the full quantum-based model, the hybrid quantum-classical model requires fewer resources to build, particularly the size of the quantum system, the number of quantum gates, etc. Meanwhile, the well-developed optimization techniques can be naturally employed, which means that we are able to make use of well-maintained open-source optimization libraries such as Tensorflow and Pytorch as the optimizing backend [14,15].

In this work, relying on the advantages of the hybrid-based quantum machine learning model, we intend to exploit it for a unitary learning problem that is an essential task in quantum computing. One reason why learning unitary matrices matters a lot is that the dynamics of a quantum system can be described by a unitary transformation, exploring the efficient way to implement the unitary transformation is fundamental. Since we are still in the early stages of quantum computing with the limitation of quantum resources for computation, it is necessary to investigate how to control or manipulate the quantum system efficiently. Besides, various tasks will benefit from the results of the study on learning unitary transformation. For instance, unitary matrix compiling, that is the task of ‘compiling’ a known unitary transformation into a quantum circuit constructed by a sequence of gates chosen from the specific quantum gate set. A large amount of existing works has been focused on how to implement unitary transformations efficiently. Such works provide theoretic analysis of approximate error and gate or time complexity for implementing target unitary transformation [16,17]. Recently, some work investigated the ideas from machine learning for simulating a desired unitary transformation [18,19]. Similar to these related literatures, we formulate the learning of a given target unitary operator as an optimization problem, and employ the hybrid-based quantum machine learning model to approximate the solution. From our numerical experiments, the results show that the proposed method provides a reasonable approximation while the number of required quantum gates is much less than those of traditional approach.

We organized this paper as follows. In Section 2, we formalize the learning problem of unitary transformation. In Section 3, we give a brief introduction of the techniques in hybrid quantum-classical framework and quantum generative adversarial networks (QGAN). In Section 4, we propose a method by which a quantum machine learning model can be used for solving the formulized unitary learning problem. In Section 5, we provide the numerical experiments which demonstrate the advantages of the proposed model. We also introduce the background of quantum computing in the appendix.

2 Learning Problem of Unitary Transformation

In quantum mechanics, we have two pictures for the dynamics of a quantum system. One is the Hamiltonian picture; another is the unitary operator picture. In the Hamiltonian picture, given the Hamiltonian, a Hermitian operator, that is the total energy function of the closed system, we can describe the evolution of the system by Schrödinger equation as follow,

$$i\hbar \frac{d|\varphi\rangle}{dt} = H|\varphi\rangle \quad (1)$$

It can be verified that the solution of Eq. (1) can be,

$$|\varphi_{t_2}\rangle = e^{-\frac{i}{\hbar}H(t_2-t_1)} |\varphi_{t_1}\rangle = U|\varphi_{t_1}\rangle \quad (2)$$

Because the Hamiltonian H is a Hermitian operator, the operator U is unitary. Thus, in the unitary operator picture, we use this unitary transformation to formulate the quantum dynamics.

$$U = e^{-\frac{i}{\hbar}H \cdot t} \quad (3)$$

In the following, our discussion is mainly under the unitary operator picture. As shown in Eq. (2), the manipulation of a quantum system is equivalent to manipulating corresponding unitary matrices. Once there is an efficient way to implement the desired unitary operator, we can effectively imitate the behavior, that is the dynamics, of a quantum mechanical system, which is an intractable task for classical methods [20]. This learning unitary problem introduced above is related to Hamiltonian simulation, which suppose given a Hamiltonian H , time t and precision $\varepsilon > 0$, the task is to find a quantum algorithm (quantum circuit) to produce operator U which is approximate to the target unitary operator $U_T = e^{-iHt}$, that is,

$$|U_T - \tilde{U}| \leq \varepsilon \quad (4)$$

Plenty of methods were proposed in past years. Lloyd [20] first invented the quantum algorithm by using Lie-Trotter product formula for simulating Hamiltonian which only includes the local interactions. Berry et al. [21] based on Trotter–Suzuki algorithm gave a method for simulating sparse Hamiltonian. As the unitary operator can be rewritten by the Taylor series expansion, Berry et al. [22] gave an algorithm for approximating the unitary operator with the truncated Taylor series in the same year. Low et al. [23] proposed a quantum signal processing for simulating. Being different from these approaches by quantum algorithm, some works intended to build a unitary operator by neural network or parametrized quantum gate sequence to approximate the target [19,24]. They formulize the approximation of the target evolution operator as an optimization problem. The loss function is the square of Frobenius norm between desired unitary operator and parametrized alternating operator sequences [19].

3 Quantum Machine Learning

In this section, we will introduce the hybrid-based quantum machine learning model. The hybrid quantum-classical framework is derived from the quantum approximate optimization algorithm (QAOA) [25]. The QAOA was proposed for solving the MAX-CUT problem, which can be viewed as minimizing the expectation value of a specific observable A that encodes the subjective function in the parametrized quantum state generated by applying alternating operators $e^{-i\beta_i \cdot B}$ and $e^{-i\gamma_i \cdot A}$ on a uniform quantum state as follow,

$$\min_{\beta, \gamma} \langle A \rangle = \text{tr}(A \cdot \rho_{\beta, \gamma}) = \langle \beta, \gamma | A | \beta, \gamma \rangle \quad (5)$$

where we are able to program the $|\beta, \gamma\rangle$ by following way

$$|\beta, \gamma\rangle = \left(e^{-i\beta_n \cdot B} e^{-i\gamma_n \cdot A} \right) \dots \left(e^{-i\beta_1 \cdot B} e^{-i\gamma_1 \cdot A} \right) |x\rangle \quad (6)$$

By calculating the derivative, in terms of parameters β and γ , of the loss function Eq. (5), the gradient-based optimizer can iteratively update these parameters and eventually approximate a solution. The fundamental of QAOA is the hybrid quantum-classical framework which consists of classical optimizer and variational quantum circuit (VQC).

The framework of hybrid-based model is shown in Fig. 1, where VQC is responsible for calculating the gradient of the loss function L respect to parameters θ , then classical optimizer updates the parameters according to the gradient information and feed back to VQC. Similar to this idea discussed above, plenty of hybrid-based quantum machine learning models are successively developed for tackling different problems such as quantum generative model [26], quantum generative adversarial network [8,9], and so on.

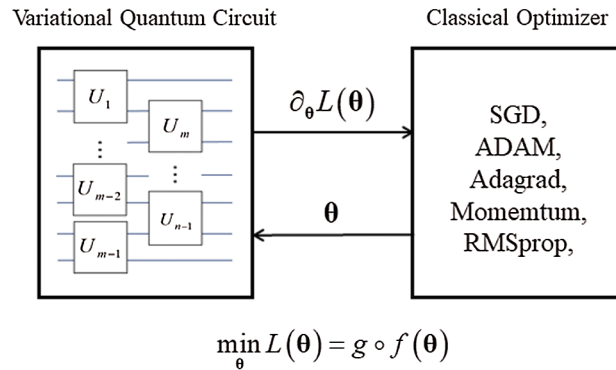


Figure 1: The framework of hybrid-based model

The works listed above form the problem as a non-convex optimization problem, such as minimizing or maximizing a composite function $g \circ f(\theta)$, where $f(\theta)$ is the expectation value. The choice of function g can refer to the loss function in classical learning like mean square error (MSE), cross entropy loss etc.

$$f(\theta) = \langle O \rangle = \text{tr}(O \cdot \rho_{\theta}) \quad (7)$$

where the observable O is the Hermitian operator of interest. In general, the density matrix ρ_θ can be described as an ensemble of pure states,

$$\rho_\theta = \sum_{j=1}^n p_j |\varphi_j\rangle\langle\varphi_j| \tag{8}$$

where $\sum_j p_j = 1$. In the most of literatures, the pure state $|\varphi_j\rangle$ is produced by applying a variational quantum circuit $U(\theta_j)$ on initial quantum state $|\varphi_0\rangle$, that is

$$|\varphi_j\rangle = U(\theta_j) |\varphi_0\rangle \tag{9}$$

By virtue of the success of QAOA, the non-convex optimization of the loss function $g \circ f(\theta)$ can also be optimized by gradient descent or heuristic algorithms.

Variational quantum circuit (VQC), known as the parametrized quantum circuit, plays the core role in the hybrid quantum-classical framework. It can be viewed as the quantum black-box model that is capable of approximating any given unitary operator with a small error like neural networks in classical machine learning [13]. Unlike neural networks using weight matrices (parameters) between layers that are adjusted by back-propagation methods, VQC achieves a similar work in the quantum setting by utilizing a sequence of fixed and parametrized quantum gates to construct a circuit model with a specific structure, that is

$$U(\theta) = U_{\theta_N} U_{\theta_{N-1}} \cdots U_{\theta_2} U_{\theta_1} \tag{10}$$

where U_{θ_j} can be regarded as a quantum gate acting on specific qubits. They are selected from a gate set that consists of parameter-free quantum gates including Pauli matrices ($\sigma_X, \sigma_Y, \sigma_Z$), controlled NOT gate (CNOT) and parametrized gates such as rotation gates ($R_X(\theta), R_Y(\theta), R_Z(\theta)$). Besides, U_{θ_j} can also represent a quantum circuit block, a similar role of the layer in neural network, that is built by quantum gates with the specific structure. The widely used ways to construct variational quantum circuits by U_{θ_j} are shown in Fig. 2.

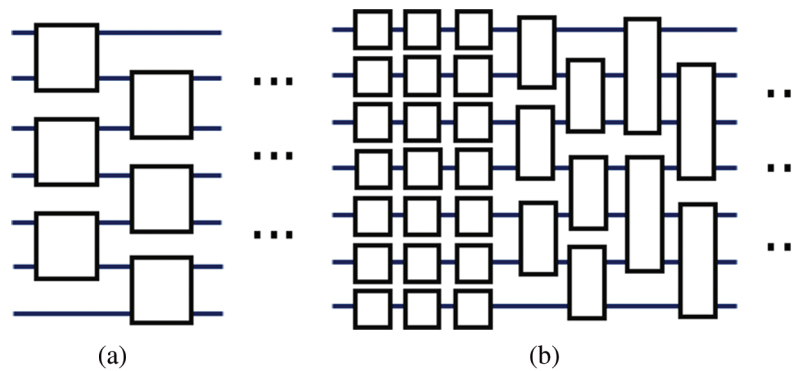


Figure 2: The structure of quantum circuit block U_{θ_j} . (a) Universal two-qubits gate, (b) one qubit gates + entanglement gates

In subgraph (a) of Fig. 2, each square represents the universal two-qubits gate which can be implemented by 15 rotation gates and 3 CNOT gates [27]. All these universal two-qubits gates are applied on nearest-neighbor qubits. If the size of the quantum system is n qubits and number of

the layer is l , this type of structure requires $15(n-1)l$ quantum gates to create. In subgraph (b) of Fig. 2, the gates of U_{θ_j} is divided into two parts, the layer of one qubit gates and the layer of entangling gates. The squares represent the one rotation gates with angle and the rectangles stand for the entangling gates such as CNOT gates or Ising (XX, YY, ZZ) coupling gates. The details of these quantum gates are listed in Appendix A.

4 Learning Unitary by QML

Different from the methods introduced to solve the problem of learning a unitary transformation in Section 2, we reformulate the problem from another perspective. According to Choi–Jamiołkowski isomorphism in quantum information theory [11], there is the correspondence between a quantum mapping \mathcal{E} and quantum states $\rho_{\mathcal{E}}$ such that,

$$\rho_{\mathcal{E}} = \mathcal{E} \otimes I(\rho_{\Phi}) = \frac{1}{d} \sum_{i,j=0}^{d-1} \mathcal{E}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \quad (11)$$

where $\rho_{\mathcal{E}}$ is known as Choi matrix which is identical to the quantum map \mathcal{E} and ρ_{Φ} is the maximally entangled state,

$$\rho_{\Phi} = \frac{1}{d} \sum_{i,j=0}^{d-1} |i, i\rangle\langle j, j| \quad (12)$$

Since the unitary transformation U_T can be regarded as a mapping \mathcal{E} from one density matrix to another one. Thus the learning unitary problem can be formed as the follow optimization problem, that is minimizing the distance between quantum state ρ_{U_T} and $\rho_{U_{\theta}}$.

$$\min_{\theta} L(\rho_{U_T}, \rho_{U_{\theta}}) \quad (13)$$

where the quantum state $\rho_{U_{\theta}}$ is

$$\begin{aligned} \rho_{U_{\theta}} &= U_{\theta} \otimes I(\rho_{\Phi}) \\ &= \frac{1}{d} \sum_{i,j=0}^{d-1} U_{\theta}(|i\rangle\langle j|) U_{\theta}^{\dagger} \otimes |i\rangle\langle j| \end{aligned} \quad (14)$$

The learning problem of Eq. (13) stands for finding the unitary operator which minimizes the average output fidelity when feed the maximally mixed state as the input to U_{θ} and U_T . Thus, it is necessary to add a regularization term to the problem Eq. (13), which also guarantees the distance between U_{θ} and U_T are still small enough. Therefore, the learning unitary problem is shown as follow,

$$\min_{\theta} L(\rho_{U_T}, \rho_{U_{\theta}}) + \lambda \|U_T - U_{\theta}\|_p \quad (15)$$

There are many choices for $L(\cdot, \cdot)$ and $\|\cdot\|_p$ in problem Eq. (15). Since the bad selection of $L(\cdot, \cdot)$ and $\|\cdot\|_p$ will affect the performance of the model, the appropriate choice of $L(\cdot, \cdot)$ and $\|\cdot\|_p$ is required. We will introduce the measure of distance between quantum states and between operators.

Given two quantum state ρ_A and ρ_B , the commonly used measure of distance for how ‘close’ between the quantum states are shown as follows,

- Trace distance:

$$\begin{aligned} L_T(\rho_A, \rho_B) &= \frac{1}{2} \text{tr} \left(\sqrt{(\rho_A - \rho_B)^\dagger (\rho_A - \rho_B)} \right) \\ &= \sup_{0 \leq P \leq I} \text{tr}(P(\rho_A - \rho_B)) \end{aligned} \quad (16)$$

It can be viewed as a generalization of the total variation distance on probability distribution that satisfies the symmetry and triangle inequality. It is also related to the maximum probability of distinguish different quantum state. The variational form of trace distance is shown in the second line of Eq. (16) [28].

- Fidelity:

$$L_F(\rho_A, \rho_B) = \text{tr} \left(\sqrt{\rho_A^{1/2} \rho_B \rho_A^{1/2}} \right) \quad (17)$$

Although the fidelity is not a metric, it has many good properties, such as it is invariant unitary transformation, the value of fidelity lies within 0 and 1. It also can be interpreted as the angle between states on a unit sphere. If ρ_A, ρ_B are pure state, the fidelity can reduced to the overlap (inner product) between states.

- Quantum Optimal Transport Distance:

$$\begin{aligned} L_{OT}(\rho_A, \rho_B) &= \min_{\pi} \text{tr}(\pi C) \\ \text{s.t. } \text{tr}_{\mathcal{X}}(\pi) &= \rho_A, \text{tr}_{\mathcal{Y}}(\pi) = \rho_B \end{aligned} \quad (18)$$

It is the quantum extension of classical optimal transport distance [10]. However, it is semi-metric which does not satisfy the triangle inequality of metric on quantum states. It inherits the nice property of classical optimal transport distance including smoothness and continues.

Because the quantum optimal transport distance can naturally be implemented by hybrid quantum-classical framework on quantum device, we select quantum optimal transport distance as $L(\cdot, \cdot)$.

Given a unitary operator U , the norm of operator U can be represented in a general form, namely Schatten p -norm, $\|U\|_p = \left[\text{tr}(U^\dagger U)^{p/2} \right]^{1/p}$.

- Trace norm ($p = 1$):

$$\|U\|_{\text{tr}} = \|U\|_1 = \text{tr} \left[\left(U^\dagger U \right)^{\frac{1}{2}} \right] \quad (19)$$

- The Frobenius norm ($p = 2$):

$$\|U\|_F = \|U\|_2 = \left[\text{tr} \left(U^\dagger U \right) \right]^{\frac{1}{2}} \quad (20)$$

- The Spectral norm ($p = \infty$):

$$\|U\|_{sp} = \|U\|_\infty = \max \{ \|Ux\| : x \in \mathcal{X}, \|x\| = 1 \} \quad (21)$$

where $\|\cdot\|$ is the Euclidean norm.

There are many nice properties of Schatten p -norm, such as the inequality between the different norms. For any operator X , if $1 \leq p \leq q \leq \infty$, then $\|X\|_q \leq \|X\|_p$ [28]. Therefore, we have the following inequality between trace norm, Frobenius norm and spectral norm of operator X ,

$$\|X\|_{sp} \leq \|X\|_F \leq \|X\|_{tr} \quad (22)$$

The related work employed the square of Frobenius norm as the loss function for learning unitary transformation [19]. In this work, we utilize the typically stronger norm, spectral norm, to estimate the error between approximate and desired unitary operator. Consequently, we can rewrite Eq. (15), the loss function of learning a unitary operator, in the following concrete way,

$$\min_{\theta, \pi} \max_{\mathbf{u}} \text{tr}(\pi C) + \lambda \|U_T - U_\theta\|_{sp} \quad (23)$$

$$s.t. \quad \text{tr}_{\mathcal{X}}(\pi) = \rho_{U_T}, \quad \text{tr}_{\mathcal{Y}}(\pi) = \rho_{U_\theta}$$

Since the min-max problem of Eq. (23) can be considered as quantum GAN with regularization term. As directly solving the problem of Eq. (23) is intractable, we intend to tackle the dual form of this problem by the Lagrange multiplier technique.

$$\min_{\theta} \max_{\mathbf{u}, \phi, \varphi} \text{tr}(\varphi \cdot \rho_{U_T}) - \text{tr}(\phi \cdot \rho_{U_\theta}) + \lambda \|U_T - U_\theta\|_{sp} \quad (24)$$

$$s.t. \quad \mathbf{I} \otimes \varphi - \phi \otimes \mathbf{I} \preceq C$$

As the spectral norm of matrix A is to find the unit vector \mathbf{x} which maximize the Euclidean norm of $A\mathbf{x}$, we use a full-connected neural network with soft-max layer to construct unit vector \mathbf{x} , i.e., $\mathbf{x} = f_{\text{softmax}}(\mathbf{W} \cdot \mathbf{x}_0 + \mathbf{b})$. Moreover, we regard the variational quantum circuit U_θ as the generator and the parametrized observables φ, ϕ as the discriminator. On account of the variational quantum circuit can naturally be used for modeling the unitary transformation U_θ , we investigate the feasibility of applying the hybrid quantum classical framework to solve the optimization problem shown as Eq. (24). By applying the gradient based optimization algorithm, we iteratively adjust the parameters of quantum circuit to approximate the target unitary operator. Inspired by the training process of classical generative adversarial network, we also apply the strategy of alternately updating the parameters of generator and discriminator. The training algorithm is shown in Algorithm 1.

Algorithm 1: The algorithm for solving the optimization problem

Input: learning rate η , maximum iteration $iter_{\max}$, target unitary operator U_T , input quantum state $\rho_\Phi = d^{-1} \sum_{i,j=0}^{d-1} |i, i\rangle \langle j, j|$, ep .

Output: the trained quantum circuit U_θ .

```

1 Initialize the parameters of variational quantum circuit  $U_\theta$ , and the parameters  $\alpha, \beta, W, t=0$ .
2 While  $t < iter_{\max}$  do
3   for  $k=0, \dots, ep$  do
4     Update the parameters of  $U_\theta$ 
5      $\theta^{t+1} \leftarrow \theta^t - \eta_\theta \cdot \nabla_\theta L(\theta, \alpha, \beta, W)$ 
6     If  $t \bmod ep = 0$  then
7       Update the parameters of  $\mathbf{x}, \phi, \varphi$ 
8        $\alpha^{t+1} \leftarrow \alpha^t + \eta_\alpha \cdot \nabla_\alpha L(\theta, \alpha, \beta, W)$ 
9        $\beta^{t+1} \leftarrow \beta^t + \eta_\beta \cdot \nabla_\beta L(\theta, \alpha, \beta, W)$ 
10       $W^{t+1} \leftarrow W^t + \eta_W \cdot \nabla_W L(\theta, \alpha, \beta, W)$ 
11    end for
12   $t \leftarrow t + 1$ 

```

5 Experimental Results

In this section, we provide the experiments of applying the quantum machine learning model for learning the unitary transformation. We apply the quantum machine learning model discussed above to the task of learning the unitary transformation of one dimensional Heisenberg spin model which is also considered in reference [29]. The target unitary operator U_T on N -qubit system is given by,

$$U_T = e^{-i \cdot H \cdot t} \quad (25)$$

In which the Hamiltonian H is described as,

$$H = \sum_{j=1}^N \vec{\sigma}_j \cdot \vec{\sigma}_{j+1} + h_j \sigma_j^z, \vec{\sigma}_j = (\sigma_j^x, \sigma_j^y, \sigma_j^z) \quad (26)$$

where $\vec{\sigma}_j$ stands for a vector of Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ on j th qubit, and the external magnetic field h_j in the z direction is randomly chosen from $-h$ to h . We investigate the following cases that the system size $N=2, 3$ and the evolution time $t=N$, and $h=1$.

In Figs. 3 and 4, we provide the changes of trace norm, Frobenius norm and spectral norm between target and generated unitary operator during the training process. The solid black line represents the average norm for 10 runs with random initialization of parametrized quantum circuit, and the shaded area refers to the range of variation of the norm. In all case, the norm will approximate to zero which means the generated unitary operator is close to the target. In addition, the small system will lead to faster convergence, which is same as we expected.

We select fidelity between the generated and target quantum states to evaluate the performance of our model. In Fig. 5, we show the fidelity between the quantum states which are generated by applying U_θ and U_T on the same random input quantum states. As shown in Fig. 5, the fidelity is eventually around 0.997 which indicates that the generated quantum circuit provides a practicable approximation.

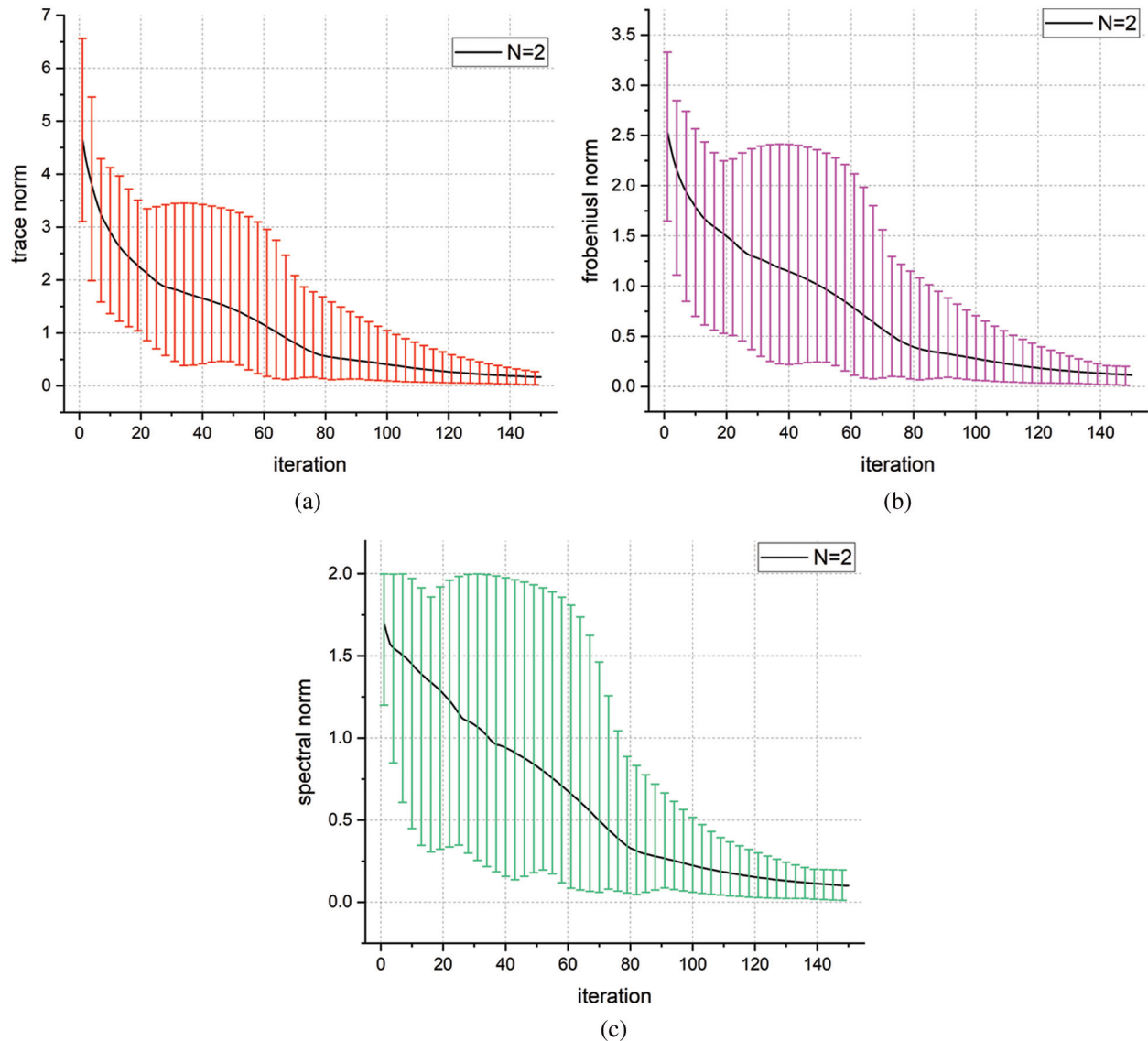


Figure 3: The trace norm, Frobenius norm and spectral norm vs. iterations ($N = 2$). (a) Trace norm ($N = 2$), (b) Frobenius norm ($N = 2$), (c) spectral norm ($N = 2$)

6 Discussion

Learning unitary transformation is an important and vexing task in quantum computing, which is related to controlling a quantum system or implementing a quantum algorithm with fewer resources. In this work, we investigate the use of promising techniques from quantum machine learning for learning a unitary transformation of a quantum system. Instead of the related works which formulate the learning problem as minimizing the norms between target and generated unitary operator, we express the problem as a quantum generative adversarial network with regularization term from the other perspective based on Choi–Jamiołkowski isomorphism. Comparing the trace norm and Frobenius norm used in related works, we add an intuitively

stronger norm, spectral norm, as the regularization term to the loss function. Our numerical experiments demonstrate that the operator generated by our proposed model can successfully approximate the desired target unitary operator. The average spectral norm error of 10-replication runs is 0.1, and the average fidelity between the states produced by applying target and generated operators on the random input is around 0.997. Meanwhile, compared to the traditional method using product formulas for Hamiltonian simulation, our proposed model can significantly reduce the number of quantum gates for implementation. There are some potential applications of our proposed model such as providing help for assisting us in implementing quantum algorithms or compiling quantum circuits.

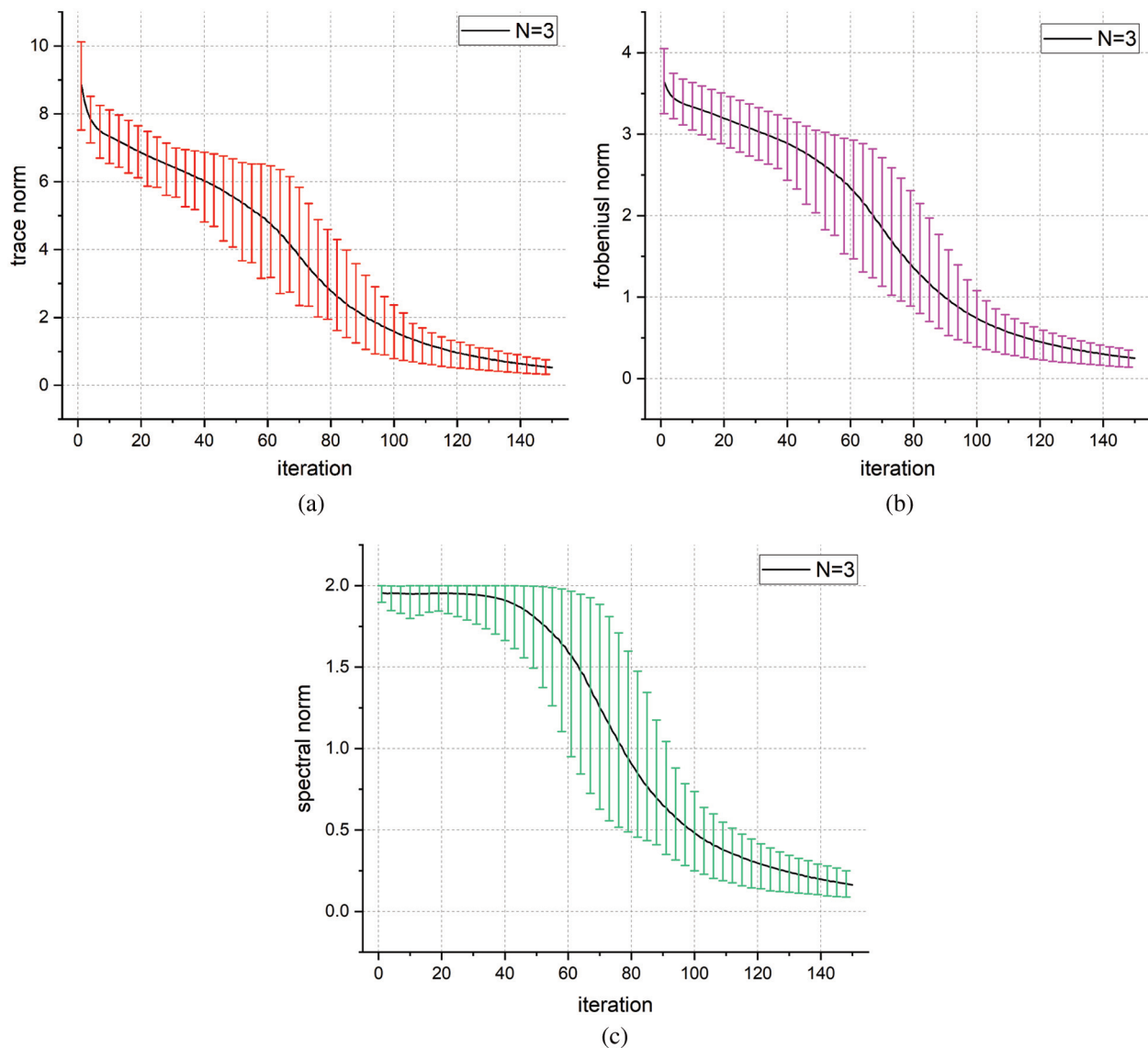


Figure 4: The trace norm, Frobenius norm and spectral norm vs. iterations ($N = 2$). (a) Trace norm ($N = 2$), (b) Frobenius norm ($N = 2$), (c) spectral norm ($N = 2$)

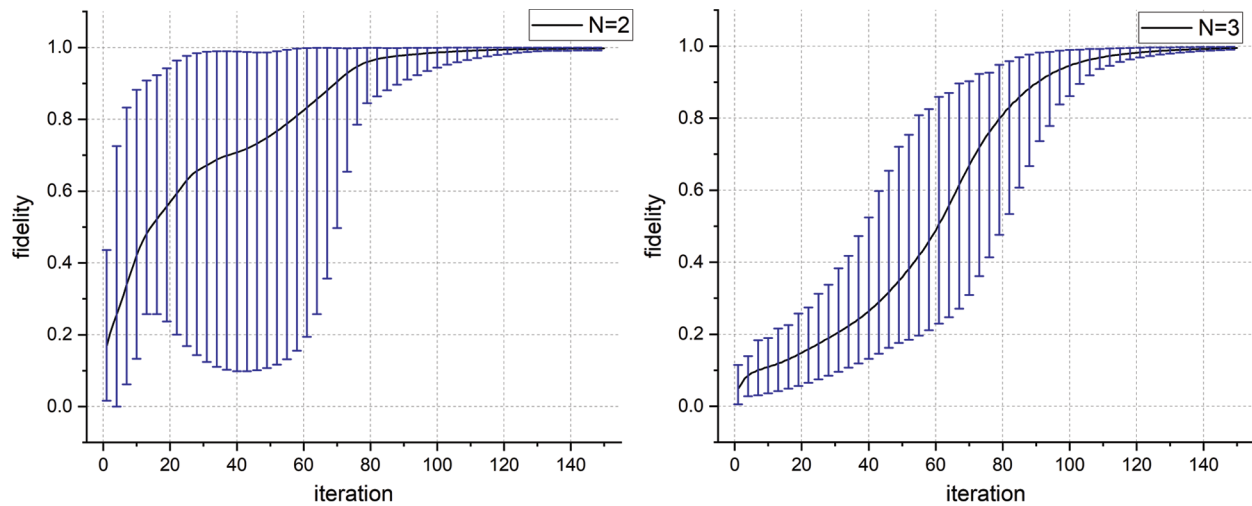


Figure 5: The fidelity of the states generated by applying U_θ on random input ($N = 2, 3$)

Acknowledgement: Thanks for constructive suggestion and helpful discussions from professor Xiaodi Wu.

Funding Statement: This work has received support from the National Key Research and Development Plan of China under Grant No. 2018YFA0306703.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] B. Indira and K. Valarmathi, "A perspective of the machine learning approach for the packet classification in the software defined network," *Intelligent Automation & Soft Computing*, vol. 26, no. 4, pp. 795–805, 2020.
- [2] A. Alhussain, H. Kurdi and L. Altoaimy, "A neural network-based trust management system for edge devices in peer-to-peer networks," *Computers, Materials & Continua*, vol. 59, no. 3, pp. 805–815, 2019.
- [3] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner *et al.*, "Identifying quantum phase transitions using artificial neural networks on experimental data," *Nature Physics*, vol. 15, no. 9, pp. 917–920, 2019.
- [4] A. Tkatchenko, "Machine learning for chemical discovery," *Nature Communications*, vol. 11, no. 1, pp. 4125, 2020.
- [5] P. Rebentrost, M. Mohseni and S. Lloyd, "Quantum support vector machine for big data classification," *Physical Review Letters*, vol. 113, no. 13, pp. 130503, 2014.
- [6] S. Lloyd, M. Mohseni and P. Rebentrost, "Quantum principal component analysis," *Nature Physics*, vol. 10, no. 9, pp. 631–633, 2014.
- [7] P. L. Dallaire-Demers and N. Killoran, "Quantum generative adversarial networks," *Physical Review A*, vol. 98, no. 1, pp. 12324, 2018.
- [8] C. Zoufal, A. Lucchi and S. Woerner, "Quantum generative adversarial networks for learning and loading random distributions," *NJP Quantum Information*, vol. 5, no. 1, pp. 1–9, 2019.
- [9] L. Hu, S. H. Wu, W. Cai, Y. Ma, X. Mu *et al.*, "Quantum generative adversarial learning in a superconducting quantum circuit," *Science Advances*, vol. 5, no. 1, pp. eaav2761, 2019.

- [10] S. Chakrabarti, H. Yiming, T. Li, S. Feizi and X. Wu, “Quantum wasserstein generative adversarial networks,” in *Proc. the Neural Information Processing Systems*, Vancouver, BC, Canada, pp. 6781–6792, 2019.
- [11] M. A. Nielsen and I. L. Chuang, “The quantum Fourier transform and its applications,” in *Quantum Computation and Quantum Information*, 10th Anniversary Edition, 2nd ed., vol. 1. New York, NY, USA: Cambridge University Press, 2010.
- [12] A. W. Harrow, A. Hassidim and S. Lloyd, “Quantum algorithm for linear systems of equations,” *Physical Review Letters*, vol. 103, no. 15, pp. 150502, 2009.
- [13] M. Benedetti, E. Lloyd, S. Sack and M. Fiorentini, “Parameterized quantum circuits as machine learning models,” *Quantum Science and Technology*, vol. 4, no. 4, pp. 43001, 2019.
- [14] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis *et al.*, “TensorFlow: A system for large-scale machine learning,” in *Proc. the Operating Systems Design and Implementation*, Savannah, GA, USA, pp. 265–283, 2016.
- [15] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury *et al.*, “PyTorch: An imperative style, high-performance deep learning library,” in *Proc. the Neural Information Processing Systems*, Vancouver, BC, Canada, pp. 8026–8037, 2019.
- [16] A. Bisio, G. Chiribella, G. M. D’Ariano, S. Facchini and P. Perinotti, “Optimal quantum learning of a unitary transformation,” *Physical Review A*, vol. 81, no. 3, pp. 32324, 2010.
- [17] S. L. Hyland and G. Rättsch, “Learning unitary operators with help from $u(n)$,” in *Proc. the National Conf. on Artificial Intelligence*, New York, NY, USA, pp. 2050–2058, 2016.
- [18] S. Lloyd and R. Maity, “Efficient implementation of unitary transformations,” *Arxiv Quantum Physics*, arXiv: 1901.03431 2019.
- [19] B. T. Kiani, S. Lloyd and R. Maity, “Learning unitaries by gradient descent,” *Arxiv Quantum Physics*, arXiv: 2001.11897, 2001.
- [20] S. Lloyd, “Universal quantum simulators,” *Science*, vol. 273, no. 5278, pp. 1073–1078, 1996.
- [21] D. W. Berry, A. M. Childs and R. Kothari, “Hamiltonian simulation with nearly optimal dependence on all parameters,” in *Proc. the Foundations of Computer Science*, Berkeley, CA, USA, pp. 792–809, 2015.
- [22] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari and R. D. Somma, “Simulating hamiltonian dynamics with a truncated Taylor series,” *Physical Review Letters*, vol. 114, no. 9, pp. 90502, 2015.
- [23] G. H. Low and I. L. Chuang, “Optimal hamiltonian simulation by quantum signal processing,” *Physical Review Letters*, vol. 118, no. 1, pp. 10501, 2017.
- [24] B. T. Kiani, “Quantum artificial intelligence: Learning unitary transformations,” M.S. Dissertation, Massachusetts Institute of Technology, USA, 2020.
- [25] E. Farhi, J. Goldstone and S. Gutmann, “A quantum approximate optimization algorithm,” *Arxiv Quantum Physics*, arXiv: 1411.4028, 2014.
- [26] J. G. Liu and L. Wang, “Differentiable learning of quantum circuit Born machines,” *Physical Review A*, vol. 98, no. 6, pp. 62324, 2018.
- [27] V. V. Shende, I. L. Markov and S. S. Bullock, “Minimal universal two-qubit controlled-NOT-based circuits,” *Physical Review A*, vol. 69, no. 6, pp. 62321, 2004.
- [28] P. J. Coles, M. Cerezo and L. Cincio, “Strong bound between trace distance and Hilbert-Schmidt distance for low-rank states,” *Physical Review A*, vol. 100, no. 2, pp. 22103, 2019.
- [29] A. M. Childs, D. Maslov, Y. S. Nam, N. J. Ross and Y. Su, “Toward the first quantum simulation with quantum speedup,” *Proc. of the National Academy of Sciences of the United States of America*, vol. 115, no. 38, pp. 9456–9461, 2018.

Appendix A. Preliminary of Quantum Computing

A.1 Quantum State

Instead of storing a certain state in one classical bit, a qubit, the counterpart of classical bit, is in a superposition of two basic state $|0\rangle$ and $|1\rangle$. Let $|\varphi\rangle$ represents an arbitrary one-qubit state, which can be described by the $|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle$ or $|\varphi\rangle = [\alpha \ \beta]^T$, where α and β are both complex number such that $|\alpha|^2 + |\beta|^2 = 1$. They are known as probability amplitude which means if we perform a measurement on state $|\varphi\rangle$, then $|\varphi\rangle$ will collapse into state $|0\rangle$ or state $|1\rangle$ with the probability $|\alpha|^2$ or $|\beta|^2$, respectively. Besides, the quantum state $|0\rangle$ and $|1\rangle$ are also orthonormal bases, which are also called the computational bases in quantum computing. For a multi-qubit system, the state of this composite system can be described by the tensor product of the state of subsystem such as $|\varphi\rangle = |\varphi_1\rangle \otimes \cdots \otimes |\varphi_m\rangle$. All the quantum states we discussed above are pure states. Besides, a mixed state is an ensemble of some pure states $\{|\varphi_i\rangle\}$ with coefficient p_i , which is generally described by a density matrix in the form of $\rho = \sum_i p_i |\varphi_i\rangle \langle \varphi_i|$, and $\sum_i p_i = 1$.

A.2 Quantum Gates and Quantum Circuit Model

The widely used model for quantum computation is the quantum circuit model that is constructed by a sequence of reversible quantum logic gates. In quantum mechanics, the time evolution of a quantum system can be described by a unitary transformation. Thus, the quantum logic gates are equivalent to the unitary transformations in the quantum circuit model. The followings are commonly used one-qubit and multiple-qubits quantum gates,

- Puali Gates,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (\text{A.1})$$

- Rotation Gates,

$$R_x(\theta) \equiv \begin{bmatrix} \cos \theta/2 & -i \sin \theta/2 \\ -i \sin \theta/2 & \cos \theta/2 \end{bmatrix}, \quad R_y(\theta) \equiv \begin{bmatrix} \cos \theta/2 & -i \sin \theta/2 \\ \sin \theta/2 & \cos \theta/2 \end{bmatrix},$$

$$R_z(\theta) \equiv \begin{bmatrix} \exp(-i\theta/2) & 0 \\ 0 & \exp(i\theta/2) \end{bmatrix} \quad (\text{A.2})$$

- Entangling Gates,

Controlled-NOT gate,

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{A.3})$$

Ising XX Coupling Gates,

$$XX(\phi) = \begin{bmatrix} \cos \phi & 0 & 0 & -i \sin \phi \\ 0 & \cos \phi & -i \sin \phi & 0 \\ 0 & i \sin \phi & \cos \phi & 0 \\ -i \sin \phi & 0 & 0 & \cos \phi \end{bmatrix}$$

An example of quantum circuit is shown in Fig. A.1, each solid line through quantum gates represents one qubit, which is corresponding to the passage of time. The square box denotes the quantum gates, i.e., the revolution of the quantum system. The quantum states are passed through the quantum gates from left to right.

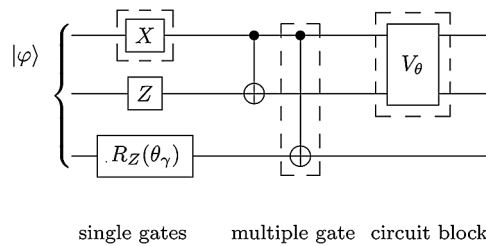


Figure A.1: An example of quantum circuit

A.3 Qubit Measurement

The classical data stored in quantum states cannot be directly read out. To extract the data, we have to perform the quantum measurement on the state. Due to the measurement involve the distraction from the exterior environment, the system will collapse after the measurement which is an irreversible action. The operation of measurement can be described by a group of measurement operator $\{M_i\}$, where $\sum_i M_i^\dagger M_i = I$. The projective measurement is simple and widely used measurement. It is described by an observable M which has the spectral decomposition, $M = \sum_m m \cdot P_m$, where $P_m = |\phi_m\rangle \langle \phi_m|$ is the projector to the corresponding eigenspace with eigenvalue m . When we perform the projective measurement on state $|\varphi\rangle$, we can get the state $|\phi_m\rangle$ with probability $\langle \phi_m | P_m | \phi_m \rangle$ and the expectation value $E(M)$ of observable M , $E(M) = \text{tr}(M\rho)$, where ρ is the density matrix of state $|\varphi\rangle$.