Reinforced Learning Explicit Circuit Representations for Quantum State Characterization from Local Measurements

Manwen Liao^{*1} Yan Zhu^{*1} Weitian Zhang² Yuxiang Yang¹

Abstract

Characterizing quantum states is essential for advancing many quantum technologies. Recently, deep neural networks have been applied to learn quantum states by generating compressed implicit representations. Despite their success in predicting properties of the states, these representations remain a black box, lacking insights into strategies for experimental reconstruction. In this work, we aim to open this black box by developing explicit representations through generating surrogate state preparation circuits for property estimation. We design a reinforcement learning agent equipped with a Transformer-based architecture and a local fidelity reward function. Relying solely on measurement data from a few neighboring qubits, our agent accurately recovers properties of target states. We also theoretically analyze the global fidelity the agent can achieve when it learns a good local approximation. Extensive experiments demonstrate the effectiveness of our framework in learning various states of up to 100 qubits, including those generated by shallow Instantaneous Quantum Polynomial circuits, evolved by Ising Hamiltonians, and many-body ground states. Furthermore, the learned circuit representations can be applied to Hamiltonian learning as a downstream task utilizing a simple linear model.

1. Introduction

Quantum state characterization is a critical task in quantum information, underpinning the development of quantum computing, quantum communication, and quantum sensing technologies. There are two main approaches to tackling this task: classical methods and quantum methods. Classical methods, such as quantum state tomography (Tóth et al., 2010; Gross et al., 2010; Cramer et al., 2010; Lanyon et al., 2017; Cotler & Wilczek, 2020), reconstructs the quantum state by measuring an informationally complete set of observables. These methods require exponentially increasing sample complexity in measurements as the size of the quantum system grows, making them impractical for systems with many qubits and thereby limiting their applicability for practical use. Quantum methods, represented by variational quantum algorithms (Cerezo et al., 2021), utilize the power of quantum circuits to learn quantum states. These methods (Peruzzo et al., 2014; Farhi et al., 2014; Du et al., 2022; Wu et al., 2023a) typically optimize a parameterized quantum circuit to approach the target quantum state. Nevertheless, due to the necessity of calculating gradients with respect to circuit parameters, where the loss landscape is highly flat, these methods often struggle with issues such as barren plateaus (McClean et al., 2018; Cerezo et al., 2021) and local minima (Anschuetz & Kiani, 2022; Huang et al., 2024), consequently affecting their performance in learning large-scale quantum systems.

In practice, the primary interest often lies in specific properties of quantum states rather than their full state representations. As a result, full quantum state tomography is not always necessary, especially when the goal is to predict certain physical observables. Motivated by this, recent approaches integrate machine learning techniques to construct "shadow" representations as a surrogate to the original states, enabling efficient quantum system characterization. These methods have shown success in quantum state learning (Carleo & Troyer, 2017; Sharir et al., 2020; Zhu et al., 2022; Zhang & Di Ventra, 2023; Tang et al., 2024a; Chen & Heyl, 2024; Du et al., 2023; Qian et al., 2024), quantum process learning (Huang et al., 2023; Torlai et al., 2023; Zhu et al., 2023), quantum property estimation (Zhang & Di Ventra, 2023; Wu et al., 2024; Lewis et al., 2024; Tang et al., 2024a), quantum state classification (Tang et al., 2024b), quantum sensing (Xiao et al., 2022; Zhou et al., 2024) and quantum verification (Wu et al., 2023b; Qian et al., 2024). By leveraging neural networks to efficiently represent quantum states, these methods signifi-

^{*}Equal contribution ¹QICI Quantum Information and Computation Initiative, Department of Computer Science, School of Computing and Data Science, The University of Hong Kong, Pokfulam Road, Hong Kong, China ²MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Shanghai, China. Correspondence to: Yuxiang Yang <yuxiang@cs.hku.hk>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

Table 1. Summary of quantum state characterization methods. #Observables: The number of observables utilized for characterizing the target quantum states. Experimental Reconstructability: The ability to construct a quantum circuit to prepare the surrogate state from its measurement data. Downstream Applicability: The capability to perform downstream tasks, such as Hamiltonian learning, based on the classical representation. Scalability: The ability to extend the learning scheme to large-scale quantum systems (e.g., N > 20 qubits).

	Methods	#Observables	Experimental Reconstructability	Downstream Applicability	Mitigate Barren Plateaus	Scalability
Ouentum	Peruzzo et al. (2014)	-	1	×	×	×
Quantum	Farhi et al. (2014)	-	✓	×	×	×
	Tóth et al. (2010); Cotler & Wilczek (2020)	$-\overline{\mathcal{O}}(2^{\overline{N}})$	×	×		×
Classical	Carleo & Troyer (2017); Chen & Heyl (2024)	$\mathcal{O}(2^N)$	×	1	-	×
Classical	Zhu et al. (2022)	$\mathcal{O}(N)$	×	✓	-	1
	Ours	$\mathcal{O}(N)$	<i>✓</i>	1	1	1

cantly reduce the measurement overhead. Approaches such as generative neural networks (GQNQ) (Zhu et al., 2022) and LLM4QPE (Tang et al., 2024a) approximate quantum states using fewer measurements by exploiting underlying patterns and correlations shared within a family of states. By learning compact and expressive representations, these techniques offer scalable solutions for quantum state characterization, making them particularly valuable for learning large, complex quantum systems where traditional methods become impractical. However, these representations are often implicit. While they capture essential features and properties of the quantum state, they do not support direct reconstruction of the state from the representation itself. This limitation poses challenges in scenarios where an explicit reconstruction of the quantum state is required, e.g., quantum phase estimation (Kitaev, 1995) and quantum simulation (Georgescu et al., 2014).

In this work, we propose a novel type of representations, the explicit circuit representations, formulated by a sequence of tokens, to characterize quantum states and design a deep reinforcement learning-based framework named QCrep to learn such representations that can experimentally reconstruct a surrogate for the target states for property estimation. A comparison of different methods for state characterization is shown in Table 1. Two main challenges for learning the circuit representation are the high measurement overhead and the barren plateaus problem. The high measurement overhead roots from the fact that exponential number of measurements is required to fully characterize an unknown quantum state. However, this can be surpassed when targeting at learning an approximation for accurate property estimation. We only use local measurements on a few neighboring sites of the quantum states to construct a local state representation for the target state. Additionally, we propose a novel Transformer-based (Vaswani et al., 2017) measurement feature aggregation block to recover global features of target states from local measurement data. To mitigate the problem of barren plateaus and local minima, we involve deep reinforcement learning that does not require computing gradients with respect to the circuit parameters. Besides, we design a novel reward function based on local fidelity, and provide a theoretical analysis on the effectiveness of reconstructing global properties given local information of the states. The contributions are:

(1) We develop a novel type of representations for quantum states, termed the explicit circuit representations. Unlike conventional implicit state representations in GQNQ (Zhu et al., 2022) and Neural Quantum State (NQS) (Carleo & Troyer, 2017; Sharir et al., 2020; Zhang & Di Ventra, 2023; Chen & Heyl, 2024), our circuit representations can be directly utilized to experimentally reconstruct the target states locally, which allows for computing the properties of interest via measuring the output states. Moreover, they possess the advantage of implicit representations that can be applied to downstream tasks.

(2) We design a reinforcement learning-based framework named QCrep to learn explicit circuit representations for specific families of quantum states using only measurement data from a small number of neighboring sites. The circuits learned by QCrep can reproduce the target states with high global fidelity, utilizing $\mathcal{O}(N)$ observables with respect to the system size N. Benefiting from our novel design, QCrep eliminates the need for gradient-based optimization of circuit parameters and mitigates the barren plateaus problem, enabling scalability to larger systems of up to 100 qubits.

(3) We experimentally demonstrate the effectiveness of our framework by learning four different families of target states and applying it to Hamiltonian learning (Wiebe et al., 2014; Wang et al., 2017) as a downstream task. Our framework achieves superior performance in learning states generated by Instantaneous Quantum Polynomial (IQP) circuits (Bremner et al., 2010), states evolved by Ising Hamiltonians, and ground states of many-body quantum systems. For the downstream task, numerical experiments reveal that the unknown parameters of Hamiltonians can be accurately learned from local measurement data of their corresponding ground states, leveraging only a linear model applied to the circuit representations. This further highlights the versatility and effectiveness of our framework.

2. Learning Explicit Circuit Representations

2.1. Task Definition

We define the task of learning explicit circuit representations for quantum states as characterizing a family of unknown quantum states $S = \{\rho_s\}_s$ by constructing quantum circuits $\mathcal{U} = \{U_s\}_s$ that can prepare these states with high local fidelity, so that the reconstructed states can be directly measured to predict quantum properties of interests. We assume that the states can only be accessed in a black-box manner, meaning one can measure the states using measurement operators \mathcal{M} but remains agnostic to the underlying circuits used for their preparation. Additionally, we assume that the measurement operators can only act on neighboring sites of the quantum states, a setup we refer to as local measurements. This measurement configuration has been widely adopted in prior works on quantum state characterization (Lanyon et al., 2017; Friis et al., 2018; Zhu et al., 2022; Kurmapu et al., 2023; Guo & Yang, 2024; Wu et al., 2024) due to its feasibility for experimental realization. For this learning task, we do not put explicit constraints on the global fidelity between the reconstructed states and the target states, but focus on maximizing the local fidelity.

Explicit Circuit Representations. Let ρ_s be an *N*-qubit quantum state that we want to characterize. The reconstructed state $\sigma_s = U_s |0\rangle \langle 0|^{\otimes N} U_s^{\dagger}$ possesses high average local fidelity with the target state ρ_s , thus can serve as a surrogate for predicting quantum properties of interest. U_s can be expressed as a product of unitary gates, i.e., $U_s = \prod_t U_{s,t}(\phi_{s,t})$, where $U_{s,t}$ represents the quantum gates applied at time step t, and $\phi_{s,t}$ denotes the corresponding parameter(s) for those gates. The explicit circuit representation of ρ_s is a sequence of $(u_{s,t}, \phi_{s,t})_t$, where $u_{s,t}$ is the classical description of the gate type of $U_{s,t}$. The correlations between the gate types and actual gates are encoded using a look-up table.

Overview. To learn the circuit representations, we first perform local measurements on the target states, which is introduced in Section 2.2. After that, we design a reinforcement learning-based framework, QCrep, to decode the measurement data into quantum circuits, and keep the classical descriptions of the circuits as the circuit representations. This is described in Section 2.3, wherein a measurement feature aggregation block is proposed to process the local measurements, and a local fidelity reward function is designed to ensure learnability. Background information on quantum computation is introduced in Appendix B.

2.2. Measurement Setup

We consider a set of measurements $\mathcal{M} = \{M_i\}_{i=0}^{N-2}$, termed local measurements, performed on neighboring sites of the unknown N-qubit quantum states ρ_s . Each measure-

ment $M_i = (M_{ij})_{j=1}^K$ is a positive operator-valued measure (POVM) acting on two neighboring qubits (i, i + 1) of ρ_s , satisfying $\sum_{j=1}^K M_{ij} = I$. Specifically, we select the measurement operators M_{ij} as the tensor product of two single-qubit Pauli operators, i.e., $M_{ij} \in \{X, Y, Z\}^{\otimes 2}$. We measure each pair of neighboring qubits using all such operators in a fixed order, taking the expectation values of the measurements to obtain the measurement output $m_i \in \mathbb{R}^K$, where K = 9. We repeat this process for all qubit pairs and record the measurement data as $m \in \mathbb{R}^{(N-1) \times K}$.

Importantly, the measurement operators are discarded when we input the measurement data into the agent. The correspondence between the operators and their expectation values is expected to be reconstructed during training. Note that although quantum states and measurement operators are represented by complex-valued numbers, the measurement expectation values are real and range from -1 to 1, since the eigenvalues of Pauli operators are either -1 or 1. This property, along with the removal of measurement operators from the neural network's input, exempts the framework from the overhead of processing complex values.

2.3. QCrep Framework

To construct circuit representations for reproducing a family of quantum states, we design a reinforcement learningbased framework, OCrep. This framework relies exclusively on local measurements and avoids performing gradient descent on circuit parameters, effectively mitigating the barren plateau problem. The overall pipeline is shown in Figure 1. A deep reinforcement learning agent utilizing a neural network policy is employed to construct the circuit representations for a family of unknown quantum states S. The environment in which the agent interacts and learns is defined as the quantum system. This environment is initialized with the quantum state to be learned, $\rho_s^{(0)} = \rho_s \in \mathcal{S}$, and is responsible for applying gates to the state as the agent iteratively learns to reconstruct the state. The observations are the local measurement values m_s obtained in Section 2.2. We define the actions that the agent can take at step t as applying a layer of quantum gates to the state. The reward function is the local fidelity reward defined in Equation 3.

Instead of directly learning U_s , the agent is trained to construct $V_s = \prod_{t=1}^T V_{s,t}(\phi_{s,t}) = U_s^{\dagger}$, which evolves ρ_s towards $|0\rangle\langle 0|^{\otimes N}$, where $V_{s,t}$ represents a layer of quantum gates chosen at step t, and $\phi_{s,t}$ is the corresponding gate parameter. This approach enables the learning of a family of quantum states, as directly learning U_s requires a fixed input state $|0\rangle\langle 0|^{\otimes N}$, which limits it to learning a single state. In contrast, by evolving towards $|0\rangle\langle 0|^{\otimes N}$, any state can be set as the input, facilitating the learning of a family of states. The target U_s can then be obtained via taking the inverse of V_s , i.e., $U_s = V_s^{\dagger} = \prod_{t=T}^1 V_{s,t}^{\dagger}(\phi_{s,t})$. Note that during



Figure 1. **QCrep framework.** Given an initial state $\rho_s^{(0)}$ sampled from an unknown quantum states family S, the agent iteratively applies quantum gates $V_t(\phi_t)$ to evolve the state towards $|0\rangle\langle 0|^{\otimes N}$. The policy is parameterized by a neural network, which includes an Attention-based measurement feature aggregation block followed by a Multilayer Perceptron (MLP). The agent is trained using the PPO algorithm with a local fidelity reward.

the circuit construction procedure, only the local fidelity between an arbitrary state and $|0\rangle^{\otimes N}$ needs to be evaluated, which can be obtained via Pauli Z measurements. This empowers the practical implementation of our framework.

The entire process of learning the circuit representation for ρ_s consists of several iterative steps. At each step t, the state $\rho_s^{(t)}$ is measured using local measurement operators and the agent takes the expectation values $\boldsymbol{m}_{s}^{(t)}$ as observations from the environment. Then, the agent selects the action $V_{s,t}(\phi_{s,t})$ according to its policy π_{α} , which is parameterized by a trainable Gaussian distribution generated from a neural network composed of a feature aggregation block followed by a Multilayer Perceptron (MLP). The action $V_{s,t}(\phi_{s,t}) = \bigotimes_k V_{s,t,i}(\phi_{s,t,i})$ is a column of single-qubit or two-qubit gates acting in parallel to every qubit i, where $V_{s,t,i}(\phi_{s,t,i}) = \exp\left(-i\phi_{s,t,i}G\right)$ are generated from the linear combination of the single-qubit and two-qubit Pauli operators, namely, $G \in \text{span}(\{X, Y, Z\} \cup \{X, Y, Z\}^{\otimes 2})$. To further reduce the search space, we adopt a task-aware fashion to select a subset of gates as the action space, which will be detailed in Section 3. After that, the environment updates the quantum state as $\rho_s^{(t+1)} = V_{s,t}(\phi_{s,t})\rho_s^{(t)}V_{s,t}^{\dagger}(\phi_{s,t})$ and the agent receives a reward $r^{(t)}$ defined in Equation 6. We repeat the above procedure until the average local fidelity $L(\rho_s^{(t)}, |0\rangle \langle 0|^{\otimes N})$, defined in Equation 3, exceeds a threshold of $1 - \epsilon$, or until the number of iterative steps reaches a predefined maximum of T. We set $\epsilon = 0.001$

in the experiments. Note that this T can be flexibly adjusted to control the accuracy of the reconstructed states or to meet hardware requirements when implemented on real quantum computers. The measurement complexity scales linearly with T, because for each t, only constant number of measurements is performed if the system size is fixed. The policy π_{α} is updated using Proximal Policy Optimization (PPO) algorithm (Schulman et al., 2017),

$$\boldsymbol{\alpha}_{k+1} = \arg \max_{\boldsymbol{\alpha}} \mathbb{E}_{(\boldsymbol{m}, V(\phi)) \sim \pi_{\boldsymbol{\alpha}_{k}}} [J(\boldsymbol{\alpha}, \boldsymbol{m}, V(\phi), \boldsymbol{\alpha}_{k})],$$
(1)

and

$$J(\boldsymbol{\alpha}, \boldsymbol{m}, V(\phi), \boldsymbol{\alpha}_{\boldsymbol{k}}) = \min\left(\frac{\pi_{\boldsymbol{\alpha}}(V(\phi)|\boldsymbol{m})}{\pi_{\boldsymbol{\alpha}_{k}}(V(\phi)|\boldsymbol{m})}A^{\pi_{\boldsymbol{\alpha}_{k}}}, \operatorname{clip}_{\delta}\left(\frac{\pi_{\boldsymbol{\alpha}}(V(\phi)|\boldsymbol{m})}{\pi_{\boldsymbol{\alpha}_{k}}(V(\phi)|\boldsymbol{m})}\right)A^{\pi_{\boldsymbol{\alpha}_{k}}}\right),$$

$$(2)$$

where $A^{\pi \alpha_k}$ is the estimated advantage function associated with reward r, and δ measures the gap between the new and old policies. Finally, we keep the sequence of classical descriptions of the quantum gates $(v_{s,t}^{\dagger}, \phi_{s,t})_{t=T}^1$ as the circuit representation of ρ_s .

Attention-based Measurement Feature Aggregation Block. We construct a novel feature aggregation block to map the quantum measurement data m to a compact vector representation p. There are two main features for this block: (1) A Transformer (Vaswani et al., 2017) module is proposed to capture the entanglement property of the quantum states from local measurement data. Due to the entangled nature of quantum states, non-local correlations exist among qubits, leading to long-range dependencies between measurement values. Therefore, we utilize self-attention to model the dependencies between different qubits. (2) An aggregation layer, implemented as global average pooling along the sequence axis (the second axis), is introduced to globally model the state. This enables transferability across quantum systems of varying sizes, allowing the framework to perform zero-shot transfer learning. Ablation studies demonstrating the effectiveness of this component compared to vanilla MLP are presented in Appendix I.

Local Fidelity-based Reward Function. Training based on global fidelity is prone to be trapped by barren plateaus (Mc-Clean et al., 2018; Cerezo et al., 2021; Bittel & Kliesch, 2021; Larocca et al., 2025). To mitigate this effect, we propose a novel reward function based on average *n*-local fidelity, inspired by the use of local cost functions to mitigate barren plateaus (Cerezo et al., 2021; Caro et al., 2023). Given two *N*-qubit quantum states ρ and σ , the average *n*-local ($1 \le n \le N$) fidelity is defined as

$$L^{(n)}(\rho,\sigma) = \frac{1}{N-n+1} \sum_{i=0}^{N-n} F(\rho_{i:i+n},\sigma_{i:i+n}), \quad (3)$$

where *F* is the (global) fidelity between the reduced density matrices $\rho_{i:i+n}$ and $\sigma_{i:i+n}$ of the original states on qubits $\{i, i + 1, ..., i + n - 1\}$. This reward is derived exclusively from local measurements. In our scenario, we set $\sigma_{i:i+n} = |0\rangle\langle 0|^{\otimes n}$ and the average *n*-local fidelity, denoted as $L^{(n)}(\rho_s^{(t)}, |0\rangle\langle 0|^{\otimes N})$, can be estimated by measuring $\rho_s^{(t)}$ using local operators $\{O_i^{(n)}\}_{i=0}^{N-n}$, where

$$O_i^{(n)} = |0\rangle \langle 0|_{i:i+n} \otimes I_{N\setminus i:i+n}, \tag{4}$$

which applies a projector $|0\rangle\langle 0|^{\otimes n}$ to qubits $\{i, \ldots, i+n-1\}$, and identity to the remaining qubits. The overall operator associated with the average local fidelity is defined as

$$O^{(n)} = \frac{1}{N-n+1} \sum_{i=0}^{N-n} O_i^{(n)}.$$
 (5)

We can compute average local fidelity between the state at step t and the target state as $L^{(n)}(\rho_s^{(t)}, |0\rangle \langle 0|^{\otimes N}) =$ $\operatorname{Tr}(O^{(n)}\rho_s^{(t)})$. The reward for the agent is defined as

$$r^{(t)} = -1 + L^{(n)}(\rho_s^{(t)}, |0\rangle \langle 0|^{\otimes N}).$$
(6)

An additional -1 term is added into the reward to encourage generating circuits with lower depth. In our experiments, we fix n = 1 for efficiency. Results on circuit learning with larger n is presented in Appendix H.1. To bound the accuracy of the circuit representation trained with this reward function, we present the following proposition (the proof is given in Appendix C):

Proposition 2.1. If the agent learns a policy that constructs an N-qubit quantum state with average n-local fidelity $L^{(n)}(\rho_s^{(T)}, |0\rangle\langle 0|^{\otimes N}) \geq 1 - \epsilon$, then the global fidelity satisfies $F(\rho_s^{(T)}, |0\rangle\langle 0|^{\otimes N}) \geq 1 - (N - n + 1)\epsilon$.

This presents a lower bound for the global fidelity when the agent learns a good approximation of the local fidelity. The bound is derived without assumptions on the entanglement of the target states. However, one expects that for low-entangled states, the gap between local and global fidelity becomes much higher than $1 - (N - n + 1)\epsilon$. This phenomena exists commonly in quantum many-body systems. We will show in experiments that in addition to predicting properties, the learned states can even achieve high global fidelity with the target states.

3. Experiments

In this section, we apply our framework to learn circuit representations for four different families of states – the states prepared by Instantaneous Quantum Polynomial (IQP) circuits, states evolved by Ising Hamiltonians, and two types of quantum many-body ground states. In addition, we use Hamiltonian learning as an example to showcase the interpretability of circuit representations learned by our model.

Our framework is compared with Transformer Quantum State (TQS) (Zhang & Di Ventra, 2023), Variational Quantum Eigensolver (VQE) (Peruzzo et al., 2014), Quantum Approximate Optimization Algorithm (QAOA) (Farhi et al., 2014) and Quantum Architecture Search (QAS) (Du et al., 2022). The metrics for evaluation are second-order Rényi entropy (Rényi, 1961), two-point correlations (Fetter & Walecka, 2003), spin-Z values (Atkins & de Paula, 2010), and square root global fidelity. The definitions of them are introduced in Appendix D. For the former three metrics, we compute the Root Mean Squared Error (RMSE) between the ground-truth values measured form the target states, and the actual values obtained form the learned representations / output states. For fair comparisons, the training objective for all methods is fidelity-global fidelity for the other methods and local fidelity reward for ours. All properties are predicted once trained without fine-tuning. The simulation details for the experiments are presented in Appendix E.

3.1. Learning Quantum States Generated by Instantaneous Quantum Polynomial Circuits

IQP circuits are frequently used to benchmark the classical simulatability of quantum circuits (Bremner et al., 2010). While general IQP circuits are classically intractable to simulate, in this experiment, we focus on a specific family of states generated from shallow circuits to demonstrate the capabilities of our framework. The output states generated

by IQP circuits are

$$|\psi\rangle_k = \bigotimes_{i=0}^{N-1} H_i Z[\alpha]_k \bigotimes_{i=0}^{N-1} H_i |0\rangle^{\otimes N}, \tag{7}$$

where $Z[\alpha]_k$ are single- or two-qubit gates that can be diagonalized in the computational basis, e.g., Z, CZ and $R_z(\alpha)$. In our setting, $Z[\alpha]_k$ contains one column of CZ gates acting on every two adjacent qubits, followed by one column of single-qubit gates randomly sampled from $R_z(\alpha)$, where $\alpha \in [-\pi/2, \pi/2]$ for each qubit. We consider a quantum system with a size of N = 50. We generate 100 different circuits and record the output states as our training set. The quantum circuits are discarded once the states are generated. We train our framework to reconstruct the circuits that prepare the target states in the training set. The action space for generating circuit representations is $\{H, CZ, R_z(\phi)\}$, where ϕ is determined by the agent. During training, we set the maximum number of iterative steps to T = 100. In each step, one gate is applied to one qubit or two nearest neighbor qubits. After training, we generate another 10 different states for evaluation. We use global fidelity and local fidelity as metrics between the reconstructed states and the target states to test the learned circuit representations. Figure 2(a)and (b) show the scaling of global and local fidelity at step t. For 4-qubit states, the local fidelity increases concurrently with global fidelity. However, for 50-qubit states, while the local fidelity monotonically increases, the global fidelity remains stable and sharply rises at the end of the period. The trajectory of the evolution of global fidelity shows that in a large region corresponding to iterative steps 0 to 70, the global fidelity is almost 0. This indicates that the variance of the loss function over the whole parametric space is small, highlighting the existence of barren plateaus. We



Figure 2. Learning quantum states generated by IQP circuits. (a) States generated by 4-qubit IQP circuits. (b) States generated by 50-qubit IQP circuits.

also compare the performance of our framework with other state characterization methods for quantum systems of size N = 10. The results in Table 2 show that our framework outperforms all others across all metrics.

Table 2. Evaluation results of learning states generated by 10-qubit IQP circuits.

Method	Rényi Entropy \downarrow	Correlations \downarrow	$\text{Spin-Z} \downarrow$	Global Fidelity \uparrow
TQS	0.5765	0.2906	0.2309	$0.6894{\pm}0.2946$
VQE	0.1665	0.1100	0.1539	$0.9174 {\pm} 0.1042$
QAOA	0.2538	0.1026	0.1429	$0.8336 {\pm} 0.1617$
QAS	0.3977	0.0895	0.2401	$0.4694{\pm}0.1500$
Ours	3.0737e-07	1.0902e-08	0.0441	0.9851±0.0208

3.2. Learning Quantum States Evolved by Transverse Field Ising Hamiltonians

Next, we consider learning the circuit representations for a family of states evolved by transverse field Ising Hamiltonians, where the exact parameters of the Hamiltonians and evolution time are agnostic to the framework. Starting with product state $|0\rangle^{\otimes N}$, the state is evolved by an Ising Hamiltonian for time τ . The target states after the evolution are defined as

$$|\psi\rangle_k = e^{-iH_{\text{Ising}}\tau}|0\rangle^{\otimes N},\tag{8}$$

where $H_{\text{Ising}} = J \sum_{i=0}^{N-2} Z_i Z_{i+1} + g \sum_{i=0}^{N-1} X_i$ is the transverse field Ising Hamiltonian, t is the evolution time. In our experiment, we set $N = 50, J = -1, g \in [-2.0, -1.0]$ and $\tau \in [0.1, 1.0]$. We sample 10 different qs and 10 τ s uniformly from the range with stride 0.1 to construct the training set of size 100. For training, we set the maximum number of iterative steps to T = 100, each corresponds to applying one gate to each qubit or every two nearest neighbor qubits. The quantum gates composing the action space are $\{\exp(-i\phi X), \exp(-i\phi Z \otimes Z)\}$. To exhibit the results, we average the performance on different parameters g for each evolution time τ . Figure 3(a) shows that the learned circuit can successfully recover the target quantum states with high fidelity. Additionally, we evaluate the circuit depth and compare it to the first-order Trotter decomposition (Suzuki, 1985), which is considered one of the most straightforward methods for simulating the dynamics of quantum systems. As shown in Figure 3, our framework can construct circuits shallower than those generated by the Trotter decomposition in general. This indicates that our framework can serve as an optimization technique for traditional quantum simulation technologies. Notably, our framework does not require prior knowledge on the Hamiltonian parameters, offering greater flexibility compared to the Trotter decomposition method when simulating the dynamics of quantum systems. Figure 3(c) shows the zero-shot transfer performance of applying the framework trained on 50-qubit systems to other quantum N-qubit systems with $N \in \{10, 30, 70, 100\}$. The output states remain high fidelity with the target states of unseen sizes, demonstrating the success of our measurement feature aggregation block. In addition, we compare the performance of our framework with other methods for quantum



Figure 3. Learning 50-qubit quantum states evolved by Ising Hamiltonians. (a) Scaling of global and local fidelity w.r.t. the evolution time. (b) Comparison of the circuit depths for simulating the dynamics between our framework and the Trotter decomposition method. (c) Zero-shot transfer performance on quantum systems of various sizes. Our framework is trained on the 50-qubit system.

systems of size N = 4. Table 3 illustrates the results of predicting different properties. Our framework outperforms other methods on all metrics.

Table 3. Evaluation results of learning 10-qubit states evolved by Ising Hamiltonians, where the evolution time $\tau \in [0.1, 1]$.

Method	Rényi Entropy↓	Correlations \downarrow	$\text{Spin-Z} \downarrow$	Global Fidelity \uparrow
TQS	0.1727	0.1037	0.0944	$0.8524{\pm}0.0957$
VQE	0.5824	0.3044	0.3619	$0.2795 {\pm} 0.2359$
QAOA	0.0324	0.0382	0.0513	$0.9637 {\pm} 0.1402$
QAS	0.3729	0.4349	0.4806	$0.5215{\pm}0.2153$
Ours	0.0108	0.0227	0.0231	$0.9979 {\pm} 0.0012$

3.3. Learning Many-body Ground States

Our third experiment is learning the circuit representations for a family of many-body ground states. We consider two families of ground states separately, the transverse-field Ising ground states, and the anisotropic Heisenberg XXZ ground states.

Learning Transverse-field Ising Ground States. In this experiment, we consider the same Ising Hamiltonians as in Section 3.2, but with the goal of learning the ground states rather than time-evolved states. The configurations are N = 50, J = -1 and $g \in [-2.0, -1.5]$. We uniformly sample 20 different parameters q and compute the corresponding ground states, storing them into the training set. Then we use the QCrep agent to learn the circuits to prepare these ground states. The action space is the same as described in Section 3.2. We set the maximum number of iterative steps to T = 200 during training. Figure 4(a) shows the scaling of global and local fidelity with the parameters g. The local fidelity almost remains stable but the global fidelity slightly drops with the increment of q. This indicates that in high-dimensional space, global fidelity is more sensitive to differences between states compared to local fidelity. Thus, global fidelity may not be a good guidance on learning quantum states, in which a relaxed metric encourages exploration and increases the chance of finding the optimal result. Figure 4(c) shows the zero-shot transfer performance of the framework trained on the 50-qubit system when applied to system sizes of $\{10, 30, 70, 100\}$. The comparison results between different methods for learning 10-qubit Hamiltonian ground states are presented in Table 4.



Figure 4. Learning 50-qubit ground states of transverse-field Ising model. (a) Scaling of global and local fidelity w.r.t. the Ising parameters. (b) Zero-shot transfer performance on learning Ising ground states of various sizes (trained on the 50-qubit system).

Table 4. Evaluation results of learning ground states of 10-qubit Ising systems.

Method	Rényi Entropy \downarrow	Correlations \downarrow	$\text{Spin-Z}\downarrow$	Global Fidelity \uparrow
TQS	0.1187	0.0958	0.0306	$0.9537 {\pm} 0.0724$
VQE	0.6442	0.1475	0.0425	$0.4773 {\pm} 0.0087$
QAOA	0.0368	0.1009	0.0229	$0.9614{\pm}0.0181$
QAS	0.2260	0.1806	0.1706	$0.8032 {\pm} 0.0450$
Ours	0.0989	0.0947	0.0309	$0.9691{\pm}0.0083$

Learning Anisotropic Heisenberg XXZ Ground States. Here, we learn the circuit representations of ground states of a family of 1-D Heisenberg XXZ Hamiltonians. The Hamiltonian is $H_{\text{Heisenberg}} = \sum_{i=0}^{N-1} J_x X_i X_{i+1} + J_y Y_i Y_{i+1} + JZ_i Z_{i+1}$. Throughout the experiment, we set $J_x = J_y =$ -1, and $J \in [-3.0, -2.0]$. To construct the training set, we uniformly sample 10 different J and generate the ground state of system size N = 10. The action space of the agent is $\{\exp(-i\phi X \otimes X), \exp(-i\phi Y \otimes Y), \exp(-i\phi Z \otimes Z)\}$. We set the maximum iterative steps to T = 100. The scaling of global and local fidelity with parameter J is shown in Figure 5(a). Besides, we evaluate the trained framework on out-of-distribution data. We generate 9 different ground states corresponding to $J \in [-1.9, -1.1]$, and use the trained framework to generate circuit representations to reproduce these states. Results in Figure 5(b) show that our framework can successfully be generalized to prepare unseen states within the same state family. The comparison



Figure 5. Learning 10-qubit Heisenberg ground states. (a) Scaling of global and local fidelity w.r.t. the Hamiltonian parameters. (b) Out-of-distribution generalization.

with other methods on 10-qubit system is shown in Table 5. Our framework can accurately recover the three properties of the target states and achieves the highest performance.

Table 5. Evaluation results of learning ground states of 10-qubit Heisenberg XXZ systems.

Method	Rényi Entropy \downarrow	Correlations \downarrow	$\text{Spin-}Z\downarrow$	Global Fidelity \uparrow
TQS	0.7071	0.0017	0.0159	$0.6288{\pm}0.1204$
VQE	0.0042	0.0816	0.7840	$0.4765 {\pm} 0.0105$
QAOA	0.0038	0.0693	0.0000	$0.5970{\pm}0.0085$
QAS	0.3379	0.1234	0.4962	$0.7613 {\pm} 0.0609$
Ours	0.0000	0.0000	0.0000	$0.9550{\pm}0.0229$

3.4. Downstream Application: Hamiltonian Learning

After learning the circuit representations for quantum states, it is natural to investigate the interpretability of these representations. To this end, we consider Hamiltonian learning as a downstream task to evaluate their effectiveness. Hamiltonian learning is a task to determine the parameters of an unknown Hamiltonian, which serves as a meaningful benchmark for assessing how well the learned representations encodes the information of the underlying physical system.

In our setting, we use the circuit representations of the ground states to learn the corresponding Hamiltonians. The quantum systems we consider are the Ising model and the Heisenberg XXZ model. Specifically, we first use QCrep to learn the circuit representations $(v_t^{\dagger}, \phi_t)_{t=T}^1$ for ground states corresponding to Hamiltonians with unknown parameters (g in Ising model and J in Heisenberg model). Next, we concatenate the representations into vectors and pad 0s at the end to ensure the same length. Finally, we employ linear regression to establish the relationship between circuit representations and Hamiltonian parameters using a small training set, and we utilize the learned framework to predict the relationship on the test set. Experimental results in Figure 6 show that, given the circuit representation of a ground state associated with a Hamiltonian with unknown parameters, these unknown parameters can be accurately predicted using only linear regression. Meanwhile, for comparison,



Figure 6. The test set performance of different methods on learning Hamiltonian parameters for 10-qubit (a) Ising and (b) Heisenberg XXZ quantum systems. The x-axis represents the parameter indices, and the y-axis shows the corresponding parameter values.

we use the circuit parameters learned from VQE and QAOA to perform Hamiltonian learning. However, the linear model fails to establish a relationship between the Hamiltonian and circuit parameters. We attribute this outcome to the QCrep learning pipeline, which effectively encodes information about the underlying Hamiltonian into the circuit parameters. This is not achievable with VQE or QAOA, as they rely on gradient-based optimization of the circuit parameters, which perturbs the parameters and hinders the preservation of Hamiltonian information.

3.5. Discussions on Practical Implementation

Impact of Finite Sampling. In the above experiments, our framework is trained using expectation values of measurement outcome computed via classical simulation, which corresponds to the ideal setting of infinite measurement samples. However, real-world experiments only allow sampling the states for finite times. To assess the robustness of our model under realistic conditions, we simulate finite-sample effects by introducing measurement inaccuracies at test time.

The framework is first trained on simulation of infinite sampling data $m = \text{Tr}(M\rho)$ given measurement operator Mand state ρ . At test time, we use finite measurement shots $k \in \{128, 256, 512, 1024\}$ to obtain the measurement data $\langle m \rangle_k$ as the input to our framework. System sizes for IQP, Ising evolution, Ising ground states, Heisenberg ground states are 10, 50, 50, 10 respectively. The results are shown in Figure 7. Inaccurate measurement has nearly no effect on learning IQP circuits, where the action space contains no continuous parameters. For the other three families of states, using only 512 measurement shots is enough for high fidelity reconstruction, demonstrating the effectiveness of our framework in practical scenarios.



Figure 7. Results under finite sampling conditions. (a) Learning states generated by IQP circuits. (b) Learning states evolved by Ising Hamiltonians. (c) Learning Ising ground states. (d) Learning Heisenberg ground states.

Impact of Circuit Noise. Real-world quantum circuits are affected by noise, which causes deviations between actual and ideal measurement outcomes. Unlike measurement inaccuracy, this gap cannot be mitigated via increasing the number of measurement shots. Therefore, it is important to investigate the impact of circuit noise to the construction procedure of circuit representations.

We evaluate the performance of our framework under the condition that the quantum circuit is affected by a global depolarizing noise. The noisy output state is $\rho =$ $\mathcal{N}(U|0\rangle\langle 0|U^{\dagger})$, where \mathcal{N} represents the noise channel, U is the noise-free circuit. We set the noise parameter associated with the noise strength of \mathcal{N} to $p \in \{0.05, 0.1, 0.15, 0.2\}$. System sizes for IQP, Ising evolution, Ising ground states, Heisenberg ground states are 10, 50, 50, 10 respectively. Figure 8 illustrates the impact of varying noise strengths on both global and local fidelity between the learned and target quantum states. Although fidelity degrades with increasing noise strength, our framework maintains consistently high performance up to a noise level of 0.2, demonstrating robustness to moderate levels of noise. To deal with strong noise, strategies like error correction (Shor, 1995; Fowler et al., 2012) or error mitigation (Giurgica-Tiron et al., 2020; Liao et al., 2025) can be employed.



Figure 8. Impact of global depolarizing noise on the performance of our framework. (a) Learning states generated by IQP circuits. (b) Learning states evolved by Ising Hamiltonians. (c) Learning Ising ground states. (d) Learning Heisenberg ground states.

4. Conclusion and Outlook

We propose the explicit circuit representations, which feature efficient property estimation and experimental reconstruction of quantum states. To learn this representation, we design a reinforcement learning framework with a Transformer feature aggregation block and a novel local fidelity reward function. The learning procedure relies exclusively on local measurement data, achieving high accuracy on predicting properties. The learned representations can further be transferred to quantum systems of varying sizes and applied to Hamiltonian learning as a downstream task using a linear model.

Outlook. Our current experiments focus on learning quantum states prepared by shallow circuits and ground states of one-dimensional Hamiltonians. In future work, we would extend our framework to more complex scenarios, such as deeper circuits, two-dimensional Hamiltonians, and critical behaviors of many-body Hamiltonians. Additionally, we would generalize our approach to quantum process learning, enabling the model to handle tasks involving varying input-output state pairs. Another promising direction is to integrate the physics knowledge from large language models into the agent, which may further enhance the learning capability and efficiency of the framework.

Acknowledgements

We thank the anonymous reviewers for their helpful feedback and suggestions to improve this work. This work is supported by the Ministry of Science and Technology, China (MOST2030) with Grant No. 2023200300600, the National Natural Science Foundation of China (NSFC)/Research Grants Council (RGC) Joint Research Scheme via Project N_HKU7107/24, the Guangdong Provincial Quantum Science Strategic Initiative (GDZX2303007 and GDZX2203001), and the Hong Kong Research Grant Council (RGC) through the General Research Fund (GRF) grant 17303923.

Impact Statement

This paper presents work whose goal is to advance the field of Quantum Computation through incorporating advanced Machine Learning techniques to characterize quantum states. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

- Aaronson, S. Shadow tomography of quantum states. In Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, pp. 325–338, 2018.
- Akhtar, A. A., Hu, H.-Y., and You, Y.-Z. Scalable and Flexible Classical Shadow Tomography with Tensor Networks. *Quantum*, 7:1026, 2023.
- Anschuetz, E. R. and Kiani, B. T. Quantum variational algorithms are swamped with traps. *Nature Communications*, 13(1), 2022.
- Atkins, P. and de Paula, J. Atkins' Physical Chemistry. OUP Oxford, 2010. ISBN 9780199543373. URL https://books.google.com.hk/books?id= BV6cAQAAQBAJ.
- Bengio, Y., Ducharme, R., Vincent, P., and Janvin, C. A neural probabilistic language model. *Journal of Machine Learning Research*, 3:1137–1155, 2003.
- Bittel, L. and Kliesch, M. Training variational quantum algorithms is NP-hard. *Physical Review Letters*, 127: 120502, 2021.
- Bremner, M. J., Jozsa, R., and Shepherd, D. J. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, 2010.
- Carleo, G. and Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science*, 355 (6325):602–606, 2017.

- Caro, M. C., Huang, H.-Y., Ezzell, N., Gibbs, J., Sornborger, A. T., Cincio, L., Coles, P. J., and Holmes, Z. Out-of-distribution generalization for learning quantum dynamics. *Nature Communications*, 14(1), 2023.
- Cerezo, M., Sone, A., Volkoff, T., Cincio, L., and Coles, P. J. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1), 2021.
- Chen, A. and Heyl, M. Empowering deep neural quantum states through efficient optimization. *Nature Physics*, 20 (9):1476–1481, 2024.
- Cotler, J. and Wilczek, F. Quantum overlapping tomography. *Physical Review Letters*, 124:100401, 2020.
- Cramer, M., Plenio, M. B., Flammia, S. T., Somma, R., Gross, D., Bartlett, S. D., Landon-Cardinal, O., Poulin, D., and Liu, Y.-K. Efficient quantum state tomography. *Nature Communications*, 1(149):1–7, 2010.
- Du, Y., Huang, T., You, S., Hsieh, M.-H., and Tao, D. Quantum circuit architecture search for variational quantum algorithms. *npj Quantum Information*, 8(1), 2022.
- Du, Y., Yang, Y., Liu, T., Lin, Z., Ghanem, B., and Tao, D. Shadownet for data-centric quantum system learning. arXiv preprint arXiv:2308.11290, 2023.
- Farhi, E., Goldstone, J., and Gutmann, S. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.
- Fetter, A. and Walecka, J. Quantum Theory of ManypArticle Systems. Dover Books on Physics. Dover Publications, 2003. ISBN 9780486428277. URL https://books.google.com.hk/books?id= 0wekf1s83b0C.
- Fowler, A. G., Mariantoni, M., Martinis, J. M., and Cleland, A. N. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3), 2012.
- Friis, N., Marty, O., Maier, C., Hempel, C., Holzäpfel, M., Jurcevic, P., Plenio, M. B., Huber, M., Roos, C., Blatt, R., et al. Observation of entangled states of a fully controlled 20-qubit system. *Physical Review X*, 8(2):021012, 2018.
- Georgescu, I. M., Ashhab, S., and Nori, F. Quantum simulation. *Reviews of Modern Physics*, 86:153–185, 2014.
- Giurgica-Tiron, T., Hindy, Y., LaRose, R., Mari, A., and Zeng, W. J. Digital zero noise extrapolation for quantum error mitigation. In 2020 IEEE International Conference on Quantum Computing and Engineering (QCE), 2020.

- Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S., and Eisert, J. Quantum state tomography via compressed sensing. *Physical Review Letters*, 105:150401, 2010.
- Guo, Y. and Yang, S. Quantum state tomography with locally purified density operators and local measurements. *Communications Physics*, 7(1), 2024.
- Huang, H.-Y., Kueng, R., and Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.
- Huang, H.-Y., Chen, S., and Preskill, J. Learning to predict arbitrary quantum processes. *PRX Quantum*, 4:040337, 2023.
- Huang, H.-Y., Liu, Y., Broughton, M., Kim, I., Anshu, A., Landau, Z., and McClean, J. R. Learning shallow quantum circuits. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pp. 1343–1351, 2024.
- Hubig, C., McCulloch, I. P., and Schollwöck, U. Generic construction of efficient matrix product operators. *Physi*cal Review B, 95:035129, 2017.
- Kitaev, A. Y. Quantum measurements and the abelian stabilizer problem. arXiv preprint arXiv:quant-ph/9511026, 1995.
- Kurmapu, M. K., Tiunova, V., Tiunov, E., Ringbauer, M., Maier, C., Blatt, R., Monz, T., Fedorov, A. K., and Lvovsky, A. Reconstructing complex states of a 20-qubit quantum simulator. *PRX Quantum*, 4(4):040345, 2023.
- Lanczos, C. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45(4):255, 1950.
- Lanyon, B. P., Maier, C., Holzäpfel, M., Baumgratz, T., Hempel, C., Jurcevic, P., Dhand, I., Buyskikh, A., Daley, A. J., Cramer, M., et al. Efficient tomography of a quantum many-body system. *Nature Physics*, 13(12): 1158–1162, 2017.
- Larocca, M., Thanasilp, S., Wang, S., Sharma, K., Biamonte, J., Coles, P. J., Cincio, L., McClean, J. R., Holmes, Z., and Cerezo, M. Barren plateaus in variational quantum computing. *Nature Reviews Physics*, 7(4):174–189, 2025.
- Lewis, L., Huang, H.-Y., Tran, V. T., Lehner, S., Kueng, R., and Preskill, J. Improved machine learning algorithm for predicting ground state properties. *Nature Communications*, 15(1):895, 2024.
- Liao, M., Zhu, Y., Chiribella, G., and Yang, Y. Noiseagnostic quantum error mitigation with data augmented neural models. *npj Quantum Information*, 11(1):8, 2025.

- McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R., and Neven, H. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9 (1), 2018.
- McMillan, W. L. Ground state of liquid he⁴. *Phys. Rev.*, 138:A442–A451, 1965.
- Mitarai, K., Negoro, M., Kitagawa, M., and Fujii, K. Quantum circuit learning. *Physical Review A*, 98(3), 2018.
- Motta, M., Sun, C., Tan, A. T., O'Rourke, M. J., Ye, E., Minnich, A. J., Brandao, F. G., and Chan, G. K.-L. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nature Physics*, 16(2):205–210, 2020.
- Nielsen, M. and Chuang, I. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press, 2010. ISBN 9781139495486. URL https://books.google. com.hk/books?id=-s4DEy70-a0C.
- Ostaszewski, M., Trenkwalder, L. M., Masarczyk, W., Scerri, E., and Dunjko, V. Reinforcement learning for optimization of variational quantum circuit architectures. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, volume 34, pp. 18182–18194, 2021.
- Perez-Garcia, D., Verstraete, F., Wolf, M. M., and Cirac, J. I. Matrix product state representations. arXiv preprint arXiv:quant-ph/0608197, 2007.
- Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A., and O'Brien, J. L. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1), 2014.
- Qian, Y., Du, Y., He, Z., Hsieh, M.-H., and Tao, D. Multimodal deep representation learning for quantum crossplatform verification. *Physical Review Letters*, 133: 130601, 2024.
- Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., and Dormann, N. Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268):1–8, 2021.
- Rényi, A. On measures of entropy and information. In Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics, volume 4, pp. 547–562, 1961.
- Sachdev, S. Quantum Phase Transitions. Cambridge University Press, 1999. ISBN 9780521582544. URL https://books.google.com.hk/books?id= K_lhQqAACAAJ.

- Sachdev, S. The quantum phases of matter. *arXiv preprint arXiv:1203.4565*, 2012.
- Scarpa, G., Molnár, A., Ge, Y., García-Ripoll, J. J., Schuch, N., Pérez-García, D., and Iblisdir, S. Projected entangled pair states: Fundamental analytical and numerical limitations. *Physical Review Letters*, 125:210504, 2020.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- Sharir, O., Levine, Y., Wies, N., Carleo, G., and Shashua, A. Deep autoregressive models for the efficient variational simulation of many-body quantum systems. *Physical Review Letters*, 124(2), 2020.
- Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52:R2493–R2496, 1995.
- Sorella, S. Green function monte carlo with stochastic reconfiguration. *Physical Review Letters*, 80:4558–4561, 1998.
- Suzuki, M. Decomposition formulas of exponential operators and lie exponentials with some applications to quantum mechanics and statistical physics. *Journal of Mathematical Physics*, 26(4):601–612, 1985.
- Tang, Y., Xiong, H., Yang, N., Xiao, T., and Yan, J. Towards LLM4QPE: Unsupervised pretraining of quantum property estimation and a benchmark. In *The Twelfth International Conference on Learning Representations*, 2024a.
- Tang, Y., Yang, N., Long, M., and Yan, J. SSL4Q: Semisupervised learning of quantum data with application to quantum state classification. In *Forty-first International Conference on Machine Learning*, 2024b.
- Torlai, G., Wood, C. J., Acharya, A., Carleo, G., Carrasquilla, J., and Aolita, L. Quantum process tomography with unsupervised learning and tensor networks. *Nature Communications*, 14(1):2858, 2023.
- Tóth, G., Wieczorek, W., Gross, D., Krischek, R., Schwemmer, C., and Weinfurter, H. Permutationally invariant quantum tomography. *Physical Review Letters*, 105: 250403, 2010.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u., and Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pp. 6000–6010, 2017.
- Wang, J., Paesani, S., Santagati, R., Knauer, S., Gentile, A. A., Wiebe, N., Petruzzella, M., O'Brien, J. L., Rarity,

J. G., Laing, A., and Thompson, M. G. Experimental quantum hamiltonian learning. *Nature Physics*, 13(6): 551–555, 2017.

- Wauters, M. M., Panizon, E., Mbeng, G. B., and Santoro, G. E. Reinforcement-learning-assisted quantum optimization. *Physical Review Research*, 2:033446, 2020.
- White, S. R. and Feiguin, A. E. Real-time evolution using the density matrix renormalization group. *Physical Review Letters*, 93:076401, 2004.
- Wiebe, N., Granade, C., Ferrie, C., and Cory, D. Hamiltonian learning and certification using quantum resources. *Physical Review Letters*, 112(19), 2014.
- Wu, W., Yan, G., Lu, X., Pan, K., and Yan, J. QuantumDARTS: differentiable quantum architecture search for variational quantum algorithms. In *Proceedings of the 40th International Conference on Machine Learning*, 2023a.
- Wu, Y.-D., Zhu, Y., Bai, G., Wang, Y., and Chiribella, G. Quantum similarity testing with convolutional neural networks. *Physical Review Letters*, 130:210601, 2023b.
- Wu, Y.-D., Zhu, Y., Wang, Y., and Chiribella, G. Learning quantum properties from short-range correlations using multi-task networks. *Nature Communications*, 15(1): 8796, 2024.
- Xiao, T., Fan, J., and Zeng, G. Parameter estimation in quantum sensing based on deep reinforcement learning. *npj Quantum Information*, 8(1):2, 2022.
- Yao, J., Lin, L., and Bukov, M. Reinforcement learning for many-body ground-state preparation inspired by counterdiabatic driving. *Physical Review X*, 11:031070, 2021.
- Zhang, S.-X., Hsieh, C.-Y., Zhang, S., and Yao, H. Differentiable quantum architecture search. *Quantum Science* and Technology, 7(4):045023, 2022.
- Zhang, Y.-H. and Di Ventra, M. Transformer quantum state: A multipurpose model for quantum many-body problems. *Physical Review B*, 107(7), 2023.
- Zhou, Z., Du, Y., Yin, X.-F., Zhao, S., Tian, X., and Tao, D. Optical quantum sensing for agnostic environments via deep learning. *Physical Review Research*, 6(4):043267, 2024.
- Zhu, Y., Wu, Y.-D., Bai, G., Wang, D.-S., Wang, Y., and Chiribella, G. Flexible learning of quantum states with generative query neural networks. *Nature communications*, 13(1):6222, 2022.
- Zhu, Y., Wu, Y.-D., Liu, Q., Wang, Y., and Chiribella, G. Predictive modelling of quantum process with neural networks. arXiv preprint arXiv:2308.08815, 2023.

A. Related Work

Tomography-based quantum state characterization. Tomography-based methods use direct measurement to characterize quantum states. To accurately characterize the full quantum state, Quantum State Tomography (Tóth et al., 2010; Gross et al., 2010; Cramer et al., 2010; Lanyon et al., 2017; Cotler & Wilczek, 2020) has been proposed, which measures the state in exponential number of basis to obtain the state vector. Other methods focus on constructing a partial knowledge of the state. For instance, Shadow Tomography (Aaronson, 2018) targets at characterizing the measurement values of 2-outcome measurements using only a few copies of the states. Classical shadow (Huang et al., 2020; Akhtar et al., 2023) utilizes randomized measurement to efficiently estimate local properties of the states. Noteworthy, there is a special family of work that uses Tensor network, e.g., Matrix Product State (MPS) (Perez-Garcia et al., 2007) and Projected Entangled Pair States (PEPS) (Scarpa et al., 2020), to approximate the state vector of a quantum state. The original high dimensional state vector is decomposed into multiple low-rank tensors with restricted bound dimension.

Variational-based quantum state characterization. Alternative to state tomography, variational quantum algorithms optimize the parameters of a variational ansatz, i.e., a parameterized quantum circuit, to approach the target state. Two representative methods are Variational Quantum Eigensolver (VQE) (Peruzzo et al., 2014) and Quantum Approximate Optimization Algorithm (QAOA) (Farhi et al., 2014). These methods update their output towards the target states, usually the ground states of a Hamiltonian, by measuring the energy and computing quantum gradient descend via, e.g., parameter shift rule (Mitarai et al., 2018). In addition to optimizing parameters, Quantum Architecture Search has been proposed to optimize the circuit ansatz. Du et al. (2022) traverse a candidate gate set and select the gate configurations that achieve the highest scores on the target objective. Wauters et al. (2020); Yao et al. (2021); Ostaszewski et al. (2021) utilize reinforcement learning to optimize the circuit while keeping quantum gradient descend to update parameters. Zhang et al. (2022); Wu et al. (2023a) propose differentiable strategy to simultaneously update the ansatz and parameters.

Machine learning-based quantum state characterization. Machine learning can be used to learn the measurement values of states, and predict state properties. The machine learning state characterization methods can mainly be categorized into two classes – Neural Quantum State (Carleo & Troyer, 2017; Sharir et al., 2020; Zhang & Di Ventra, 2023; Chen & Heyl, 2024) and Neural State Representation (Zhu et al., 2022; Tang et al., 2024a; Qian et al., 2024). The Neural Quantum State represents a quantum state as a neural network, where sampling the neural network corresponds to measuring the state. Parameters of the neural network can be updated via Variational Monte Carlo (McMillan, 1965) and Stochastic Reconfiguration (Sorella, 1998) methods. The Neural State Representation compresses the quantum state into a classical description, usually a low dimensional vector, via pretraining. Zhu et al. (2022) adopt a self-supervised manner to predict the measurement values of some measurement operators given other operators. Tang et al. (2024a) use language modeling (Bengio et al., 2003) as the pretraining strategy. In Qian et al. (2024), the vector pretrains the representation by fitting the inner product to fidelity. Afterwards, the pretrained representation can be fine-tuned for downstream tasks, such as predicting the properties of quantum states.

Different from previous machine learning-based methods, we decode the state representation into a novel circuit representation instead of low dimensional vector to support experimental reconstruction ability. Our representation is suitable for downstream applications like Hamiltonian learning. Unlike the reinforcement learning for quantum architecture search, our framework circumvents the need of calculating gradients with respect to the circuit parameters, and possesses the ability to characterize a family of states rather than one specific state.

B. Preliminaries

We review some of the key concepts in quantum computation. For a more comprehensive overview, please refer to Nielsen & Chuang (2010).

Quantum states are quantum counterparts of classical bits. They can be mathematically represented as vectors in Hilbert space, i.e., state vectors, denoted as $|\psi\rangle \in \mathbb{C}^{2^N}$, satisfying $|||\psi\rangle||_2 = 1$, where N is the system size or the number of qubits. The notation $|\cdot\rangle$ is just used to emphasize that ψ is a (column) vector. Its dual (row vector) is given by $\langle \cdot | \equiv |\cdot\rangle^{\dagger}$, where " \dagger " is the notation for conjugate transpose. The standard basis for quantum states is the computational basis $\{|i\rangle\}_{i=0}^{2^{N-1}}$, where " \dagger " is the vector whose *i*-th element is 1 and others are 0. For example, $|0\rangle = (1, 0, 0, \cdots, 0)$. Alternatively, we can use the mixed state to describe a probability ensemble of quantum states $\{p_i, |\psi_i\rangle\}$. p_i is the probability of the quantum system being in the state $|\psi_i\rangle$. This can be represented as density matrix $\rho \in \mathbb{C}^{2^N \times 2^N}$, where $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$. Clearly, for pure state $|\psi\rangle$, the corresponding density matrix is $|\psi\rangle \langle \psi|$. Multiple quantum states can be combined to form a compositional system, which is

represented by the tensor product (Kronecker product) denoted as " \otimes ". For two states $|\psi\rangle$, $|\phi\rangle \in \mathbb{C}^{2^N}$, their composition is given by $|\psi\rangle \otimes |\phi\rangle \in \mathbb{C}^{2^{2N}}$. We use the notation $|\cdot\rangle^{\otimes N}$ to denote an N-qubit product state, e.g., $|0\rangle^{\otimes N} \equiv |0\rangle \otimes \cdots \otimes |0\rangle$.

The similarity between two quantum states can be quantified by (global) fidelity and trace distance. In this paper, we focus exclusively on the global fidelity. Given two density matrices ρ and σ , the global fidelity is defined as

$$F(\rho,\sigma) = \left(\operatorname{Tr}\left(\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right)\right)^2.$$
(9)

If the two states are pure states $|\psi\rangle$ and $|\phi\rangle$, the fidelity simplifies to $|\langle\psi|\phi\rangle|^2$, which is closely related to the cosine similarity between two vectors.

Quantum states can be measured, causing them to collapse into classical bits. Measurement is described by a set of measurement operators $\{M_j\}$, where each M_j is a Hermitian matrix, i.e., $M_j^{\dagger} = M_j$. In the case of projective measurements, the operators are projectors that satisfy $\sum_j M_j = I$ and $M_j M_k = \delta_{j,k} M_j$. The measurement outcomes, which correspond to classical bits, are associated with the index j. When measuring a state ρ , the probability of obtaining outcome j is given by $p(j) = \text{Tr}(M_j\rho)$. The observable $M = \sum_j jM_j$ describes the overall measurement results, and the expectation value of the measurement on the state ρ is $m = \sum_j jp(j) = \text{Tr}(M\rho)$. Additionally, measurement operators can be composed using tensor products to form new measurements for larger quantum systems.

Quantum states can be evolved by quantum gates, analogous to classical logical gates, which are represented by unitary matrices U that satisfy $U^{\dagger}U = UU^{\dagger} = I$. A unitary matrix can be generated from a Hamiltonian H – a Hermitian matrix – using a parameter ϕ , and is expressed as $U(\phi) = \exp(-iH\phi)$. A special group of unitary matrices are Pauli matrices – X, Y, and Z, where

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$
 (10)

The Pauli matrices form the single-qubit Pauli gates. Besides these gates, other typical quantum gates are single-qubit rotation gates $R_x(\theta) = \exp(-iX\theta/2)$, $R_y(\theta) = \exp(-iY\theta/2)$, $R_z(\theta) = \exp(-iZ\theta/2)$, and two-qubit gates $CX = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X$, $CZ = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Z$. More general quantum gates can be decomposed into these single-qubit and two-qubit gates.

C. Proof of Proposition 2.1

Proposition 2.1 states that if the agent learns a policy that constructs an N-qubit quantum state with average n-local $(1 \le n \le N)$ fidelity $L^{(n)}(\rho_s^{(T)}, |0\rangle\langle 0|^{\otimes N}) \ge 1 - \epsilon$, then the global fidelity satisfies $F(\rho_s^{(T)}, |0\rangle\langle 0|^{\otimes N}) \ge 1 - (N - n + 1)\epsilon$. Before proving this proposition, we present the following lemma.

Lemma C.1. The observable $O^{(n)}$ associated with the *n*-local fidelity has the largest eigenvalue $\lambda_0 = 1$ and the second largest eigenvalue $\lambda_1 = 1 - 1/(N - n + 1)$.

Proof. The local operator $O_i^{(n)}$ acting on qubits $\{i, \ldots, i+n-1\}$ can be expressed as

$$O_i^{(n)} = \underbrace{I \otimes \cdots \otimes I}_i \otimes |0\rangle \langle 0|_{i:i+n} \otimes \underbrace{I \otimes \cdots \otimes I}_{N-i-n}$$
(11)

$$= \operatorname{diag}(\underbrace{1, 1, \dots, 1}_{2^{i}}, \underbrace{0, 0, \dots, 0}_{2^{i+n} - 2^{i}}) \otimes \operatorname{diag}(\underbrace{1, 1, \dots, 1}_{2^{N-i-n}})$$
(12)

$$= \operatorname{diag}(\underbrace{\mathbf{1}_{2^{i}}, \mathbf{0}_{2^{i+n}-2^{i}}, \mathbf{1}_{2^{i}}, \mathbf{0}_{2^{i+n}-2^{i}}, \dots, \mathbf{1}_{2^{i}}, \mathbf{0}_{2^{i+n}-2^{i}}}_{2^{N-i-n}}).$$
(13)

Now that $O_i^{(n)}$ is a diagonal matrix, the elements 1s and 0s are the eigenvalues. Next, we are interested in the eigenvalues of $O^{(n)}$, which is defined as

$$O^{(n)} = \frac{1}{N - n + 1} \sum_{i=0}^{N - n} O_i^{(n)}.$$
(14)

 $O^{(n)}$ is also a diagonal matrix that has eigenvalues $0 \le \lambda_j \le 1$, with corresponding eigenvectors $|j\rangle$. Each of λ_j is a sum of N - n + 1 items of the corresponding *j*-th entries on the diagonal of local operators $O_i^{(n)}$, denoted by $O_i^{(n)}[j]$

 $(0 \le j \le 2^N - 1)$. Since the entries $O_i^{(n)}[j]$ are either 1 or 0, the value of λ_j depends on the number of 1s in $O_i^{(n)}[j]$. Set j = 0, we obtain the largest eigenvalue as all entries $O_i^{(n)}[0]$ are 1, thus $\lambda_0 = 1$. Next, let j = 1, the entry $O_0^{(n)}[1]$ is 0 while others are 1, so the second largest eigenvalue $\lambda_1 = (N - n)/(N - n + 1) = 1 - 1/(N - n + 1)$.

Corollary C.2. λ_0 is a unique eigenvalue of $O^{(n)}$.

Proof. Consider $j = 2^i - 1$ for any 0 < i < N - n + 1, we notice that $O_i^{(n)}[j] = 1$ and $O_{i-1}^{(n)}[j] = 0$. Thus there cannot exist a j > 0 in which the entries $O_i^{(n)}[j]$ are 1 for all i. This means that j = 0 generates the unique largest eigenvalue $\lambda_0 = 1$ with eigenvector $|\lambda_0\rangle = |0\rangle^{\otimes N}$.

Now we prove Proposition 2.1. We use spectral decomposition on the local fidelity observable $O^{(n)}$ to construct the relation between average local fidelity $\text{Tr}(O^{(n)}\rho)$ and fidelity F as follows

$$\operatorname{Tr}(O^{(n)}\rho) = \operatorname{Tr}\left(\sum_{k\geq 0} \lambda_k |\lambda_k\rangle \langle \lambda_k | \rho\right)$$
(15)

$$= \langle 0|^{\otimes N} \rho |0\rangle^{\otimes N} + \operatorname{Tr}\left(\sum_{k \ge 1} \lambda_k |\lambda_k\rangle \langle \lambda_k |\rho\right)$$
(16)

$$=F + \sum_{k>1} \lambda_k \langle \lambda_k | \rho | \lambda_k \rangle \tag{17}$$

$$\leq F + \lambda_1 \sum_{k>1} \langle \lambda_k | \rho | \lambda_k \rangle \tag{18}$$

$$= F + \lambda_1 (1 - \langle \lambda_0 | \rho | \lambda_0 \rangle) \tag{19}$$

$$=F+\lambda_1(1-F).$$
(20)

Lemma C.1 tells us that $\lambda_1 = 1 - \frac{1}{N-n+1}$. Suppose $L^{(n)}(\rho, |0\rangle\langle 0|^{\otimes N}) = \text{Tr}(O^{(n)}\rho) \ge 1 - \epsilon$, then $F \ge 1 - \frac{\epsilon}{1-\lambda_1} = 1 - (N-n+1)\epsilon$.

D. Quantum Properties of Interest in the Experiments

In the experiment section, we consider 3 different properties along with global fidelity for performance evaluation, namely the second-order Rényi entropy (Rényi, 1961), two-point correlations (Fetter & Walecka, 2003) and spin-Z values (Atkins & de Paula, 2010). These are important quantities that characterize quantum states from different perspectives. Rényi entropy is a non-linear property, while the two-point correlation and the spin-Z are linear properties.

Second-order Rényi entropy. This quantity is used to characterize the subsystem (some of the qubits) entanglement of a quantum state. Denote ρ_A as the reduced density matrix of quantum state ρ on its subsystem A, i.e., $\rho_A = \text{Tr}_A(\rho)$. The Rényi entropy quantifies the entanglement strength of A, which is computed by

$$S_{\alpha}(\rho_A) = \frac{1}{1-\alpha} \log \operatorname{Tr}(\rho_A^{\alpha}), \tag{21}$$

where α is the order, which is set to 2 in our experiments. We consider the average value of N-1 qubit subsystems.

Two-point correlation. The correlation function describes the relationships between different parts of the quantum system. This is useful for characterizing quantum phases of matter (Sachdev, 2012) and studying critical behavior (Sachdev, 1999). We consider the two-point correlation defined as follows

$$\mathcal{C}_{0,j} = \operatorname{Tr}(Z_0 Z_j \rho). \tag{22}$$

We take the average of all correlation values for $0 \le j < N$.

Spin-Z value. This quantity describes the angular momentum of a many-body quantum state. In our experiments, we consider the spin-Z value, namely the angular momentum in the Z direction, which is defined as

$$s = \operatorname{Tr}\left(\sum_{i} Z_{i}\rho\right). \tag{23}$$

To evaluate the performance of different methods in predicting the aforementioned properties, we first compute the true properties of the target states. Next, we apply the benchmarked methods to predict these properties. Finally, we calculate the root mean squared error (RMSE) between the actual and predicted properties as the evaluation metric.

E. Simulation of Quantum Systems

To simulate large-scale quantum systems, we use the Matrix Product State (MPS) (Perez-Garcia et al., 2007) to represent quantum states, rather than directly using the full state vector. MPS decomposes the state vector into a chain of low-rank tensors through methods such as singular value decomposition, truncating the singular values to compress the state from $\mathcal{O}(2^N)$ to $\mathcal{O}(N\chi^2 d)$ scale, where d is the physical dimension (typically d = 2 for qubit systems), and χ is the bond dimension, which represents the number of singular values retained and is related to the degree of entanglement. For product states, $\chi = 1$, while for maximally entangled state, χ scales exponentially with the system size. Since the quantum states we consider exhibit a low degree of entanglement, e.g., the Ising ground states, the Heisenberg ground states, and states prepared by shallow circuits, we restrict $\chi \leq 16$ throughout our experiments.

Afterwards, to simulate the evolution of states, we apply Matrix Product Operators (MPO) (Hubig et al., 2017) to MPS. The evolution of quantum states can be viewed as applying unitaries to the states, which is equivalent to applying MPO to MPS. For single-qubit gates, the MPO is simply the gate itself. For multi-qubit gates, the corresponding MPO can be derived through tensor decomposition similar to MPS. To simulate the time evolution of a state $|\psi\rangle$ governed by a Hamiltonian $H = \sum_{l} H_l$, where H_l are local Pauli terms, we first apply the first-order Trotter decomposition (Suzuki, 1985) to approximate $e^{-iH\tau}$. This yields

$$e^{-iH\tau} \approx \prod_{k=i}^{N} \prod_{l} e^{-iH_{l}\delta_{\tau}},\tag{24}$$

where δ_{τ} is the time step and $N = \tau/\delta_{\tau}$. In the Ising evolution experiment, we set $\delta_{\tau} = 0.1$. Following this, we use the Time-Evolving Block Decimation (TEBD) algorithm (White & Feiguin, 2004) to simulate the evolution. The Hamiltonian terms are divided into even and odd components, and a series of brickwork MPOs are applied to the MPS to perform the time evolution.

For simulating the ground states, we use the DMRG algorithm. First, the Hamiltonian is decomposed into MPO. Then each tensor of MPS is iteratively updated, sweeping from left to right and from right to left. For each tensor, Lanczos method (Lanczos, 1950) is applied to find the eigenvalues and eigenvectors, and the tensor is optimized to the eigenvector with the minimum eigenvalue. This procedure is repeated until the energy converges. In our implementation, the MPS is randomly initialized. We set the maximum dimension of Krylov space to 10, and the maximum sweep steps to 200. The iteration stops if the energy difference between to updates is smaller than 10^{-4} . Note that for Hamiltonians with degenerate eigenspace, the ground states found by DMRG can be different for different initialization of MPS and different parameter specification. Therefore, we turn to imaginary-time evolution (Motta et al., 2020) to simulate the Heisenberg ground states, which is steered by TEBD algorithm with the time being an imaginary number. This guarantees deterministic ground states if the initial MPS, the time step δ_{τ} and total steps N are fixed. We set the initial MPS to $|0\rangle$, $\delta_{\tau} = 0.01$ and N = 10.

F. Additional Implementation Details

The policy network of our agent consists of a Transformer measurement feature aggregation block as encoder followed by an MLP for decision making. The encoder comprises two Transformer encoder layer. The positional encoding follows the standard procedure in (Vaswani et al., 2017). The embedding dimension is set to 128. The number of heads for the MHA is 4. For the decision making MLP, we use 3 linear layers with ReLU activation, and the feature dimension is 512. We discard any dropout layer.

To train the policy network, we use Adam optimizer with learning rate 0.001. For the implementation of PPO, we use Stable

Baselines3 package (Raffin et al., 2021). The batch size is set to 1000. We use a cutoff KL divergence 0.05 between two updates of the policy to enhance training stability.

G. Resource Requirement for Training and Inference of QCrep

Table 6 and Table 7 detail the resources required in training and inference for each experiment we conducted respectively. "System Size" denotes the number of qubits of the target state family. "#Iterations" denotes the total number of iterations required for the RL agent to learn the family of states from beginning until convergence, where each iteration is an episode of maximum length T = 200 for Ising ground states and 100 for others. "#Observables" is the number observables required for measurements in each iteration.

Experiment	System Size	#Iterations	#Observables
IQP	50	610	441
Evolve Ising	50	1240	441
Ground Ising	50	1880	441
Ground Heisenberg	10	2040	81

Table 6. Resource requirement for training.

Table 7. Resource requirement for inference.

Experiment	System Size	Circuit Depth	#Observables
IQP	50	2	441
Evolve Ising	50	10	441
Ground Ising	50	22	441
Ground Heisenberg	10	28	81

H. Additional Experiment Results

H.1. Learning States Using n-local Fidelity with Larger n

With the increment of entanglement, the local fidelity and global fidelity diverges but can still be lower-bounded by Proposition 2.1. In practice, one can increase n to obtain higher global fidelity. We consider 50-qubit quantum states prepared by 10 layers of randomly selected single-qubit and two-qubit gates. The gates satisfy $V = \exp(-i\phi G)$, where $\phi \in [-\pi/2, \pi/2]$ and G are sampled from universal single- and two-qubit Pauli operators. The model is trained using n-local fidelity. Figure 9 shows the scaling of global and local fidelity with respect to n. The local fidelity remains high, indicating that the circuit representation can still be faithfully utilized to predict local properties of interest. Whereas if training on 1-local fidelity, the global fidelity drops to nearly 0. However, this can be improved if we increase n, which attains a relatively high value at n = 4. The result further demonstrates the versatility of our model on learning various types of quantum states.

H.2. Learning Rotated GHZ States

To achieve faithful global fidelity during circuit construction for states with higher entanglement, it is typically necessary to use *n*-local fidelity with larger *n*. However, we demonstrate that in certain cases, high global fidelity can be attained using only 1-local fidelity, even for highly entangled states, by carefully designing the action space. Specifically, we consider a family of Z-rotated GHZ states as the target states, which are known to be maximally entangled. We set the number of qubits N = 50. The action space is chosen as $\{H, \text{CNOT}, R_z(\theta)\}$, with $\theta \in (-\pi/4, \pi/4)$. As shown in Table 8, both the local and global fidelity between the reconstructed and target states indicate that the agent can still achieve strong performance, despite the high entanglement.



Figure 9. Learning 50-qubit states prepared by 10-layer universal circuits using n-local fidelity.

Table 8. Learning 50-qubit Z-rotated GHZ states.

Local Fidelity	Global Fidelity
0.9927 ± 0.0151	0.9275 ± 0.1511

H.3. Universal Gate Set as Action Space and Mixture State Family

In our experiments, we focus on restricted action spaces. They are constructed by utilizing prior knowledge of the underlying physical system of the target state family. It is an interesting question to explore how the agent performs when a universal gate set is considered, and when the state family is not restricted to one particular physical system.

Here we consider a mixture state family – the ground states of Ising model together with the ground states of Heisenberg model. The coefficients of the Hamiltonian are chosen the same as in Experiment Section 3.3. We set the number of qubits to 4. The gate set is chosen as $g = \exp(i\theta G)$, where $G = \{X, Y, Z\} \cup \{X, Y, Z\}^{\otimes 2}$ takes all possible combinations of single- and two-qubit Pauli operators, which form universal 2-local gates. The parameters $\theta \in [-\pi/2, \pi/2]$. Table 9 shows that our model can also perform well using a universal gate set. We highlight that in many practical scenarios, some prior information is available to inform the choice of action space. For example, it is often possible to learn the ground states of a Heisenberg-interaction many-body system without knowing the interaction coefficients but knowing the skeleton of the Hamiltonian.

Table 9. Learning a mixture state family using universal 2-local gates.

Experiment	System Size	Fidelity	Rényi Entropy	Two-point Correlations	Spin-Z
Mixture family	4	$0.9587 {\pm} 0.0130$	0.0745	0.0128	0.0434

I. Ablation Study

In our framework, the agent employs a Transformer network as a feature extractor for the measurement data. A simpler alternative is to use an MLP. In this section, we investigate the impact of architectural choices on entanglement capture and zero-shot transferability by considering the task of learning 50-qubit Ising ground states. First, we replace the Transformer block with an MLP. The input measurement data is flattened and passed through the MLP with 49×9 input neurons. Unlike the Transformer, this architecture is not size-agnostic, and the number of input neurons increases with system size. As a

result, the MLP architecture does not naturally support transfer to systems of different sizes. To enable zero-shot transfer, we pad the measurement data with zeros when applying the MLP-trained model to 10-qubit system.

Results in Table 10 shows that, while the local fidelity achieved by the Transformer and MLP are comparable, the Transformer outperforms the MLP in terms of global fidelity. This suggests two key advantages of the Transformer architecture in our framework: (1) it effectively captures entanglement as long-range dependencies in the measurement data, without increasing the number of network parameters with system size, and (2) it enables zero-shot transfer learning across quantum systems of different sizes.

Model	System Size	Local Fidelity	Global Fidelity
Transformer	50	0.9986 ± 0.0003	0.9673 ± 0.0083
Transformer (transfer)	10	0.9876 ± 0.0032	0.9391 ± 0.0160
MLP	50	0.9985 ± 0.0002	0.9560 ± 0.0041
MLP (transfer)	10	0.7311 ± 0.0271	0.1617 ± 0.0456

Table 10. Architectural comparison of the feature extraction layer.