Journal Name

Crossmark

RECEIVED
dd Month yyyy

REVISED
dd Month yyyy

ARTICLE TYPE

Quantum State Preparation using Quantum Neural Network

Asif Akhtab Ronggon[1], Tuhin Hossain[2], Sheikh Abu Al Raihan[3], and F. M. Mridha[4,*]

[1]Department of Electrical and Electronics Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

[2]Department of Computer Science and Engineering, Jahangirnagar University, Savar, Bangladesh

[3]Department of Electrical and Electronics Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

[4]*Department of Computer Science, American International University-Bangladesh, Dhaka, Bangladesh

E-mail: firoz.mridha@aiub.edu

Keywords: Quantum State Preparation(QSP), Quantum Neural Networks(QNNs), Variational Quantum Algorithms(VQA), Quantum Native Loss Function, Quantum Simulation, Quantum Optimization, Quantum Algorithms, Gradient-Based Optimization in Quantum Circuits, and Variational Ansatz Design.

Abstract

Quantum state preparation(QSP) is fundamental to the success of quantum computing(QC), enabling high-fidelity execution of quantum algorithms across diverse applications, including quantum simulation, quantum machine learning(QML), and quantum communication. This paper introduces a novel Variational Quantum Algorithm(VQA) framework that integrate Quantum Neural Networks(QNNs) for efficient state preparation. By leveraging a uniquely structured variational ansatz and incorporating quantum-native loss functions, such as trace distance, Fubini-Study distance, and Bures distance, the proposed framework optimizes the preparation of quantum states while respecting the intrinsic geometry of quantum spaces. The use of an enhanced parameter shift rule for exact gradient computation and a quantum-aware adaptive optimizer with momentum-based updates addresses challenges such as barren plateaus and computational overhead, ensuring robust and scalable QSP. The experimental results show that this strategy obtained the highest fidelities of 0.999933 for 2-qubits, 0.99990 for 3-qubits, 0.997691 for 4-qubits, and 0.937918 for 5-qubits, training time increased from 14.40s for 2-qubits to 211.09s for 5-qubits when using Trace Distance. Similarly, the fidelities obtained with the Fubini Study distance are 0.999922 for 2-qubits, 0.999896 for 3-qubits, 0.999761 for 4-qubits, 0.921357 for 5-qubits, and the training time increases from 14.46s for 2-qubits to 199.54s for 5-qubits. In addition, with Bures Distance, the fidelities obtained were 0.999940 for 2-qubits, 0.999930 for 3-qubits, 0.999261 for 4-qubits, and 0.942159 for 5-qubits, and the training time increased from 19.29s for 2-qubits to 200.02s for 5-qubits. When we use the Bures distance as a loss function, it shows superior performance over conventional methods, highlighting the framework's resilience to noise and hardware imperfections. This study presents a significant advancement in QSP, offering a scalable, noise-resilient solution with transformative implications for quantum algorithms, quantum simulations, QML, and quantum communication.

1 Introduction

Quantum computing(QC) has emerged as a transformative field with the potential to revolutionize various domains, including cryptography, optimization, machine learning(ML), and algorithm design [1] [2] [3]. A critical challenge in QC is the preparation of quantum states, which form the foundation for the execution of quantum algorithms [4]. QSP refers to the process of initializing a quantum system in a specific, desired state, and it plays a pivotal role in applications ranging from quantum simulation and QML to quantum communication and cryptography. The fidelity with which these quantum states can be prepared directly influences the performance and accuracy of quantum algorithms, making efficient and high-fidelity state preparation essential for the advancement of quantum technology [5].

Beyond QC, QSP is of paramount importance in quantum biology and quantum chemistry. In quantum biology, the processes by which biological systems operate on the molecular and atomic

scales, such as photosynthesis, enzyme catalysis, and protein folding, may be based on quantum mechanical effects [6] [7] [8]. Understanding and manipulating these quantum states can lead to breakthroughs in the development of bioinspired quantum technologies, thereby opening new frontiers in biotechnology and medicine. Similarly, in quantum chemistry, preparation of specific quantum states is crucial for simulating molecular systems and chemical reactions. QSP enables accurate modeling of complex chemical processes, which is vital in fields such as drug discovery, material science, and energy storage [9] [10].

As we continue to explore the broader applications of QSP, its role extends to QML and quantum communication. In QML, the precise preparation of quantum states allows efficient encoding of data in quantum circuits, facilitating tasks such as classification and regression [11]. This enables quantum algorithms to leverage quantum parallelism, thereby leading to faster computation and improved model accuracy. Similarly, in quantum communication, accurate state preparation is essential for protocols such as quantum key distribution(QKD), which relies on entangled states to ensure secure communication. The ability to prepare these states with high fidelity is fundamental to the success of quantum communication systems, including encryption, quantum teleportation, and the development of quantum repeaters, all of which are key to building secure and scalable quantum networks [12] [13] [14].

Traditional QSP methods, such as gradient ascent pulse engineering(GRAPE) and chopped random-basis optimization(CRAB), have achieved notable successes in the preparation of simple quantum states [15] [13]. However, these techniques are often limited by issues such as susceptibility to local optima, scalability to large quantum systems, and sensitivity to noise in practical settings. Moreover, conventional approaches require complex pulse sequences and high computational overhead, which become increasingly impractical as the number of qubits in a quantum system increases.

Recent advances in ML and Deep learning(DL) have introduced a new paradigm for solving complex optimization problems in QC. Specifically, QNNs have gained significant attention as powerful tools for QSP. QNNs leverage the expressive power of DL techniques to optimize quantum circuits in a manner that exceeds the limitations of traditional methods [16] [17]. By learning the optimal QSP protocols through training on data, QNNs are capable of approximating complex quantum states with high fidelity while also being more robust to noise and hardware imperfections.

The application of QNNs to QSP offers several distinct advantages. First, QNNs are highly adaptable to the intricate nature of quantum systems, where the Hilbert space grows exponentially with the number of qubits. This adaptability allows QNNs to navigate a vast, high-dimensional parameter space more efficiently than traditional optimization methods. Second, QNNs are capable of handling noisy intermediate-scale quantum(NISQ) devices, where errors and decoherence limit the performance of the quantum circuits. By incorporating noise resilience into the training process, QNNs can achieve high-fidelity state preparation even in the presence of imperfections in quantum hardware [18]. Despite the promising potential of QNNs, several challenges remain in their practical application to QSP [19]. The design of neural network architectures that can effectively represent complex quantum states is nontrivial, and the optimization of such models requires novel techniques to balance expressivity and trainability. In addition, ensuring that QNNs generalize across different quantum systems and can be scaled to larger systems with many qubits remains an ongoing area of research. Furthermore, the computational cost of training these models in large quantum systems, particularly in the presence of noise, remains a significant barrier.

Building on the foundation of QNNs, this study introduces the first VQA [20] framework that explicitly integrates a QNN for high-fidelity QSP on NISQ devices. Our approach features a novel ansatz with linear parameter scaling, inspired by neural network architectures, which enhances expressivity and scalability. The primary goal of this study was to develop a robust, scalable, and theoretically rigorous methodology that overcomes the limitations of traditional QSP methods. These traditional approaches are hindered by issues such as susceptibility to local optima, poor scalability to larger systems, and inefficiency in handling noise, particularly for NISQ devices.

By addressing key the challenges in QSP, including scalability, noise resilience, and expressivity, our research provides a robust and scalable framework for high-fidelity QSP on near-term quantum devices. This innovative solution not only fills the gaps in existing methodologies but also sets the stage for transformative applications in quantum algorithms, quantum simulation, QML, and quantum communication. The main contributions of our work are as follows:

(i) We introduced a variational ansatz with layers of single-qubit rotations and bidirectional entangling layers, balancing expressivity and trainability while scaling linearly with the qubit count for efficient state preparation on NISQ devices.

(ii) This work integrates three quantum native loss functions: the trace distance, Bures distance, and Fubini–Study distance within a unified optimization framework. This makes it feasible for the optimization process to follow the natural geometry of the quantum state space and improve its convergence when quantum constraints are present.

(iii) Implementing the improved parameter shift algorithm for precise gradient calculation, our study made it possible to optimise complex variational circuits using gradients in an effective manner while lowering computational costs and improving performance.

(iv) A momentum-based optimizer with adaptive learning rate scheduling is proposed to accelerate convergence and handle barren plateaus and local minima in quantum optimization landscapes.

This section introduces the concept of QSP, its importance, and the challenges involved in achieving a high-fidelity state preparation for QC applications. The latter section reviews the existing QSP methods and identifies their limitations in terms of scalability and noise resilience. Section 3 outlines the proposed VQA framework, detailing the integration of QNN, quantum native loss functions, and enhanced gradient computation techniques. Section 4 presents experimental results and, validates the effectiveness of the proposed framework across various quantum systems. Section 5 concludes the paper with a summary of the contributions and suggestions for future research directions.

2 Related Work

Chao-Chao et al. [21] presented a stochastic prediction(SP) approach to better prepare the quantum states in silicon quantum dots. The SP algorithm uses random predictions to improve fidelity and overcome local optima, unlike conventional approaches such as gradient ascent pulse engineering(GRAPE), chopped random-basis optimization(CRAB), and the greedy algorithm(GA). This neural network has several important parameters, such as the total evolution time (4π for a single qubit and 10π for two qubits), action duration ($\pi/5$ for a single qubit and $\pi/2$ for two qubits), maximum time step (20 for both types), number of allowed actions (eight for a one-qubit and sixteen for two-qubits), and a learning rate of 0.0005 for a single qubit and 0.001 for two qubits. With a greater fidelity of 0.97 for a single qubit in noiseless situations, the SP algorithm routinely outperforms competing approaches. The resilience of the algorithm is demonstrated by the fact that fidelity drops in noisy settings but remains consistent(approximately 0.93) when trained with noise. According to the simulator preparing states, the SP delivers high fidelity with shorter pulse design durations.

Ahmed et al. [22] presented a conditional generative adversarial network(QST-CGAN) designed for the reconstruction of optical quantum states. The network is composed of a generator that reconstructs physically valid density matrices using custom layers, and a discriminator that assesses the quality of these reconstructions. Both networks employ LeakyReLU activations, utilize Adam optimization with a learning rate of 2×10^{-4}, and implement binary cross-entropy loss. The QST-CGAN demonstrates markedly improved fidelity compared to conventional maximum likelihood estimation techniques, surpassing 0.99 in noiseless scenarios and consistently maintaining values above 0.95 in the presence of moderate noise. In addition, considerably fewer iterations and measurements were required.

Melnikov et al. [23] proposed an efficient approach for preparing quantum states using Riemannian gradient algorithms to enhance tensor networks and specific quantum circuits built for hardware compatibility. They simplify quantum encoding considerably by representing input vectors with Matrix Product States(MPS) sampling analytical functions. To confront optimization challenges, such as desolate plateaus, they propose the innovative "Cut Once, Measure Twice" strategy, which divides circuits into manageable segments to facilitate effective initialization and training. The simulation results showed that the system has a very high level of preparation fidelity (¿99.9%) and polynomial runtime scaling, indicating that it can be used in QC, computational finance, ML, and simulations.

Li et al. [24] experimentally demonstrated quantum state reconstruction using generative models based on recurrent neural networks(RNNs). This method is utilized on a superconducting transmon qubit platform to prepare and measure the Greenberger-Horne-Zeilinger(GHZ) and Bell states involving up to five qubits, employing informationally complete Pauli-4 POVMs. The experimental validation indicated that the generative model-based approach surpassed the standard QST in terms of scaling efficiency. The authors presented the fidelities of the reconstructed quantum states, attaining high classical fidelities exceeding 99% and quantum fidelities of

approximately 95%, which are comparable to or slightly lower than those of traditional quantum state tomography methods. The proposed method exhibits approximately linear scaling with the number of qubits, in stark contrast to the exponential scaling characteristic of standard quantum state tomography(QST). The fidelities obtained using standard QST were 98.0% for 2-qubits, 97.9% for 3-qubits, 93.3% for 4-qubits, and 89.4% for five qubits. In contrast, the generative model demonstrated consistent performance with a markedly reduced number of measurement samples.

Carrasquilla et al. [25] proposed a scalable and efficient method for quantum state reconstruction that integrates generative ML models with tensor network structures and neural network-based probabilistic models. This method simplifies quantum state tomography into an unsupervised learning problem by utilizing models such as Restricted Boltzmann Machines(RBMs) and Recurrent Neural Networks(RNNs) to replicate measurement distributions derived from informationally complete POVMs, such as tetrahedral and Pauli-based measurements. This method effectively reconstructs noisy GHZ states and ground states of complex many-body systems, such as the 1D transverse-field Ising model and the 2D triangular lattice Heisenberg model. Classical and quantum fidelities approach unity, whereas RNN-based models exhibit linear scaling in the measurement requirements as the system size increases. This framework is conducive to experimentation, scalable to large systems, and adaptable to diverse generative architectures.

Sierra-Sosa et al. [26] used TensorFlow Quantum to investigate the effects of QSP on QML. They tested three different models, namely, QCNN, angle hybrid, and amplitude hybrid, and found that the latter produced better results. By the end of the eighth epoch, the amplitude hybrid model outperformed the angle hybrid model by as much as 8.9 percentage points in terms of accuracy, with the former exceeding 90% accuracy at centroid distances, whereas the latter fell short of this mark. Amplitude encoding is an excellent tool to make learning more consistent and unpredictable. The amplitude hybrid model began to exhibit overfitting symptoms after 50 epochs, whereas the angle hybrid model continued to exhibit instability. Although amplitude encoding improves the accuracy and early convergence, it increases the circuit complexity.

Nicolas et al. [27] presented a method that uses neural networks to prepare arbitrary quantum states in qubit-coupled quantum harmonic oscillators(QHO). The neural network was used to forecast the pulse parameters required to reach a target state, and the oscillator and qubit were both subjected to a series of square pulses as part of the suggested procedure. As the number of pulses in the sequence increases, the model's performance improves, demonstrating excellent accuracy in state preparation with 99.9% fidelity for qubit states and 97% for qutrit states. More precisely, qubits exhibited optimal outcomes when pulsed with seven pulses, whereas qutrits exhibited optimal outcomes when pulsed with eight pulses. The created states exhibited high purity according to the model, as demonstrated in, with qubit states achieving fidelities of up to 99.99%.

Krisnanda et al. [28] introduced a new method for creating and concentrating quantum resource states using QNN. This method employs linear optical mixing, incoherent pumping, coherent pumping, and fermionic nodes in a quantum reservoir network. To produce high-fidelity quantum states, including cluster, discorded, NOON, and maximally entangled states, the network was trained using a hybrid random-genetic approach that optimized the mixing coefficients. For the maximally entangled NOON, W, and cluster states, the experiments show very high success rates of more than 99%, even when there is noise such as dephasing and depolarization. This technique achieves near-perfect entanglement by combining inadequately entangled states, effectively focusing on the entanglement.

Wang et al. [29] presented a neural network-based technique for QSP in noisy situations. The method was evaluated using one- and two-qubit systems, with neural network-enhancing control pulses to mitigate dynamic and quasi-static noise. According to the experimental results, the method outperforms conventional techniques, such as ADAM optimization, achieving fidelities of 99.5% for single-qubit and 98.6% for two-qubit state preparations. Under dynamic noise, the model demonstrated 94% fidelity, and under quasistatic noise, 97% fidelity, along with shorter preparation times and faster convergence. Based on these findings, adaptive denoising can effectively handle noisy quantum systems while maintaining their efficiency and excellent quality.

Torlai et al. [30] examined the application of restricted Boltzmann machines(RBMs) for the reconstruction of quantum states from noisy experimental data produced by a Rydberg quantum simulator. The RBM was trained on bit-string data with an additional noise layer to mitigate measurement errors, thereby improving the precision of state reconstruction. Experimental findings indicate that the RBM attains fidelities of 99% in reconstructing quantum states inside an 8-atom system and fidelities exceeding 95% in a 9-atom system. The results demonstrate that the RBM model effectively manages noise and precisely reconstructs significant physical measurements, such as quantum mutual information and correlation functions. Compared with conventional techniques,

RBM markedly enhances the efficiency and accuracy of state reconstruction, demonstrating the promise of amalgamating ML with quantum simulators for resilient quantum state tomography under noisy conditions.

Arrazola et al. [31] presented automation for designing quantum circuits using ML. It employs a continuous variable QNN optimized by gradient descent using TensorFlow in the Strawberry Fields simulator, which is composed of optical gates (interferometers, displacements, squeezers, and Kerr gates). This technique yielded a wide range of high-fidelity quantum states. For state preparation, single-photon states had 99.998% fidelities, ON states 99.93%, GKP states 99.60%, random states 99.82%, and NOON states 99.90%. The networks used up to 10,000 optimization steps and 8-25 layers (40-125 gates), with cutoff dimensions ranging from 6 to 50. For the GKP state, circuit characteristics such as maximum displacement, squeezing, and Kerr strength were specifically chosen: displacement ≈ 0.259, squeezing ≈ 0.2793, and Kerr ≈ 0.2484. The circuits are short and efficient; however, the simulations are based on idealized and noiseless conditions.

Hanrui et al. [32] introduced RobustState, an innovative method that improves the fidelity of quantum state preparation on noisy quantum devices. The RobustState was assessed across several quantum state preparation tasks, including arbitrary quantum states, partial differential equation(PDE) states, QML states, and quantum error correction(QEC) codewords. The experimental validation encompassed ten unique quantum machines that exhibited substantial performance improvements relative to the baseline methods. Experimental findings demonstrate substantial fidelity enhancements, with a 50% increase for 4-qubit states and a 72% increase for 5-qubit states, in addition to a reduction in coherence errors by as much as 7.1 times relative to current methodologies.

Sanjib et al. [33] presented QRST, which is a new approach for quantum state reconstruction using quantum reservoir networks. QRST streamlines conventional quantum state tomography by eliminating the requirement for numerous measurement bases and instead uses a quantum reservoir to receive input quantum states and measure occupation numbers at a single location. This method works for quantum systems with continuous variables as well as those with finite dimensions. By training the system to perform best, the approach provides scalable state reconstruction with excellent fidelity, even in noisy environments.

He et al. [34] presented a Revised Greedy(RG) algorithm for universal quantum state preparation, aimed at mitigating the shortcomings of the standard greedy(SG) algorithm by resolving its propensity to become ensnared at local maxima. This technique facilitates the generation of arbitrary quantum states by dynamic control pulses, implemented in 1-qubit and 2-qubit systems, encompassing semiconductor quantum dots(DQDs) and superconducting circuits. The RG algorithm eliminates the need for significant training, in contrast to ML techniques, and offers superior fidelity and efficiency relative to conventional optimization methods such as GRAPE and CRAB. The results demonstrate that RG succeeds in achieving fidelities as high as 0.999 for single-qubit states in superconducting circuits and 0.977 in DQDs. In both systems, the preparation of two-qubit states led to fidelities exceeding 0.9.

Haimeng et al. [35] analyzed the sample complexity of neural network approaches for mixed-state reconstruction using Neural Quantum States(NQS). This study compared two NQS encodings: the Neural Density Operator(NDO) and Positive Operator-Valued Measurement NQS(POVM-NQS). As the state becomes more mixed, NDO loses some of its edges, but still works better for almost pure states. Conversely, POVM-NQS demonstrates uniform scaling, similar to standard shadow techniques, irrespective of the mixedness of the state. Additionally, the authors presented a control variate technique that lowers training noise and increases reconstruction efficiency and accuracy.

Sanjib et al. [36] presented a quantum neuromorphic platform for QSP that employs a quantum reservoir to transform conventional optical inputs into specific quantum states, including single-photon, Schrödinger's cat states, and two-mode entangled states. The system uses a 2D nonlinear quantum reservoir lattice, which is stimulated by classical fields, and processes the resultant quantum field using a linear optical configuration. The output modes were optimized to align with the target quantum states. The findings indicate that QRSP can effectively generate diverse quantum states, with single-photon levels displaying pronounced antibunching, cat states achieving enhanced fidelity as the reservoir size increases, and entangled states exhibiting greater entanglement with larger reservoir sizes.

3 Methodology

The preparation of arbitrary quantum states within controlled quantum systems is a fundamental problem in quantum information science. This preparation enables applications in quantum

simulations, QML, and algorithm initialization. VQAs have emerged as a practical approach for approximating target quantum states using parameterized quantum circuits by leveraging hybrid quantum-classical optimization. This section provides a comprehensive theoretical and mathematical treatment of the variational framework, detailing the ansatz construction, formulation of the quantum native cost function, evaluation of the gradient, and optimization strategy. The workflow methodology of the proposed approach is illustrated in Figure 1.

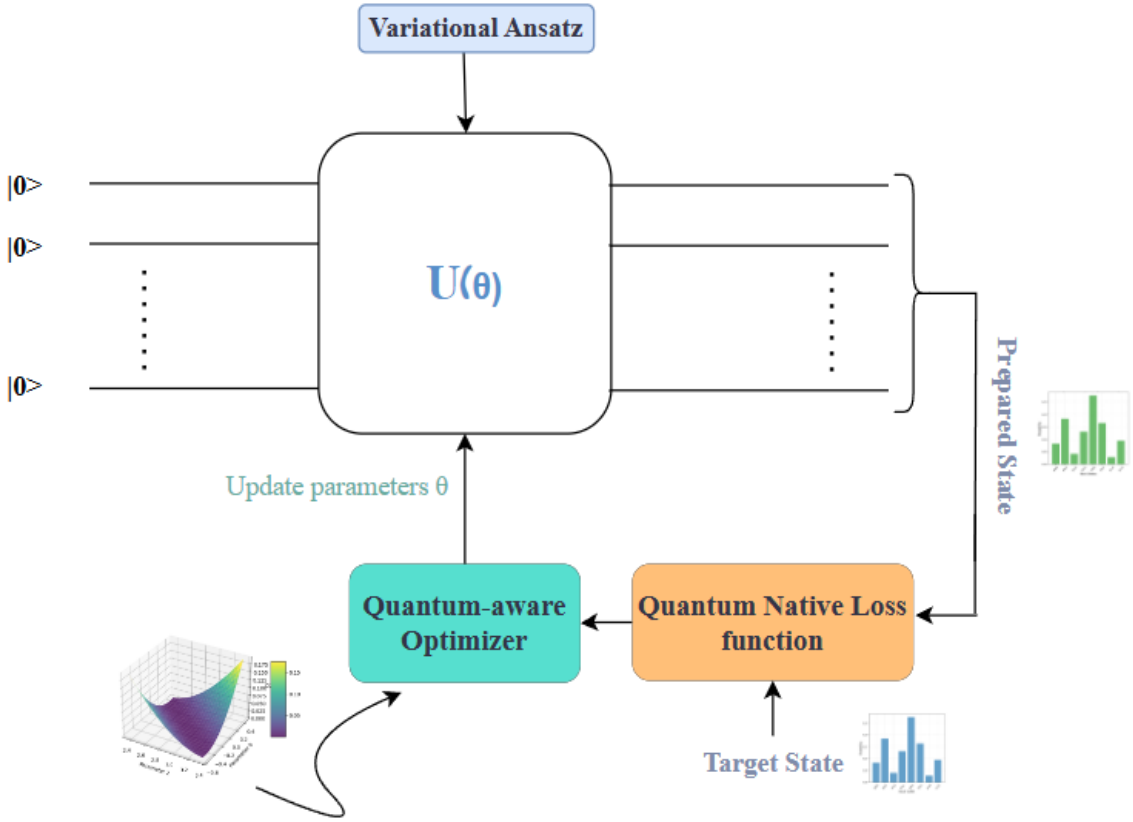

Figure 1: Workflow methodology.

3.1 ***Quantum System and Parameterized State Preparation***

Let $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ denote the Hilbert space of an n-qubit quantum system, where each qubit is associated with a two-dimensional complex vector space $\mathbb{C}^2$. The dimension of this composite Hilbert space grows exponentially with the number of qubits, specifically,

$$\dim(\mathcal{H}) = d = 2^n \tag{1}$$

The computational basis $\{|i\rangle\}_{i=0}^{d-1}$, where each basis vector corresponds to an n-bit binary string $i = i_1 i_2 \dots i_n$ with $i_j \in \{0, 1\}$, forms an orthonormal basis spanning $\mathcal{H}$. These basis states satisfy the following conditions:

$$\langle i|j\rangle = \delta_{ij}, \quad \forall i, j \in \{0, \dots, d-1\} \tag{2}$$

where δ_{ij} is the Kronecker delta [4].

An arbitrary pure quantum state $|\psi\rangle \in \mathcal{H}$ can be expressed as a complex linear combination of these basis vectors:

$$|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle \tag{3}$$

where the complex amplitudes $\alpha_i \in \mathbb{C}$ satisfy the normalization condition as follows:

$$\sum_{i=0}^{d-1} |\alpha_i|^2 = 1. \tag{4}$$

The global phase of $|\psi\rangle$ is physically irrelevant, and the state is fully characterized by set $\{\alpha_i\}$.

Our primary goal is to design a parameterized family of unitary transformations $U(\boldsymbol{\theta}) : \mathcal{H} \to \mathcal{H}$, where $\boldsymbol{\theta} \in \mathbb{R}^M$ represents the vector of the real values of the tunable parameters, such that

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|0\rangle^{\otimes n} \tag{5}$$

approximates a predetermined target quantum state $|\psi_{\text{target}}\rangle$ with high fidelity.

The unitary $U(\boldsymbol{\theta})$ corresponds to a variational quantum circuit, or ansatz, composed of a sequence of parameterized quantum gates. The dimension M of the parameter space depends on the ansatz structure and depth, which determines the expressivity of the circuit in the approximation of complex states.

This framework is fundamental for variational quantum algorithms, where $\boldsymbol{\theta}$ is optimized using classical procedures to minimize a suitable cost function that quantifies the distance to $|\psi_{\text{target}}\rangle$.

3.2 ***Variational Ansatz Design***

We consider a parameterized quantum circuit $U(\boldsymbol{\theta})$ acting on the Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ of an n-qubit system, where the dimension of $\mathcal{H}$ is $d = 2^n$. The goal is to prepare a variational quantum state $|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|0\rangle^{\otimes n}$ that approximates a target quantum state $|\psi_{\text{target}}\rangle$ with high fidelity. Here, $\boldsymbol{\theta} \in \mathbb{R}^M$ denotes the vector of variational parameters to be optimized.

The ansatz structure was designed to balance the expressivity and trainability by alternating layers of single-qubit rotations and entangling operations. Specifically, the circuit consists of three layers of single-qubit rotations, each parameterized by rotations about the three Pauli axes X, Y, and Z. These are interleaved with two entangling layers composed of controlled-NOT (CNOT) gates arranged to maximize entanglement and connectivity among the qubits. Figures 2, 3, 4, and 5 illustrate the diverse ansatz structures with varying qubit counts employed in this study.

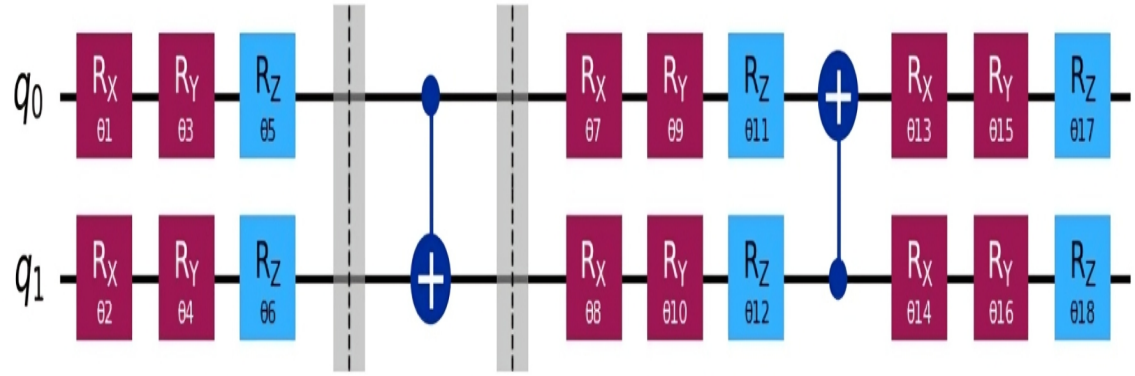

Figure 2: Random 2-qubit ansatz architecture.

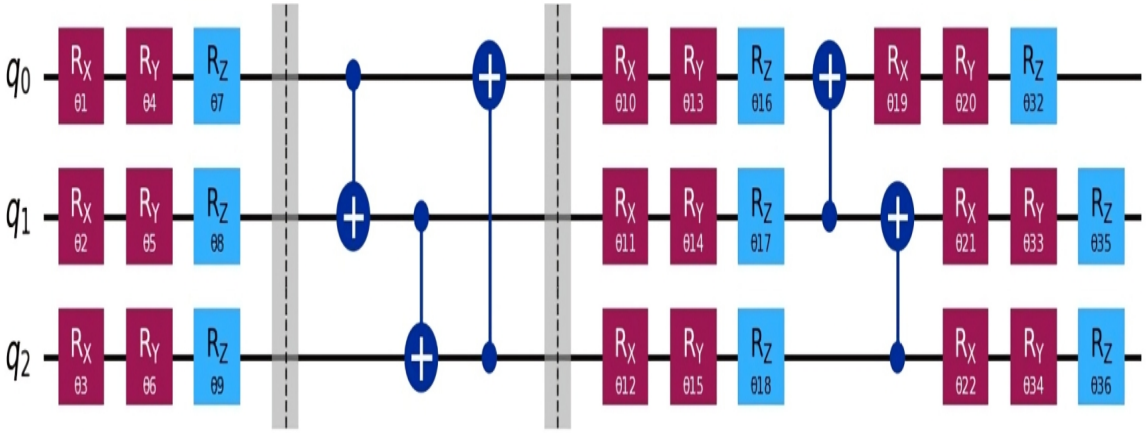

Figure 3: Random 3-qubit ansatz architecture.

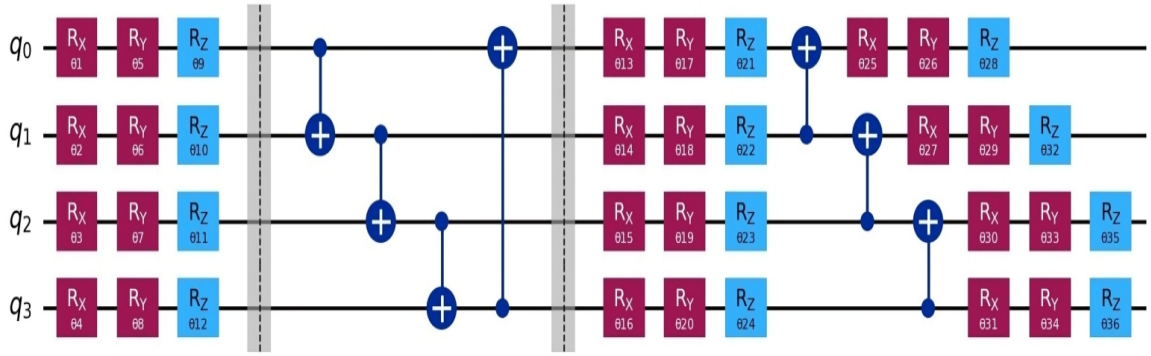

Figure 4: Random 4-qubit ansatz architecture.

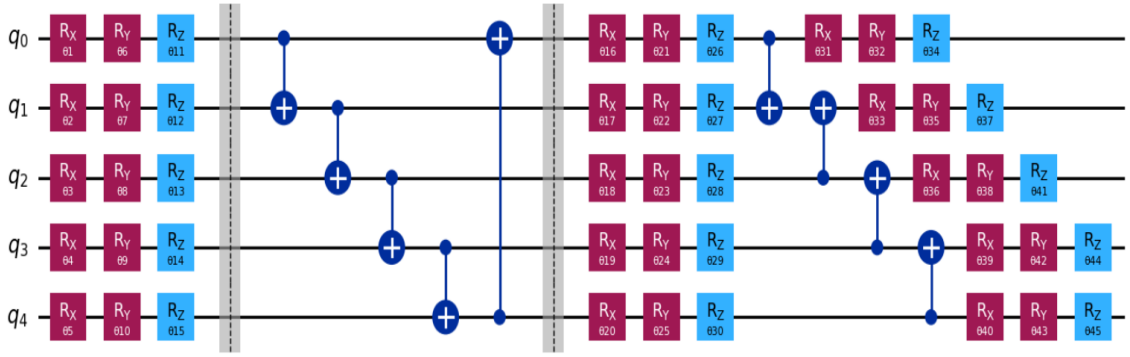

Figure 5: Random 5-qubit ansatz architecture.

Mathematically, the unitary transformation implemented by the ansatz can be expressed as

$$
\begin{aligned}
U(\boldsymbol{\theta}) = & \left(\bigotimes_{i=0}^{n-1} R_X(\theta_{3i}) R_Y(\theta_{3i+1}) R_Z(\theta_{3i+2}) \right) E_1 \\
& \left(\bigotimes_{i=0}^{n-1} R_X(\theta_{3n+3i}) R_Y(\theta_{3n+3i+1}) R_Z(\theta_{3n+3i+2}) \right) E_2 \\
& \left(\bigotimes_{i=0}^{n-1} R_X(\theta_{6n+3i}) R_Y(\theta_{6n+3i+1}) R_Z(\theta_{6n+3i+2}) \right)
\end{aligned} \tag{6}
$$

where each single-qubit rotation is defined as:

$$
R_\mu(\phi) = \exp\left(-i \frac{\phi}{2} \sigma_\mu \right), \quad \mu \in \{X, Y, Z\}
$$

where σ_μ denotes the Pauli operator. The full parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{9n}$ comprises the rotation angles applied to each qubit in each layer.

The entangling layers E_1 and E_2 are composed of CNOT gates arranged to improve qubit connectivity and entanglement generation. The first entangling layer E_1 consists of a linear chain of nearest-neighbor CNOT gates along with an additional CNOT connecting the last and first qubits, forming a circular entanglement pattern as follows:

$$
E_1 = \prod_{j=0}^{n-2} \mathrm{CNOT}_{j \to j+1} \cdot \mathrm{CNOT}_{n-1 \to 0} \tag{7}
$$

The second entangling layer E_2 reverses the direction of the CNOT gates to

$$
E_2 = \prod_{j=0}^{n-2} \mathrm{CNOT}_{j+1 \to j} \tag{8}
$$

further enriches the circuit entanglement structure.

This architectural choice is motivated by the need to generate a complex, multipartite entanglement, which is essential for accurately representing arbitrary quantum states. The rotations about all three Pauli axes guarantee that any single-qubit unitary can be approximated through successive layers by leveraging the Euler decomposition theorem. The layered and bidirectional CNOT configuration fosters a rich entanglement topology, allowing the ansatz to explore a wide subspace of the full unitary group on n qubits.

Importantly, the parameter count scales linearly with the number of qubits, $M = 9n$, offering a practical balance between expressive power and resource requirements, making it suitable for implementation in near-term quantum devices.

Empirically, this ansatz demonstrates high fidelity in the preparation of complex random quantum states, confirming its suitability and flexibility for variational quantum state engineering tasks.

3.3 ***Quantum Native Loss Functions***

The optimization of variational quantum circuits for state preparation necessitates the careful selection of loss functions that effectively capture the quantum geometric structure of the underlying Hilbert space. Traditional classical optimization approaches often fail to account for the unique properties of quantum states, including their normalization constraints, phase invariance, and the non-Euclidean geometry of the quantum state space. In this study, we implemented and analyzed four distinct quantum-native loss functions, each designed to exploit different geometric properties of quantum state spaces for enhanced convergence and optimization performance.

3.3.1 ***Trace Distance-Based Loss Function:*** The trace distance provides an alternative quantum-geometric measure that captures the distinguishability between quantum states [4]. For pure states, the trace distance can be expressed in terms of fidelity as follows:

$$D_{\text{tr}}(\boldsymbol{\theta}) = \sqrt{1 - F(\boldsymbol{\theta})} = \sqrt{1 - |\langle\psi_{\text{target}}|\psi_{\text{prepared}}(\boldsymbol{\theta})\rangle|^2} \tag{9}$$

The trace distance-based loss function is defined as:

$$L_{\text{trace}}(\boldsymbol{\theta}) = D_{\text{tr}}(\boldsymbol{\theta}) = \sqrt{1 - F(\boldsymbol{\theta})} \tag{10}$$

This formulation offers different optimization characteristics compared to the fidelity-based approach. The square root transformation amplifies small differences near the optimal solution, providing enhanced sensitivity in regions where the fidelity approaches unity. The improved gradient properties arise from the derivative structure, which differs from the fidelity loss, potentially avoiding certain local minima that plague standard fidelity optimization. Furthermore, this loss function maintains metric consistency by directly corresponding to the trace distance metric in the space of quantum states, thereby ensuring that the optimization procedure respects the natural geometric structure of the quantum state manifold.

3.3.2 ***Bures Distance-Based Loss Function:*** The Bures distance represents a Riemannian metric in the space of quantum states, providing a geometrically natural measure for quantum state comparison. For pure quantum states, the Bures distance is given by:

$$d_{\text{Bures}}(\boldsymbol{\theta}) = \sqrt{2\left(1 - \sqrt{F(\boldsymbol{\theta})}\right)} = \sqrt{2\left(1 - \sqrt{|\langle\psi_{\text{target}}|\psi_{\text{prepared}}(\boldsymbol{\theta})\rangle|^2}\right)} \tag{11}$$

The corresponding loss function is formulated as follows:

$$L_{\text{Bures}}(\boldsymbol{\theta}) = 2\left(1 - \sqrt{F(\boldsymbol{\theta})}\right) = 2\left(1 - \sqrt{|\langle\psi_{\text{target}}|\psi_{\text{prepared}}(\boldsymbol{\theta})\rangle|^2}\right) \tag{12}$$

The distance-based loss function of bures exhibits several distinctive properties that make them particularly suitable for quantum optimization. The Riemannian structure respects the natural geometry of quantum state manifolds, ensuring that the optimization path follows geodesics in the appropriate metric space. The scale invariance of the loss function provides balanced sensitivity across the optimization landscape, avoiding the potential issues of oversensitivity near the optimum or undersensitivity in intermediate regions. Additionally, this distance measure is related to the quantum Fisher information metric, which governs the fundamental bounds in quantum parameter estimation theory, thereby connecting the optimization process to information-theoretic principles.

3.3.3 ***Fubini Study Distance Based Loss Function:*** The Fubini-Study metric represents the natural Riemannian metric on complex projective spaces, making it particularly suitable for quantum state optimization where global phases are irrelevant. The Fubini-Study distance between quantum states is defined as:

$$d_{\mathrm{FS}}(\boldsymbol{\theta}) = \arccos(|\langle\psi_{\mathrm{target}}|\psi_{\mathrm{prepared}}(\boldsymbol{\theta})\rangle|) \tag{13}$$

The corresponding loss function is the following:

$$L_{\mathrm{FS}}(\boldsymbol{\theta}) = \arccos^2(|\langle\psi_{\mathrm{target}}|\psi_{\mathrm{prepared}}(\boldsymbol{\theta})\rangle|) \tag{14}$$

This loss function incorporates several mathematically elegant properties that distinguish it from other native quantum approaches. Projective geometry naturally accounts for the projective structure of quantum state spaces, recognizing that quantum states are elements of the complex projective Hilbert space rather than the Hilbert space itself. Phase invariance explicitly ignores the global phase differences between states, focusing on the optimization of physically significant differences. For single-qubit systems, this loss function represents the squared geodesic distance on the Bloch sphere, providing a direct geometric interpretation of the optimization landscape.

3.4 ***Gradient Estimation via Parameter-Shift Rule***

In variational quantum algorithms, efficient and hardware-compatible gradient evaluation is crucial for optimizing parameterized quantum circuits. To this end, we employed the parameter shift rule, a native analytic gradient estimation method for QC that yields exact derivatives for a wide range of gates, including all those utilized in our ansatz [37].

For a cost function $\mathcal{L}(\boldsymbol{\theta})$ that depends on a parameterized unitary transformation $U(\boldsymbol{\theta})$, the derivative with respect to an individual parameter θ_k is given by:

$$\frac{\partial\mathcal{L}}{\partial\theta_k} = \frac{1}{2}\left[\mathcal{L}(\boldsymbol{\theta}_+^{(k)}) - \mathcal{L}(\boldsymbol{\theta}_-^{(k)})\right] \tag{15}$$

where $\boldsymbol{\theta}_\pm^{(k)} = (\theta_1, \ldots, \theta_k \pm \frac{\pi}{2}, \ldots, \theta_M)$ denotes the parameter vector with a shift applied to the k-th component.

This formula holds exactly for gates generated by Pauli operators, such as R_X, R_Y, and R_Z, which are employed in our ansatz. The parameter-shift rule eliminates the need for numerical finite differences or complex backpropagation, thereby offering both theoretical rigor and practical compatibility with quantum hardware. Unlike classical automatic differentiation techniques, this approach provides exact gradients without approximation errors, making it particularly valuable for quantum optimization, for which precision is paramount.

The theoretical foundation of the parameter-shift rule stems from the eigenvalue structure of Pauli generators. For a parameterized gate $U(\theta_k) = e^{-i\theta_k P_k/2}$ where P_k is a Pauli operator with eigenvalues ±1, the expectation value of any observable exhibits a sinusoidal dependence on θ_k. This periodic structure enables the exact computation of derivatives through strategic parameter changes, thereby avoiding the accumulation of numerical errors inherent in finite-difference approximations [38].

We implemented an enhanced version of the parameter-shift rule to improve the computational efficiency and scalability. Our implementation computes gradients for all parameters through a systematic evaluation of the cost function at strategically shifted parameter point $2M$, where M represents the total number of variational parameters. This batched approach enables efficient parallelization and reduces the total number of quantum circuit evaluations required compared to sequential parameter-by-parameter gradient estimation. Computational complexity scales linearly with the number of parameters, making this approach feasible for large-scale variational quantum circuits.

The gradient computation procedure involves the creation of parameter vectors with systematic shifts applied to the individual components. For each parameter θ_k, we constructed two shifted versions: $\boldsymbol{\theta}_+^{(k)}$ with θ_k increased by $\pi/2$, and $\boldsymbol{\theta}_-^{(k)}$ with θ_k decreased by $\pi/2$. The partial derivative is then computed as the scaled difference between the cost function evaluations at these shifted points. This process is repeated for all parameters to construct the complete gradient vector $\nabla\mathcal{L}(\boldsymbol{\theta})$.

3.5 ***Quantum-Aware Optimization Strategy***

To optimize the variational parameter $\boldsymbol{\theta}$, we introduce a custom quantum-aware optimizer that integrates momentum-based updates with adaptive learning rate scheduling. This method is

specifically designed to navigate the complex, nonconvex landscapes characteristic of quantum cost functions, which often exhibit multiple local minima, saddle points, and regions of vanishing gradients known as barren plateaus [39].

The optimization strategy builds upon the classical momentum-based gradient descent while incorporating quantum-specific adaptations [20]. Let $\eta(t)$ denote the learning rate in iteration t, and $\mu \in [0, 1)$ the momentum coefficient. The update rule for the parameter vector $\boldsymbol{\theta}(t)$ follows a two-step process:

$$\mathbf{v}(t+1) = \mu\mathbf{v}(t) - \eta(t)\nabla\mathcal{L}(\boldsymbol{\theta}(t)), \quad \boldsymbol{\theta}(t+1) \quad = \boldsymbol{\theta}(t) + \mathbf{v}(t+1) \tag{16}$$

where $\mathbf{v}(t)$ represents the velocity vector that accumulates momentum from the previous gradient evaluations. The momentum term $\mu\mathbf{v}(t)$ enables the optimizer to maintain direction through regions of inconsistent gradient information, helping to escape shallow local minima and navigate saddle points that frequently appear in quantum optimization landscapes.

The adaptive learning rate mechanism is a crucial component of our quantum-aware approach [40]. Traditional fixed-learning-rate schedules often prove inadequate for quantum optimization, where the appropriate step size can vary dramatically across different regions of the parameter space. Our adaptive scheme adjusts the learning rate based on the immediate progress in cost minimization:

$$\eta(t+1) = \begin{cases} 1.05 \cdot \eta(t), & \text{if } \mathcal{L}(t+1) < \mathcal{L}(t), \\ 0.95 \cdot \eta(t), & \text{otherwise.} \end{cases} \tag{17}$$

This adaptive mechanism promotes rapid learning when the cost function shows a consistent decrease, indicating that the current optimization direction is productive. By contrast, when the cost stagnates or increases, the learning rate is conservatively reduced to prevent overshooting and maintain stability. Multiplicative factors (1.05 and 0.95) were chosen to provide gradual adaptation while avoiding oscillatory behavior that could destabilize the optimization process.

The learning rate adaptation strategy addresses several challenges specific to quantum optimization. In regions where gradients are large and the cost function behaves well, a higher learning rate accelerates the convergence. In contrast, near local minima or in regions with small gradients, the reduced learning rate prevents the optimizer from making erratic updates that could degrade performance. This dynamic adjustment is particularly valuable when dealing with the multiscale nature of quantum cost landscapes, where different parameters may require different learning rates for optimal convergence.

To ensure numerical stability and prevent divergence, we implemented bounds for adapting of the learning rate. The learning rate was constrained to remain within a specified range $[\eta_{\min}, \eta_{\max}]$, where $\eta_{\min} = 0.1\eta_0$ and $\eta_{\max} = 5\eta_0$ for the initial learning rate η_0. This constraint prevents the adaptive mechanism from reducing the learning rate to ineffectively small values or from increasing it to potentially unstable levels.

The quantum-aware optimizer incorporates additional features to enhance robustness and performance. Early stopping criteria based on fidelity thresholds prevent unnecessary computations when the quality of preparation of the target state is achieved. Specifically, the optimization terminates when the fidelity exceeds 0.9999, which corresponds to a nearly perfect state preparation. This criterion is particularly relevant for QSP tasks, where achieving exact fidelity is the primary objective.

The combination of parameter shift gradients and adaptive momentum-based optimization offers a resilient and scalable framework for training variational quantum circuits. This approach is particularly effective in hardware-aware simulation environments and NISQ applications, where the combination of exact gradient computation and adaptive learning provides both theoretical rigor and practical performance [41]. Empirical validation demonstrated that this quantum-aware optimizer consistently outperformed the standard gradient descent methods and fixed learning rate approaches, achieving faster convergence and improved final solution quality across diverse QSP tasks.

4 Result Analysis

This section presents the results obtained from the QML model evaluated using various distance metrics, including trace distance, Fubini-Study distance, and Bures distance. The results highlight the trade-offs between the qubit count, training time, fidelity, and the number of parameters required for optimization. In particular, we examined the impact of increasing the qubit count on the model's performance and optimization efficiency.

No. of Qubit	**No. of Parameter**	**Training Time(s)**	**No. of Iteration**	**Fidelity**
2	18	14.40	73	**0.999933**
3	27	28.45	72	0.999902
4	36	135.59	200	0.997691
5	45	211.09	200	**0.937918**

Table 1: Performance evaluation of the QNN model with different qubit counts using trace distance.

Table 1 illustrates the efficiency with which the QNN model operates with different numbers of qubits using the trace distance. With an increase in the qubit count from two to five, there was a substantial increase in both the number of parameters and the training duration. The training time increases from 14.4s for the two qubits to 211.09s for 5-qubits. The number of repetitions remained the same for four and five qubits, whereas smaller systems converged more quickly. The fidelity exhibited a slight decrease as the number of qubits increased, transitioning from 0.999933 for 2-qubits to 0.937918 for 5-qubits, thereby illustrating a trade-off between model complexity and performance.

No. of Qubit	**No. of Parameter**	**Training Time(s)**	**No. of Iteration**	**Fidelity**
2	18	14.46	73	**0.999922**
3	27	29.03	71	0.999896
4	36	135.14	200	0.999761
5	45	199.54	200	**0.921357**

Table 2: Performance evaluation of the QNN model with different qubit counts using Fubini Study distance.

Table 2 analyzes the performance of the QNN model with different numbers of qubits using the Fubini Study Distance. From two to five qubits, the number of parameters and training times increased dramatically. The computational demand for larger quantum systems is increasing training time from 14.46s for 2-qubits to 199.54s for 5-qubits. The training iterations stabilized at 200 for four and five qubits, but dropped to 73 and 71 iterations, respectively, for two and five qubits. As the system size increases, the model loses accuracy, with fidelity dropping to 0.921357 for 5-qubits from 0.999922 for 2-qubits.

No. of Qubit	**No. of Parameter**	**Training Time(s)**	**No. of Iteration**	**Fidelity**
2	18	19.29	73	**0.999940**
3	27	27.76	70	0.999930
4	36	131.49	200	0.999261
5	45	204.02	200	**0.942159**

Table 3: Performance evaluation of the QNN model with different qubit counts using Bures distance.

Table 3 displays the performance of the QNN model with different qubit counts, assessed by the Bures distance. As the number of qubits increases from two to five, there is a substantial increase in both the number of parameters and the duration of training. The training duration escalates from 19.29s for 2-qubits to 204.02s for 5-qubits, indicating the increased computational expense associated with additional qubits. The number of training iterations decreases for smaller qubit systems, with 73 iterations for 2-qubits and 70 for three qubits, whereas it stabilizes at 200 for both four and five qubits. The fidelity is exceptionally high for smaller qubit systems, measuring 0.999940 for 2-qubits and 0.999930 for 3-qubits. However, there is a significant reduction in the fidelity to 0.942159 for the 5-qubits, signifying a decrease in the accuracy of the model as the system becomes increasingly complicated.

As shown in Tables 1, 2, and 3, there is a clear trend that as the number of qubits increases, both the number of parameters and the training time also increase. Although the number of iterations required for convergence remains relatively stable, the fidelity of the model decreases as the qubit count increases. This highlights the inherent complexity of scaling quantum systems and the corresponding challenges in maintaining a high fidelity in more complex models.

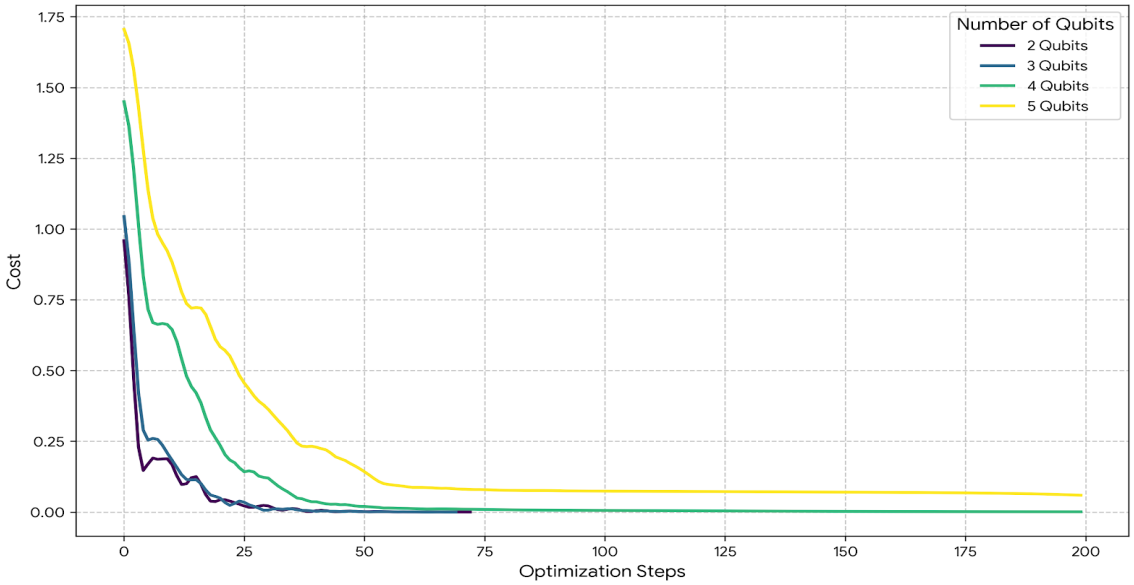

Figure 6: Cost graph for Bures distance.

The plot in Figure 6 shows that the cost function decreases as the optimization process progresses for various qubit systems. The rate of convergence becomes slower as the number of qubits increases, with the 2-qubit system achieving the quickest convergence, and the 5-qubit system achieving the slowest convergence.

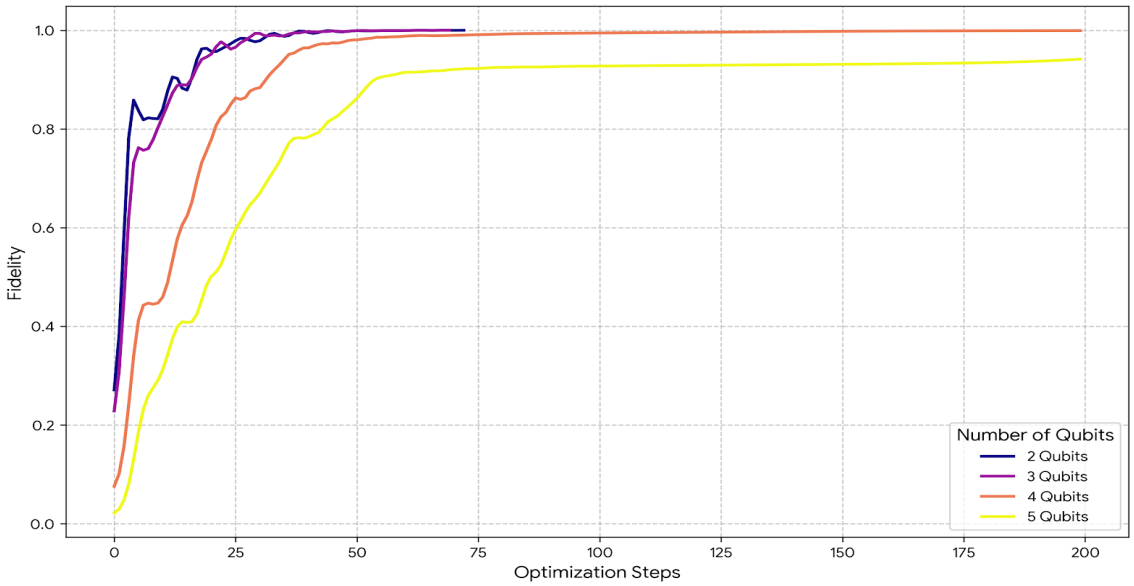

Figure 7: Fidelity graph for Bures distance.

Figure 7 shows the evolution of fidelity in optimization steps for quantum systems with different qubit counts. The 2-qubit system reaches values near 1.0 in the first few optimization stages and quickly converges to high fidelity. With a 5-qubit system that exhibits the slowest approach to high fidelity, the convergence becomes slower with increasing qubit number.

In this study, the performance of the QNN model was evaluated using several distance metrics, namely the trace distance, Fubini-Study distance, and Bures distance. Among these, the Bures

distance demonstrated superior performance in terms of fidelity and optimization efficiency, particularly as the number of qubits increased. Therefore, the subsequent analysis focuses solely on the results obtained using Bures distance. The fidelity behavior for different qubit systems under the Bures distance is shown in the following figures: 8, 9, 10, and 11. These graphs illustrate the evolution of fidelity across various optimization stages for quantum systems with 2, 3, 4, and 5 qubits.

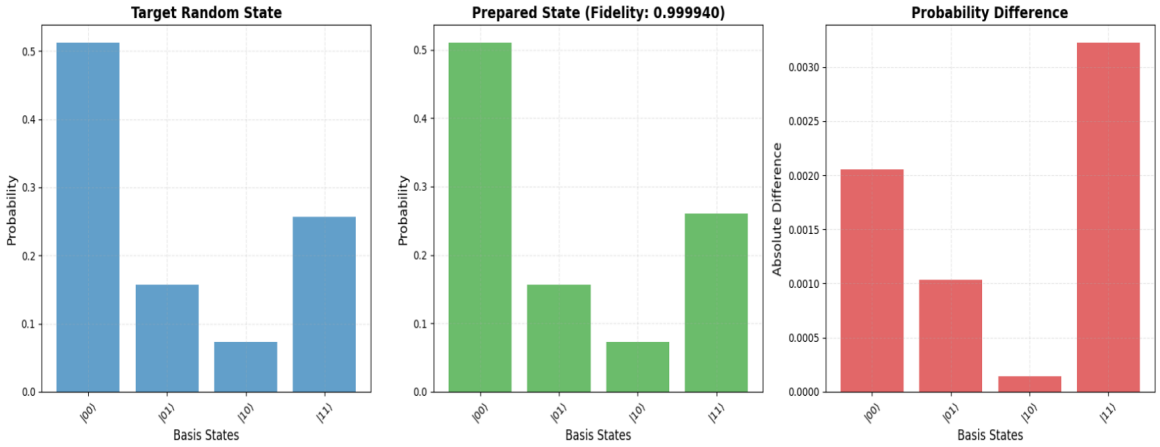

Figure 8: Fidelity graph for Bures distance for a 2-qubit system.

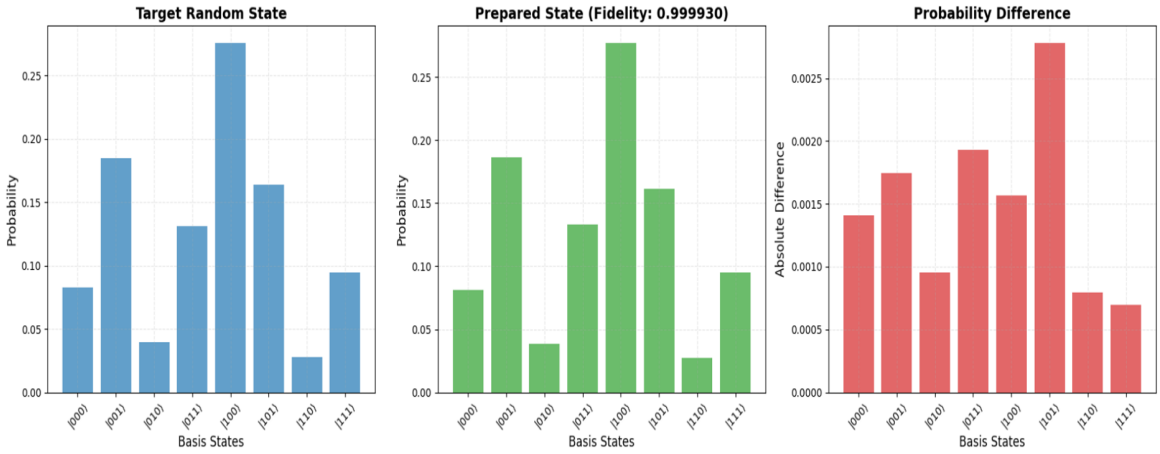

Figure 9: Fidelity graph for Bures distance for a 3-qubit system.

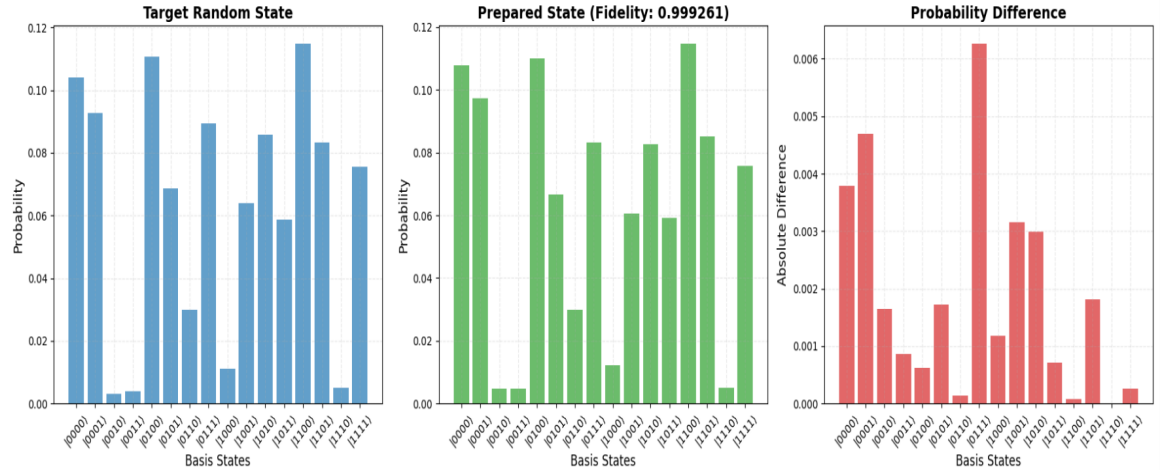

Figure 10: Fidelity graph for Bures distance for a 4-qubit system.

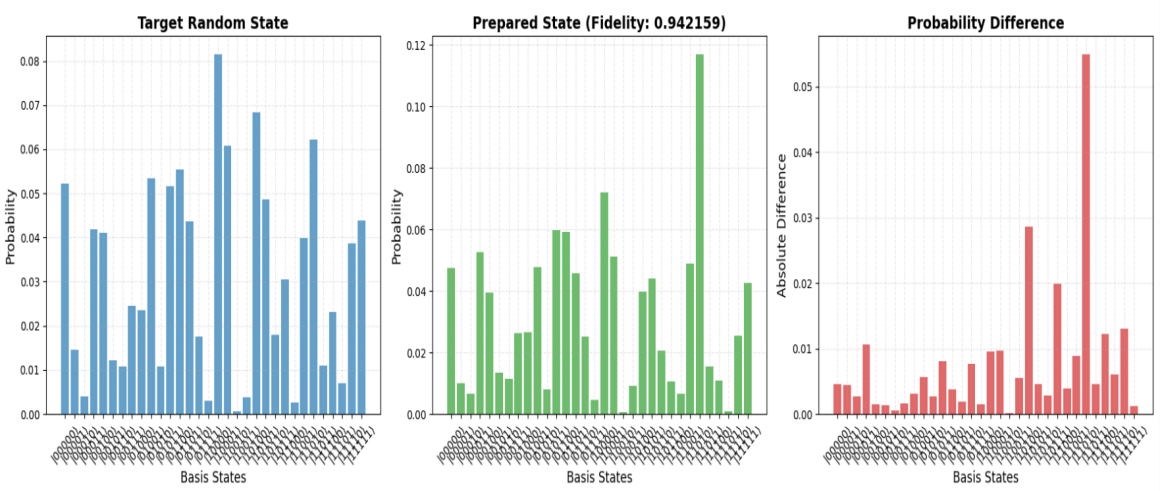

Figure 11: Fidelity graph for Bures distance for a 5-qubit system.

The probability distributions for randomly generated quantum states with 2 two 5-qubits are shown in Figure 8-11. Figure 8 indicates that the 2-qubit state has an extremely high fidelity of 0.999940, which means that there are almost no variations between the target and the prepared states and that they are almost identical. In Figure 9, the 3-qubit state, where the prepared state is similar to the target state, has a fidelity of 0.999930, and no significant differences are visible. With relatively few changes in some of the basis states, the 4-qubit state in Figure 10 attains a fidelity of, indicating a very precise preparation. Figure 11 shows a 5-qubit state with a fidelity of 0.942159. Greater differences in particular basis states indicate areas in which state preparation should be strengthened. In general, the data show that state preparation is of high quality, with a few exceptions in higher-qubit states.

5 Conclusion

This paper introduces a QSP framework that leverages a QNN in a VQA architecture to prepare high-fidelity quantum states for NISQ devices. Our novel variational ansatz with linear parameter scaling balances circuit expressivity with trainability, making it scalable for large quantum systems and high-performance on NISQ technology. To improve the convergence speed of the optimization processes and overall robustness, quantum native loss functions such as trace distance, Bures distance, and Fubini-Study distance are used. These functions optimize the preparation of quantum states by considering the intrinsic geometry of quantum state spaces. In addition, the enhanced parameter shift rule for precise gradient computation and the quantum-aware adaptive optimizer with momentum-based updates reduce the processing overhead in QSP problems' complicated and nonconvex optimization landscape tasks. Among the three distance metrics evaluated, the Bures distance had the best fidelity and optimization efficiency. Experimental results show that training time grows with qubit count, from 19.29s for 2-qubits to 204.02s for 5-qubits. The fidelity of 0.999940 for 2-qubits, 0.999930 for 3-qubits, and 0.999261 for 4-qubits shows that smaller qubit systems still have good quantum state preparation. The fidelity lessens with increasing qubits, reaching 0.942159 for 5-qubits. Because of its superior performance, noise control capabilities, and scalability to larger systems, Bures distance has been proposed as the best cost function for our quantum state preparation technique.

Although the framework has outstanding potential, it must be improved in various areas. To improve the optimization and state preparation, we repeated this experiment using real quantum hardware. In addition, to reduce quantum noise and increase short-term device performance, we propose the use of quantum error correction approaches, including surface coding and topological error mitigation. Another approach involves using quantum feature maps to enhance expressivity, particularly when generating complex quantum states in fields such as quantum chemistry and drug development. Additionally, our method can be combined with quantum optimization techniques such as QAOA to enhance state preparation.

References

[1] C. Easttom, "Quantum computing and cryptography," in *Modern Cryptography: Applied Mathematics for Encryption and Information Security.* Springer, 2022, pp. 397–407.

[2] V. Hassija, V. Chamola, A. Goyal, S. S. Kanhere, and N. Guizani, "Forthcoming applications of quantum computing: peeking into the future," *IET Quantum Communication*, vol. 1, no. 2, pp. 35–41, 2020.

[3] A. O. Pittenger, *An introduction to quantum computing algorithms*. Springer Science & Business Media, 2012, vol. 19.

[4] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge university press, 2010.

[5] A. W. Harrow and A. Montanaro, "Quantum computational supremacy," *Nature*, vol. 549, no. 7671, pp. 203–209, 2017.

[6] L. Marchetti, R. Nifosì, P. L. Martelli, E. Da Pozzo, V. Cappello, F. Banterle, M. L. Trincavelli, C. Martini, and M. D'Elia, "Quantum computing algorithms: getting closer to critical problems in computational biology," *Briefings in Bioinformatics*, vol. 23, no. 6, p. bbac437, 2022.

[7] A. K. Fedorov and M. S. Gelfand, "Towards practical applications in quantum computational biology," *Nature Computational Science*, vol. 1, no. 2, pp. 114–119, 2021.

[8] H.-P. Cheng, E. Deumens, J. K. Freericks, C. Li, and B. A. Sanders, "Application of quantum computing to biochemical systems: a look to the future," *Frontiers in chemistry*, vol. 8, p. 587143, 2020.

[9] M. Motta and J. E. Rice, "Emerging quantum computing algorithms for quantum chemistry," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 12, no. 3, p. e1580, 2022.

[10] A. Pyrkov, A. Aliper, D. Bezrukov, Y.-C. Lin, D. Polykovskiy, P. Kamya, F. Ren, and A. Zhavoronkov, "Quantum computing for near-term applications in generative chemistry and drug discovery," *Drug Discovery Today*, vol. 28, no. 8, p. 103675, 2023.

[11] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.

[12] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, "Remote preparation of quantum states," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 56–74, 2005.

[13] N. Peters, J. Barreiro, M. Goggin, T.-C. Wei, and P. Kwiat, "Remote state preparation: arbitrary remote control of photon polarizations for quantum communication," in *Quantum Communications and Quantum Imaging III*, vol. 5893. SPIE, 2005, pp. 52–61.

[14] D. Wu, S. Chen, P. Cui, J. Ma, W. Chen, J. Zhang, Y. Wang, J. Li, and J. An, "Quantum state preparation for quantum key distribution using plc module," *IEEE Photonics Journal*, vol. 15, no. 6, pp. 1–8, 2023.

[15] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, "Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms," *Journal of magnetic resonance*, vol. 172, no. 2, pp. 296–305, 2005.

[16] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," *Quantum science and technology*, vol. 4, no. 4, p. 043001, 2019.

[17] Y. Du, M.-H. Hsieh, T. Liu, S. You, and D. Tao, "Learnability of quantum neural networks," *PRX quantum*, vol. 2, no. 4, p. 040337, 2021.

[18] S. Sajadimanesh, H. A. Rad, J. P. L. Faye, and E. Atoofian, "Nr-qnn: Noise-resilient quantum neural network," *IEEE Access*, vol. 13, pp. 40 185–40 197, 2025.

[19] Y. Kwak, W. J. Yun, S. Jung, and J. Kim, "Quantum neural networks: Concepts, applications, and challenges," in *2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2021, pp. 413–416.

[20] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, no. 9, p. 625–644, Aug. 2021. [Online]. Available: http://dx.doi.org/10.1038/s42254-021-00348-9

[21] C.-C. Li, R.-H. He, and Z.-M. Wang, "Enhanced quantum state preparation via stochastic predictions of neural networks," *Physical Review A*, vol. 108, no. 5, p. 052418, 2023.

[22] S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, "Classification and reconstruction of optical quantum states with deep neural networks," *Physical Review Research*, vol. 3, no. 3, p. 033278, 2021.

[23] A. A. Melnikov, A. A. Termanova, S. V. Dolgov, F. Neukart, and M. Perelshtein, "Quantum state preparation using tensor networks," *Quantum Science and Technology*, vol. 8, no. 3, p. 035027, 2023.

[24] X. Li, W. Jiang, Z. Hua, W. Wang, X. Pan, W. Cai, Z. Lu, J. Han, R. Wu, C.-L. Zou *et al.*, "Experimental demonstration of reconstructing quantum states with generative models," *arXiv preprint arXiv:2407.15102*, 2024.

[25] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, "Reconstructing quantum states with generative models," *Nature Machine Intelligence*, vol. 1, no. 3, pp. 155–161, 2019.

[26] D. Sierra-Sosa, M. Telahun, and A. Elmaghraby, "Tensorflow quantum: Impacts of quantum state preparation on quantum machine learning performance," *IEEE Access*, vol. 8, pp. 215 246–215 255, 2020.

[27] V. Vargas-Calderón, H. Vinck-Posada *et al.*, "Arbitrary state preparation in quantum harmonic oscillators using neural networks," *arXiv preprint arXiv:2502.04598*, 2025.

[28] T. Krisnanda, S. Ghosh, T. Paterek, and T. C. Liew, "Creating and concentrating quantum resource states in noisy environments using a quantum neural network," *Neural Networks*, vol. 136, pp. 141–151, 2021.

[29] Z.-M. Wang and T.-Z. Chen, "Adaptive denoising quantum state preparation in a dynamic environment," *Physical Review Research*, vol. 6, no. 4, p. 043195, 2024.

[30] G. Torlai, B. Timar, E. P. Van Nieuwenburg, H. Levine, A. Omran, A. Keesling, H. Bernien, M. Greiner, V. Vuletić, M. D. Lukin *et al.*, "Integrating neural networks with a quantum simulator for state reconstruction," *Physical review letters*, vol. 123, no. 23, p. 230504, 2019.

[31] J. M. Arrazola, T. R. Bromley, J. Izaac, C. R. Myers, K. Brádler, and N. Killoran, "Machine learning method for state preparation and gate synthesis on photonic quantum computers," *Quantum Science and Technology*, vol. 4, no. 2, p. 024004, 2019.

[32] H. Wang, Y. Liu, P. Liu, J. Gu, Z. Li, Z. Liang, J. Cheng, Y. Ding, X. Qian, Y. Shi *et al.*, "Robuststate: Boosting fidelity of quantum state preparation via noise-aware variational training," *arXiv preprint arXiv:2311.16035*, 2023.

[33] S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. Liew, "Reconstructing quantum states with quantum reservoir networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 3148–3155, 2020.

[34] R.-H. He, H.-D. Liu, S.-B. Wang, J. Wu, S.-S. Nie, and Z.-M. Wang, "Universal quantum state preparation via revised greedy algorithm," *Quantum Science and Technology*, vol. 6, no. 4, p. 045021, 2021.

[35] H. Zhao, G. Carleo, and F. Vicentini, "Empirical sample complexity of neural network mixed state reconstruction," *Quantum*, vol. 8, p. 1358, 2024.

[36] S. Ghosh, T. Paterek, and T. C. Liew, "Quantum neuromorphic platform for quantum state preparation," *Physical review letters*, vol. 123, no. 26, p. 260404, 2019.

[37] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Physical Review A*, vol. 98, no. 3, Sep. 2018. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.98.032309

[38] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Physical Review A*, vol. 99, no. 3, Mar. 2019. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.99.032331

[39] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nature Communications*, vol. 9, no. 1, Nov. 2018. [Online]. Available: http://dx.doi.org/10.1038/s41467-018-07090-4

[40] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, pp. 2121–2159, 2011.

[41] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.