Quantum Principal Component Analysis

Vandana Preethi Annavaram
Carnegie Mellon University
Pittsburgh, PA 15213
vannavar@andrew.cmu.edu

Prabh Simran Singh Baweja
Carnegie Mellon University
Pittsburgh, PA 15213
pbaweja@andrew.cmu.edu

Joseph Konan
Carnegie Mellon University
Pittsburgh, PA 15213
jkonan@andrew.cmu.edu

Chaoran Zhang
Carnegie Mellon University
Pittsburgh, PA 15213
chaoranz@andrew.cmu.edu

Fan Zhou
Carnegie Mellon University
Pittsburgh, PA 15213
fanzhou@andrew.cmu.edu

Abstract

Quantum computers are gaining increased prominence and are expected to be the mainstream by 2026. Quantum algorithms provide theoretically exponential speedup over classical algorithms when applied to Machine Learning algorithms for Signal Processing. However quantum computers are limited by their decoherence and ability to measure relative entropy or distinguishability of quantum states. Principal Component Analysis(PCA) is one of the prominent algorithms in the Machine Learning and Signal Processing domain. The main motivation behind this project is that the Quantum PCA has exponential speed up with state preparation assumptions. In this project we intend to explore whether PCA could be deployed on quantum computers and contextualize it's limitations.

1 Problem Statement

1.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is one of the most widely used data dimension reduction algorithm. The main idea of PCA is to map n - dimensional features to k - dimensional features, where $k \ll n$. This k-dimensional space is constructed by a set of mutually orthogonal basis, referring as principal components, which preserves most of data's variation.

In fact, by calculating the covariance matrix of the data matrix and then obtaining the eigenvalues and eigenvectors of the covariance matrix, the matrix composed of the corresponding eigenvectors of the k features with the largest eigenvalue (i.e., the largest variance) is selected. In this way, the data matrix can be transformed into a new space to achieve dimensionality reduction of data features. We can also realize this process by singular value decomposition of data matrix.

$$T = XW$$

$$X^T X = W \Lambda W^T$$

X is the data matrix. $X^T X$ is the covariance matrix. W is the weights whose columns are the eigen vectors of $X^T X$.

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

1.2 Quantum Principal Component Analysis (QPCA)

Principal component analysis (PCA) is widely used for dimensionality reduction of data. However, the use of eigenvalue decomposition could easily cause its time complexity becomes $O(N^3)$, where N is the dimension of data, making the algorithm becomes difficult to be used when the dimensionality of data is large. On the contrary, quantum Principal Component Analysis (qPCA) are beenefited from quantum computer's parallel characteristics (1) deriving time complexity $O(Nploy())$, which is much better than the one performed on classical one.

The problem for qPCA is similar to its classical counter-part, i.e. given a matrix $A_0 \in \mathcal{C}^{p\times q}$, and $A = A_0 A_0^\dagger$, we would like to retrieve

$$|\psi'_A\rangle = \sum_{k=1}^{t} \sigma_k \, |\lambda_k\rangle \, |u_k\rangle \, |v_k\rangle$$

, where $t \ll r$, the totall number of components of A, σ_k are the singular values of A_0 and u_k and v_k are left and right singular vectors respectively.

2 Literature Review

The Quantum Principal Component Analysis (qPCA) is firstly introduced in (2) to retrieve quantum form eigenvectors and corresponding eigenvalues of unknown low-rank matrix. It deviates from conventional quantum state tomography (3), which reveal the properties of unknown quantum state by performing measurements of different observable and then measure its result statistically. Though both of these methods assume multiple copies, i.e. one assumes multiple copies of a quantum system with density matrix ρ and another assumes multiple copies of a quantum system in that state, their ideas come from different aspect, one uses state function as an energy operator while another let the state plays a passive role only.

In (2), qPCA outputs all the eigenvalues and eigenvectors of the data:

$$|\psi_A\rangle = \sum_{k=1}^{r} \lambda_k \, |\lambda_k\rangle \, |u_k\rangle$$

, and then obtain the top-k principal components by sampling. This motivates (4) to use a quantum state contains the approximate top-k largest eigenvalues instead:

$$|\psi'_A\rangle \approx \sum_{k=1}^{t} \sigma_k \, |\lambda_k\rangle \, |u_k\rangle \, |v_k\rangle$$

, where σ_k denotes the singular value of A and u_k, v_k are left and right singular vectors respectively. With this, the time complexity is reduced from $O(rpoly(\log(p)))$ to $O(tpoly(\log(p)))$. In (4), a quantum Singular Value Thresholding (qSVT) (5), together with a modified quantum Singular Value Thresholding (MqSVT) module is used filter out small σ_ks approximately and utilizing a controlled rotation operation $R_y(\alpha)$ can be adjusted to improve the success probability of fidelity of the algorithm.

In (4), the qSVT module is further modified by using a a controlled X operation at the end, where only one approximation is made in total, leading to better accuracy and less quantum gates are used, achieving the state-of-the-art result. Their method are composed of a phase estimation module (6) to extract eigenvalues to the quantum register, an unitary operation is used to filter out small eigenvalues from quantum register by numerically approximate $\frac{1}{\lambda_k}$ with Newton's iteration and then apply the thresholding using quantum Fourier Transform (QFT) (7) arithmic, a controlled operation module (8) to validate the eigenvalues obtained have their corresponding eigenvectors and, finally, reverse operations of the unitary operation and phase estimation the unnecessary registers, followed by measurements.

3 Quantum operations

For familiarize readers not familiar with quantum computing, this section introduces some basic of quantum operations.

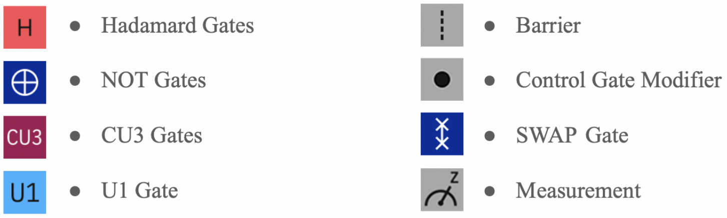

Figure 1: Gates

3.1 Quantum Circuits

Quantum operations could be represented in circuit format. By convention, quantum operations should be unitary operations.

3.1.1 Hadamard Gate

The Hadamard gate and T-gate forms universal gate set for single-qubit gates. The H or Hadamard gate rotates the states $|o\rangle$ and $|1\rangle$ to $|1\rangle$ and $|0\rangle$ respectively. It is useful for making superposition. If we have a universal gate set on a classical computer and add the Hadamard gate, it becomes a universal gate set on a quantum computer.

Hadamard gate, expressed in matrix form is:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

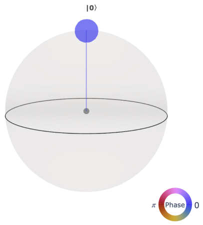

Figure 2: Q Sphere for Hadamard Gate

3.1.2 T-Gate

The T-Gate and the above Hadamard Gate together forms universal gate set for single-qubit gates. T-Gate, expressed in matrix form is:

$$T = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix}$$

3.1.3 NOT Gate

The NOT gate, also known as the Pauli X gate, flips the $|0\rangle$ state to $|1\rangle$, and vice versa. The NOT gate is equivalent to RX for the angle π or to HZH, where Z is defined as:

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = T^4$$

3.1.4 Barrier

To make your quantum program more efficient, the compiler will try to combine gates. The barrier is an instruction to the compiler to prevent these combinations being made. Additionally, it is useful for visualizations.

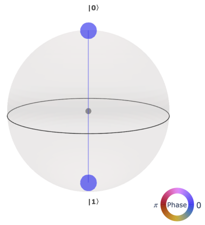

Figure 3: Q Sphere for NOT Gate

3.1.5 Control Gate Modifier

A control modifier yields a gate whose original operation is now contingent on the state of the control qubit. When the control is in the $|1\rangle$ state, the target qubit(s) undergo the specified unitary evolution. In contrast, no operation is performed if the control is in the $|0\rangle$ state. If the control is in a superposition state, then the resulting operation follows from linearity.

Drag the control modifier to a gate in order to add a control to it. Dots will appear above and below the gate, on the qubit wires that can be targets that control; click on one or more dots to assign the target to one or more qubits. You can also assign a control by right-clicking a gate.

Controlled gates mathematically could be expressed as $\Lambda(G)$, and it follows:

$$\Lambda(G)(\alpha|0\rangle + \beta|1\rangle)|\psi\rangle = \alpha|0\rangle|\psi\rangle + \beta|1\rangle G|\psi\rangle$$

Controlled not gate (CNOT) is an example for controoled gate, which is defined by following matrix:

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3.1.6 SWAP Gate

The SWAP gate swaps the states of two qubits.

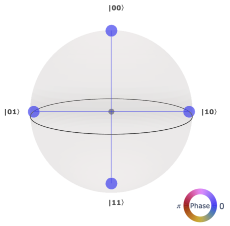

Figure 4: Q Sphere for Swap Gate

3.1.7 U1 Gate

The U1 gate renamed as Phase gate, applies a phase of – to the $|1\rangle$ state. For certain values of θ, it is equivalent to other gates. For example, P(π) = Z, P(π/2) = S, and P(π/4) = T. Up to a global phase of –, it is equivalent to RZ(θ).

3.1.8 CU3 Gate

This gate is an addition of control gate and U3 gate which allows the construction of any single qubit

3.1.9 Measurement

Measurement in the standard basis, also known as the z basis or computational basis. Can be used to implement any kind of measurement when combined with gates. It is not a reversible operation.

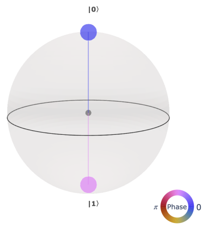

Figure 5: Q Sphere for U1 Gate

3.2 Dirac Notation

Though column vector notation performs well in linear algebra, it becomes a bit cumbersome for quantum computing, especially when we are considering multiple qubits. Dirac notation is widely used for quantum computing and we make a short introduction here.

3.2.1 Bra $\langle\cdot|$ and Ket $|\cdot\rangle$

Given a unit column vector ψ, we could denote it by $|\psi\rangle$, and similarly, given a unit row vector $\psi^\dagger$, we could denote it by $\langle\psi|$, where $\psi^\dagger$ is the transpose of ψ applied element-wise complex conjugation.

This notation implies that the inner product of a unit vector, denoted by 1, is denoted by $\langle\psi|\,|\psi\rangle$

There are some convention to follow:

$$|0\rangle = [1, 0]^\dagger$$

,

$$|1\rangle = [0, 1]^\dagger$$

,

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

3.2.2 Tensor product

Tensor product is widely used, with $\bigotimes$ notation is always omitted. Following is an example:

$$|0\rangle \bigotimes |0\rangle = |0\rangle\,|0\rangle = |00\rangle = [1, 0, 0, 0]^\dagger$$

4 Implementation

4.1 Methodology

In qPCA, for a matrix $A_0 \in C^{p\times q}$, we could represent the eigendecomposition of $A = A_0 A_0^+$ as $A = \sum_{k=1}^{r} \lambda_k u_k u_k^T$ and thus our goal is to obtain its t-largest eigenvalue and corresponding eigenvectors $|\psi_A^{'}\rangle \approx \sum_{k=1}^{t} \sigma_k |\lambda_k\rangle|u_k\rangle|u_k\rangle$.

To do so, we first apply phase estimation as in (6), recalling that with phase estimation module U_{PE}, one could extract the eigenvalues of A. That is, the $e^{2\pi i A}$ is a unitary operator with eigenvalues $e^{2\pi i \lambda_k}$, left eigenvectors u_k and right eigenvectors v_k. With that, one could estimate the eigenvalue λ_k as

$$|0\rangle\,|\psi_{A_0}\rangle \xrightarrow{U_{PE}} \sum_{k=1}^{r} \lambda_k\,|\lambda_k\rangle\,|u_k\rangle$$

, where $U_{PE}(A) = (QFT^\dagger \bigotimes I)(e^{2\pi i A} \bigotimes I)(H \bigotimes I)$.

Similar to classical PCA pipeline, after obtaining eigenvalues, we would like to filter out those with small eigenvalues. It is achieved in the similar way from quantum Singular Value Thresholding (qSVT) (5), where we convert $y_k = (1 - \frac{\tau}{\lambda_k})_+ = \max 1 - \frac{\tau}{\lambda_k}, 0$, and τ is the threshold for filtering. We use Newton's iteration to obtain $z_k = \frac{1}{\lambda_k}$, where $z_k^{(i+1)} = 2z_k^{(i)} - (z_k^{(i)})^2\lambda_k$. Then, $y_k = (1 - \frac{\tau}{\lambda_k})_+$ is performed following QFT arithmic. In summary this whole unitary operation $U_{\lambda,\tau}$ performs:

$$|0\rangle \sum_{k=1}^{r} \sigma_k |\lambda_k\rangle |u_k\rangle |v_k\rangle \xrightarrow{U_{\lambda,\tau}} \sum_{k=1}^{r} \sigma_k |y_k\rangle |\lambda_k\rangle |u_k\rangle |v_k\rangle$$

.

After obtaining eigenvalues, we apply unitary controlled operation $CU(8)$ to validate that these eigenvalues have principal components correspondingly. We perform this by reversing the auxiliary qubit when a qualified eigenvalue candidate, i.e. $y_k > 0$, presents, and do nothing for unqualified eigenvalues. The process is:

$$|0\rangle \sum_{k=1}^{r} \sigma_k |y_k\rangle |\lambda_k\rangle |u_k\rangle |v_k\rangle \xrightarrow{CU} |1\rangle \sum_{k=1}^{t} \sigma_k |y_k\rangle |\lambda_k\rangle |u_k\rangle |v_k\rangle + |0\rangle \sum_{k=t+1}^{r} \sigma_k |y_k\rangle |\lambda_k\rangle |u_k\rangle |v_k\rangle$$

.

Having achieved an expression with large and smalle eigenvalues separated, we then remove undesired components, i.e. $|y_k\rangle$ and $|\lambda_k\rangle$, by reverse operation $U^\dagger_{\lambda,\tau}$ and $U^\dagger_{PE}$. It will

$$\sum_{k=1}^{t} \sigma_k |y_k\rangle |\lambda_k\rangle |u_k\rangle |v_k\rangle + |0\rangle \sum_{k=t+1}^{r} \sigma_k |y_k\rangle |\lambda_k\rangle |u_k\rangle |v_k\rangle \xrightarrow{U^\dagger}$$

$$|1\rangle |0\rangle |0\rangle \sum_{k=1}^{t} \sigma_k |u_k\rangle |v_k\rangle + |0\rangle |0\rangle |0\rangle \sum_{k=t+1}^{r} \sigma_k |u_k\rangle |v_k\rangle$$

To measure the result, we first measure the auxiliary qubit. When it collapse to 1, it evidents states of remaining qubits is our desired $|\psi'_{A_0}\rangle = |1\rangle |0\rangle |0\rangle \sum_{k=1}^{t} \sigma_k |u_k\rangle |v_k\rangle$, and to obtain $|\psi'_A\rangle = |1\rangle |0\rangle |0\rangle \sum_{k=1}^{t} \sigma_k |\lambda_k\rangle |u_k\rangle |v_k\rangle$, we use another phase estimation.

4.2 Quantum State Preparation

An entire branch of quantum machine learning algorithms encode the dataset into the amplitudes of a quantum state.

$$\begin{aligned} |\psi_D\rangle &= \sum_{m=1}^{M} \sum_{i=1}^{N} x_i^m |i\rangle |m\rangle \\ &= \sum_{m=1}^{M} |\psi_{x^m}\rangle |m\rangle \end{aligned}$$

This quantum state has an amplitude vector of dimension NM that is constructed by concatenating all training inputs, $\alpha = (x_1^1, ..., x_1^N, ..., x_M^1, ..., x_M^N)^T$.(9) The dataset has to be normalised so that the absolute square of the amplitude vector is one, $|\alpha|^2 = 1$. The training outputs can either be basis encoded in an extra qubit $|y^m\rangle$ entangled with the $|m\rangle$ register, or encoded in the amplitudes of a separate quantum register,

$$|\psi_y\rangle = \sum_{m=1}^{M} y^m |m\rangle$$

Amplitude encoding of datasets therefore requires the ability to prepare an arbitrary state

$$|\psi\rangle = \sum_i \alpha_i |i\rangle$$

both efficiently and robustly. The main advantage of amplitude encoding is that we only need $n =$ qubits to encode a dataset of M inputs with N features each.

For efficient preparation of quantum states, we can also use Multi-controlled Rotations to help state preparation keeping in linear time, and apply the parallelism-based approach on it.

4.3 Circuit

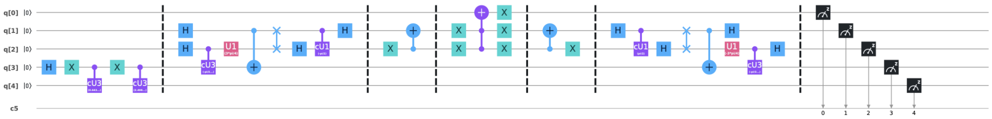

Figure 6: Circuit Diagram of qPCA of 2X2 matrix with $\tau = 1$

The qubit q[0] ia an ancillary qubit. Before the first barrier of quantum circuit, qubits q[3-4] are used to initialise the quantum state of the input of quantum circuit. Between the first barrier and the second barrier, the qubits q[1-2] are used to save the eigen values from the phase estimation. Between the third and fourth barrier is the controlled operation. The remaining quantum circuit does inverse operations and measurement.

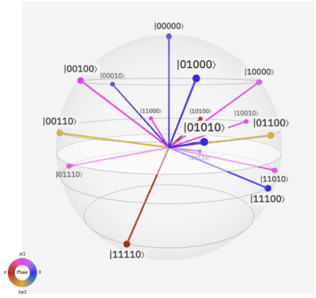

Figure 7: This is an Intermediate Q-Sphere of the circuit. Each color represents different gate

The quantum states in a system are represented by Q-sphere. The Q-sphere represents the state of a system of one or more qubits by associating each computational basis state with a point on the surface of the sphere. N qubits operate in 2^N space.

5 Experiment Results

5.1 Principal Component Analysis

We conducted the experiment on 2000 random sample points using classical PCA. The results of classical PCA are shown in the Figure 8 and the red orthogonal bases maximizing variance. We want to first understand this baseline, then achieve similar results on a quantum computer. We start with covariance matrix

$$A = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$$

and calculate its orthonormal base which is selected based on τ

We obtain the result separately: when $\tau = 1.0$, the eigenbase is $[0.7071, 0.7071]^T$, when $\tau = 0.8$, the eigenbases are $[0.7071, 0.7071]^T$ and $[-0.7071, 0.7071]^T$

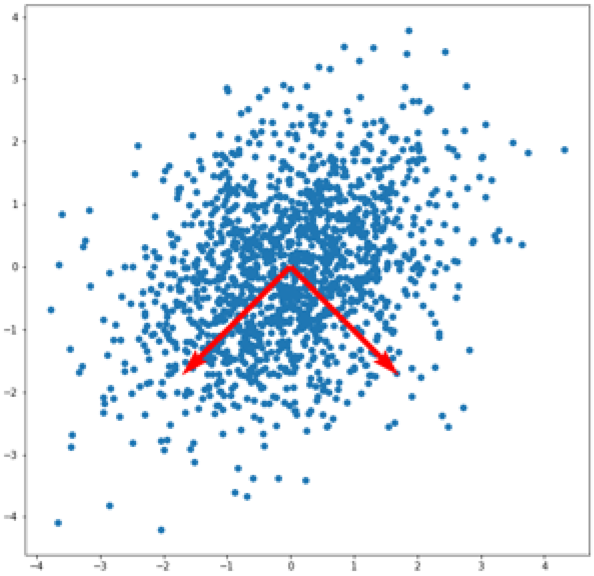

Figure 8: Classical PCA implements results on 2D data sets

5.2 quantum Principal Component Analysis

Circuit as shown in Figure 6, is implemented on IBM Quantum. It is a qPCA circuit for 2 × 2 matrix A with $\tau = 1$. We intended to test the circuit with the same matrix A in PCA, i.e., $A = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$, and thus, as in (10), we obtain initial quantum state as:

$$|\psi_A\rangle = [0.6708, 0.2236, 0.2236, 0.6708]$$

Recaling in this case, from classical PCA, we should have:

$$\lambda_1 = 2, u_1 = [0.7071, 0.7071]^T$$

$$\lambda_2 = 1, u_2 = [-0.7071, 0.7071]^T$$

We obtained the probability histogram as in Figure 9. The correspondance of the figure and the circuit location is the rightmost digit in the figure indicating the 0-th auxiliary qubit.

Recalling from Section 4.1, the first qubit is equal to 1 indicates the circuit successfully measured the eigen values, and the first qubit is equal to 0 indicates failure. From the figure, when first qubit is equal to 1, we could obtain $|\psi'_A\rangle = [0.5, 0.5, 0.5, 0.5]^T$.

On the contrary, recalling that theoretically, we should obtain $\frac{\lambda_1 |u_1\rangle |u_1\rangle}{\sqrt{\lambda_1^2}} = [0.5, 0.5, 0.5, 0.5]^T$, which validates our result.

6 Conclusion and Future Works

In summary, we have shown that PCA is possible on quantum computers but with some limitations. Quantum computers are accessible now a days. But setting up the quantum state is a very careful process. Our approach is theoretically scalable but is limited by current quantum computer constraints such as noise.

Our experiment can be further explored with signed numbers by using extra qubits as a signed indicators and perform PCA. ———How does noise effect coherence of quantum states——— The performance of higher order matrices is yet to be explored and building a 100,000 qubits quantum computer is hard, but would be ideal for testing practical Machine Learning solutions.

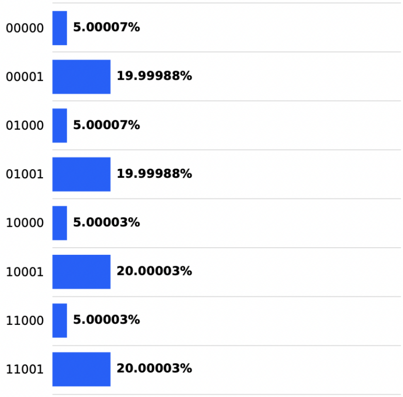

Figure 9: Probability Histogram of qPCA in reported from IBM Quantum

References

[1] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," 2002.

[2] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum principal component analysis," *Nature Physics*, vol. 10, no. 9, pp. 631–633, 2014.

[3] G. M. D'Ariano, M. G. Paris, and M. F. Sacchi, "2 quantum tomographic methods," *Quantum State Estimation*, pp. 7–58, 2004.

[4] J. Lin, W.-S. Bao, S. Zhang, T. Li, and X. Wang, "An improved quantum principal component analysis algorithm based on the quantum singular threshold method," *Physics Letters A*, vol. 383, no. 24, pp. 2862–2868, 2019.

[5] B. Duan, J. Yuan, Y. Liu, and D. Li, "Efficient quantum circuit for singular-value thresholding," *Physical Review A*, vol. 98, no. 1, p. 012308, 2018.

[6] W. van Dam, G. M. D'Ariano, A. Ekert, C. Macchiavello, and M. Mosca, "Optimal quantum circuits for general phase estimation," *Physical review letters*, vol. 98, no. 9, p. 090501, 2007.

[7] L. Ruiz-Perez and J. C. Garcia-Escartin, "Quantum arithmetic with the quantum fourier transform," *Quantum Information Processing*, vol. 16, no. 6, p. 152, 2017.

[8] D. P. DiVincenzo, "Quantum gates and circuits," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1969, pp. 261–276, 1998.

[9] M. Schuld and F. Petruccione, *Supervised learning with quantum computers*. Springer, 2018, vol. 17.

[10] P. Nimbe, B. A. Weyori, and A. F. Adekoya, "Models in quantum computing: a systematic review," *Quantum Information Processing*, vol. 20, no. 2, pp. 1–61, 2021.