Introduction to Quantum Machine Learning: Concepts

Slides Prepared by Dr Marwa Sorour

Presented by Dr Ahmed El-Mahdy

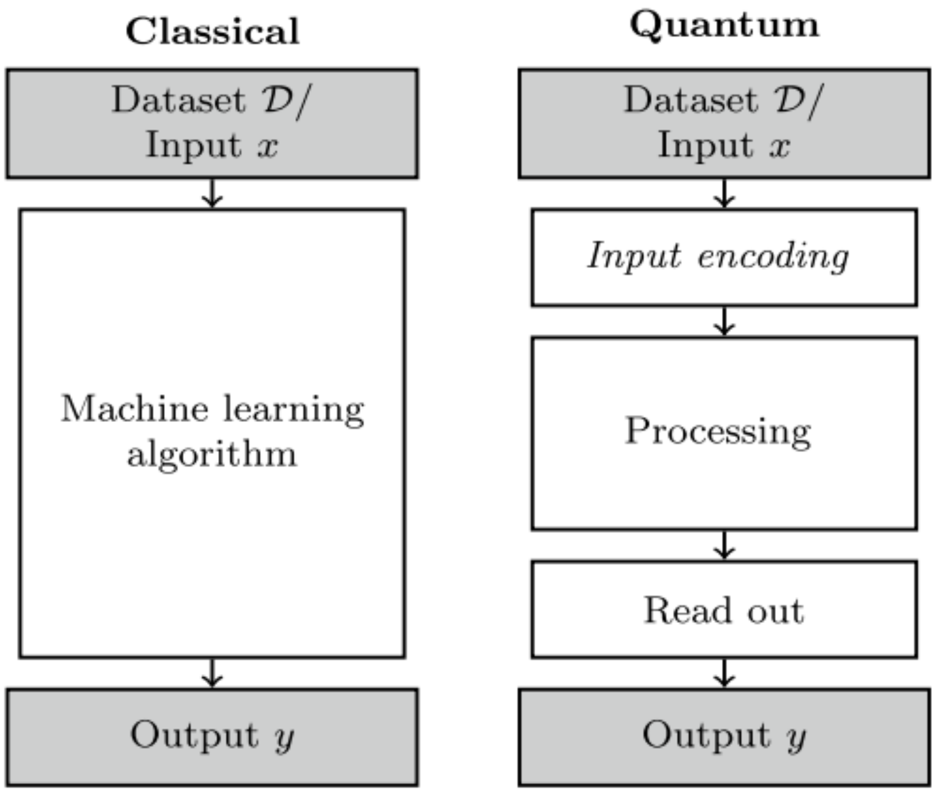
Outline

- What is Quantum Machine Learning?
- Variational Quantum Circuits
- The Variational Principle
- Data Encoding
- Parametrized Gates (Ansatz)
- Observables and Measurement

Quantum Machine Learning (QML)

- By the mid-90s, Peter Shor and Lov Grover developed algorithms that harnessed the power of quantum computing to solve factorization and database searching problems, demonstrating significant quantum speedups.
- These pioneering contributions established the foundation for applying quantum mechanics to various computational problems, including machine learning.
- QML combines quantum computing principles with classical machine learning algorithms, enabling the tackling of complex real-world problems that are challenging for classical computers.
- The advancement of quantum hardware has made it possible to test QML algorithms and demonstrate quantum advantage in certain tasks.
- With ongoing advancements in quantum hardware, the scope of QML applications is rapidly broadening. (Ranga et al., 2024)

Quantum Machine Learning



Classical versus Quantum machine learning (Schuld and Petruccione, 2021)

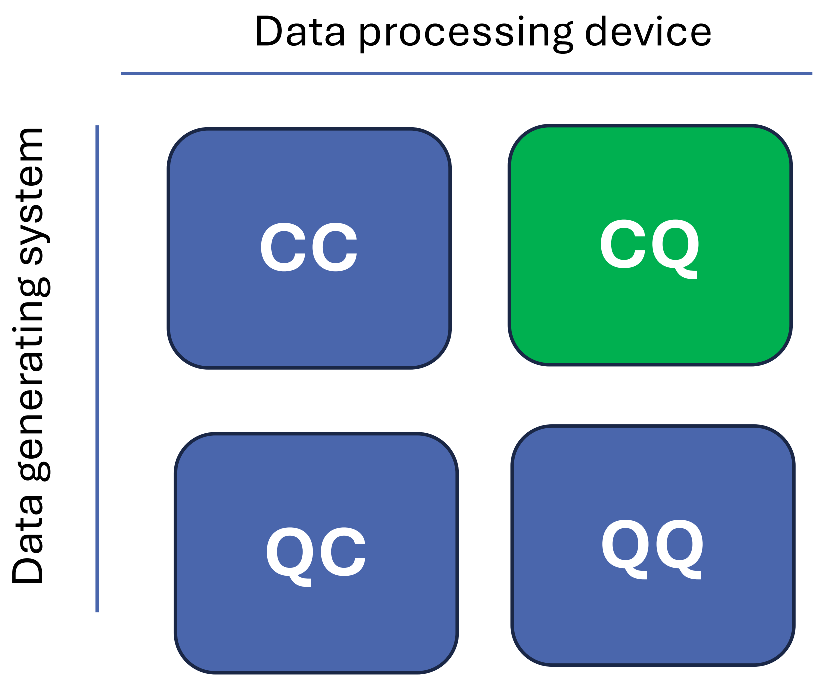
Quantum Machine Learning

CC: Classical data being processed classically

QC: Quantum data, classical processing

CQ: Classical data, quantum processing (Quantum Machine Learning)

QQ: Quantum data, quantum processing



Combining quantum computing and machine learning, C classical, Q quantum, (Schuld and Petruccione, 2021)

Quantum Machine Learning

- Quantum Machine Learning (QML) often leverages Variational Quantum Algorithms (VQAs), which use parameterized quantum circuits optimized through classical feedback loops.

- This hybrid approach enables QML models to harness quantum advantages while adapting to noisy intermediate-scale quantum (NISQ) hardware constraints including limited numbers of qubits and noise processes that limit circuit depth.

- VQAs use parametrized quantum circuits to be run on the quantum computer while relying on classical optimizers for parameter optimization. This keeps the quantum circuit shallow and hence mitigates noise.

(Cerezo et al., 2021)

Variational Quantum Algorithms

- **VQA** is a hybrid quantum-classical optimization procedure that uses a **V**ariational **Q**uantum **C**ircuit **(VQC)** inside a classical loop to solve a problem.

- The **V**ariational **Q**uantum **A**lgorithm **(VQA)** is based on the **Variational Principle** from quantum mechanics.

What is the Variational Principle?

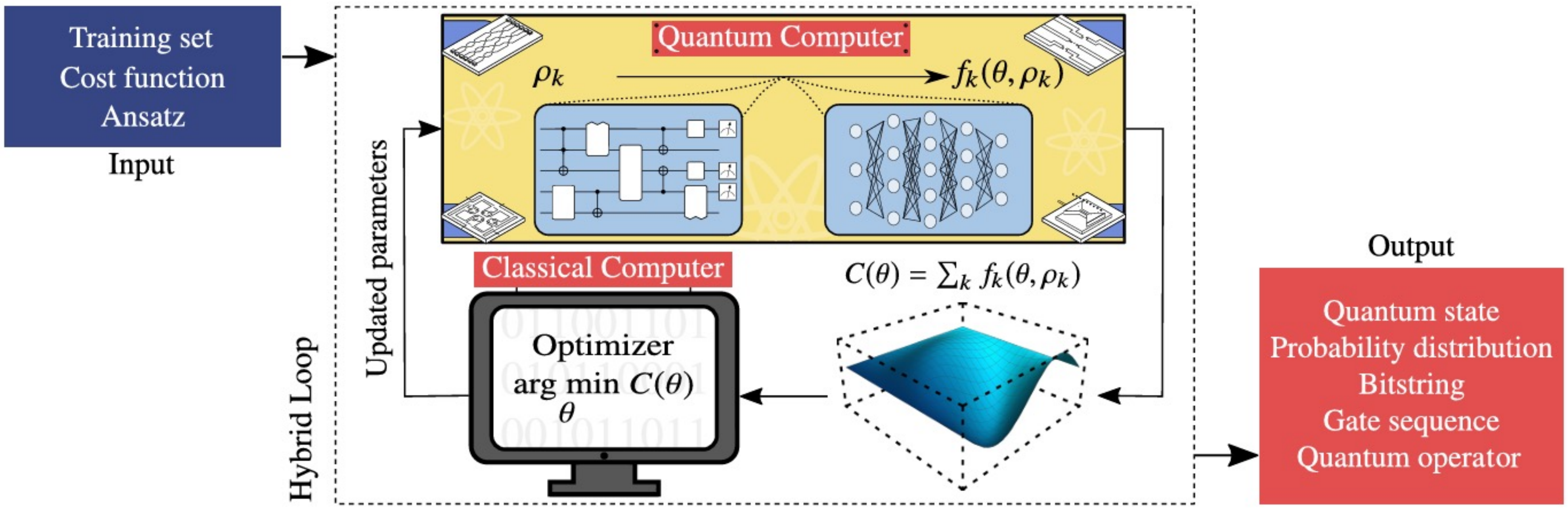
- The variational principle of quantum mechanics is used to find an **upper bound** for the **ground state energy** of the system.
- **The Hamiltonian,** $\widehat{H}$, is an operator representing the energy of the quantum system.
- The time evolution of the system is governed by the Schrödinger equation: $i\hbar \frac{d}{dt}|\psi\rangle = \widehat{H}|\psi\rangle$
- The **ground state** of the system is the **smallest eigenvalue of $\widehat{\boldsymbol{H}}$** corresponding to the ground state $|\psi_0\rangle$:

$$\widehat{H}\ |\psi_0\rangle = E_0\ |\psi_0\rangle$$

- The cost function is written as the expectation of the Hamiltonian: $C(\theta) = \langle\psi(\theta)\widehat{H}|\psi(\theta)\rangle$.
- A parametrized ansatz with the parameters θ, is used to prepare a state $|\psi(\theta)\rangle$ such that instead of finding the ground state $|\psi_0\rangle$, we find the parameters θ that **minimizes** $C(\theta)$, and hence the energy of the system ⟶ the **best approximation** $|\boldsymbol{\psi}(\boldsymbol{\theta}^*)\rangle$ to the **ground state** $|\boldsymbol{\psi}_{\mathbf{0}}\rangle$.

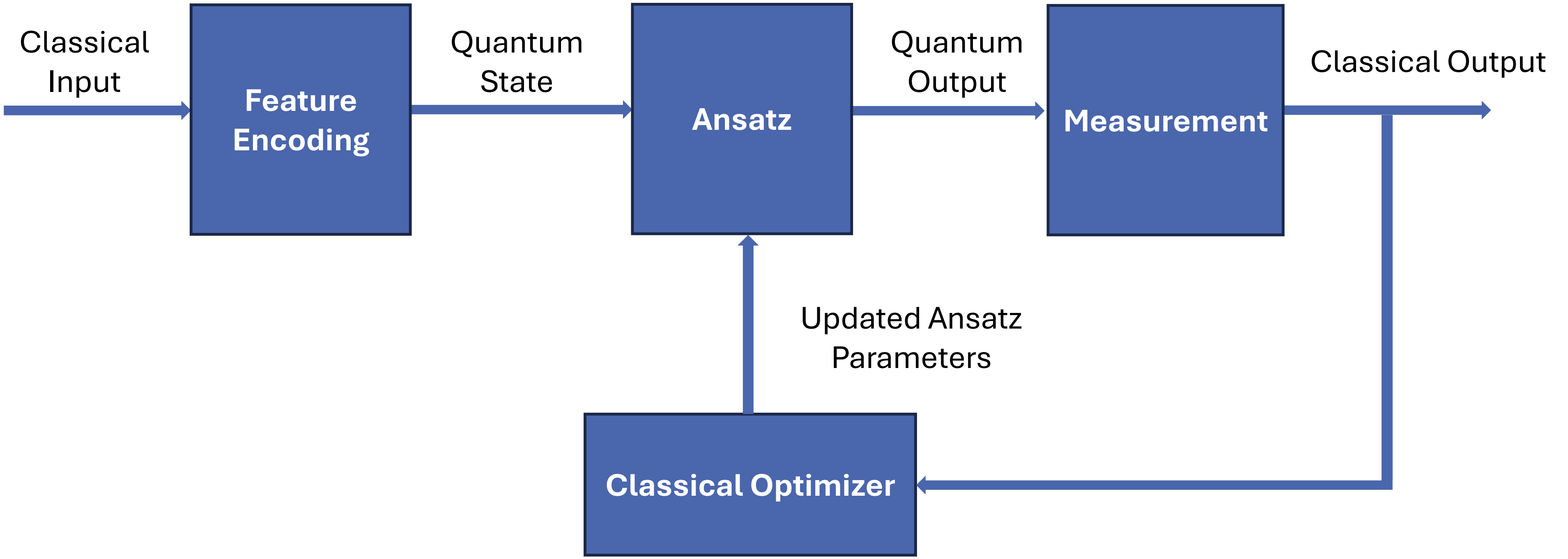
(Schuld and Petruccione, 2021)

Variational Quantum Algorithm (VQA)



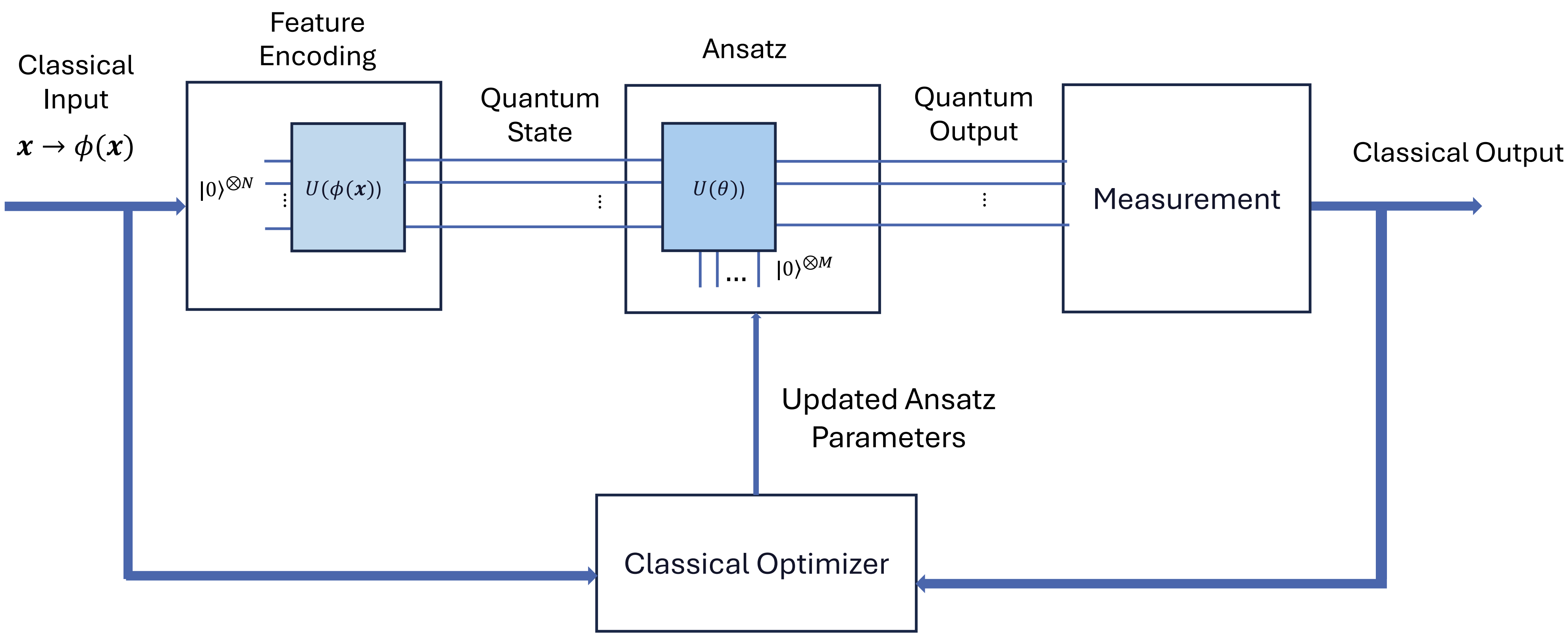
Schematic Diagram of a Variational Quantum Algorithm (Cerezo et al., 2021)

Variational Quantum Algorithm (VQA)



Illustrative Diagram of a Variational Quantum Algorithm.

Variational Quantum Algorithm (VQA)

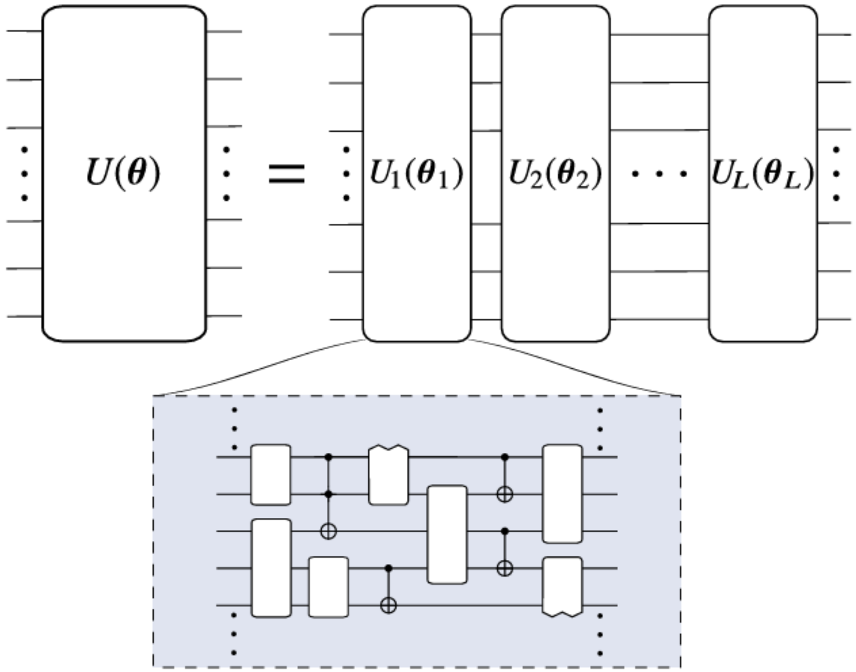


Illustrative Diagram of a Variational Quantum Algorithm.

Ansatz

- $U(\theta)$ can be expressed as a product of L unitaries, $U_l(\theta_l)$ acting sequentially on an input quantum state.
- Each unitary $U_l(\theta_l)$ can be decomposed into a sequence of parameterized and unparametrized gates.
- There are **problem-inspired** ansatz (use information about the problem to tailor an ansatz) and **problem-agnostic** ansatz (generic ansatz)

(Benedetti et al., 2019)



Schematic diagram of an ansatz (Cerezo et al., 2021)

Observable

Classical physics:

- An observable is a real valued such as position, momentum or energy…, which is retrieved from the current state of a system.

- The observable is a function F that takes state S and outputs **real number** x which corresponds to measured quantity, that is: $F(S) = x$.

.

Quantum physics:

- In quantum physics an observable is not a real-valued function, but it is represented by a matrix O that acts on a quantum state$|\psi\rangle$.

- The matrix O allows us to retrieve quantity from a system.

- The result of measurement is **discretized.**

- The eigenvalues λ_i of the matrix O are the only possible values observable can take after being measured.

$$\boldsymbol{O\ |\psi\rangle \to \lambda_i\ |a_i\rangle}$$

- The eigenvectors $|a_i\rangle$ can be interpreted as states in which the system is left after measuring the associated eigenvalue λ_i. The collapse of the state $|\psi\rangle$ to the state $|a_i\rangle$.

Observables

- Each quantity we want to retrieve from quantum state $|\psi\rangle$ is associated with a **different observable,** such as **position and momentum** having state.
- The observable is a Hermitian matrix that is the matrix is equal to its conjugate transpose, and it has real eigenvalues with orthonormal eigenvectors.
- Each Observable has eigenvalues that are the only possible values of measured quantity.
- Any state $|\psi\rangle = \sum_{i=1}^{n} c_i \, |a_i\rangle$ can be written as a linear combination of the orthonormal eigenvectors spanning C^n.
- Each measurement can yield one of the basis states $|a_i\rangle$ with different probabilities $|c_i|^2$.
- The distribution of possible outcomes λ_i is governed by probabilities $|c_i|^2$, $\boldsymbol{O \, |\psi\rangle \rightarrow \lambda_i \, |a_i\rangle}$
- Making multiple measurements of a particle in the same state $|\psi\rangle$, gives the expectation value of the observable O.

Measurement

- Measurement allows us to extract classical information from quantum systems.
- It is through measurement that quantum states are observed, and their properties inferred.

Measurements in the Computational Bases

- $|\psi\rangle = \sum_{i=1}^{n} \langle x_i|\psi\rangle \; |x_i\rangle$, where $|x_i\rangle$s are the computational basis and forms an orthonormal basis.

- $|\psi\rangle$ collpases to $|x_i\rangle$ with a probability of $|\langle x_i|\psi\rangle|^2$.

- **Example:** The Pauli Z, σ_Z= $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is an observable corresponding to a computational basis measurement which has eigenvalue of λ_1 =1 corresponding to the state $|0\rangle$, and the eigenvalue $\lambda_2 = -1$ corresponding to the state $|1\rangle$.

- An expectation value of the σ_Z observable $\langle O\rangle$= $\langle\psi|\sigma_Z|\psi\rangle = \; |\langle 0|\psi\rangle|^2\lambda_1 - \; |\langle 1|\psi\rangle|^2\,\lambda_2$
  $= \; |\langle 0|\psi\rangle|^2 - \; |\langle 1|\psi\rangle|^2$

- In practice, the expectation is estimated by running the circuit with S shots, and probabilities are calculated as follows: $\langle O\rangle = \langle\psi|\sigma_Z|\psi\rangle = \; |\langle 0|\psi\rangle|^2\lambda_1 - \; |\langle 1|\psi\rangle|^2\,\lambda_2$

$\frac{\text{How many times the state collapsed to } |0\rangle}{S}$ $\qquad$ $\frac{\text{How many times the state collapsed to } |1}{S}$

Measurements in bases other than the computational bases

- $|\psi\rangle = \sum_{i=1}^{n} \langle a_i|\psi\rangle \ \ |a_i\rangle$, as $|a_i\rangle$ (eigenvectors of O) forms an orthonormal basis of O. (O is Hermitian)

- **Measurement:** **The state $|\boldsymbol{\psi}\rangle$ collapses nto an eigenvector $|a_i\rangle$ of the observable** with a probability of $|\langle a_i|\psi\rangle|^2$.

- The result of the measurement is uncertain and is one of the eigenvalues λ_i of the observable O with probability $|\langle a_i|\psi\rangle|^2$.

- In principle this is done , yet measurement in an arbitrary basis is not efficient.

- The state has to be written in **computational basis** to collapse to one of the **classic bits,** So, we change the basis from the eigenvectors of the observable to the computational basis to calculate probabilities by running the circuit a number of shots.

Measurements in bases other than the computational bases

- Implement more complicated observables by applying a circuit U just before measuring. This pre-measurement circuit acts as a basis transformation of the quantum state which effectively implements the more complicated observable.

- **Example:** Pauli X Observable σ_X= $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, with eigenvalues 1 and -1 of and corresponding eigenvectors $|+\rangle$ and $|-\rangle$

- We change the basis from the eigenvectors of the observable to the computational basis by applying a Hadamard gate $H|+\rangle$= $|0\rangle$, and $H|-\rangle$= $|1\rangle$.

- Any observable (Hermitian operator) on n qubits can be expressed as a sum of Pauli matrix products.

- This decomposition is key for measurement and simulation on quantum hardware.

Feature Map (Data Encoding)

- To map classical data onto a quantum circuit, the real feature vector $\boldsymbol{x}$ must be embedded in a quantum state $|\psi_{\boldsymbol{x}}\rangle$.

$$|\psi_{\boldsymbol{x}}\rangle = U(\boldsymbol{x})|\psi_{\mathbf{0}}\rangle$$

Encoded state Encoding operator Initial state $|\psi_{\mathbf{0}}\rangle = |\mathbf{0}\rangle^{\otimes n}$

- The encoding parameter is a parameterized operator that encodes data into quantum states for processing by quantum machine learning algorithms
- The encoding determines how many qubits are required.
- It has direct consequences for the efficiency and performance of quantum algorithms.

Amplitude Encoding

- It associates classical information such as a real vector with quantum amplitudes.

- It encodes a real or complex valued input vector $\boldsymbol{x} \in C^N$ into the amplitudes of a quantum state $|\psi_{\boldsymbol{x}}\rangle \in \mathcal{H}$,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \leftrightarrow |\psi_{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{\sum_{i=1}^{N} |x_i|^2 = 1}} \sum_{i=1}^{n} x_i |i\rangle, \text{(Padding: } x_i = 0 \text{ for } i = N+1, \ldots 2^N\text{)}$$

- **Benefit:** You can encode $\boldsymbol{N}$ features in just $\boldsymbol{n}$ qubits, where $\boldsymbol{n} \geq \log_2 N$.

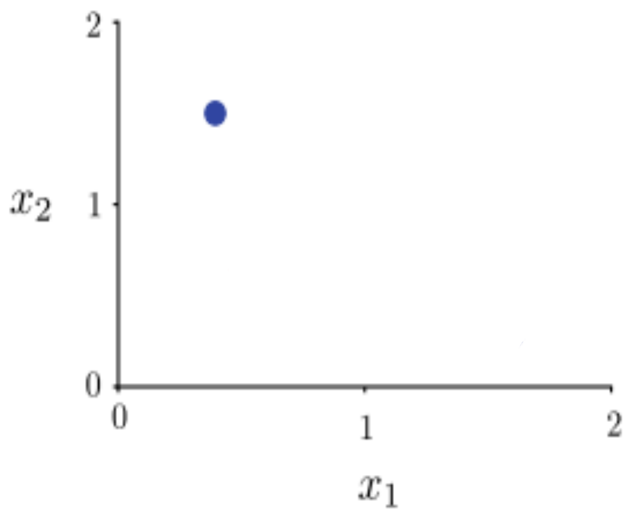
- It requires arbitrary state preparation.

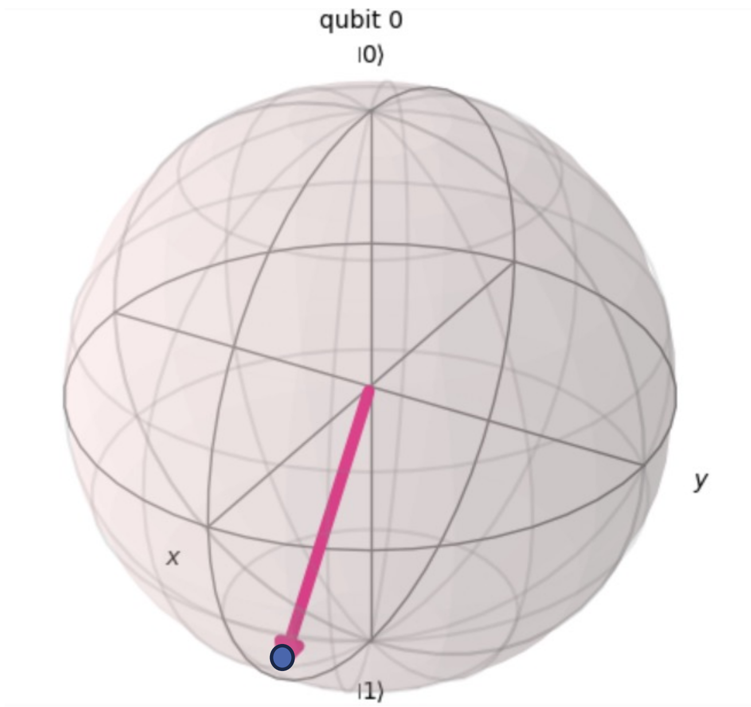
- **Note that $\boldsymbol{x}$** has to be first **normalized** as per the condition imposed by quantum mechanics and then **padded with zeros** to a dimension that is a **power of 2.**

Amplitude Encoding

Example

$$\boldsymbol{x} = (0.5, 1.5) \leftrightarrow |\psi_x\rangle = 0.316228|0\rangle + 0.948683|1\rangle$$





Amplitude Encoding

Example

$$\boldsymbol{x} = (0.1, -0.6, 1) \leftrightarrow |\psi_{\boldsymbol{x}}\rangle = 0.073|00\rangle - 0.438|01\rangle + 0.73|10\rangle + \mathbf{0}|11\rangle$$

$\boldsymbol{x}$ is represented by a quantum state of **2** qubits

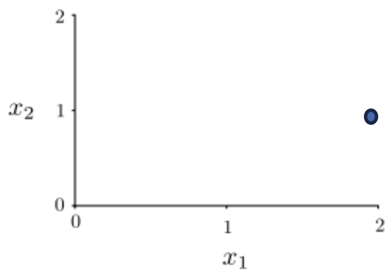
Basis Encoding

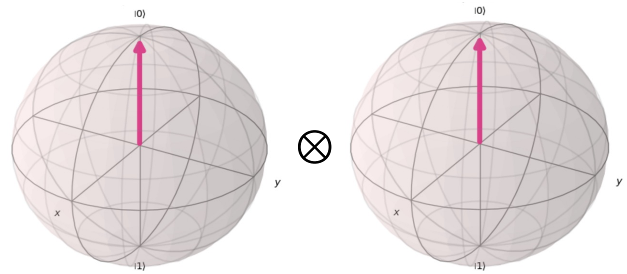
- Basis encoding encodes a classical *P*-bit string into a computational basis state of a *P*-qubit system.
- Each feature is represented by a *P*-bit string.
- A single feature $x_k^{(j)} = (b_1, b_2, \dots, b_P)$ is mapped to $\left|x_k^{(j)}\right\rangle = (b_1, b_2, \dots, b_P\rangle$, $b_i \in \{0, 1\}$, for $i = 1, \dots, P$.
- An N −dimensional feature vector $\boldsymbol{x}^{(j)}$ is mapped to a quantum state as the superposition of all the computational basis states describing the features of that vector: $\left|\boldsymbol{x}^{(j)}\right\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \left|x_k^{(j)}\right\rangle$.

Example: $\boldsymbol{x} = (2, 1)$:$2 \longrightarrow |10\rangle$ and $1 \longrightarrow |01\rangle$

Using the binary representation of the number from 0 to 2 requires **2 bits,** so we have $\mathbf{2^2}$ computational basis:

$$|\boldsymbol{x}\rangle = \frac{1}{\sqrt{2}}(0\,|00\rangle + |01\rangle + \ |10\rangle + \ 0\,|11\rangle\,)$$



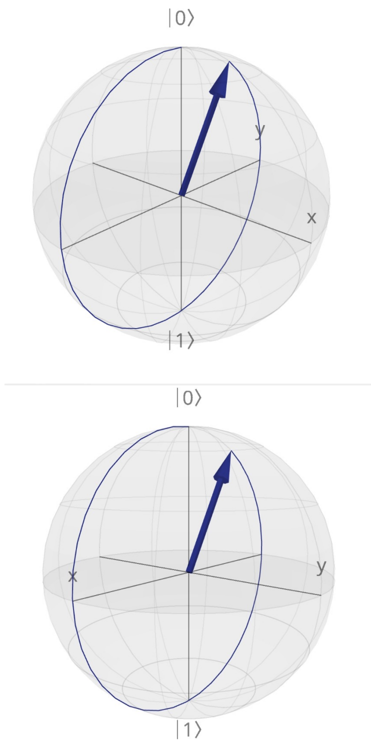


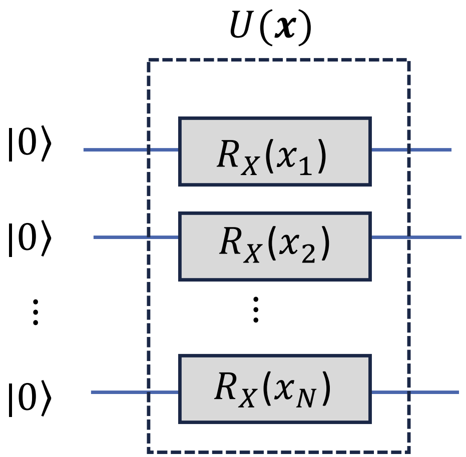
Angle Encoding

- Each feature x_i is encoded as a rotation angle of a gate around Z-axis of the Bloch sphere (a rotation angle in a parameterized circuit $R_X(x_i)$ or $R_Y(x_i)$).

- This transformation encodes the data in the phase of the quantum state without introducing entanglement between qubits.

- $\boldsymbol{n}$ qubits are required to encode $\boldsymbol{N}$ features.





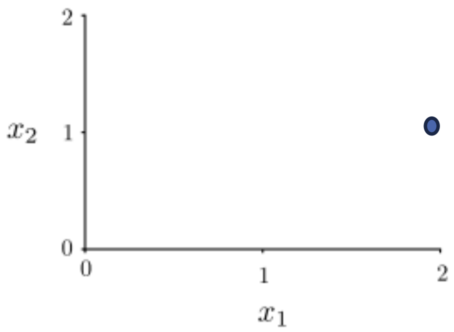
- Rescale features such that each feature $x_i \in (0, 2\pi]$ preventing information loss due to modulo- 2π effect of encoding to a qubit phase angle.

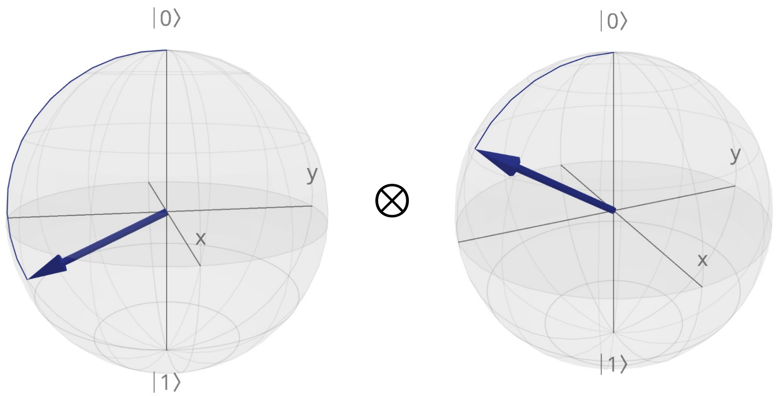
- The encoding operator $U(\boldsymbol{x}) = R_X(x_i) \otimes \ldots \otimes R_X(x_N)$.

Angle Encoding

- **Example**

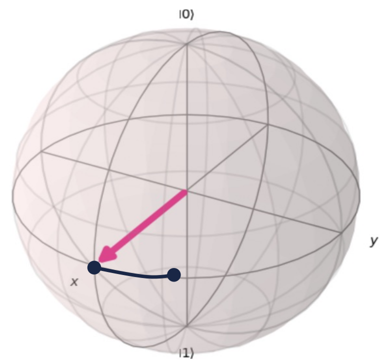
$$(\mathbf{2}, \mathbf{1}) \longrightarrow R_X(2)|\mathbf{0}\rangle \otimes R_X(1)|\mathbf{0}\rangle$$





Phase Encoding (Z Feature Map)

- The phase angle of a qubit is a real-valued angle ϕ about the Z-axis from the X −axis.
- $R_Z(\theta)\,|\,0\rangle = \; e^{i\frac{\theta}{2}} \;\; |\,0\rangle$ Global phase shift is unobservable; measurement remains equal to $|0\rangle$.
- $R_Z(\theta)|+\rangle = e^{-i\frac{\theta}{2}}\,|0\rangle + e^{i\frac{\theta}{2}}\,|1\rangle = e^{-i\frac{\theta}{2}}\,(|0\rangle + e^{i\theta}\,|1\rangle)$



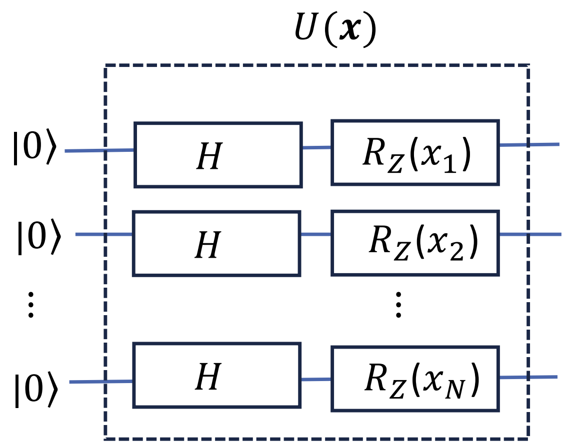
$|+\rangle$ after applying $R_Z(\theta)$

- Rescale features such that each feature $x_i \in (0, 2\pi]$.

- The encoding operator

$$U(\boldsymbol{x}) = (R_Z(x_i) \otimes \ldots \otimes R_Z(x_N))(H \otimes \ldots \otimes H)$$

- We may have r reptitions of this unitary operator.



ZZ feature map

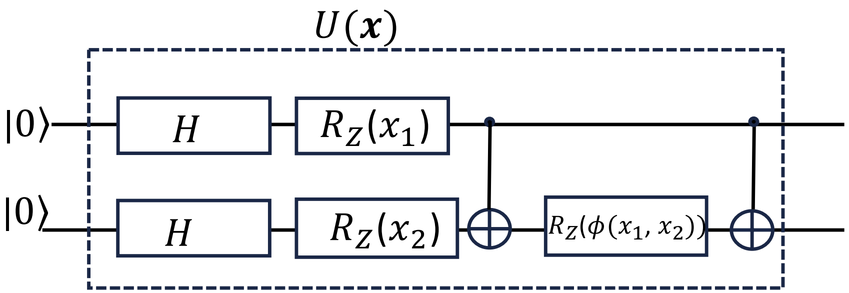
- Combines rotation and entangling gates to embed data **nonlinearly**.

- Through entanglement pairwise complex correlations between features is captured — a key ingredient for quantum advantage.

- The circuit starts with Hadamard gates on each qubit, followed by parameterized controlled-Z gates between pairs of qubits capturing pairwise feature interactions .

- For two qubits:

$\phi(x_1, x_2) = (\pi - x_1)(\pi - x_2)$



- The number of repetitions and the entanglement pattern can be customized, providing flexibility in the encoding process.

- The introduction of entanglement makes the ZZ feature map more expressive but also more vulnerable to quantum noise. (Linear , circular, or full entanglement)

(Havlíček et al., 2019)

References

- 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, pp. 625–644, Sep. 2021.
- M. Schuld and F. Petruccione, *Machine Learning with Quantum Computers*, 2nd ed. Cham, Switzerland: Springer, 2021.
- V. Havlíček, A. D. C ´orcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, “Supervised learning with quantum enhanced feature spaces,” Nature, vol. 567, no. 7747, pp. 209–212, 2019
- D. Ranga, A. Rana, S. Prajapat, P. Kumar, K. Kumar, and A. V. Vasilakos, “Quantum machine learning: Exploring the role of data encoding techniques, challenges, and future directions,” *Mathematics*, vol. 12, no. 21, p. 3318, Oct. 2024, doi: 10.3390/math12213318.
- 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, Oct. 2019, doi: 10.1088/2058-9565/ab4eb5.
- Qiskit textbook, https://qiskit.org/textbook/