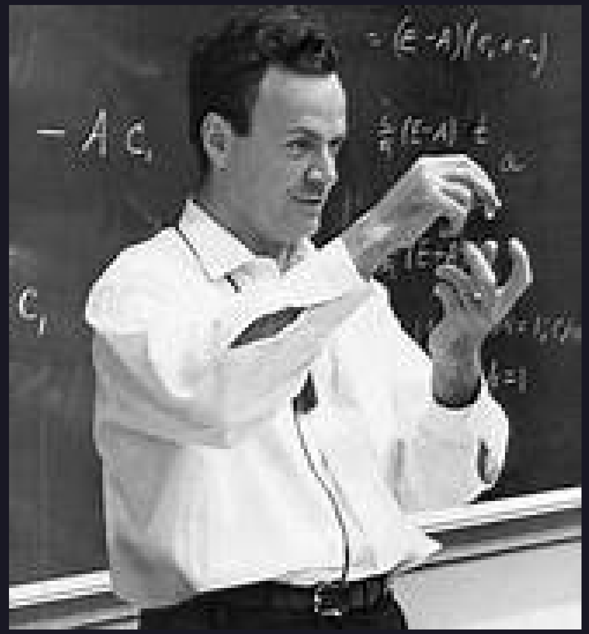
Quantum Machine Learning- An Overview

What is a quantum computer?

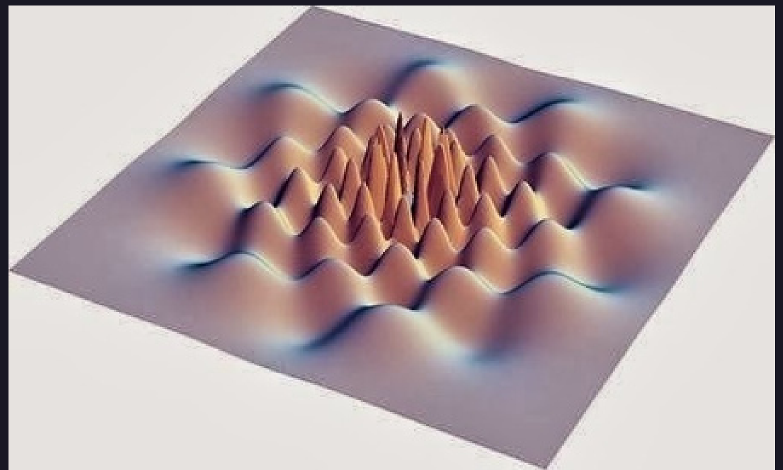
- 1982 - Feynman proposed the idea of creating machines based on the laws of quantum mechanics instead of the laws of classical physics.



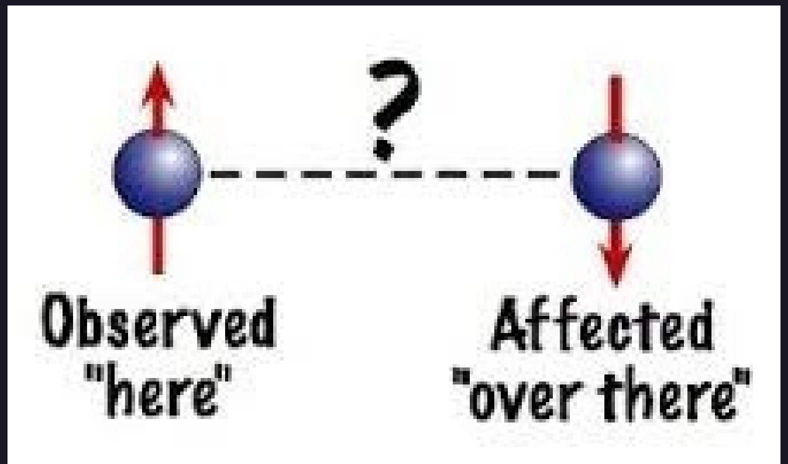
- A quantum computer is a machine that performs calculations based on the laws of quantum mechanics, which is the behavior of particles at the sub-atomic level.

- These laws are weird and counter-intuitive. “I think I can safely say that nobody understands quantum mechanics” - Feynman

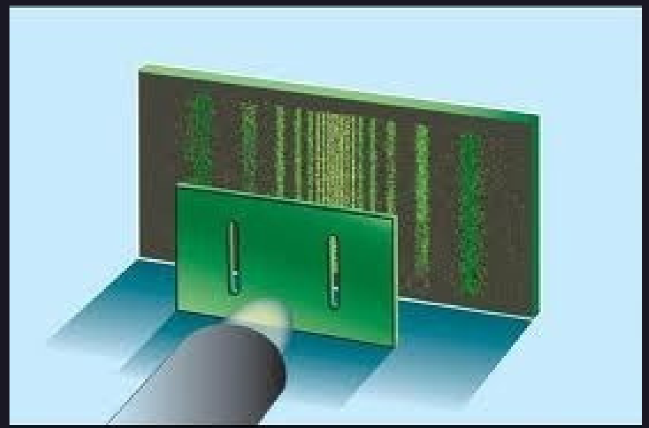
- *wave-particle duality*



- *quantum entanglement*



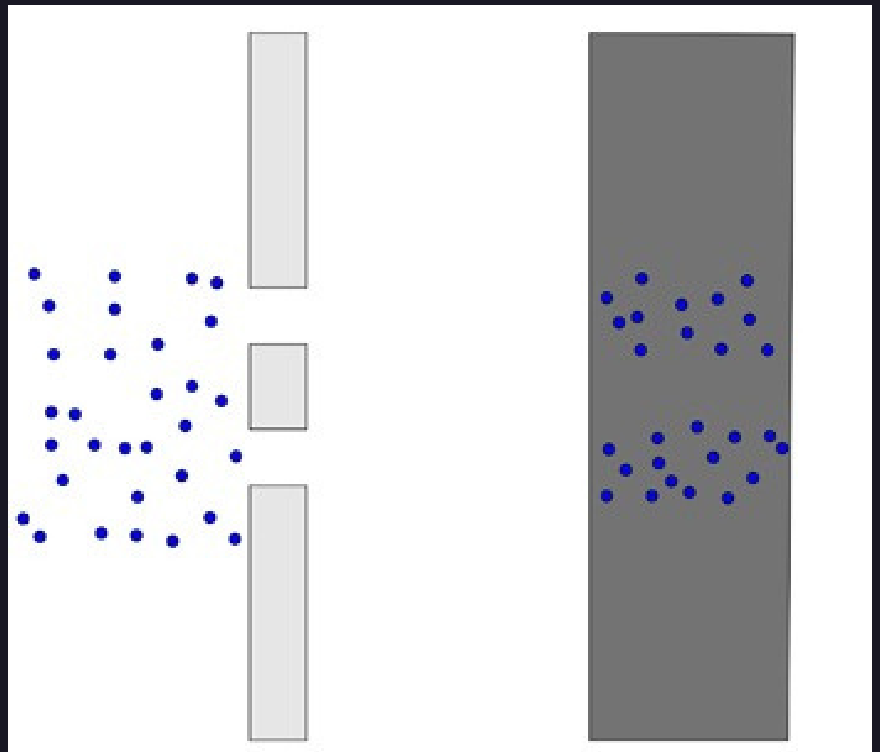
- *quantum super-position*



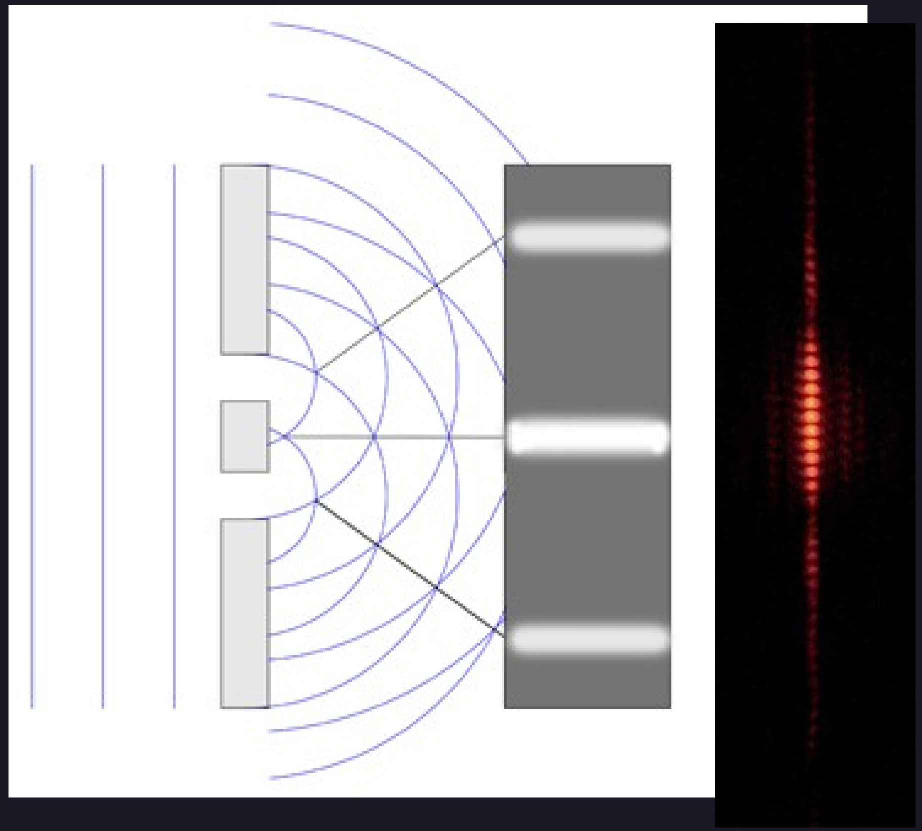
Double slit experiment

One of the most famous experiments in physics is the **double slit** experiment. It demonstrates, with unparalleled strangeness, that little particles of matter have something of a wave about them,

and suggests that the very act of observing a particle has a dramatic effect on its behavior.



The pattern you get from particles.

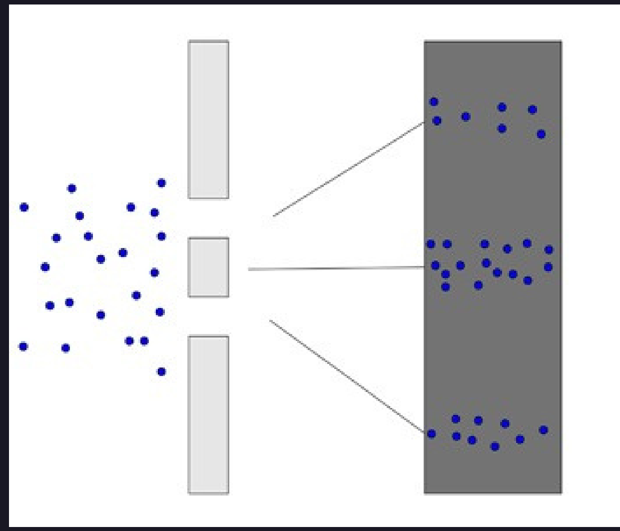


An interference pattern.

Let's go Quantum

Imagine firing electrons at our wall with the two slits, but **block** one of those slits off for the moment. You'll find that some of the electrons will pass through the open slit and strike the second wall **just as tennis balls would**: the spots they arrive at form a strip roughly the same shape as the slit.

Now **open the second slit**. You'd expect two rectangular strips on the second wall, as with the tennis balls, but what you actually see is very different: the spots where electrons hit build up to **replicate the interference pattern from a wave**.



An interference pattern.

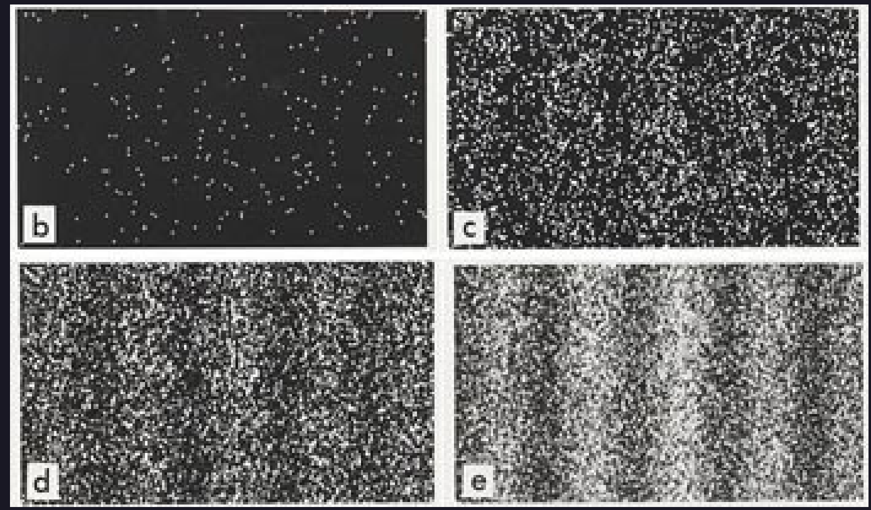


image of a real double slit experiment with electrons. The individual pictures show the pattern you get on the second wall as more and more electrons are fired.

One possibility might be that the electrons **somehow interfere with each other**, so they don't arrive in the same places they would if they were alone.

However, the interference pattern remains even when you fire the electrons one by one, so that they have no chance of interfering. Strangely, each individual electron contributes one dot to an overall pattern that looks like the interference pattern of a wave.

Could it be that each electrons somehow splits, passes through both slits at once, interferes with itself, and then recombines to meet the second screen as a single, localised particle?

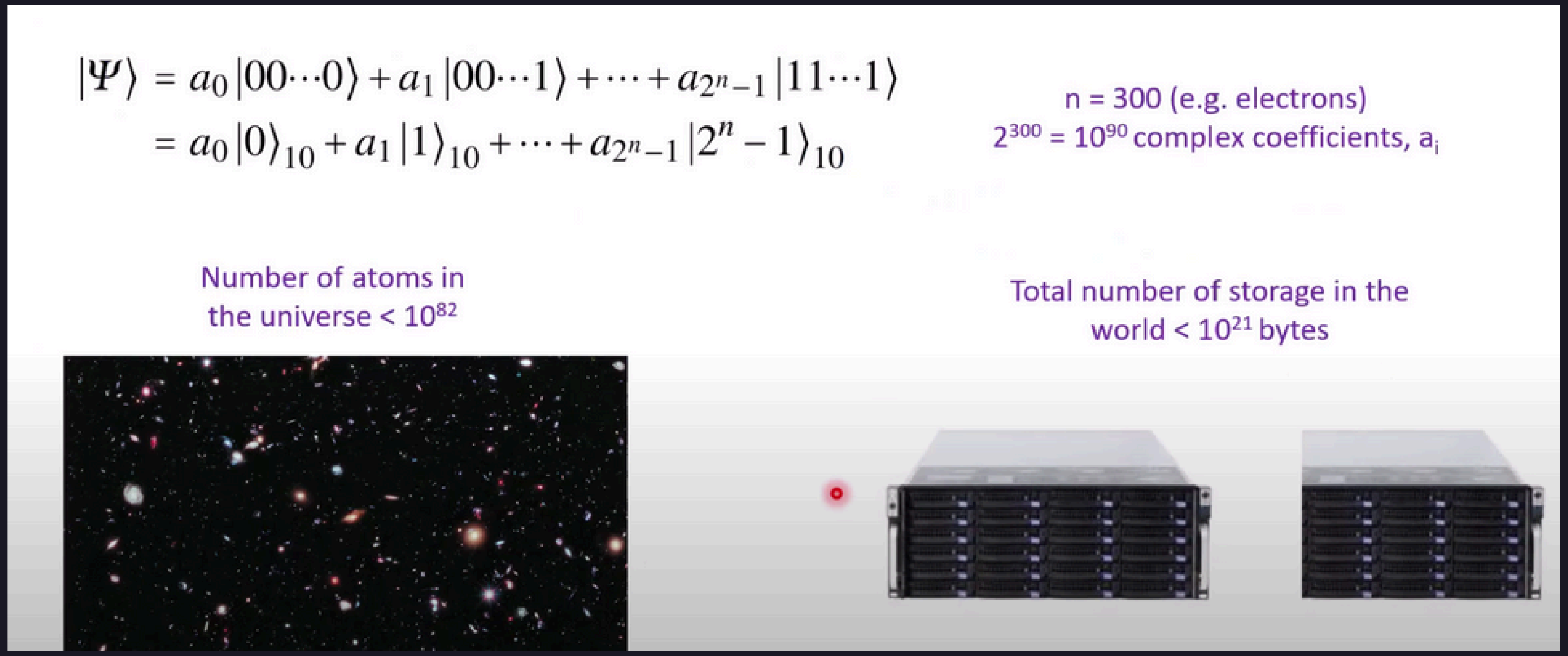
To find out, you might place a detector by the slits, to see which slit an electron passes through. And that's the really weird bit. If you do that, then the pattern on the detector screen turns into the particle pattern of two strips
Somehow, the very act of looking makes sure that the electrons travel like well- behaved little tennis balls. It's as if they knew they were being spied on and decided not to be caught in the act of performing weird quantum shenanigans.

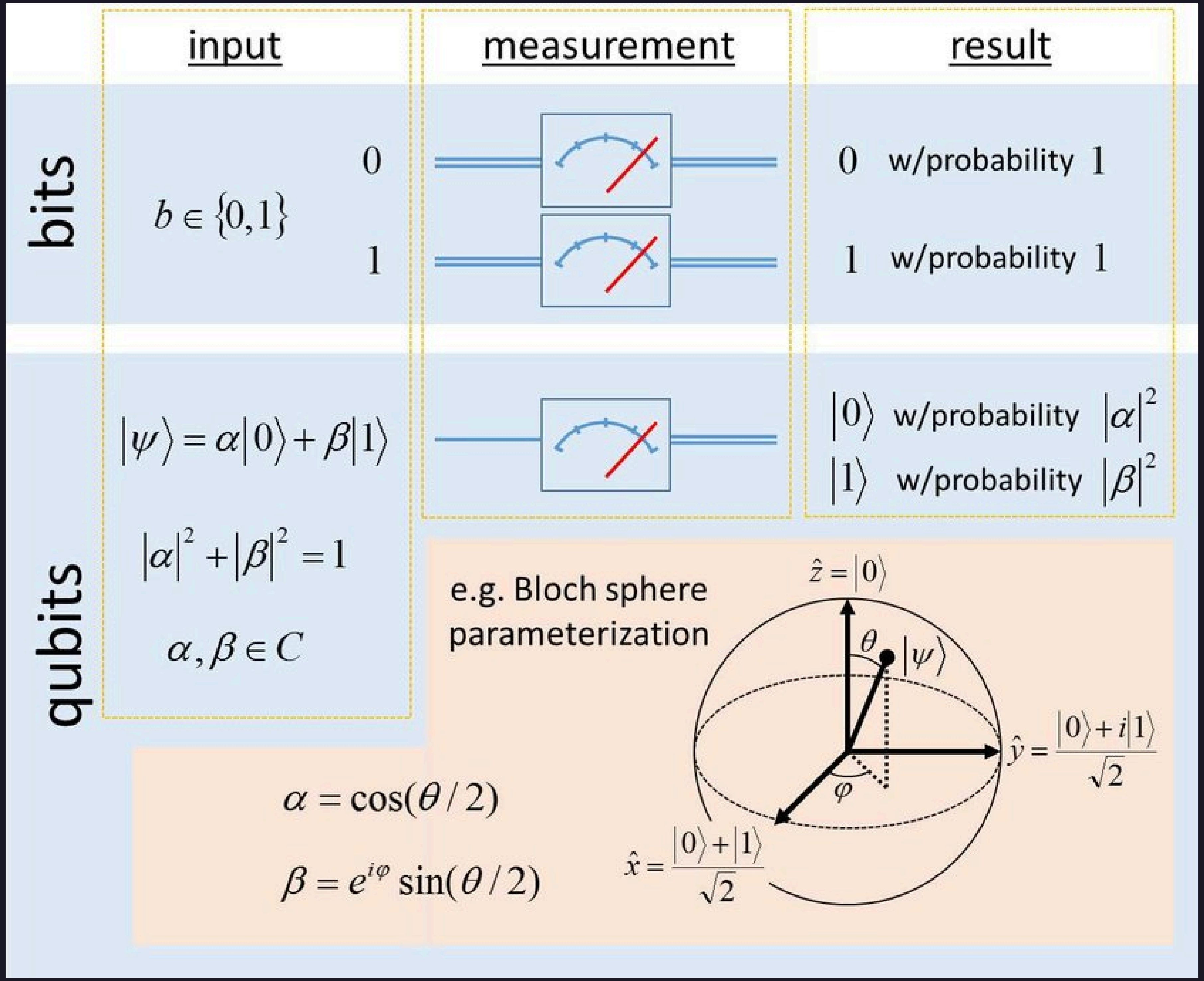
That's the famous *wave particle duality* of quantum mechanics

CONCEPTUAL FRAMEWORK

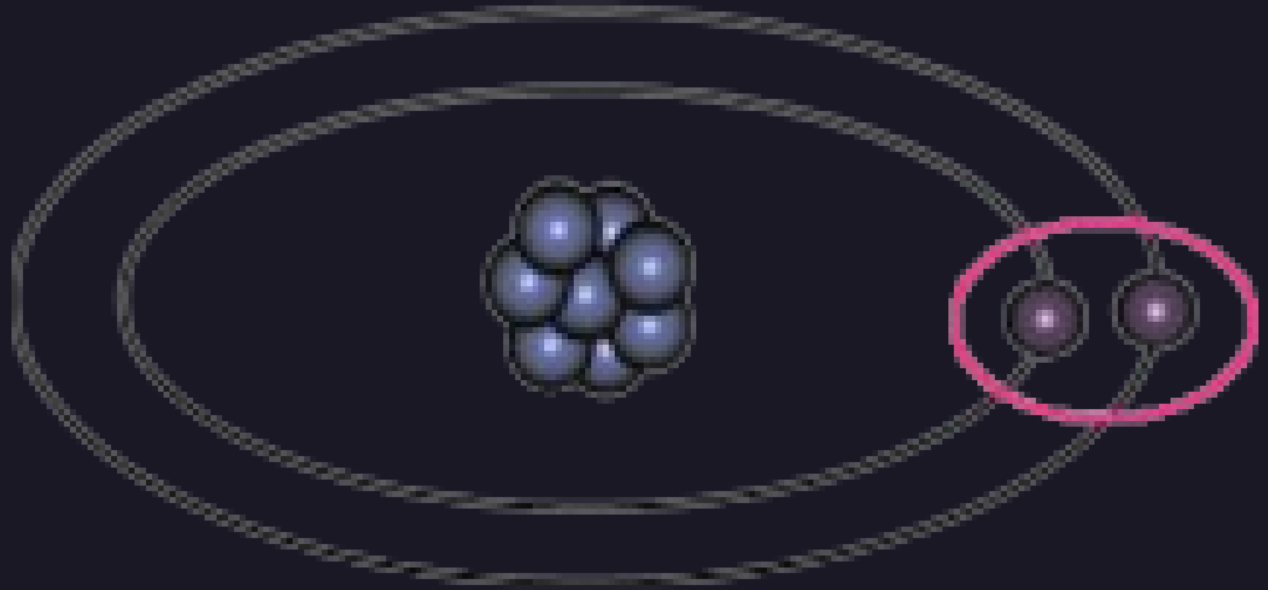
The term "qubit" is attributed to American theoretical physicist Benjamin Schumacher. Qubits are generally, although not exclusively, created by manipulating and measuring quantum particles such as photons, electrons, trapped ions, superconducting circuits, and atoms.

Enabled by the unique properties of quantum mechanics, quantum computers use qubits to store more data than traditional bits, vastly improve cryptographic systems, and perform very advanced computations that would take thousands of years (or be impossible) for even classical supercomputers to complete.





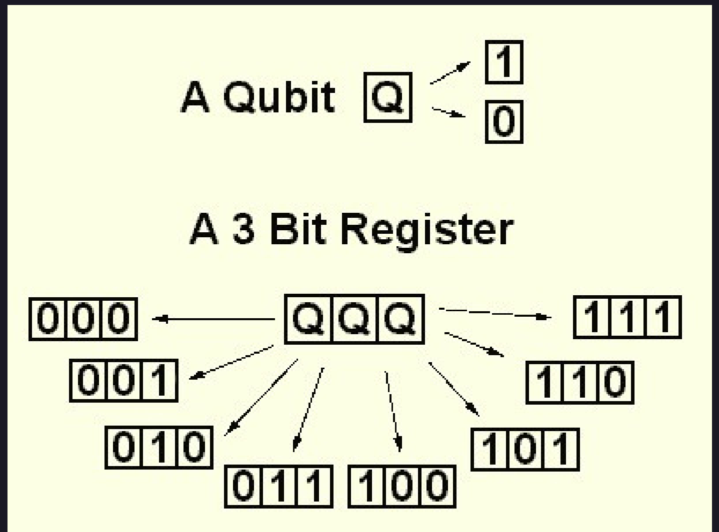
Quantum Superposition



State |0> +|1>

Electrons have a wave property which allows a single electron to be in two orbits simultaneously. In other words, the electron can be in a superposition of both orbits

qubits	*stores simultaneously*	*total number*
1	(0 and 1)	$2^1 = 2$
2	(0 and 1)(0 and 1)	$2x2 = 2^2 = 4$
3	(0 and 1)(0 and 1)(0 and 1)	$2x2x2 = 2^3 = 8$
:	:	:
300	(0 and 1)(0 and 1)........(0 and 1)	$2x2......x2 = 2^{300}$

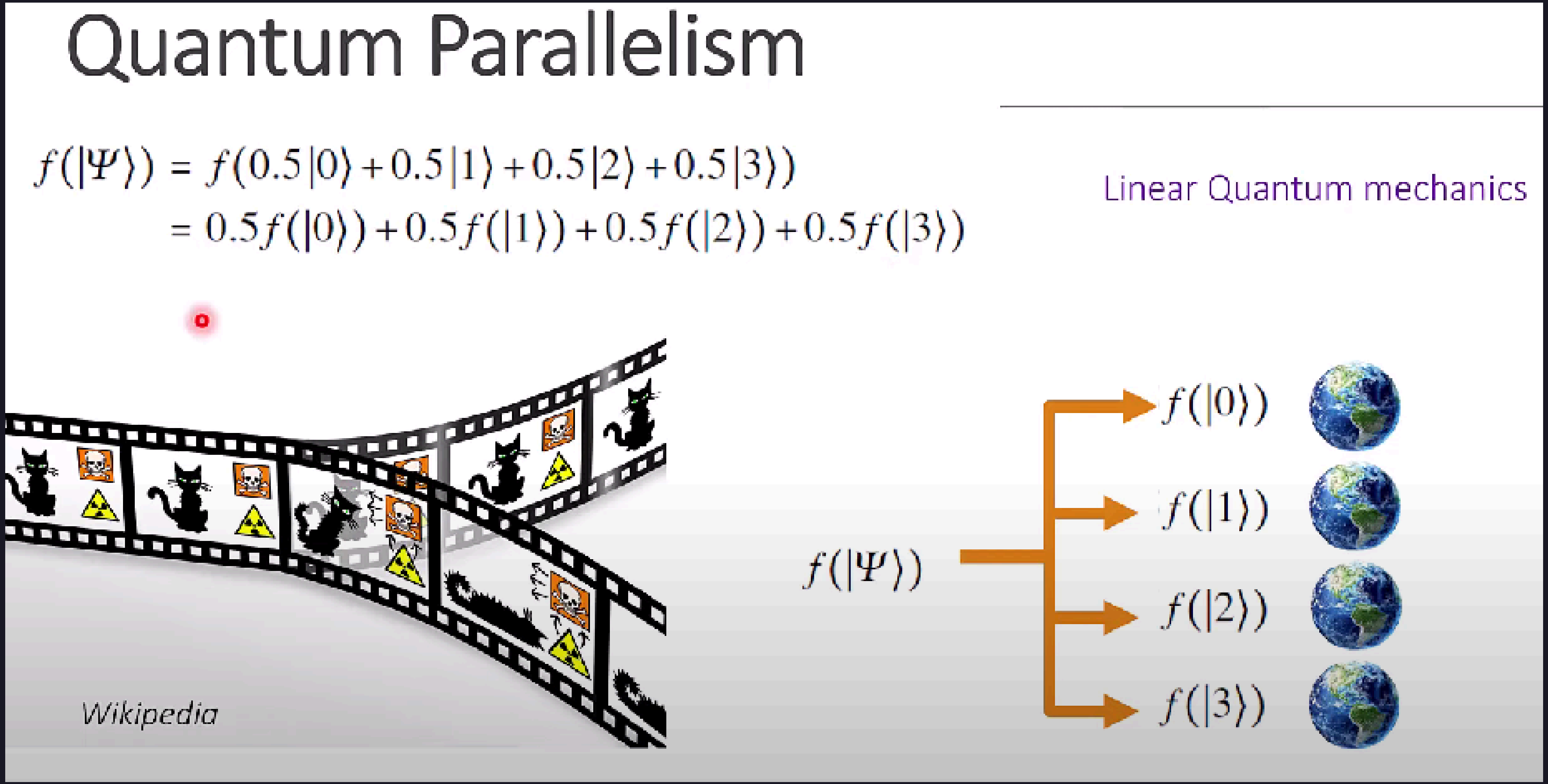


Bigger number than the number of atoms in the universe, and calculations can be performed simultaneously on each of these numbers

For every extra qubit you get, you can store twice as many numbers

QUANTUM PARALALISM

Classical computing requires bit doubling, but QC merely requires a bit increase to expand computational space.
Qubits are the fundamental unit of information in a quantum computer, and they must be isolated and managed. They can be in a condition of superposition until they are isolated.
Quantum computation is intrinsically parallel, since it deals with quantum states and their superposition. It relies on Qubits forming a basis of states that scales up with 2 n, where n is the number of Qubits employed.
The Qubits itself can be created by using two level systems ($|0\rangle$ or $|1\rangle$), such as light polarization, electronic spin, atomic nuclear spin, optical lattices, quantum dots, superconductor circuits, and others.



QUANTUM ENTANGLEMENT

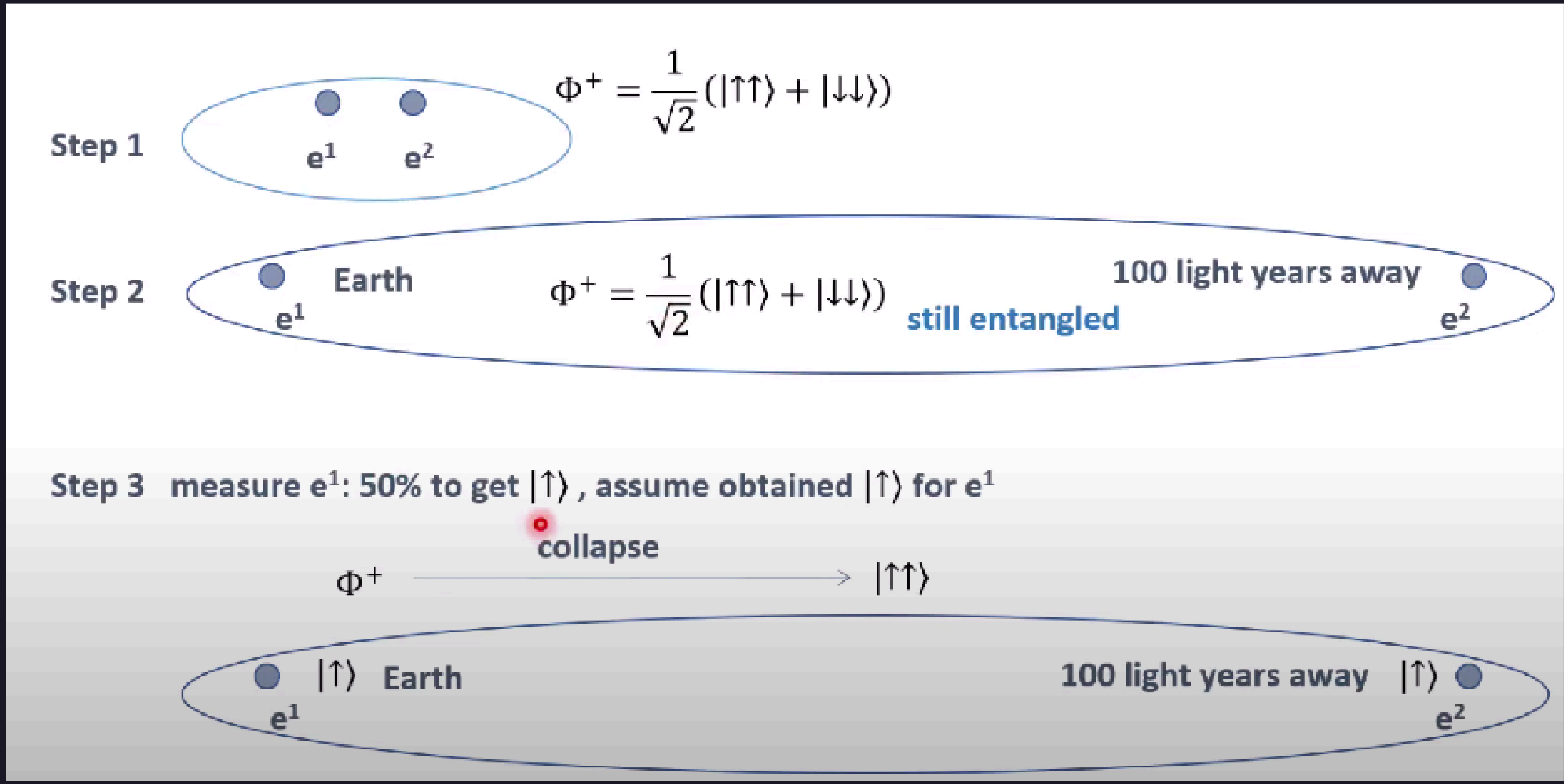
Quantum entanglement is the new way of communication with different approach from traditional communication.

Unentangled State

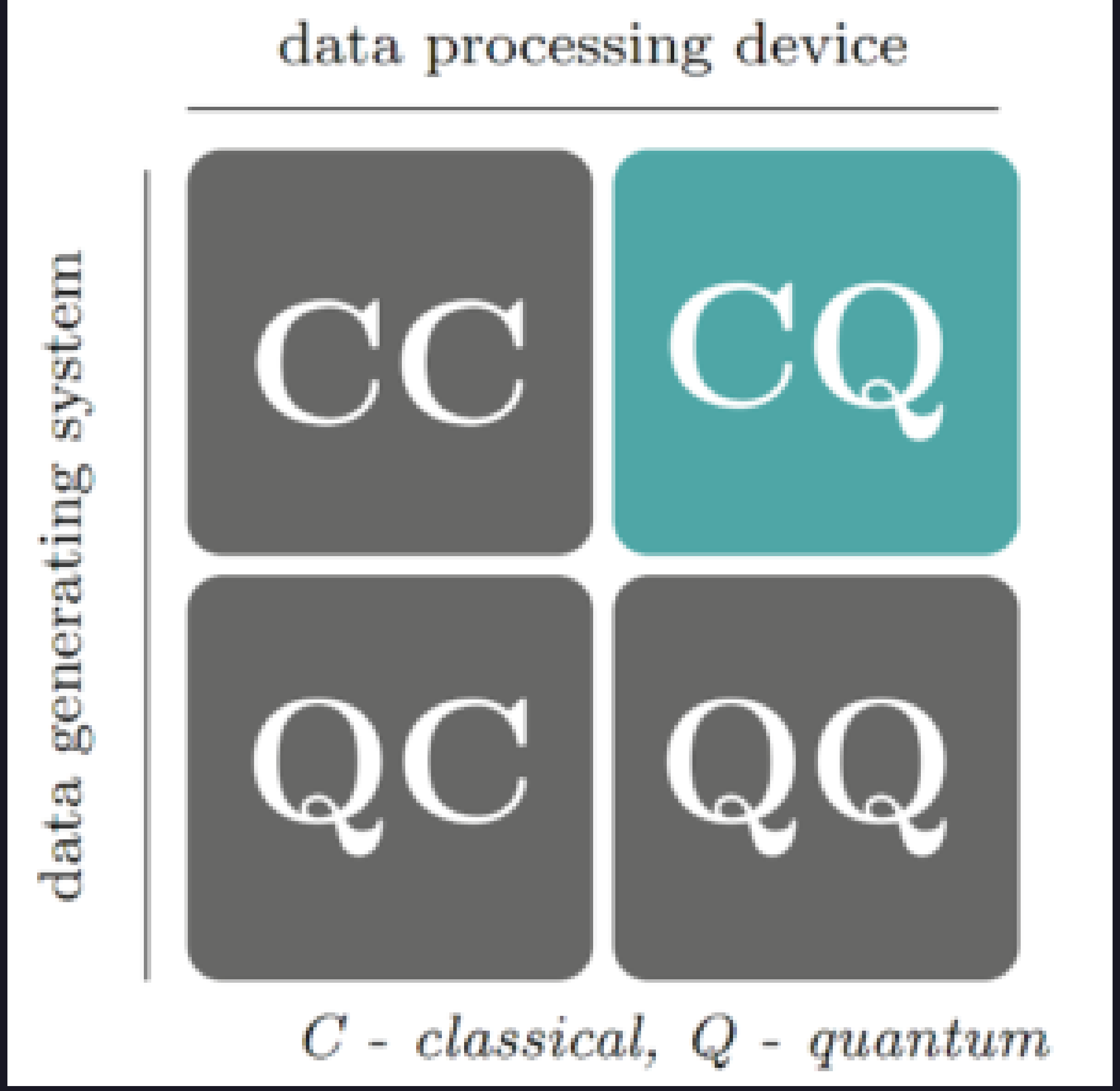
$$\frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) = \underbrace{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)}_{\text{Electron 1}} \otimes \underbrace{\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)}_{\text{Electron 2}}$$

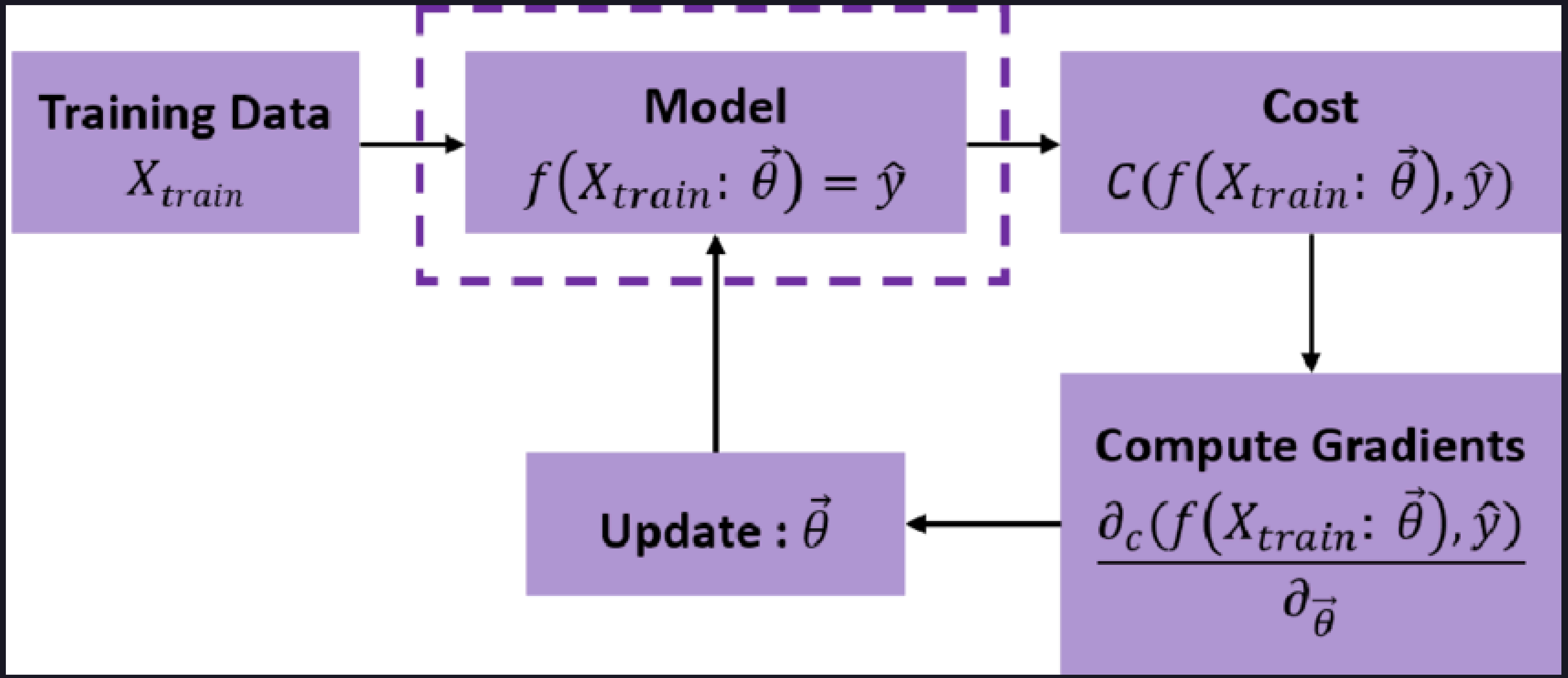
Entangled State: Used in quantum computing algorithms and also quantum communications

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \neq |Electron\ 1\rangle \otimes |Electron\ 2\rangle$$



MACHINE LEARNNG AND QUANTUM COMPUTING



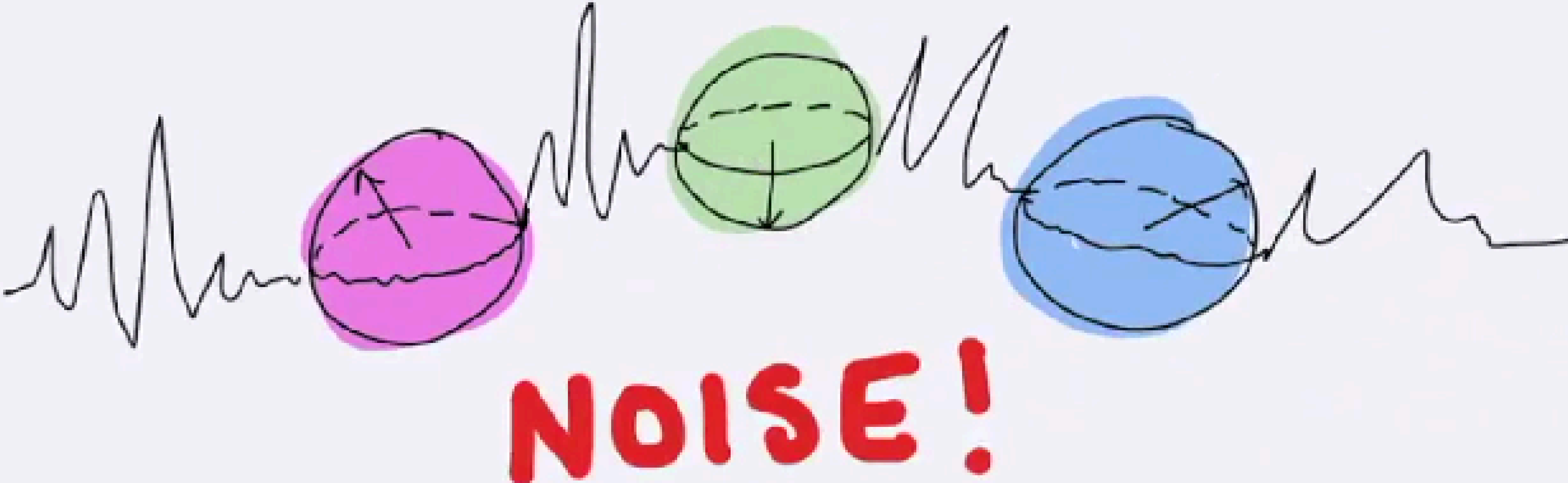


NEAR TERM & FAULT TOLERANT SYSTEMS

We mostly have near-term versus fault-tolerant quantum computation and quantum computers available to us. So the near-term devices are those that we have available today, right, those are the smaller devices that are typically noisier, these devices suffer from a lot of errors. And then the opposite end of that is the fault-tolerant machine and so when people design and think about quantum machine learning algorithms, there are really kind of two parties. The first try to design and run and optimize algorithms that run on these near-term devices,

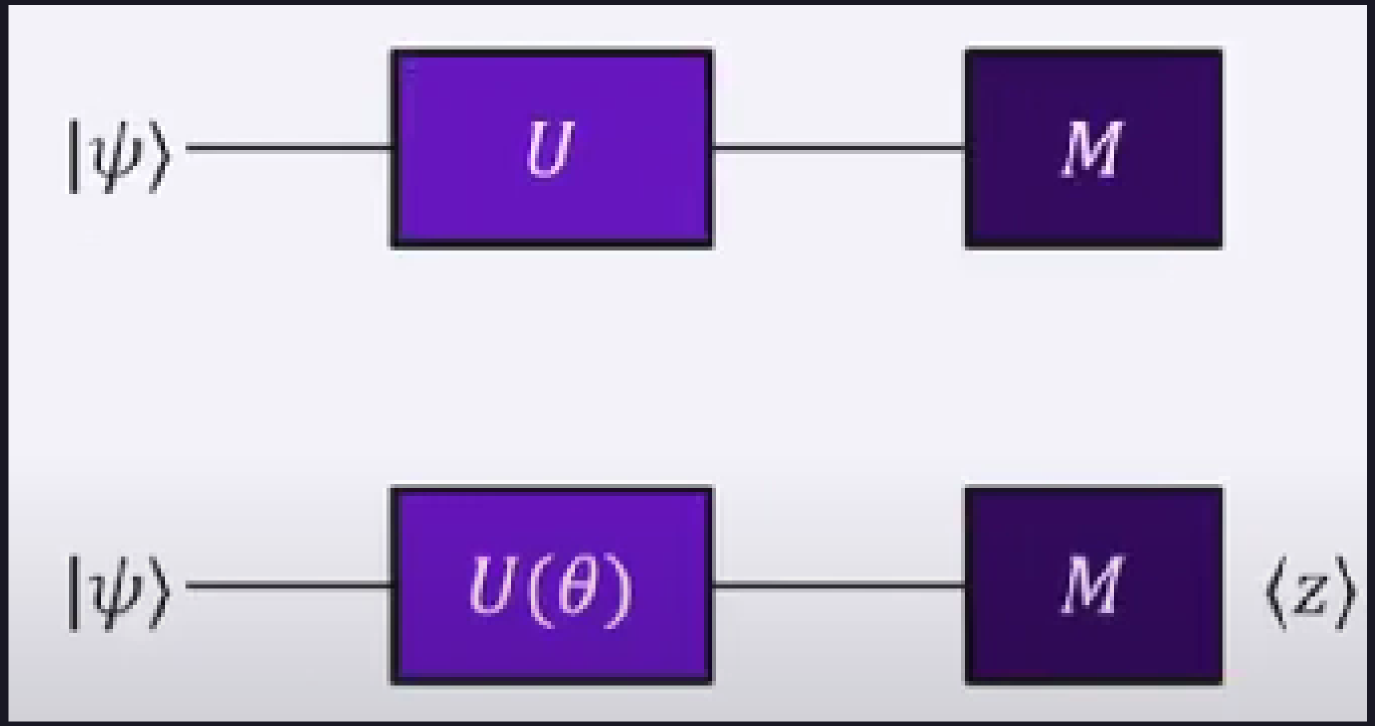
Noisy, error-prone, small devices

What can we do now?



VARIATIONAL MODELS

Variational models can very, very easily be thought of as just a quantum circuit that contains some parameters in it that we want to train and optimize and tweak.

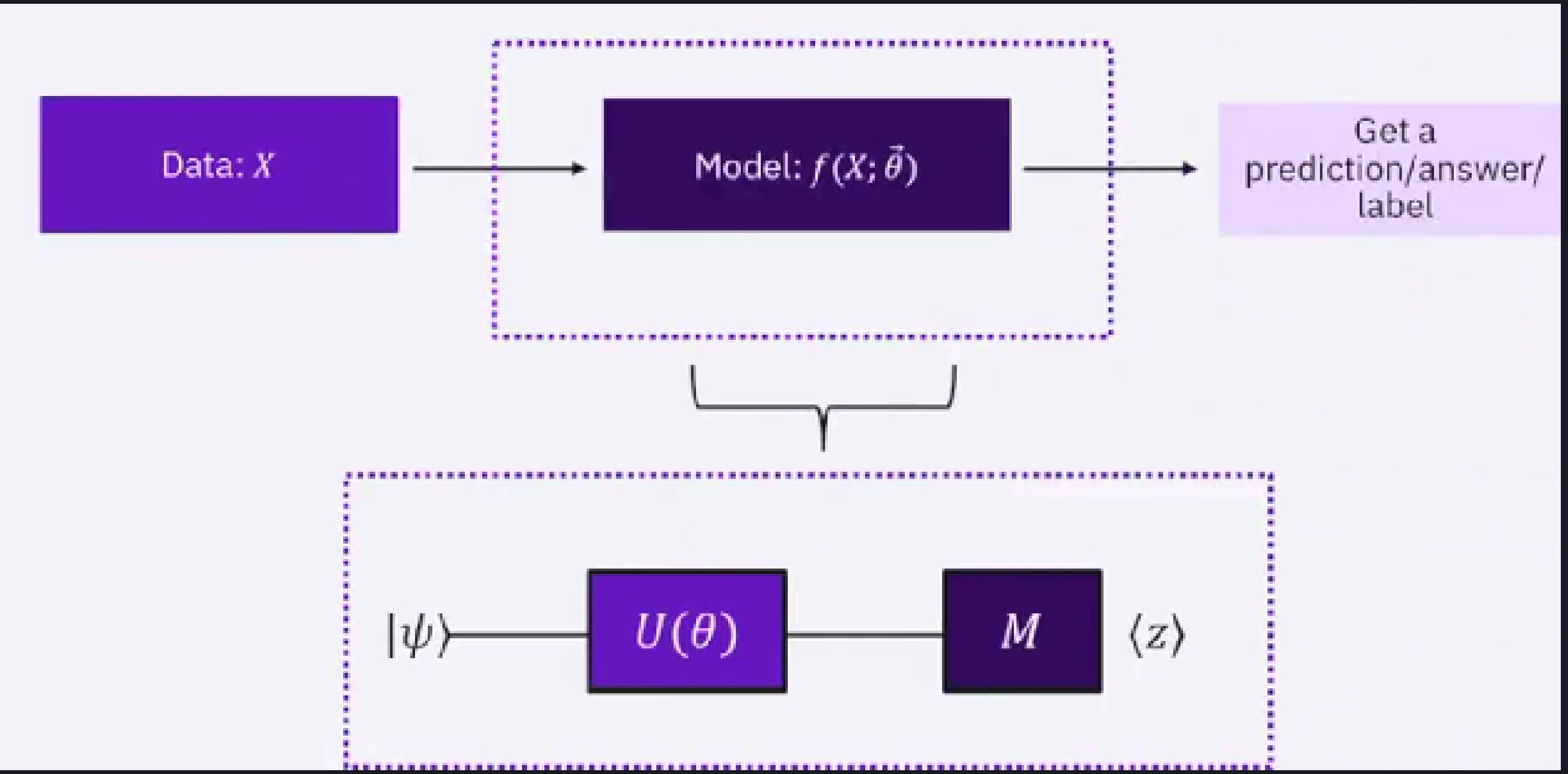


can depend on some parameters, they can depend on parameters that we can specify, that we can train and we can optimize.

when we measure quantum systems, the output is stochastic in nature, and so typically what we do is we repeat measurements multiple times to get an expectation value of the output, so we get a probability distribution of possible basis states when we do measurements of a quantum circuit.

VARIATIONAL CIRCUIT AS A CLASSIFIER

TASK: Train a quantum variational circuit, on labeled samples, so that we can predict labels for new data,



Step1: Encode Data

Step2: Apply Paramiterised Model

Step3: Measure to extract Labels

Step4: Apply Optimisation Techniques to Update Model Parameters

ENCODING

Basis Encoding

Basis encoding is a straightforward method of representing classical data in a quantum system. Here, classical data is encoded into computational basis states of the quantum system.

$$|00\rangle\,,|01\rangle\,,|10\rangle\,,|11\rangle$$

For $x1 = 01$, the quantum state is $|01\rangle$.

For $x2 = 11$, the quantum state is $|11\rangle$.

The quantum circuit is set up such that the data points are encoded into their respective basis states.
The likelihood of other basis states is set to zero, ensuring the encoded data corresponds precisely to the desired states.

Amplitude Encoding

In classical computing, data is typically represented in binary form. However, in quantum computing, we can use the amplitudes of quantum states to encode classical data. This is known as amplitude encoding.

$$|\psi\rangle = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Here, $a_0, a_1, a_2,$ and a_3 are the probability amplitudes.

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \xrightarrow{U(\beta)} |\psi\rangle = \begin{pmatrix} 0.8 \\ 0.1 \\ 0 \\ 0.1 \end{pmatrix}$$

Angle encoding

Angle encoding is a method to represent classical data using the rotation angles of qubits. This technique encodes each feature of the classical data as a rotation angle applied to a corresponding qubit.

First Qubit: Rotate by θ_1 (dependent on $x1$)

$$R_z(\theta_1)\,|0\rangle \rightarrow |\psi_1\rangle$$

Second Qubit: Rotate by θ_2 (dependent on $x2$)

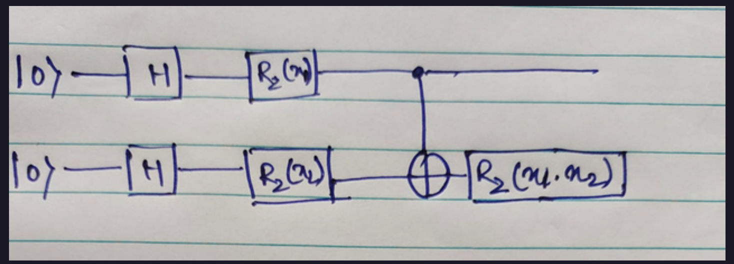
$$R_z(\theta_2)\,|0\rangle \rightarrow |\psi_2\rangle$$

Quantum State After Encoding:

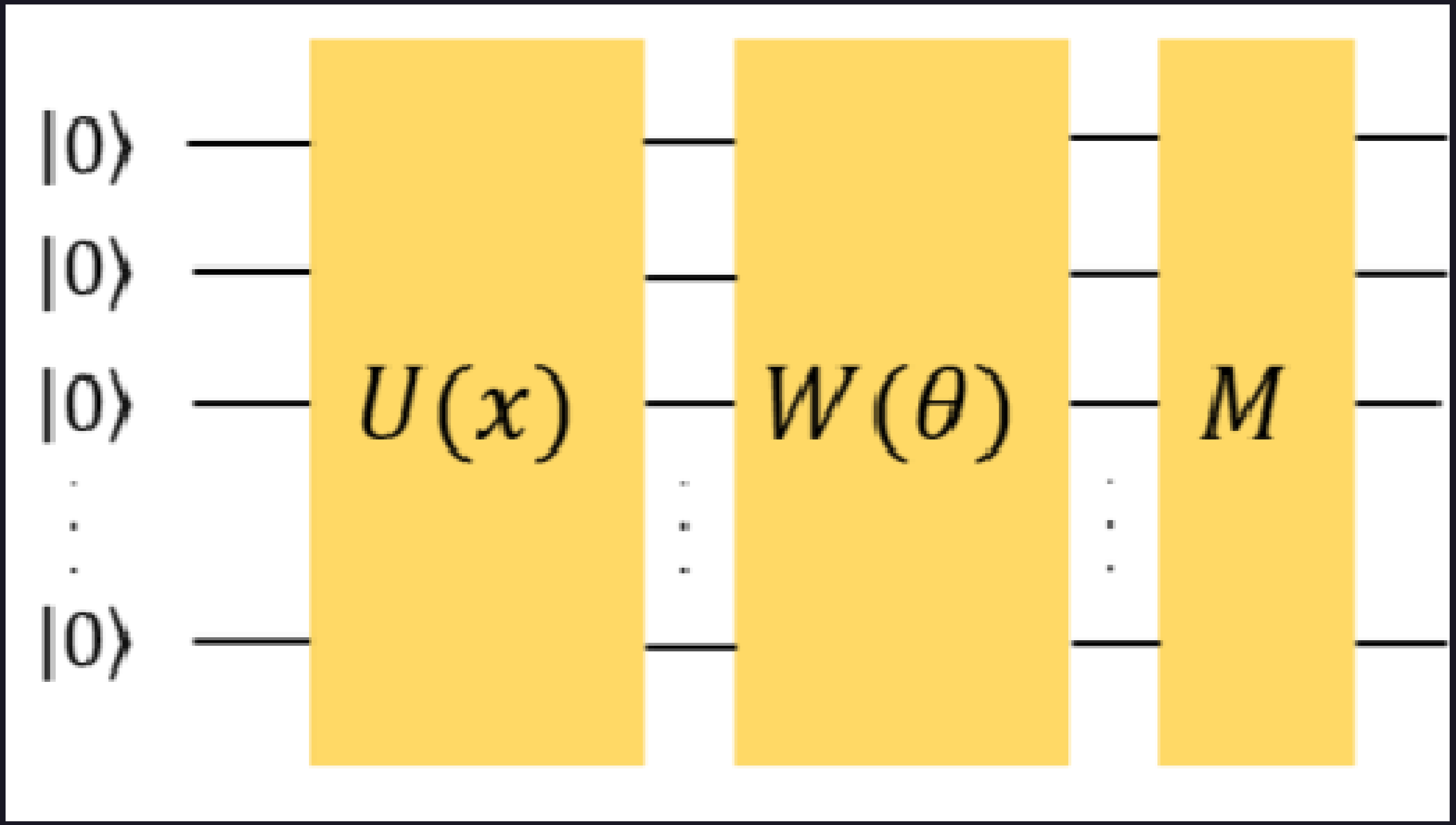
$$|\psi\rangle = R_z(\theta_1)\,|0\rangle \otimes R_z(\theta_2)\,|0\rangle$$

Higher Order Encoding

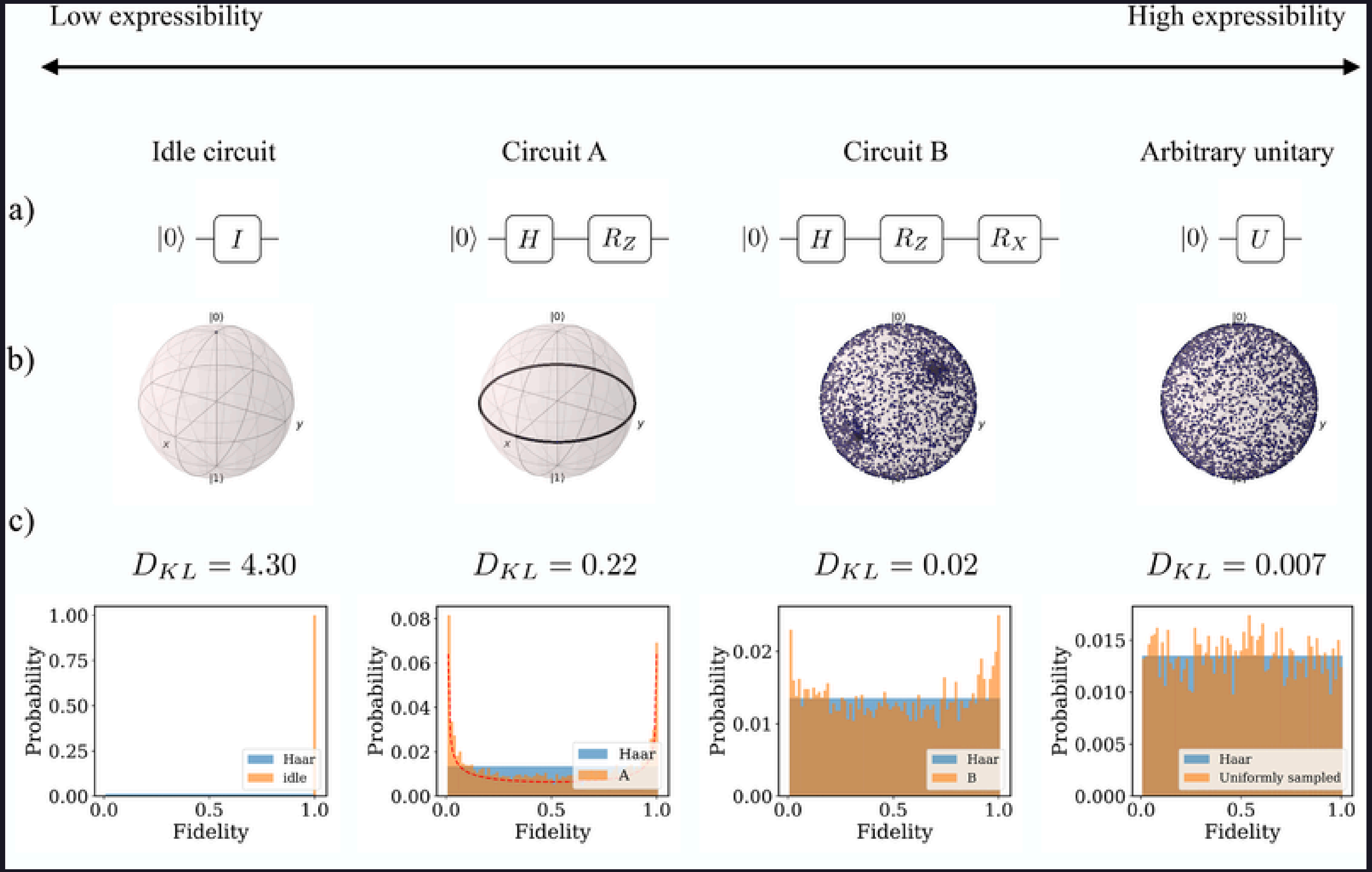
Higher-order encoding builds upon the principles of angle encoding but introduces additional steps to encode more complex relationships within the data, such as the product of feature values.

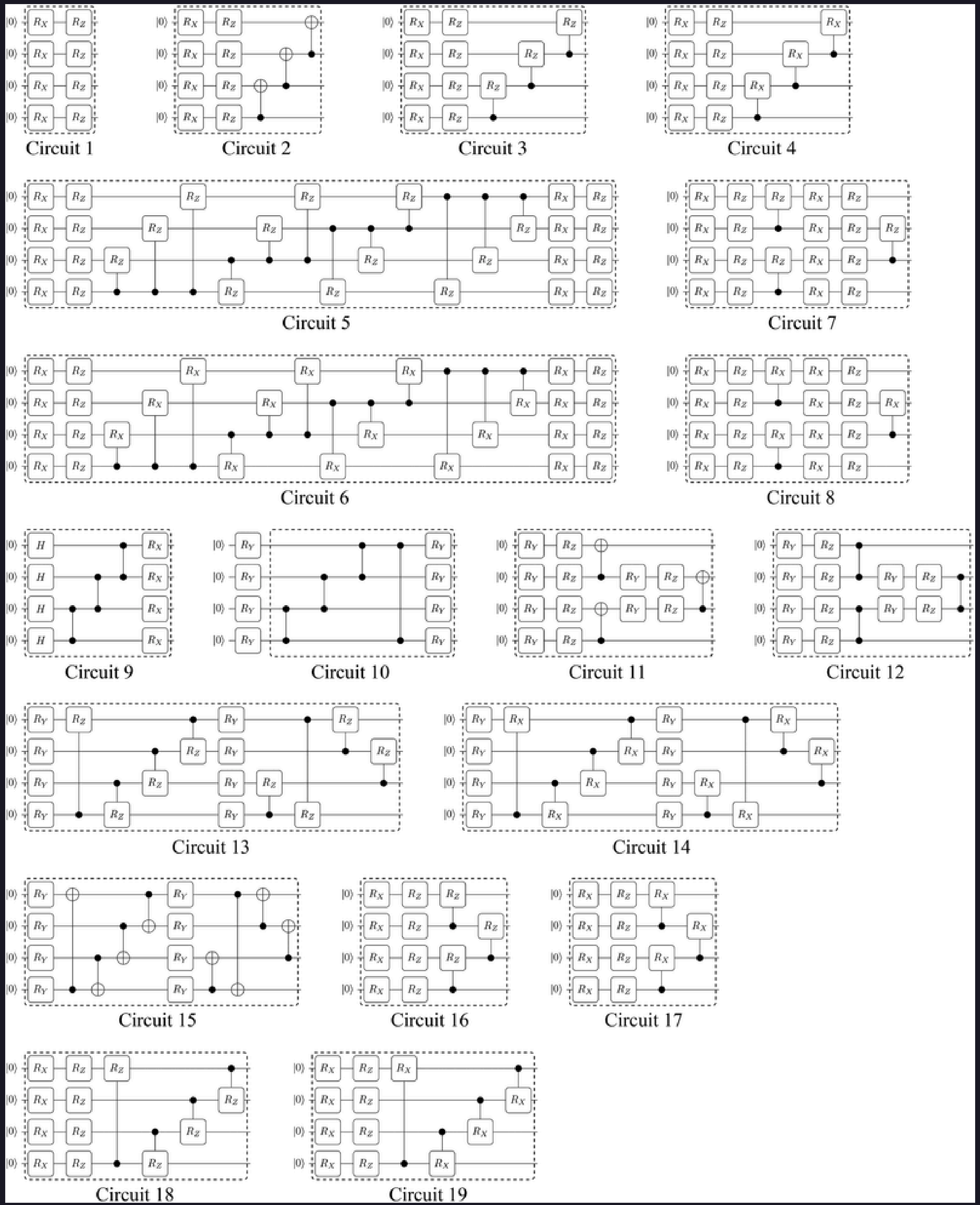


Applying Variational Model

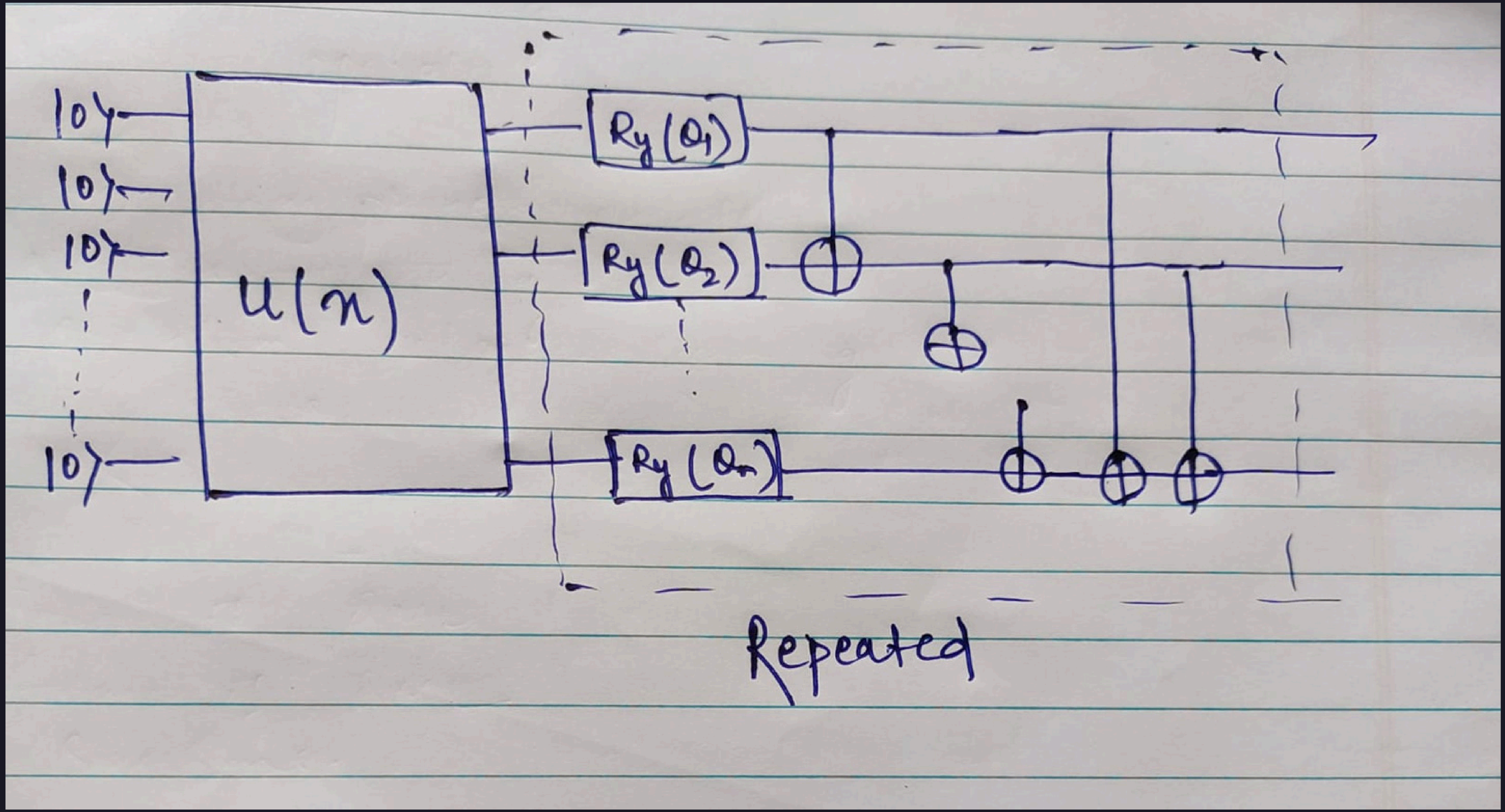


Applying Variational Model





Applying Variational Model



Extracting Labels

Parity Post-Processing

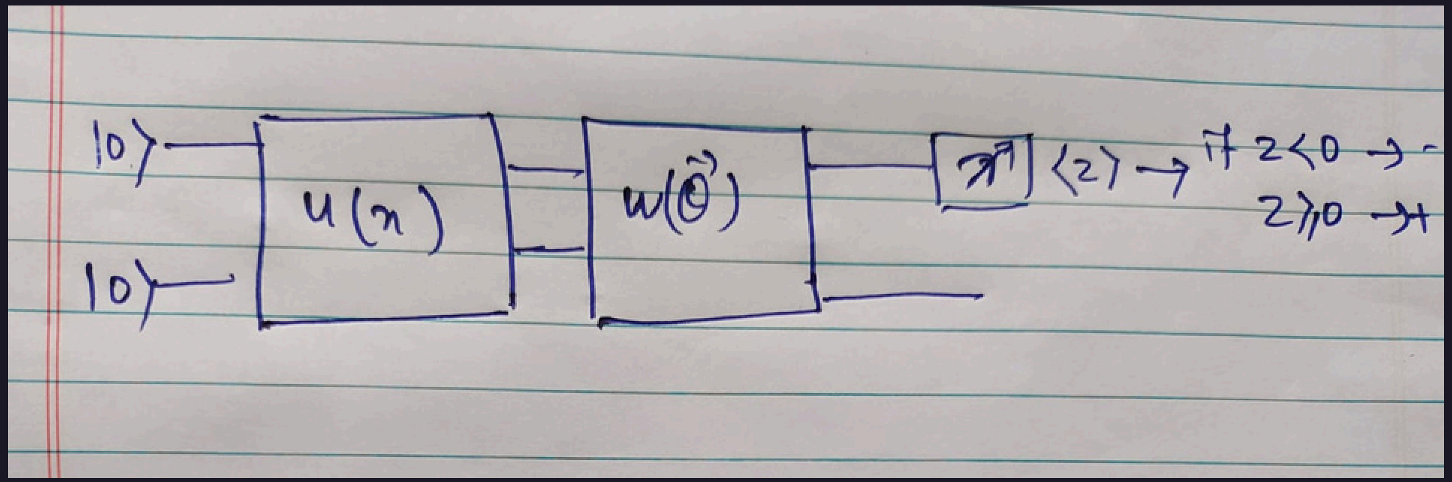
Based on the parity of the basis states, probabilities are summed to classify the data point into class 1 or class -1.

Possible Basis States:00,01,10,11

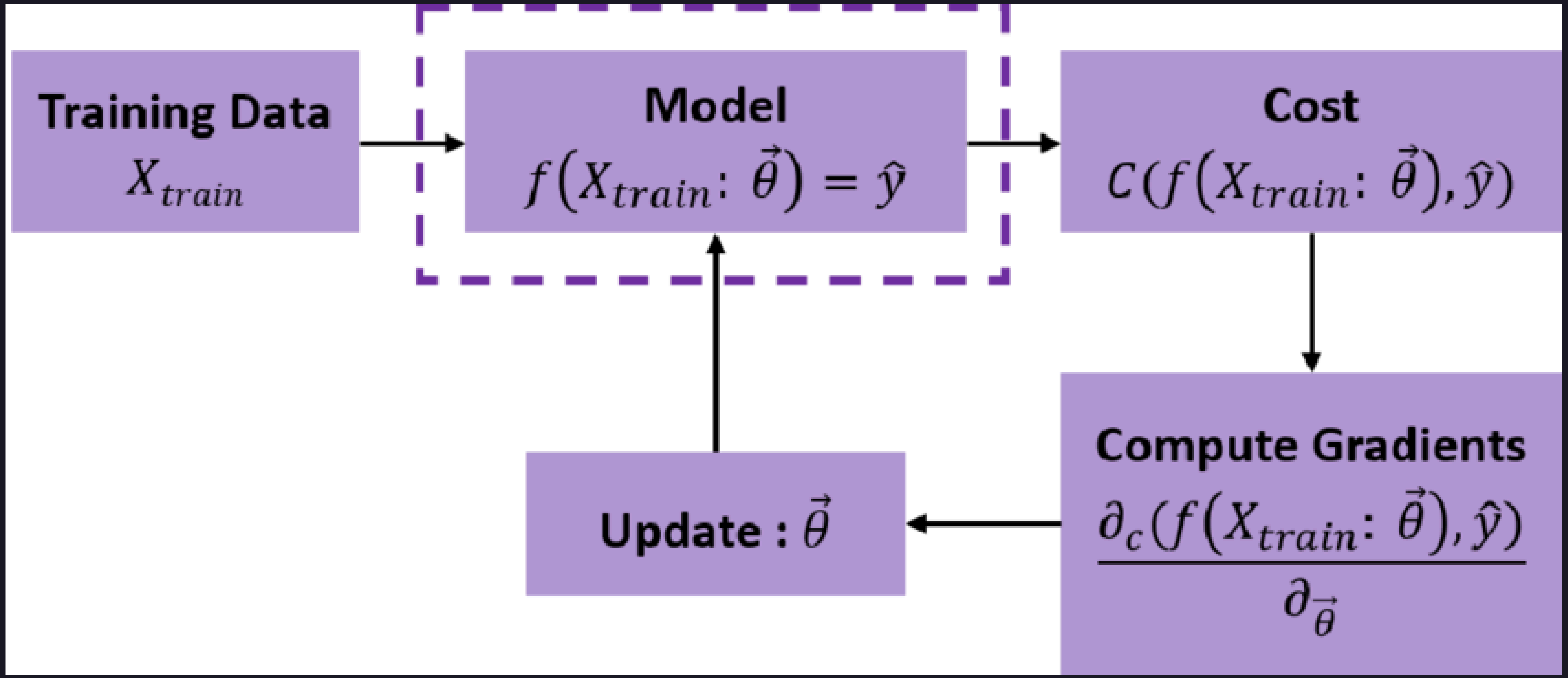
Measurement Probabilities:
Probability of 00: 0.8
Probability of 01: 0.1
Probability of 10: 0.0
Probability of 11: 0.1

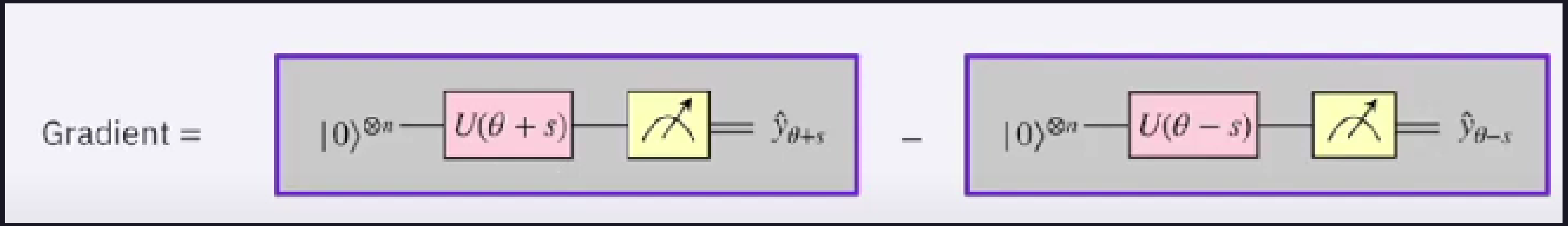
Probability Summation:
Probability of class 1:
Even parity states: 00 (0.8) + 11 (0.1) = 0.9
Probability of class -1:
Odd parity states: 01 (0.1) + 10 (0.0) = 0.1

Measuring First Qubit



Optimisation





Benifit

Assume we have two-dimensional classical data: $\vec{x} = [x_1, x_2]$.

To encode this data into a quantum state using n qubits:

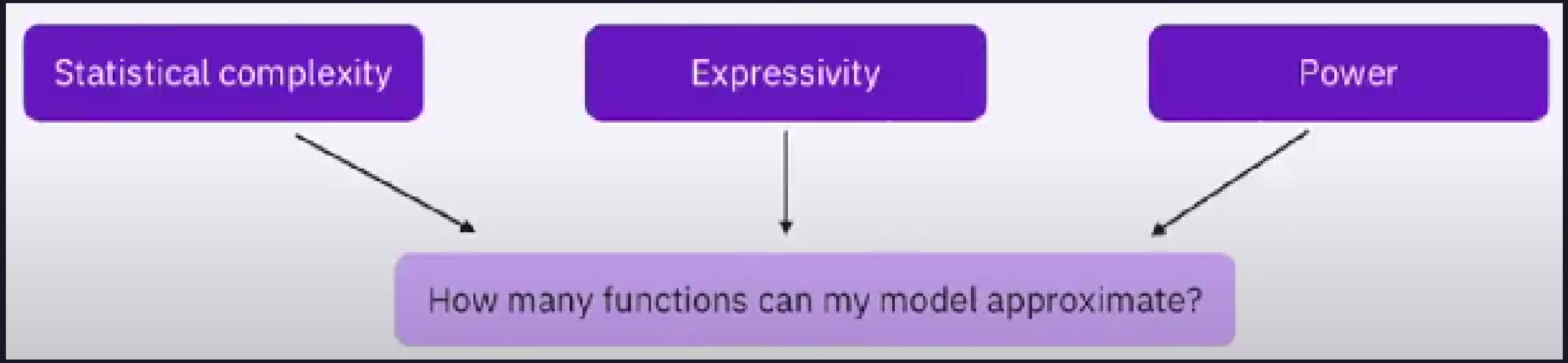
The quantum state $|\psi_x\rangle$ can be represented as a vector with 2^n entries.

For example, using $n = 2$ qubits, the quantum state $|\psi_x\rangle$ would be a four-dimensional vector.

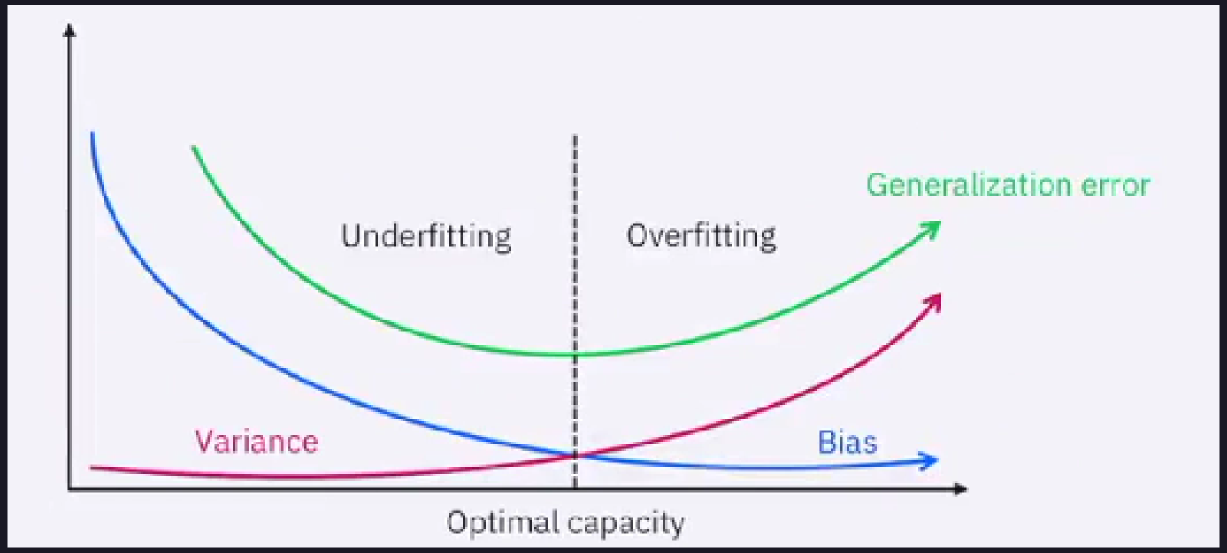
Original data: $\vec{x} = [x_1, x_2]$

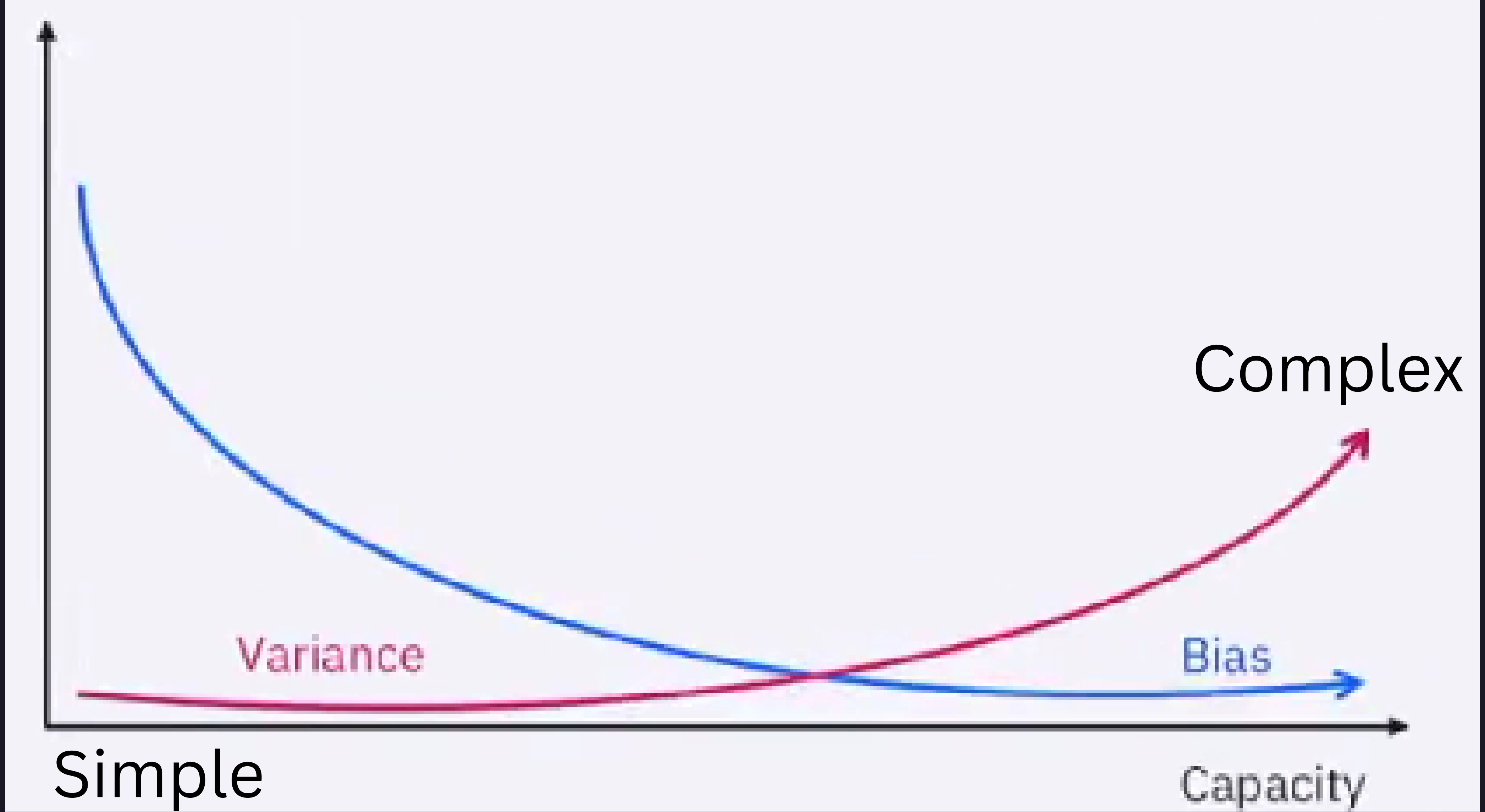
Quantum state: $|\psi_x\rangle = [x'_1, x'_2, x'_3, x'_4]$

CAPACITY AND POWER OF QML



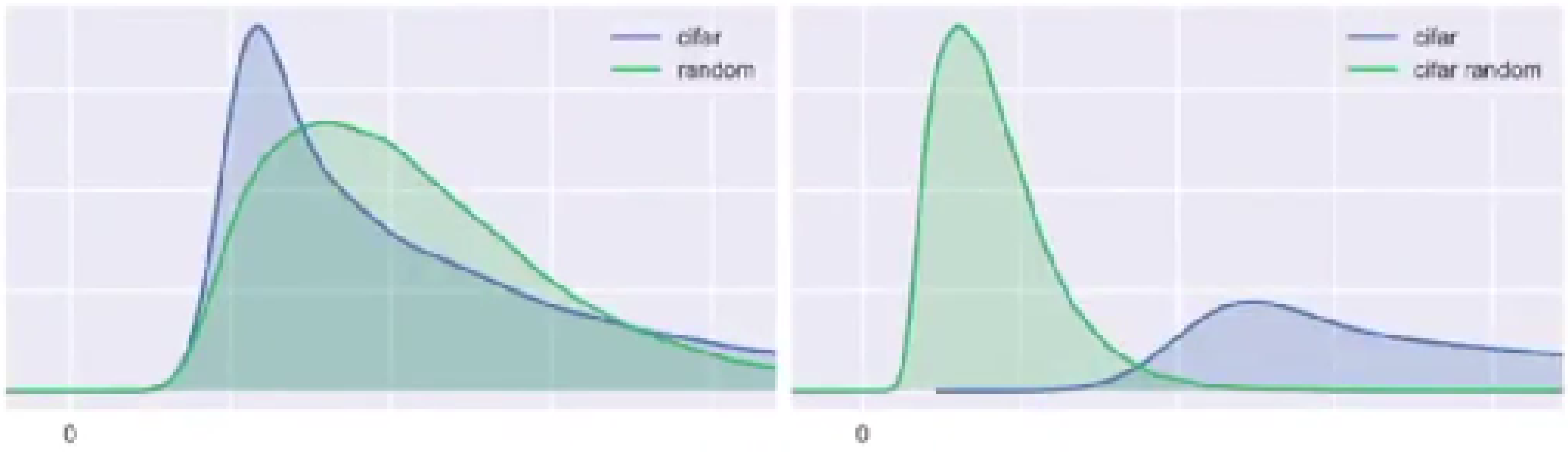
Generalisation:





Margins:

Looking at the margin alone, is not sufficient *(Bartlett, P., Foster, D. J., & Telgarsky, M. (2017). Spectrally-normalized margin bounds for neural networks. arXiv preprint arXiv:1706.08498.)*

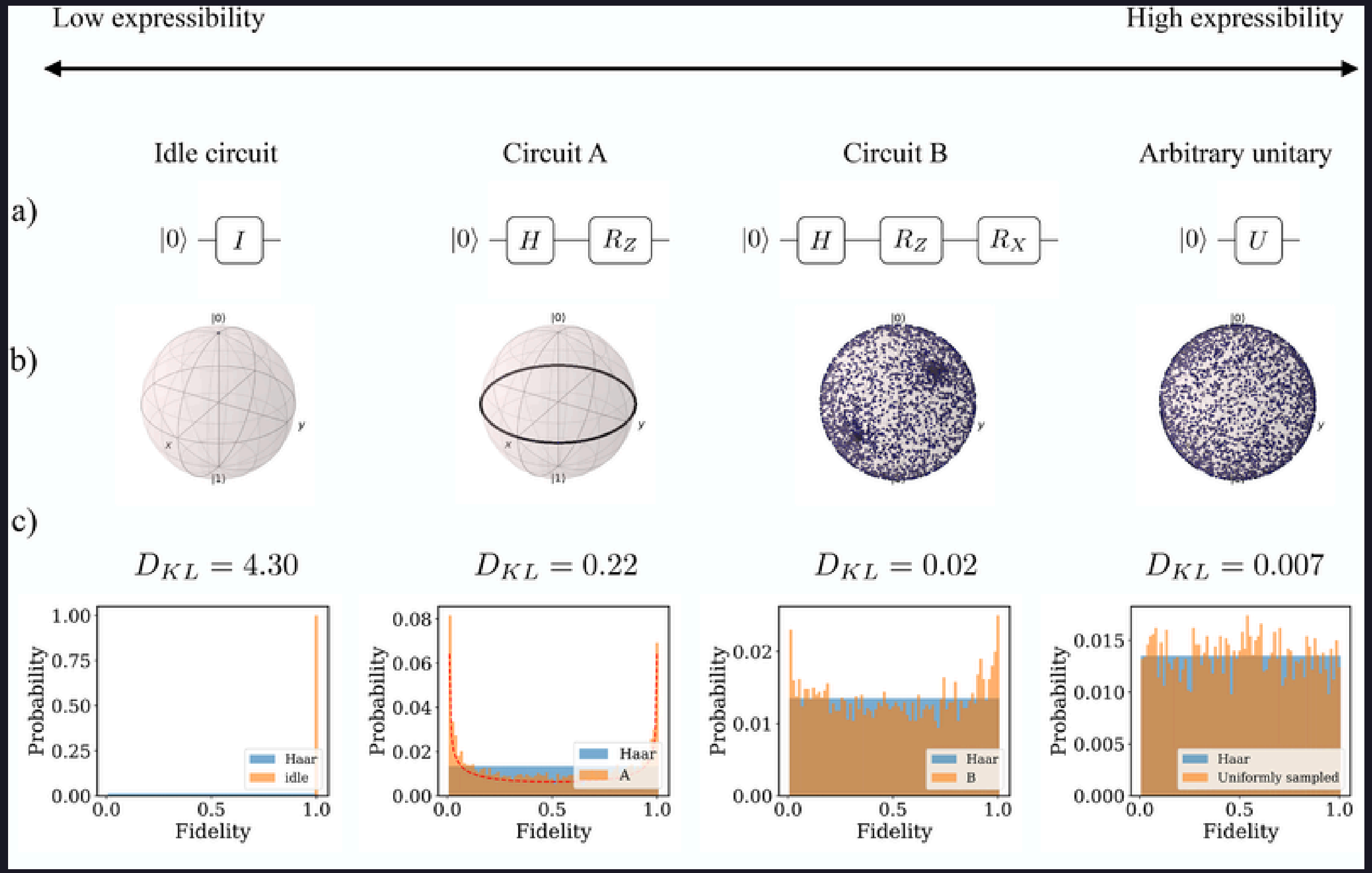


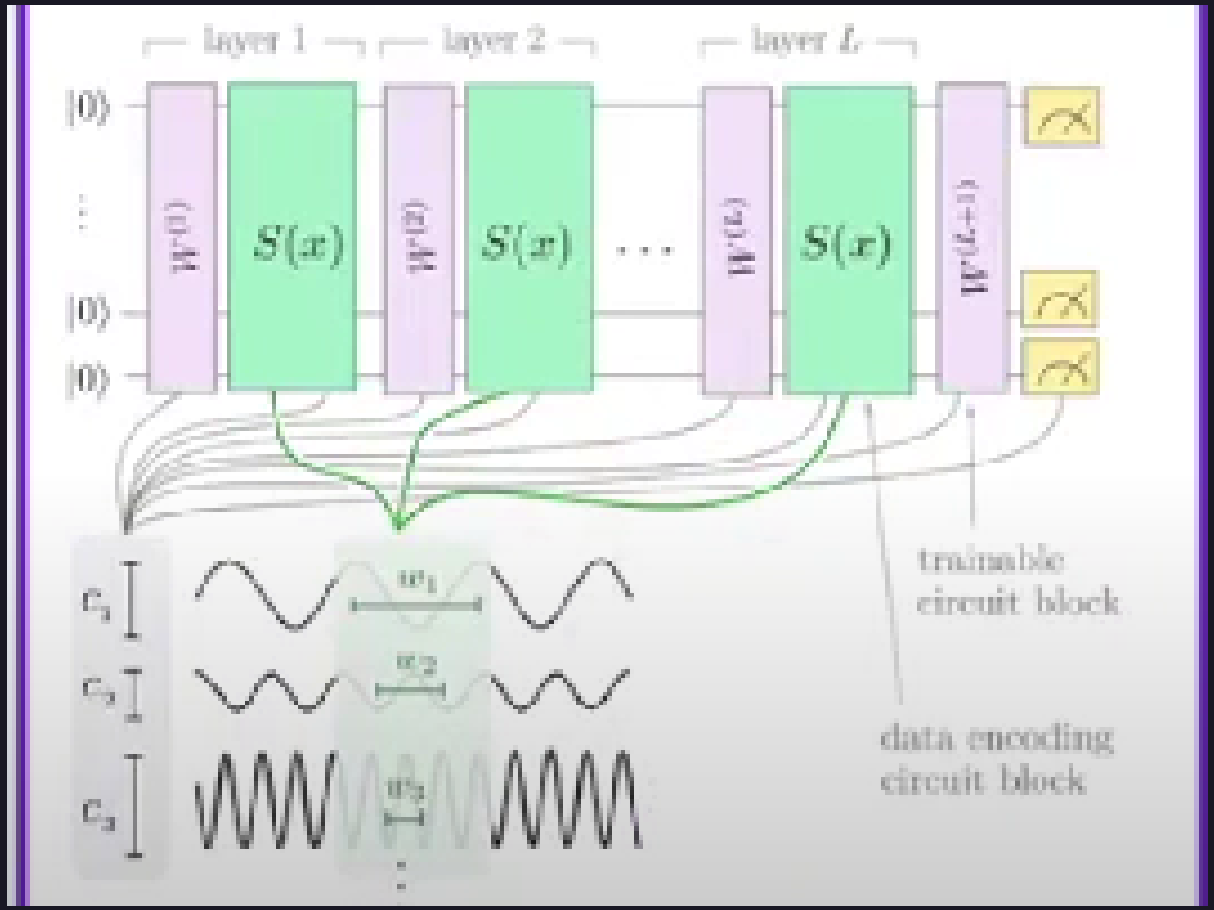
(a) Margins. (b) Normalized margins.

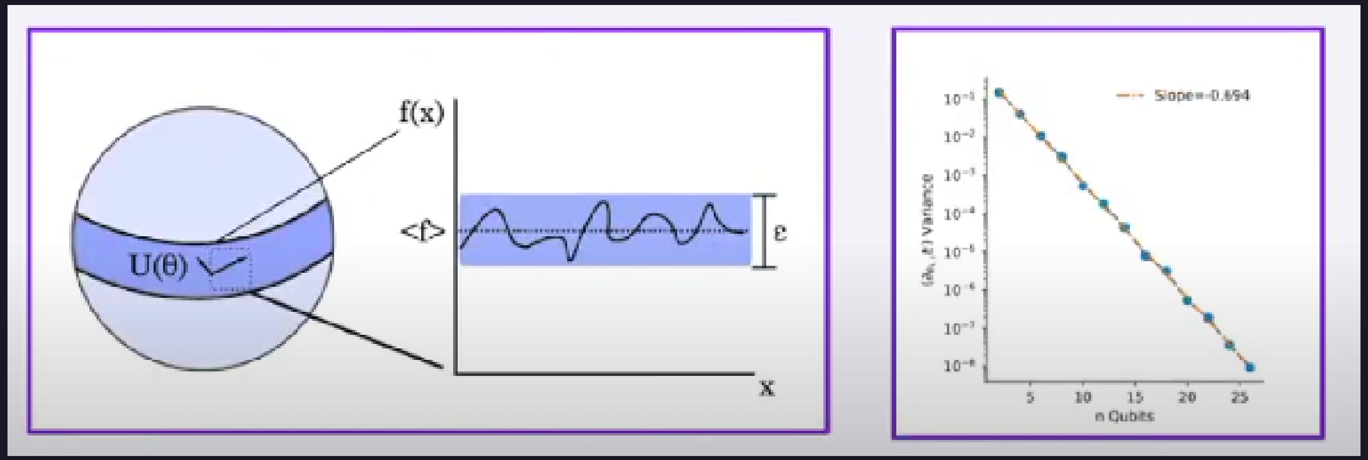
In general, finding an appropriate capacity measure is still a relevant research question

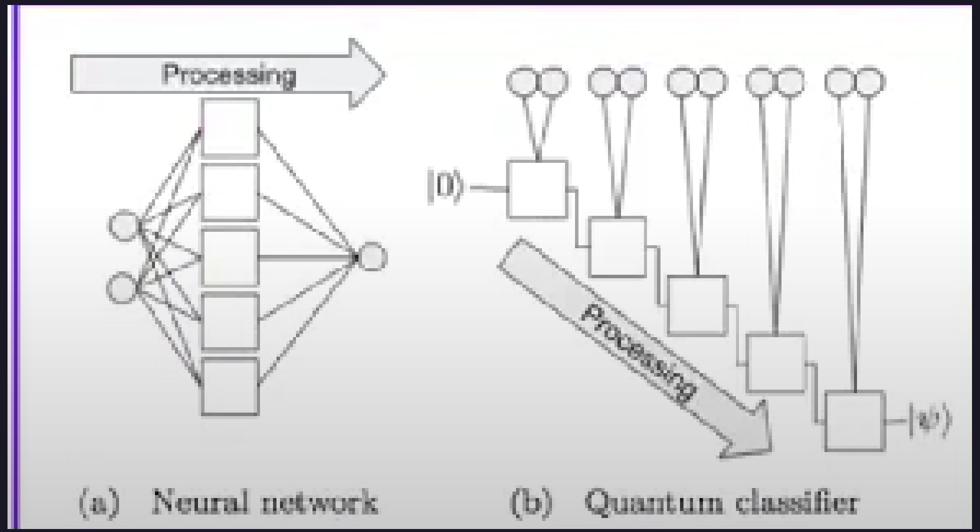
Scale-sensitive? | Data-dependent? | Calculable? | Correlates with generalization error?

Expresibility:

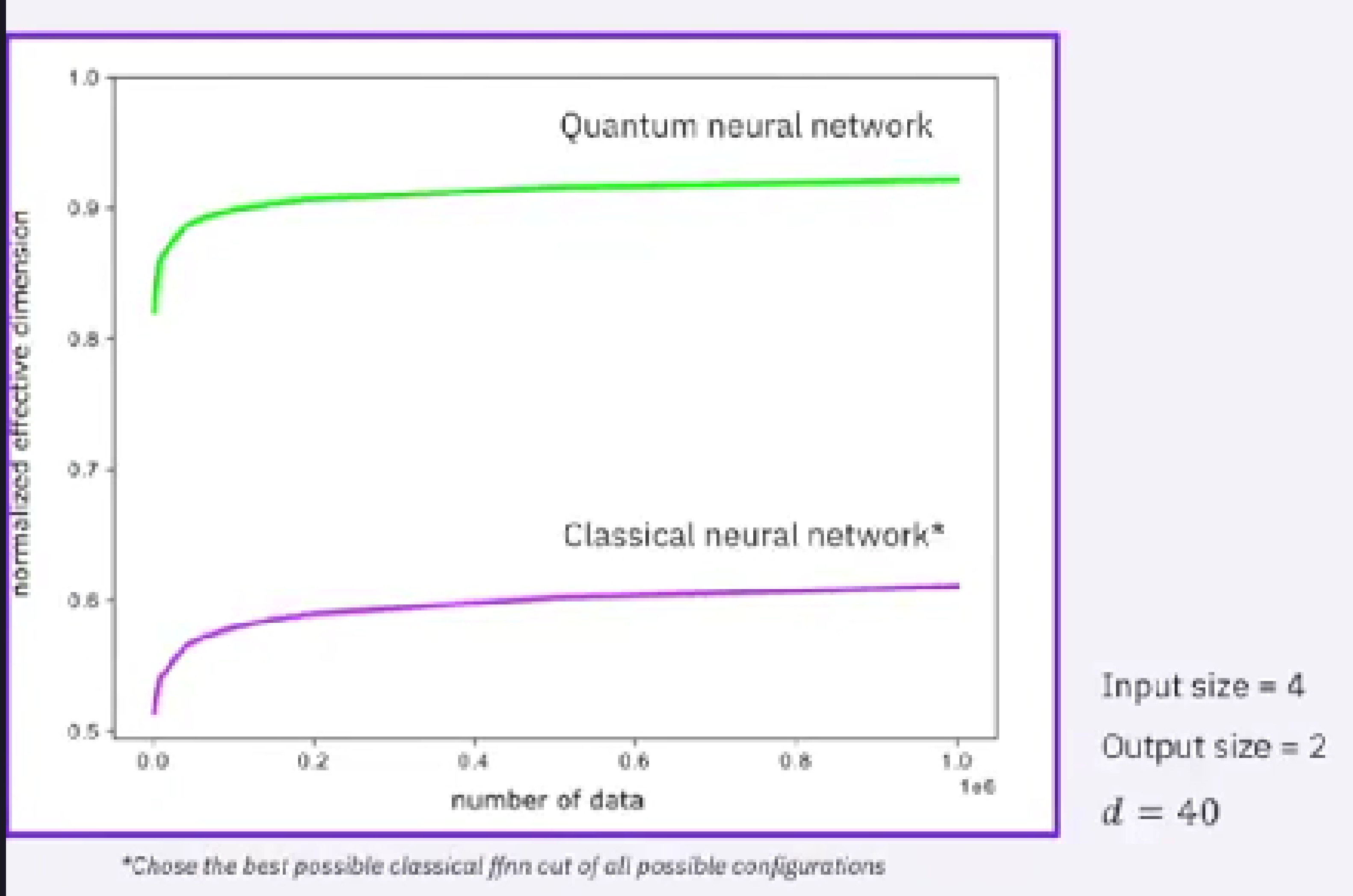


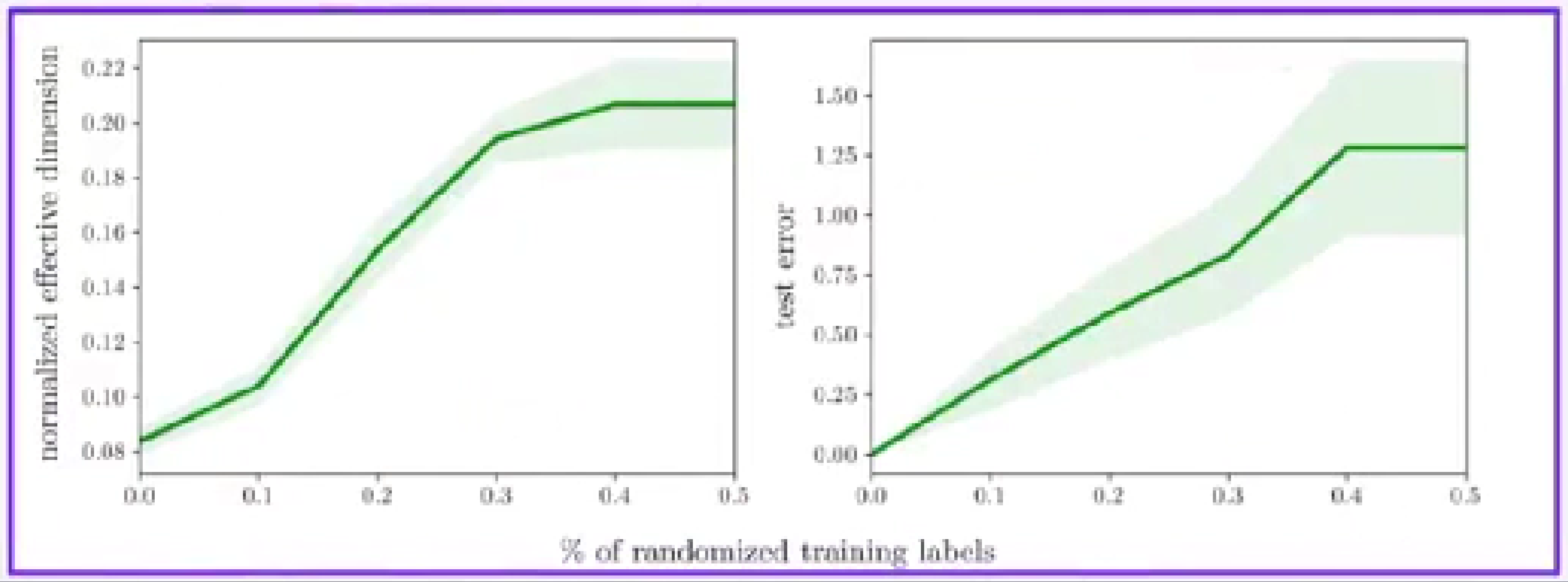






Effective Dimension:





Conclusion:

Capacity?	Generalization?	Data?
Computational?	Quantum kernels?	Training?
Statistical?	Variational circuits?	Applications?

References

- Supervised Learning with Quantum Computers: https://link.springer.com/book/10.1007/978-3-319-96424-9
- Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms: https://arxiv.org/abs/1905.10876
- Supervised learning with quantum-enhanced feature spaces:https://www.nature.com/articles/s41586-019-0980-2
- Circuit-centric quantum classifiers: https://arxiv.org/abs/1804.00633
- The effect of data encoding on the expressive power of variational quantum machine learning models: https://arxiv.org/pdf/2008.08605