

**ICECUBE**  
SOUTH POLE NEUTRINO OBSERVATORY



The NSF Institute for  
Artificial Intelligence and  
Fundamental Interactions



# Intro to Quantum Computing ML for Neutrino Astronomy

August 24th, 2023

NPML 2023

Tufts University



**Pavel Zhelnin**  
**With Jeff L. Et al.**  
**Harvard University**  
[pzhelnin@g.harvard.edu](mailto:pzhelnin@g.harvard.edu)

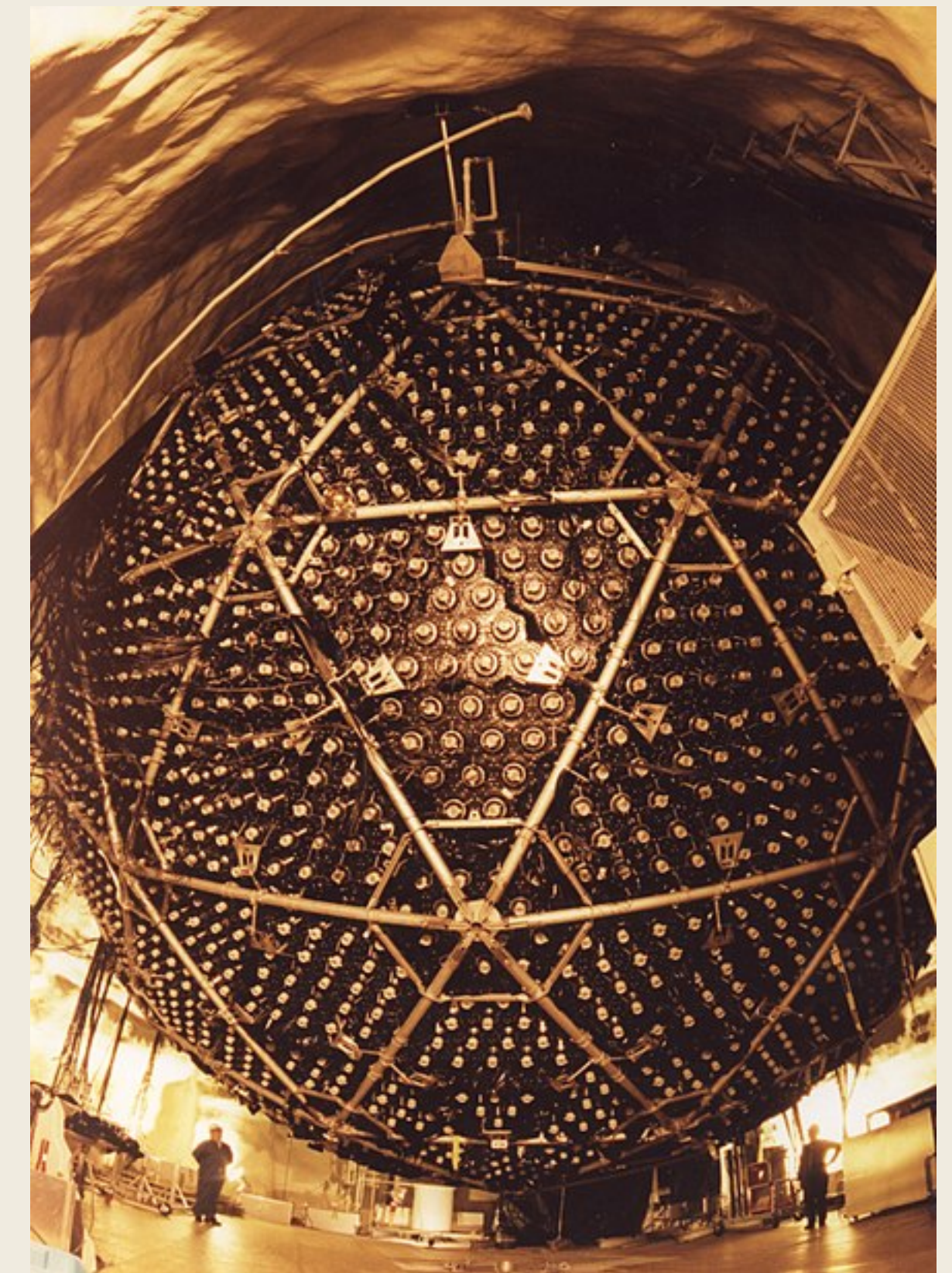
# Machine Learning

- In the 1990s the initial exploration of ML in particle physics began: SNO experiment
- At the beginning, these neural networks did not outperform other statistical techniques but they did demonstrate capabilities
- However as expertise grew ML techniques began to surpass traditional reco
- Now ML has played a role in nearly every particle physics discovery and measurement since

## Observation of high-energy neutrinos from the Galactic plane

ICECUBE COLLABORATION, R. ABBASI, M. ACKERMANN, J. ADAMS, J. A. AGUILAR, M. AHLERS, M. AHRENS, J. M. ALAMEDDINE, A. A. ALVES JR., [...], AND P. ZHELNIN

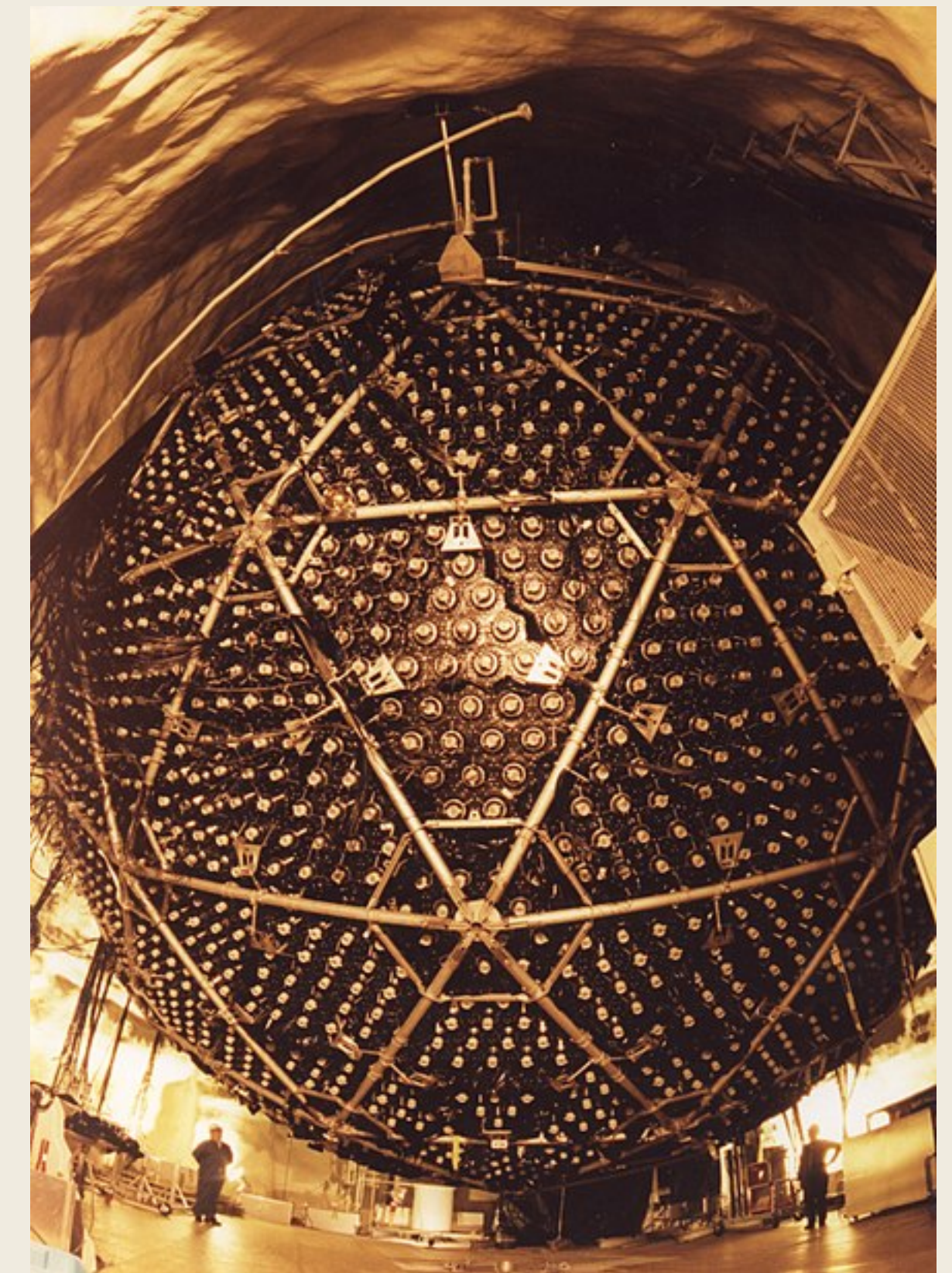
+380 authors [Authors Info & Affiliations](#)



[https://en.wikipedia.org/wiki/Sudbury\\_Neutrino\\_Observatory](https://en.wikipedia.org/wiki/Sudbury_Neutrino_Observatory)

# Machine Learning

- In the 1990s the initial exploration of ML in particle physics began: SNO experiment
- At the beginning, these neural networks did not outperform other statistical techniques but they did demonstrate capabilities
- However as expertise grew ML techniques began to surpass traditional reco
- Now ML has played a role in nearly every particle physics discovery and measurement since



[https://en.wikipedia.org/wiki/Sudbury\\_Neutrino\\_Observatory](https://en.wikipedia.org/wiki/Sudbury_Neutrino_Observatory)

Observation of high energy neutrinos from the Galactic plane

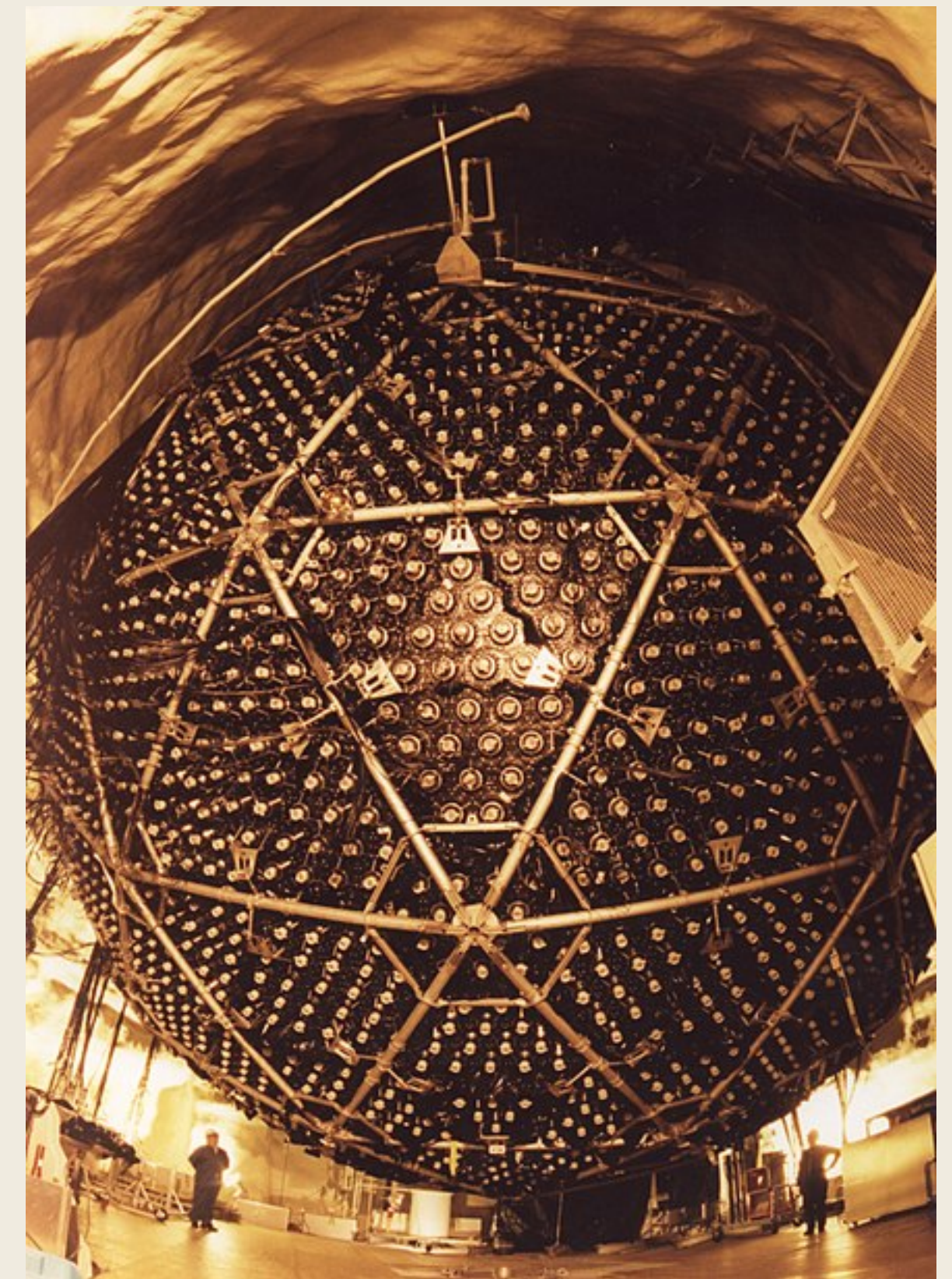
ICECUBE COLLABORATION, R. ABBASI, M. ACKERMANN

+380 authors [Authors Info & Affiliations](#)

Measurement of the properties of Higgs boson production at  $\sqrt{s} = 13$  TeV in the  $H \rightarrow \gamma\gamma$  channel using  $139 \text{ fb}^{-1}$  of  $pp$  collision data with the ATLAS experiment

# Machine Learning

- In the 1990s the initial exploration of ML in particle physics began: SNO experiment
- At the beginning, these neural networks did not outperform other statistical techniques but they did demonstrate capabilities
- However as expertise grew ML techniques began to surpass traditional reco
- Now ML has played a role in nearly every particle physics discovery and measurement since
- What's the next iteration?



[https://en.wikipedia.org/wiki/Sudbury\\_Neutrino\\_Observatory](https://en.wikipedia.org/wiki/Sudbury_Neutrino_Observatory)

Observation of high energy neutrinos from the Galactic plane

ICECUBE COLLABORATION, R. ABBASI, M. ACKERMANN

+380 authors [Authors Info & Affiliations](#)

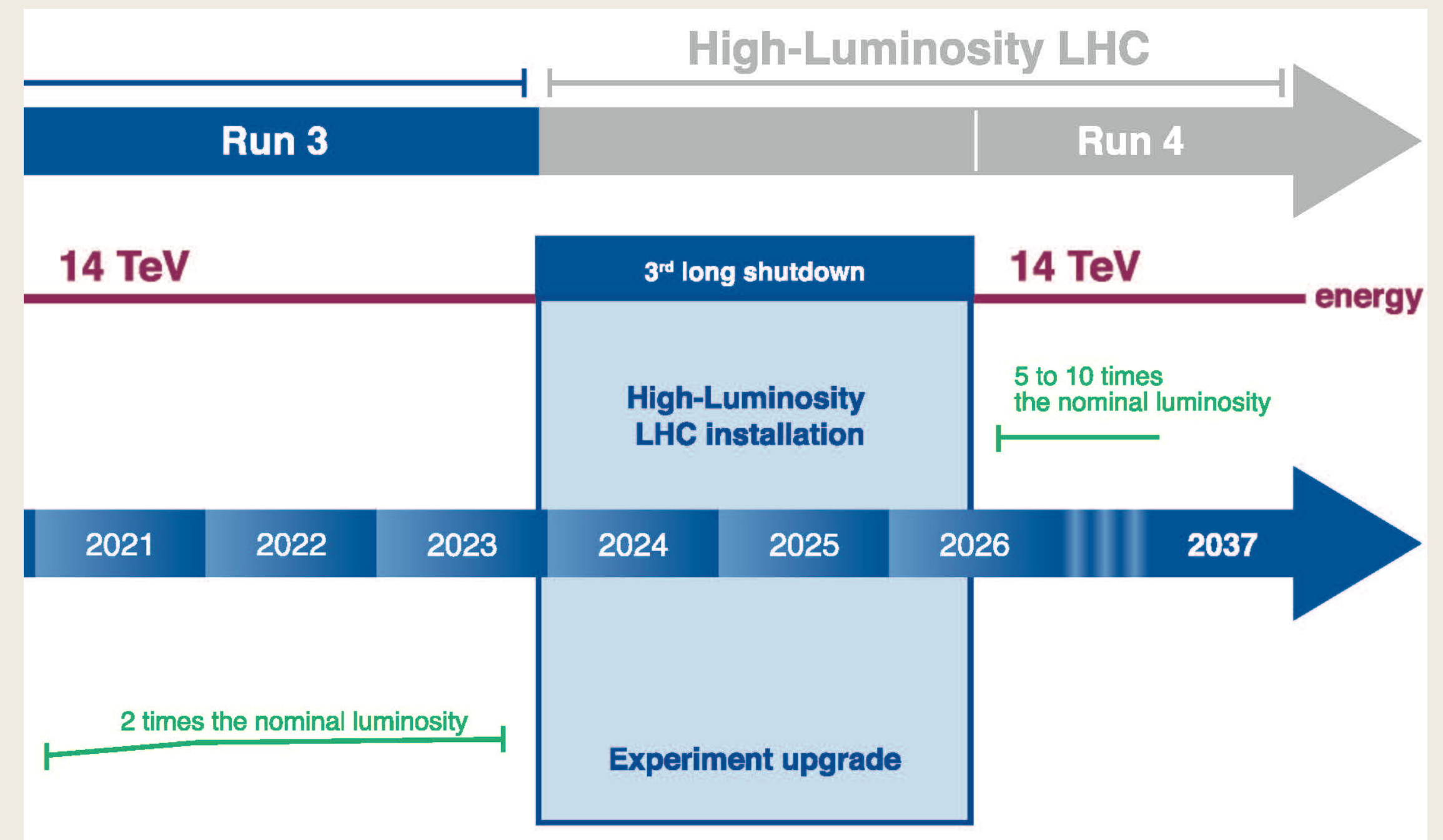
Measurement of the properties of Higgs boson production and decay in the  $gg \rightarrow H \rightarrow b\bar{b}$  channel using the ATLAS experiment

The Results of a Neural Network Statistical Event Class Analysis

S. Brice • Published 1996 • Physics

# A Growing Data Challenge

- ML is essential in analyzing a commonality among experiments now: large data size
- In fact cuts are needed to manage modern experiments
- Even after cuts, datasets are huge
  - CERN produces > 300 TB of data per day
  - IceCube produces ~ 1 TB
- Templates based on our current understanding filters data
- Furthermore, next generation experiments will increase data output by an order of magnitude
- Could new physics be hiding in cut data?



---

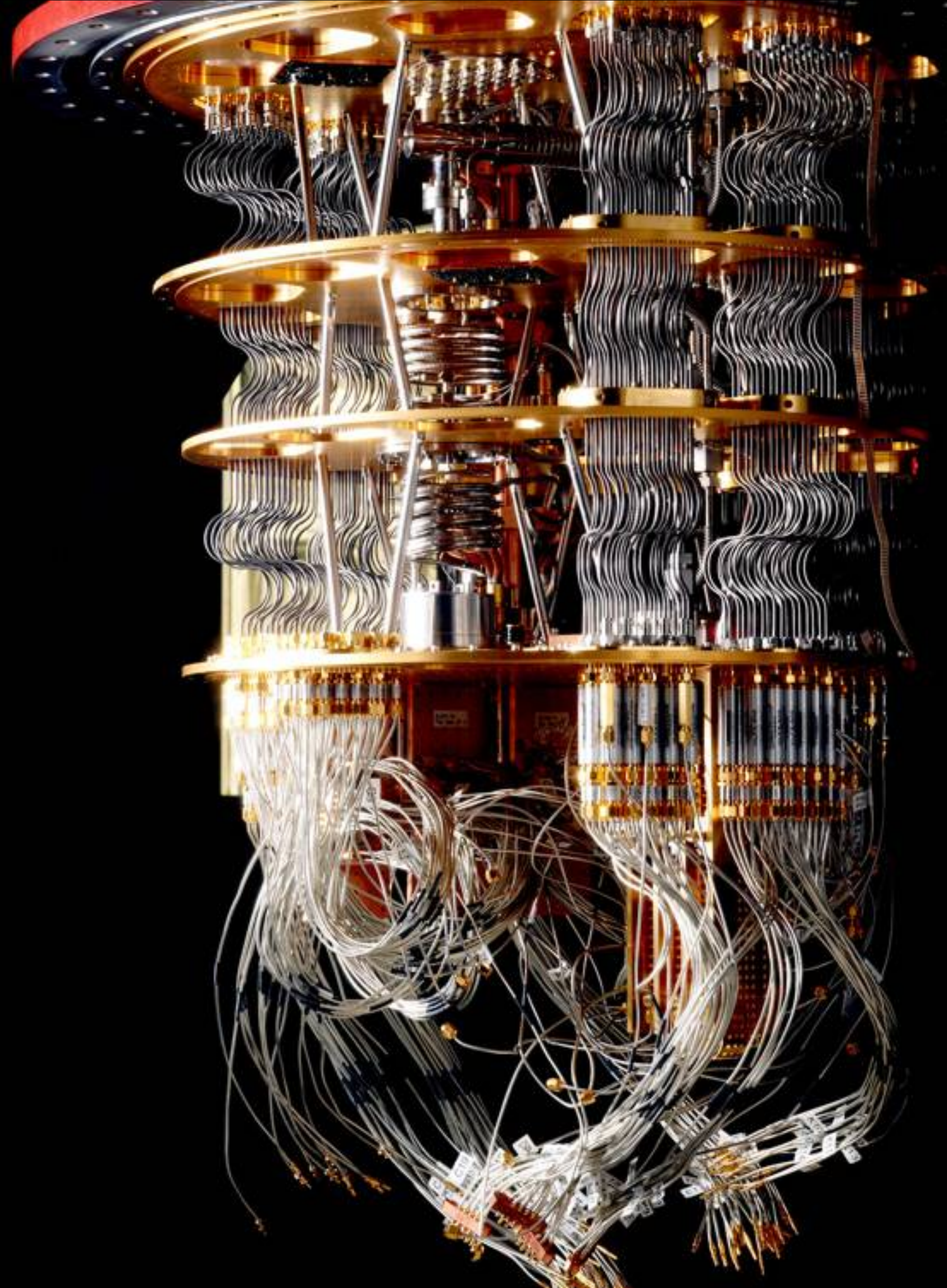
# A Growing Challenge cont.

- We could be missing new physics due to un-modeled interactions (**streetlight effect**)
- Allowing additional data may be necessary for new physics
- For this, a paradigmatic shift is needed in data management to process trigger-level data



# Quantum Computing

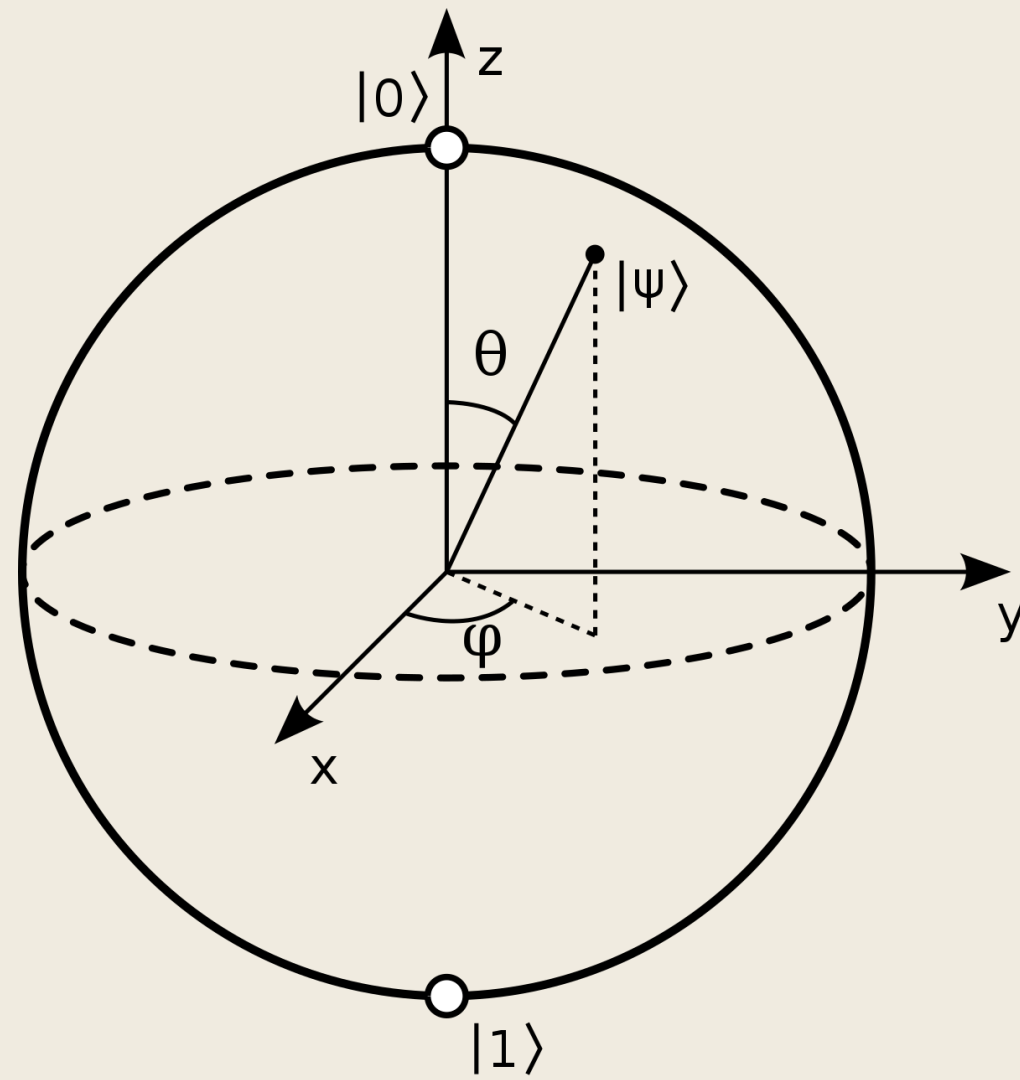
- A computer whose computations can only be described with the laws of quantum theory
  - **Exponentially large Hilbert space**
  - Entanglement
  - Superposition
  - Interference
- $2^N$  advantage over classical computers
- E.g. 8 classical bits  $\rightarrow$  3 "qubits", 64 bits  $\rightarrow$  6 "qubits", can store all of Google Drive cloud storage in  $\sim$ 60 qubits



---

# Basics of Quantum Computing

- **Qubits** = basic unit of information in a QC (akin to a bit)
- Often represented by a Bloch sphere

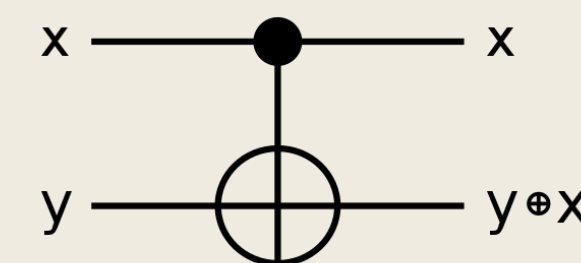
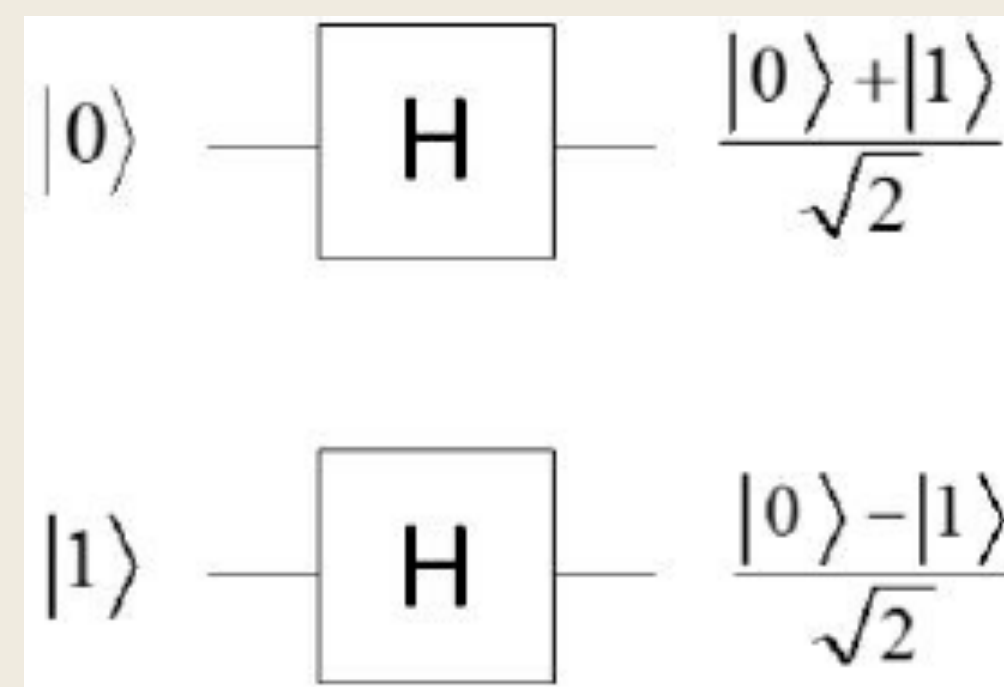
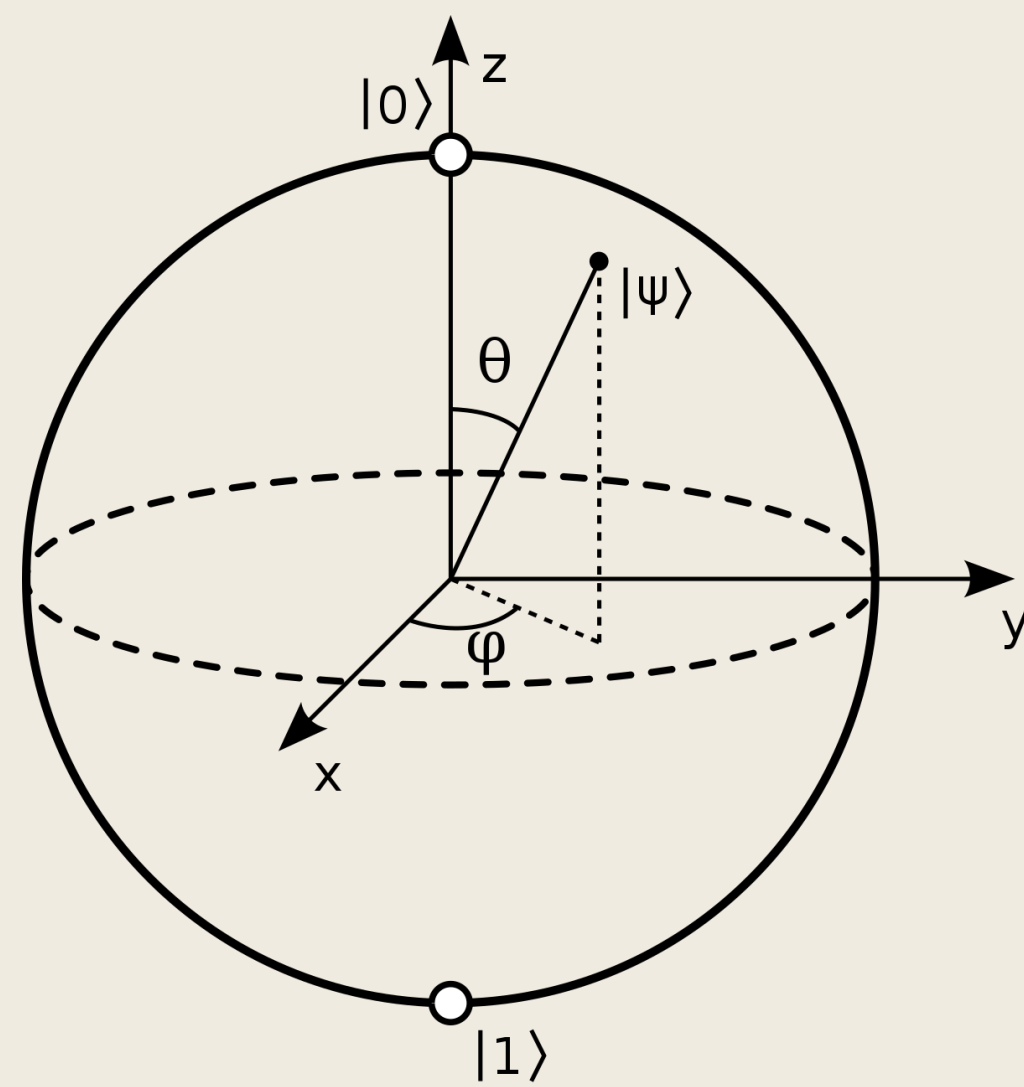




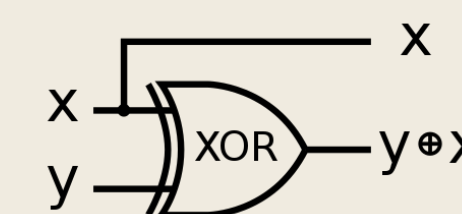
# Basics of Quantum Computing

- **Qubits** = basic unit of information in a QC (akin to a bit)
- Often represented by a Bloch sphere

- **Quantum gates** = most basic operation that can be performed on a qubit (or set of qubits)
- Two basic quantum gates: Hadamard/CNOT
- Hadamard creates superpositions
- CNOT entangles
- Combination makes a Bell State



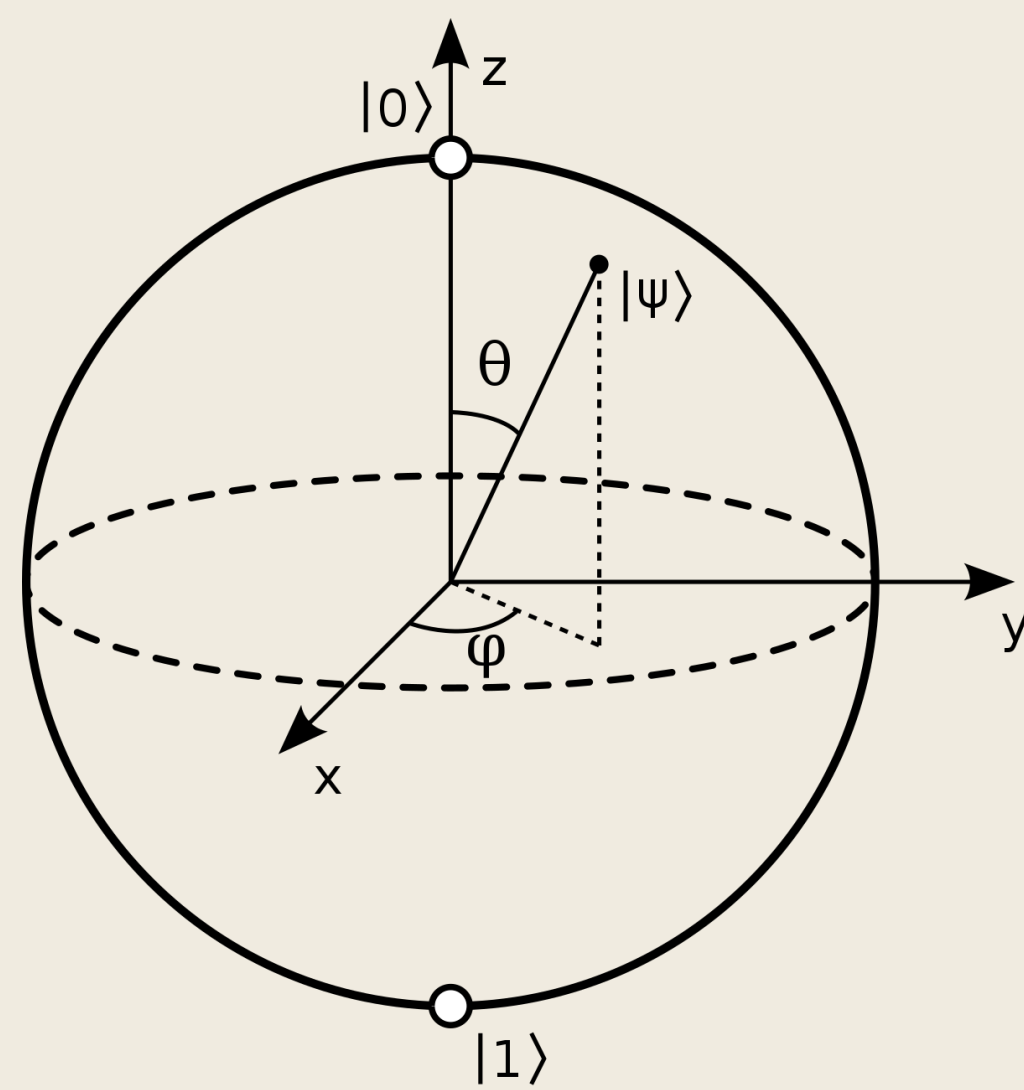
input		output	
x	y	x	y+x
0>	0>	0>	0>
0>	1>	0>	1>
1>	0>	1>	1>
1>	1>	1>	0>



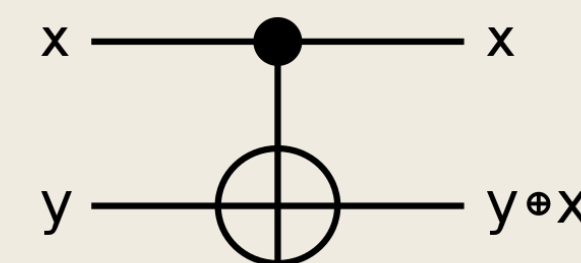
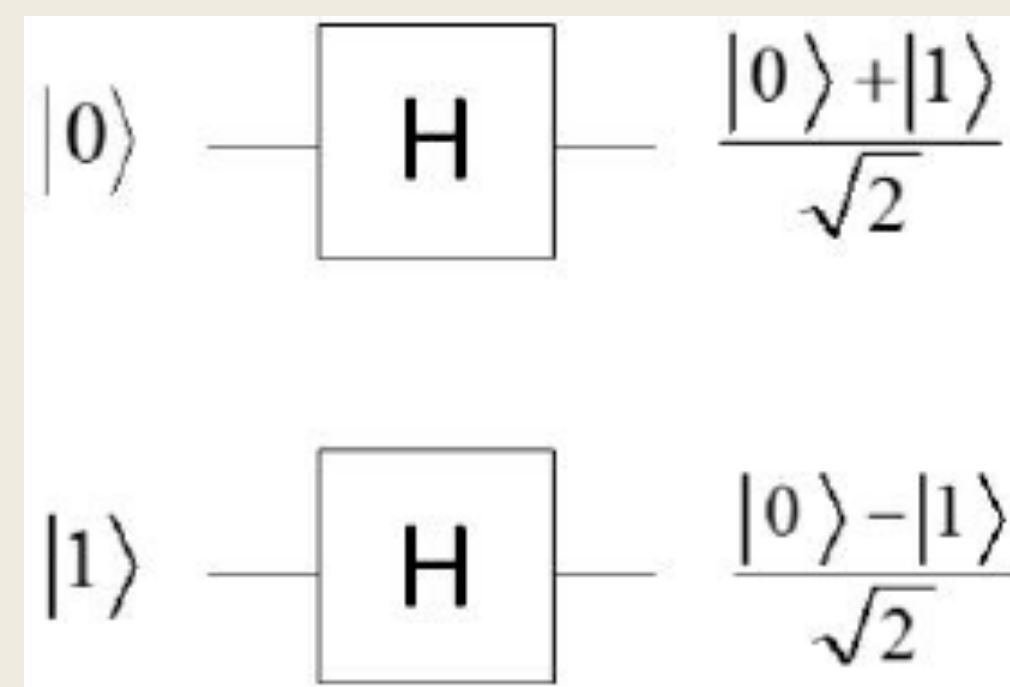
input		output	
x	y	x	y+x
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

# Basics of Quantum Computing

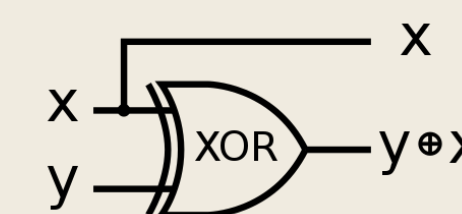
- **Qubits** = basic unit of information in a QC (akin to a bit)
- Often represented by a Bloch sphere



- **Quantum gates** = most basic operation that can be performed on a qubit (or set of qubits)
- Two basic quantum gates: Hadamard/CNOT
- Hadamard creates superpositions
- CNOT entangles
- Combination makes a Bell State

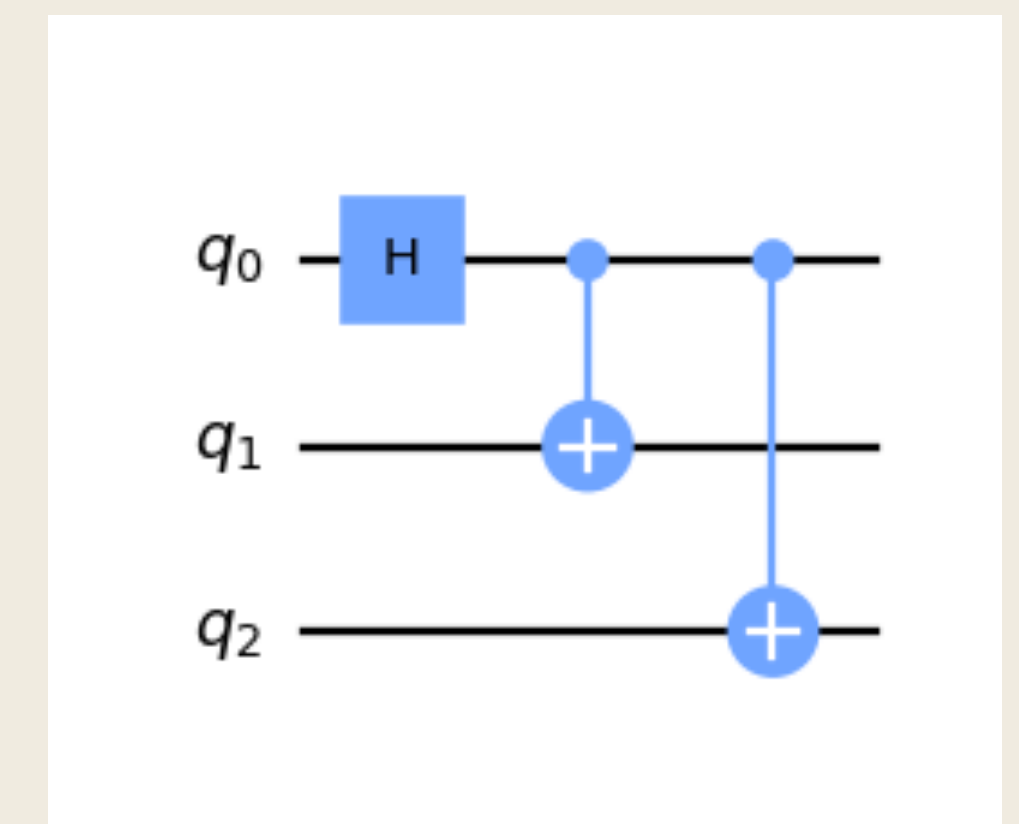


input		output	
x	y	x	y+x
0>	0>	0>	0>
0>	1>	0>	1>
1>	0>	1>	1>
1>	1>	1>	0>



input		output	
x	y	x	y+x
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

- **Quantum circuit** = a model for quantum computation in which a sequence of quantum gates are applied to a set of n qubits

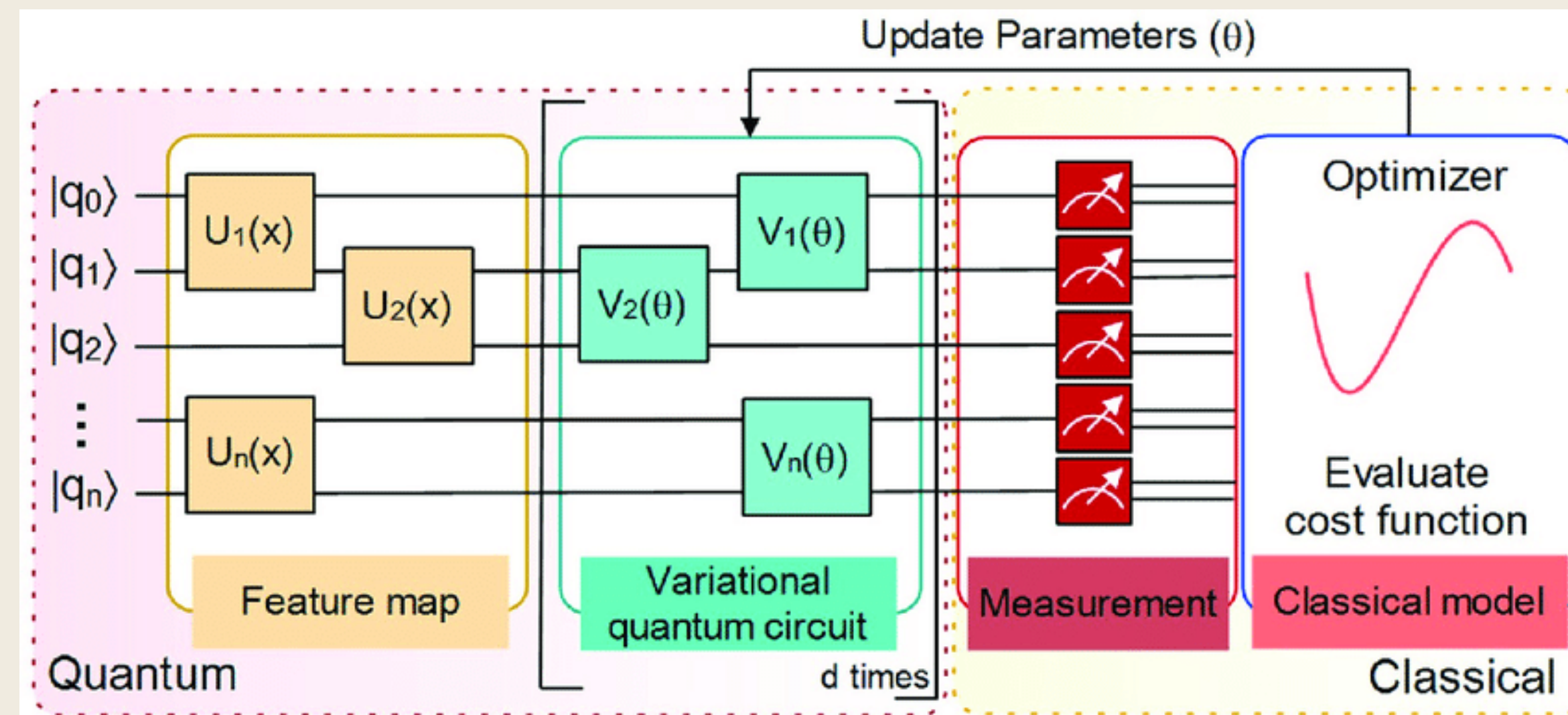


# A Data Processing Pipeline Using QC

- We don't want to just store data on a QC, we want to process it as well
- Its runtime is costly to do: classical  $\rightarrow$  quantum or quantum  $\rightarrow$  classical transfers of data
- We want a fully "quantum pipeline", no classical preprocessing

E.g. ...

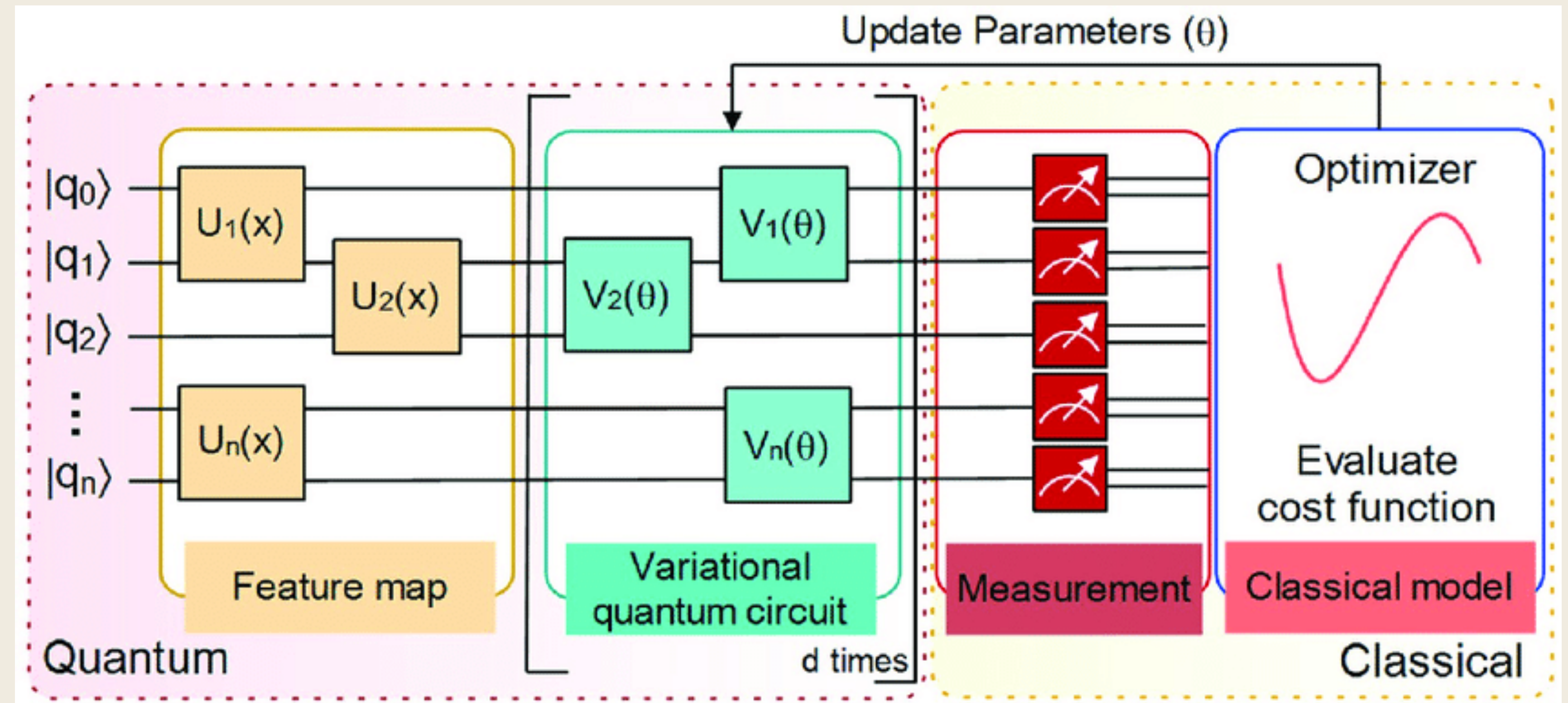
Classical Data  $\rightarrow$



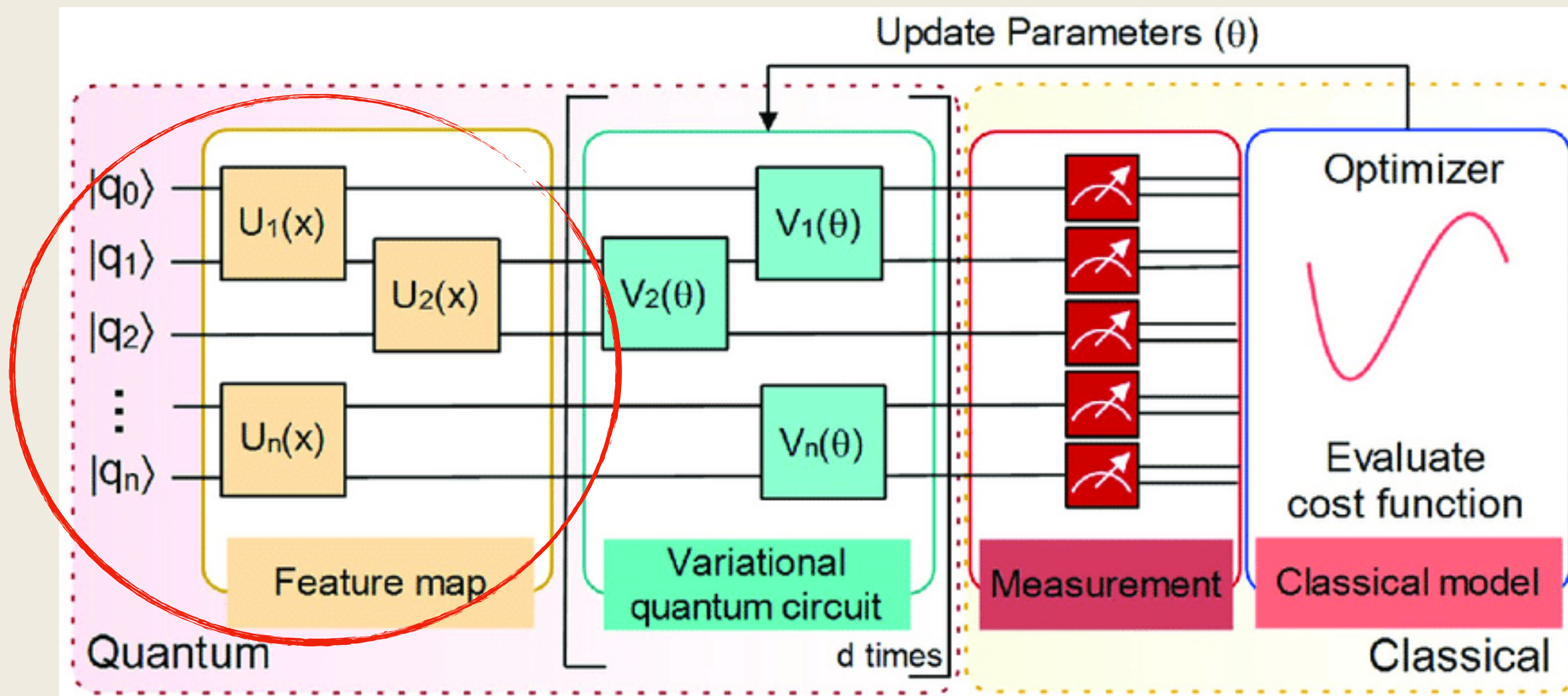
$\rightarrow$  Classification

# Rest of the talk: QML with a Variational Quantum Circuit

A VQC is a low depth, low width choice suitable for ML applications on current quantum computers.



# Rest of the talk: QML



---

# Data Encoding

- “Data encoding is often the most crucial step in QML with classical data: it influences potential quantum advantage, learning performance and runtime.”
- Most other QML encoding schemes involve some classical preprocessing then using either amplitude/basis encoding (arXiv:2012.11560, arXiv.1907.00397, arXiv.2010.07335)
- We want to avoid classical preprocessing while still working within the constraints of Near-Intermediate Scale Quantum (NISQ) computers

# Background: Data Encoding

- **Amplitude encoding** can store information with  $2^N$  efficiency
- Susceptible to decoherence

$$\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$
$$|\psi_{\mathbf{x}}\rangle = \sum_{i=0}^{N-1} x_i |i\rangle.$$
$$|\mathbf{x}\rangle = \frac{1}{2} |00\rangle + \frac{1}{2} |01\rangle - \frac{1}{2} |10\rangle - \frac{1}{2} |11\rangle$$

# Background: Data Encoding

- **Amplitude encoding** can store information with  $2^N$  efficiency
  - Susceptible to decoherence
- **Basis encoding** is the simplest encoding
  - No quantum advantage, a 1 to 1 mapping

$$\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$|\psi_{\mathbf{x}}\rangle = \sum_{i=0}^{N-1} x_i |i\rangle.$$

$$x = 1011 \rightarrow |1011\rangle$$

$$|\mathbf{x}\rangle = \frac{1}{2} |00\rangle + \frac{1}{2} |01\rangle - \frac{1}{2} |10\rangle - \frac{1}{2} |11\rangle$$



# Background: Data Encoding

- **Amplitude encoding** can store information with  $2^N$  efficiency
  - Susceptible to decoherence
- **Basis encoding** is the simplest encoding
  - No quantum advantage, a 1 to 1 mapping
- **Angle encoding**
  - Rotations around principle axes of Bloch sphere
  - Principle encoding scheme by others in QML HEP

$$\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$|\psi_{\mathbf{x}}\rangle = \sum_{i=0}^{N-1} x_i |i\rangle.$$

$$x = 1011 \rightarrow |1011\rangle$$

$$|\mathbf{x}\rangle = \frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle - \frac{1}{2}|10\rangle - \frac{1}{2}|11\rangle$$

$$|\mathbf{x}\rangle = \bigotimes_i^n R(\mathbf{x}_i) |0^n\rangle$$

Rotation by pi around y axis on Bloch sphere

$$\mathbf{x} = \begin{bmatrix} \pi \\ \pi \\ \pi \end{bmatrix} \rightarrow |111\rangle$$

# Recap so far

- We want to use quantum computers because they can handle computational challenges of increase data loads of upcoming experiments
- This way we can investigate more data
- We don't want to reduce the complexity of our data (no PCA, no classical dimension reduction)
- Traditional quantum encoding schemes either don't use quantum advantage or are overly susceptible to decoherence
- So we want a near lossless quantum encoding scheme with quantum advantage

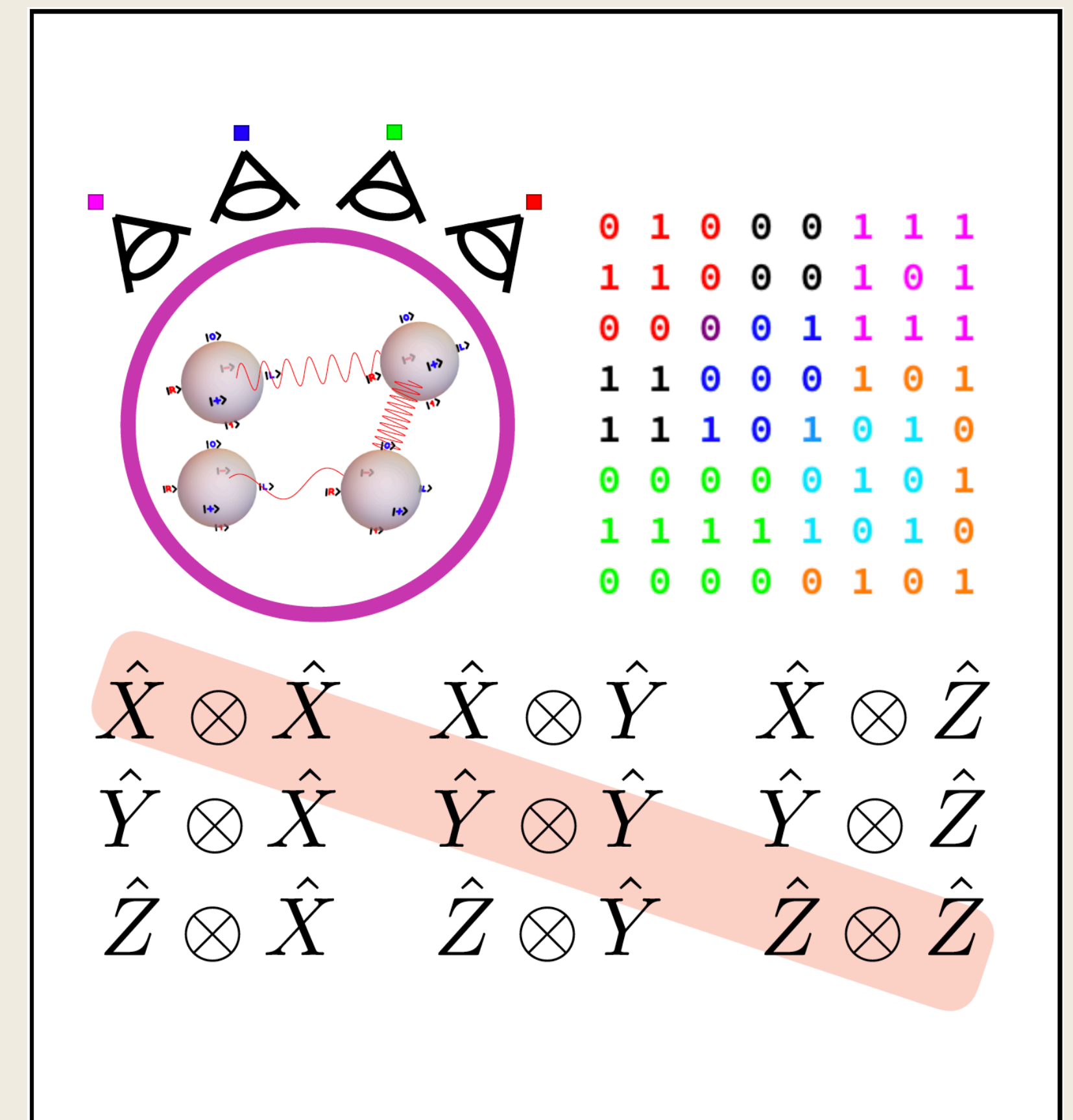


# Quantum Random Access Codes (QRAC)

- We want a resilient data encoding scheme that still occupies some advantage over classical systems
- Encodes digital information in correlations between qubits
- $\sim a 1.5^N$  advantage over classical systems
- It's resilient: for  $N \geq 18$  nearly lossless (0.999) recovery rate
- For  $N \geq 14$  QRAC has greater success than classical counterparts

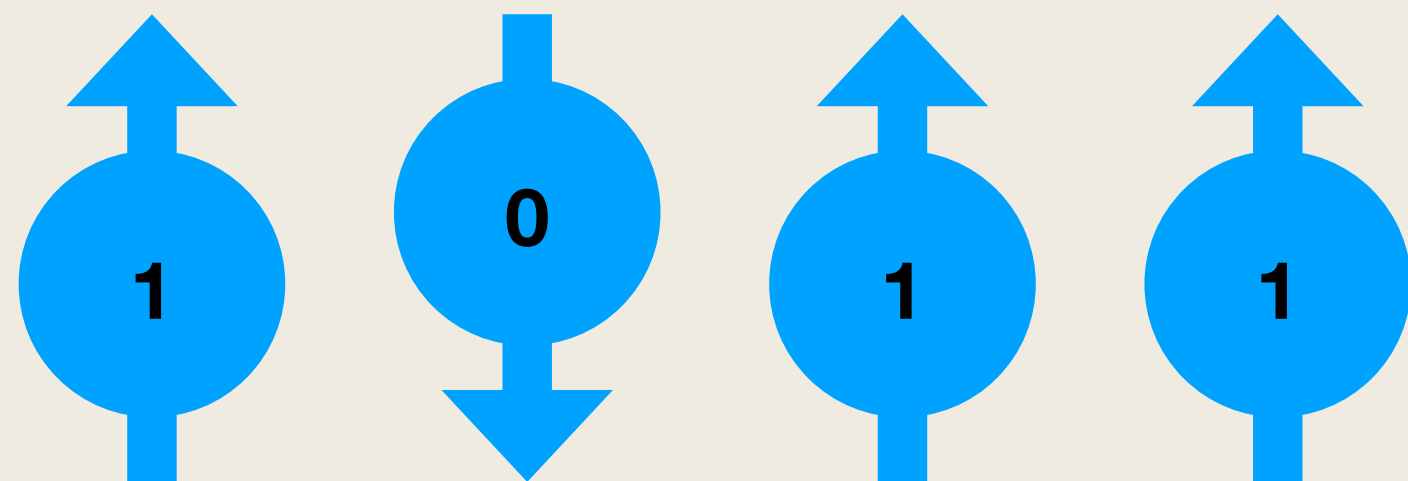
Random access codes via quantum contextual redundancy

Giancarlo Gatti,<sup>1,2,\*</sup> Daniel Huerga,<sup>1,†</sup> Enrique Solano,<sup>1,3,4,5,‡</sup> and Mikel Sanz<sup>1,4,5,§</sup>

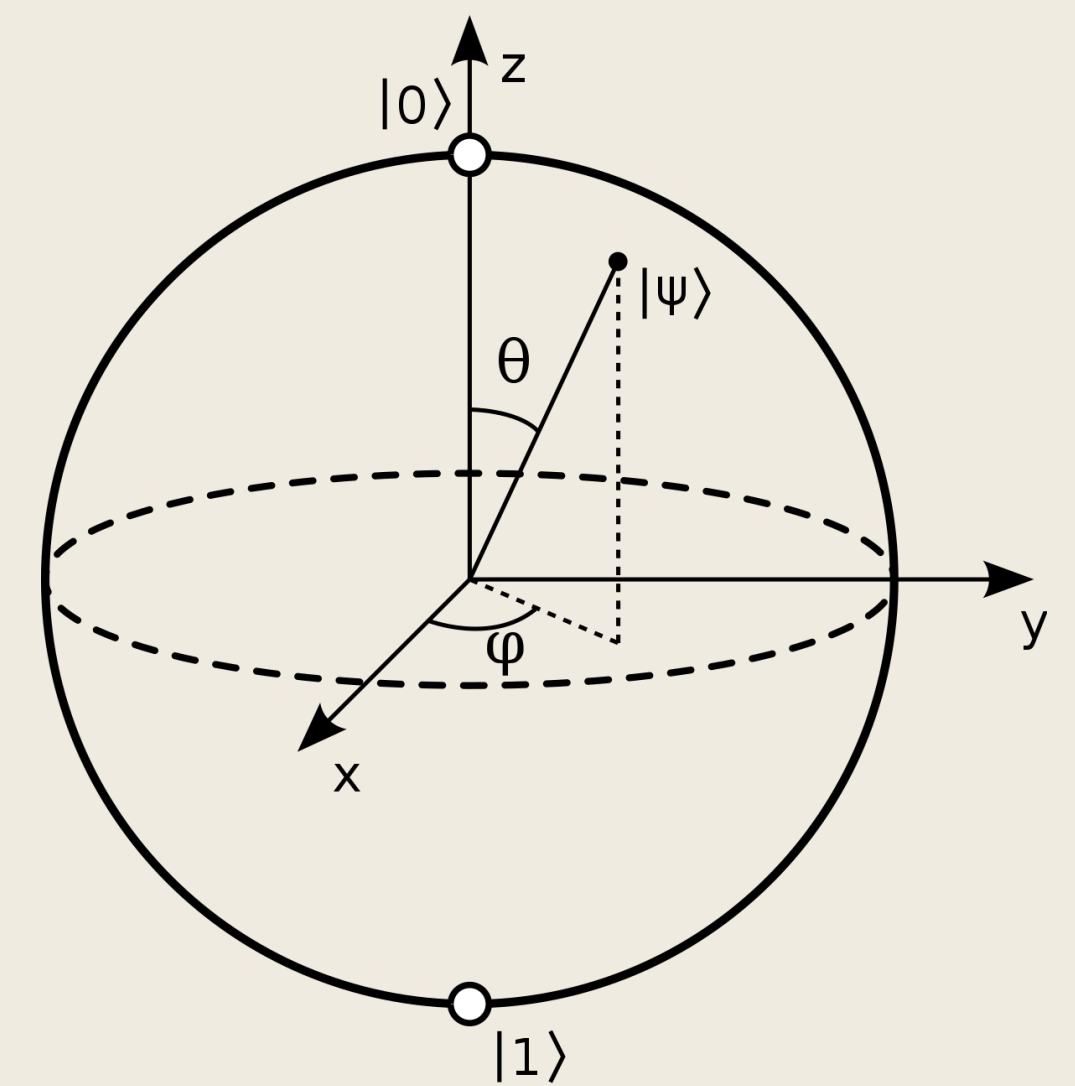
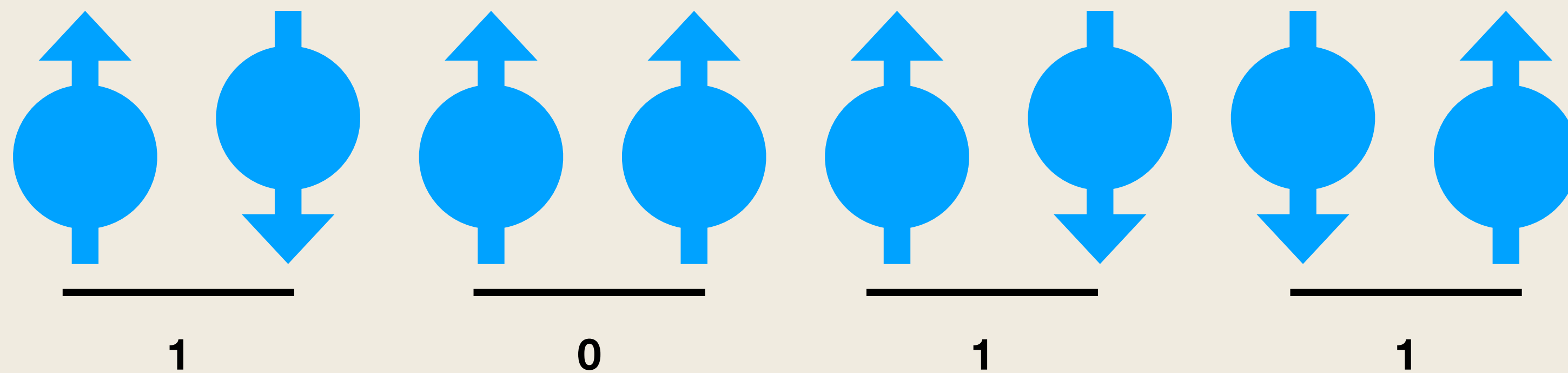


# Digital Quantum Encoding

- For example, I want to encode the bit string '1011'
- Option 0: One-to-one mapping to z-spin



- Option 1: Two-to-one mapping to z-spins (have two options to encode)



# Correlation-based digital encoding

- The set  $\{X, Y, Z\}^{\otimes N}$  are parity observables (POs) where N is number of qubits
- Measuring using POs always yield a  $\pm 1$
- Instead of assigning a bit to each PO, assign a bit to each pair of POs,  $0 \leftrightarrow =$  and  $1 \leftrightarrow \neq$
- We create eigenstates of sets of commuting POs, these are our compressed states that when measured later recover our input data

$b^{\text{target}}$	1	0	1	1					
PO relationship	$\neq$	$=$	$\neq$	$\neq$	ZZ				
POs	XX	XY	XZ	YX	YY	YZ	ZX	ZY	NA
Option 1	+1	-1	+1	+1	+1	-1	+1	-1	NA
Option 2	-1	+1	-1	-1	-1	+1	-1	+1	NA

**Example for encoding '1011' in two qubits**

**0** Let us use n-qubit systems  
(and draw figures for n=4)

**1** Alice has m bits of data

1	1	0	0	1	0	0	1
0	0	0	0	1	1	1	1
1	1	0	1	1	0	1	0
1	0	1	0	0	0	1	0
0	0	1	0	0	0	0	1

that she wants Bob to randomly access  
 $m = (3^n - 1)/2$

**2** Alice maps her data to outcomes of n-body Pauli observables

- 1  $\left\{ \begin{array}{l} X_1 X_2 X_3 X_4 \text{ yields } +1 \\ X_1 X_2 X_3 Y_4 \text{ yields } -1 \end{array} \right.$  (more than -1)
  - 1  $\left\{ \begin{array}{l} X_1 X_2 X_3 Z_4 \text{ yields } -1 \\ X_1 X_2 Y_3 X_4 \text{ yields } +1 \end{array} \right.$  (more than +1)
  - 0  $\left\{ \begin{array}{l} X_1 X_2 Y_3 Y_4 \text{ yields } -1 \\ X_1 X_2 Y_3 Z_4 \text{ yields } -1 \end{array} \right.$  (more than +1)
- $\vdots 3^n$  observables


**3a** Alice prepares a group of n-qubit states which collectively have those outcome preferences

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} (|0000\rangle + |1111\rangle)$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} (|0001\rangle + i |1110\rangle)$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}} (|000+\rangle - |111-\rangle)$$

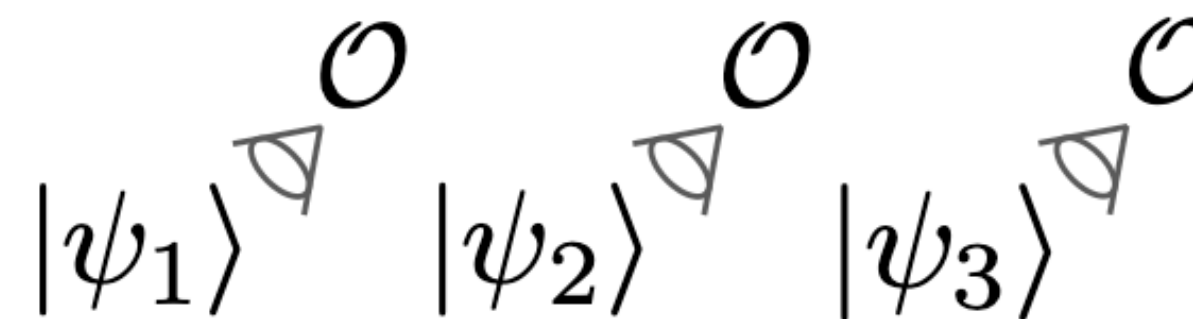
$\vdots O(n(3/2)^n)$  states

 She sends a few copies of each state to Bob ( $k/n$  states in total)

**3b** E.g., the state  $|\psi_1\rangle = \frac{1}{\sqrt{2}} (|0000\rangle + |1111\rangle)$  yields

+1 in  $X_1 X_2 X_3 X_4$ , -1 in  $X_1 X_2 Y_3 Y_4$ ,  
 -1 in  $X_1 Y_2 X_3 Y_4$ , -1 in  $X_1 Y_2 Y_3 X_4$ ,  
 -1 in  $Y_1 X_2 X_3 Y_4$ , -1 in  $Y_1 X_2 Y_3 X_4$ ,  
 -1 in  $Y_1 Y_2 X_3 X_4$ , +1 in  $Y_1 Y_2 Y_3 Y_4$   
 and +1 in  $Z_1 Z_2 Z_3 Z_4$ , with probability 1

**4** Bob measures the states with some of the observables



He finds their preferred outcomes

**5** Bob reconstructs a fragment of Alice's original data

1	1	0
---	---	---

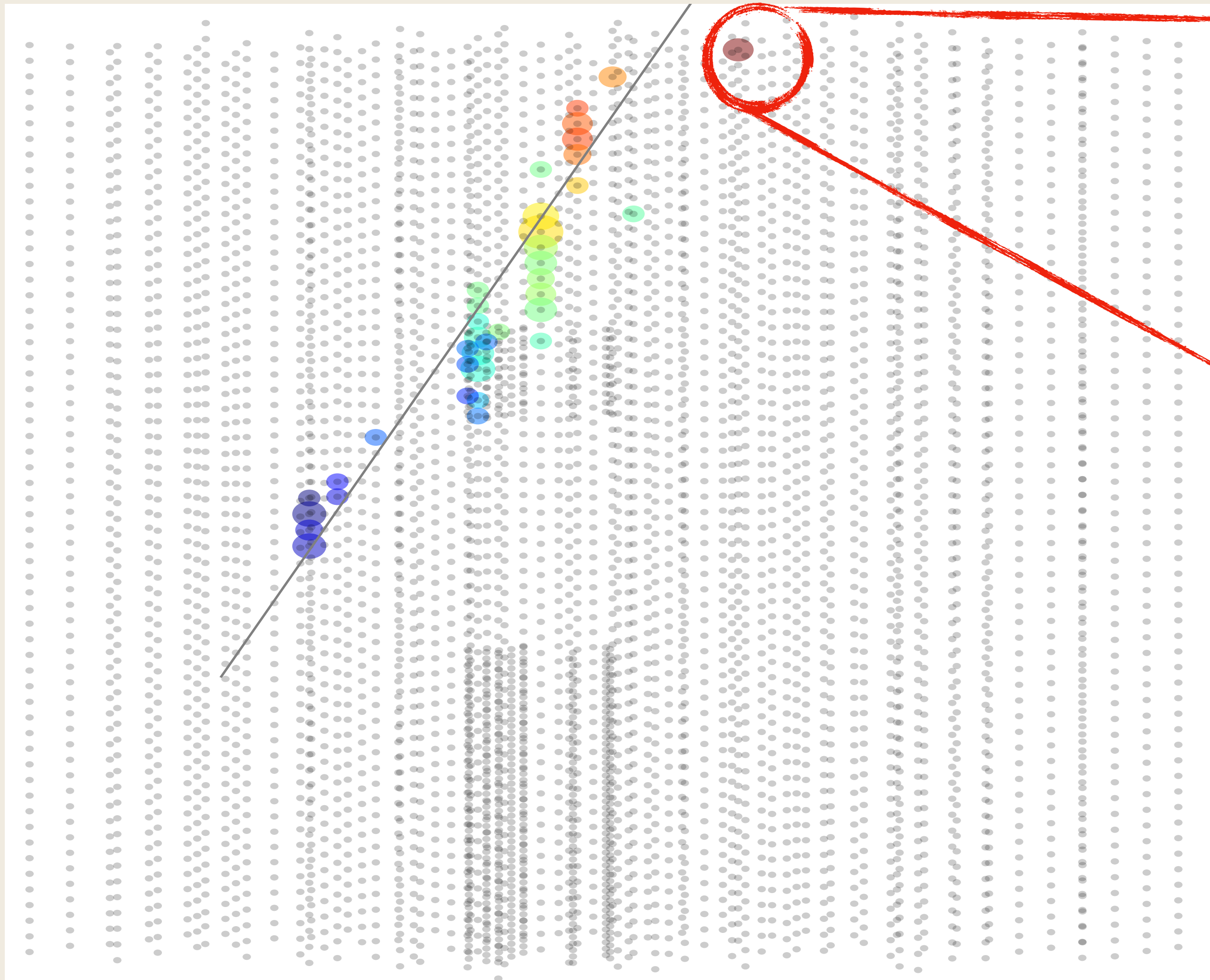
$m \rightarrow k$  QRAC

---

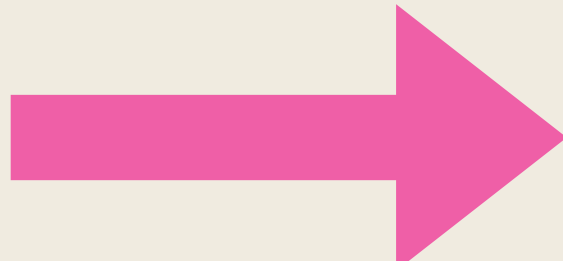
# Optimizations

- Which set of preferred parities to choose from out of  $2^N$  choices where  $N$  are number of couples of POs
  - First compute all compatible well-defined outcomes for all possible eigenstates
  - Compare eigenvalues from this optimization to compatible preferred parities
- Which eigenstates best represent that chosen preferred parity order
  - We want a low sampling requirement
  - Least number of states

# Neutrino Astronomy: IceCube Events



$x$	95.3 m	01000010101111101001100110011010
$y$	75.8 m	11000010100101111001100110011010
$z$	484.6 m	01000011111100100100110011001101
$Q_{\text{tot}}$	2.84 PE	01000000001101011100001010001111
$\bar{t}$	26.2 ns	01000001110100011001100110011010



```
0100001010111110100110011001101011000010100
1011110011001100110100100001111110010010011
0011001101010000000011010111000010100011110
1000001110100011001100110011010
```

**Digitization Scheme:** takes Optical Module (OM) position, light and time information and converts to binary. Each circle in image is a OM, size of circle indicates amount of light, color indicates time (red → purple)

This is our input to our QRAC





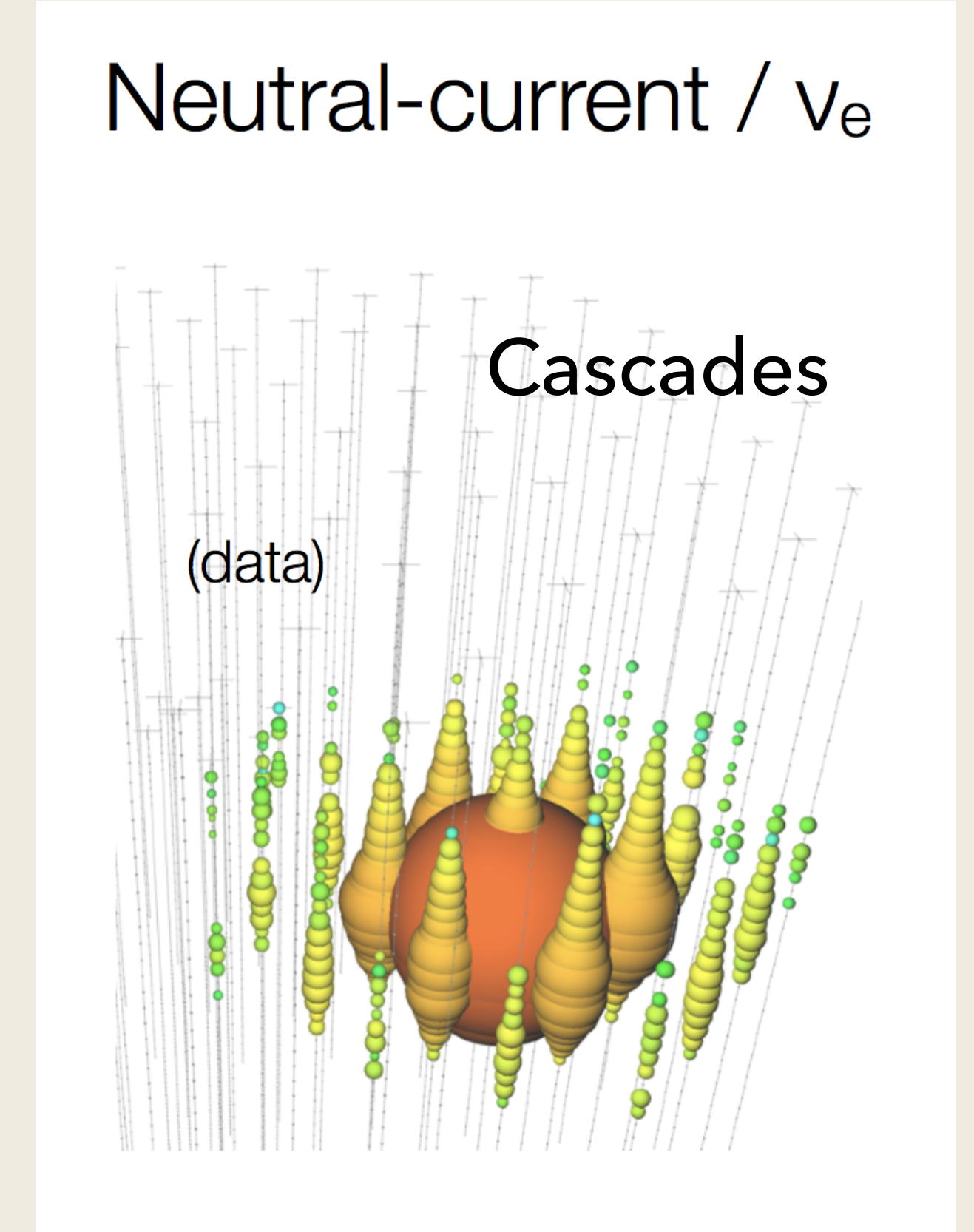
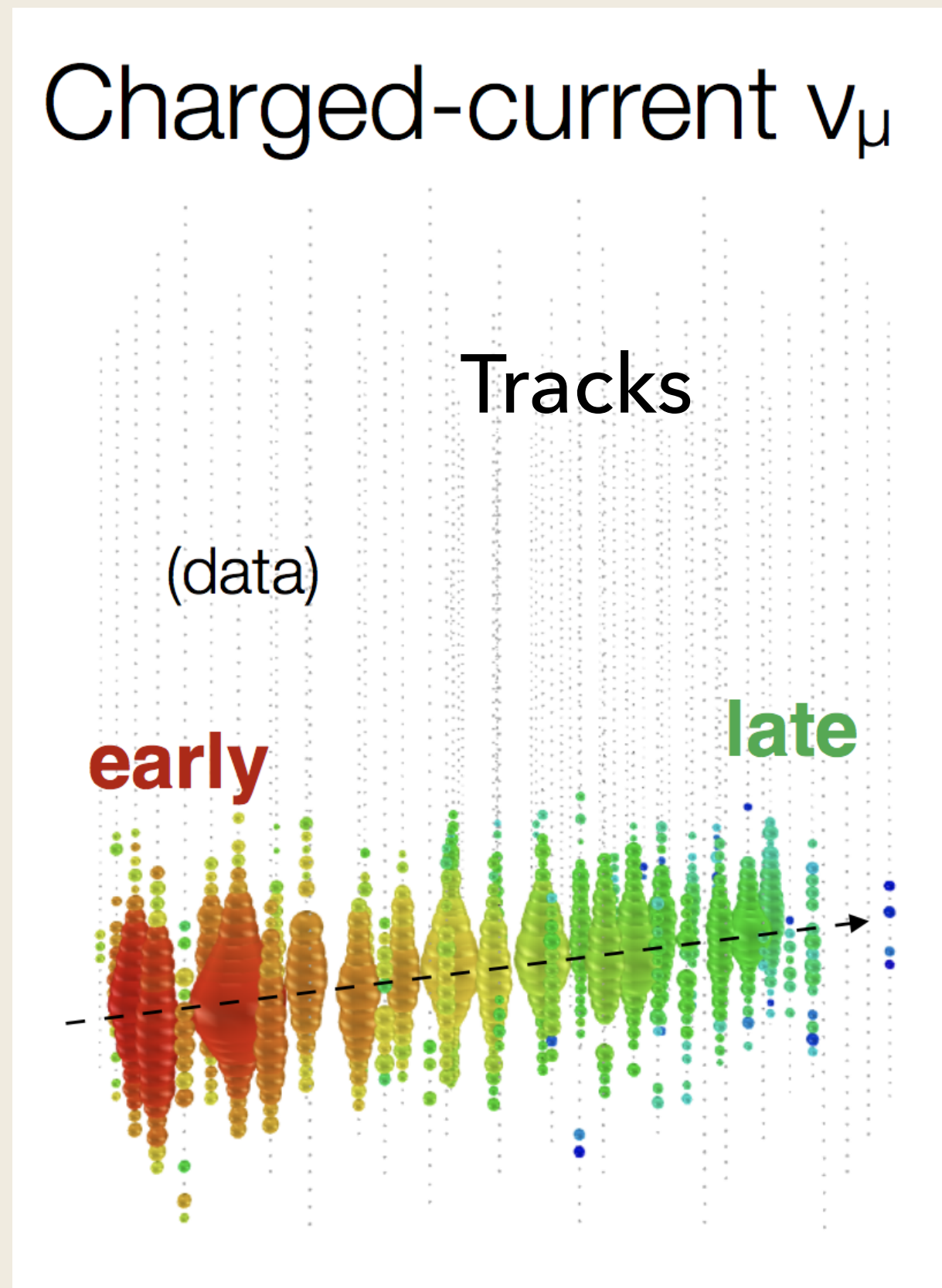
---

# QRAC: What has been done so far

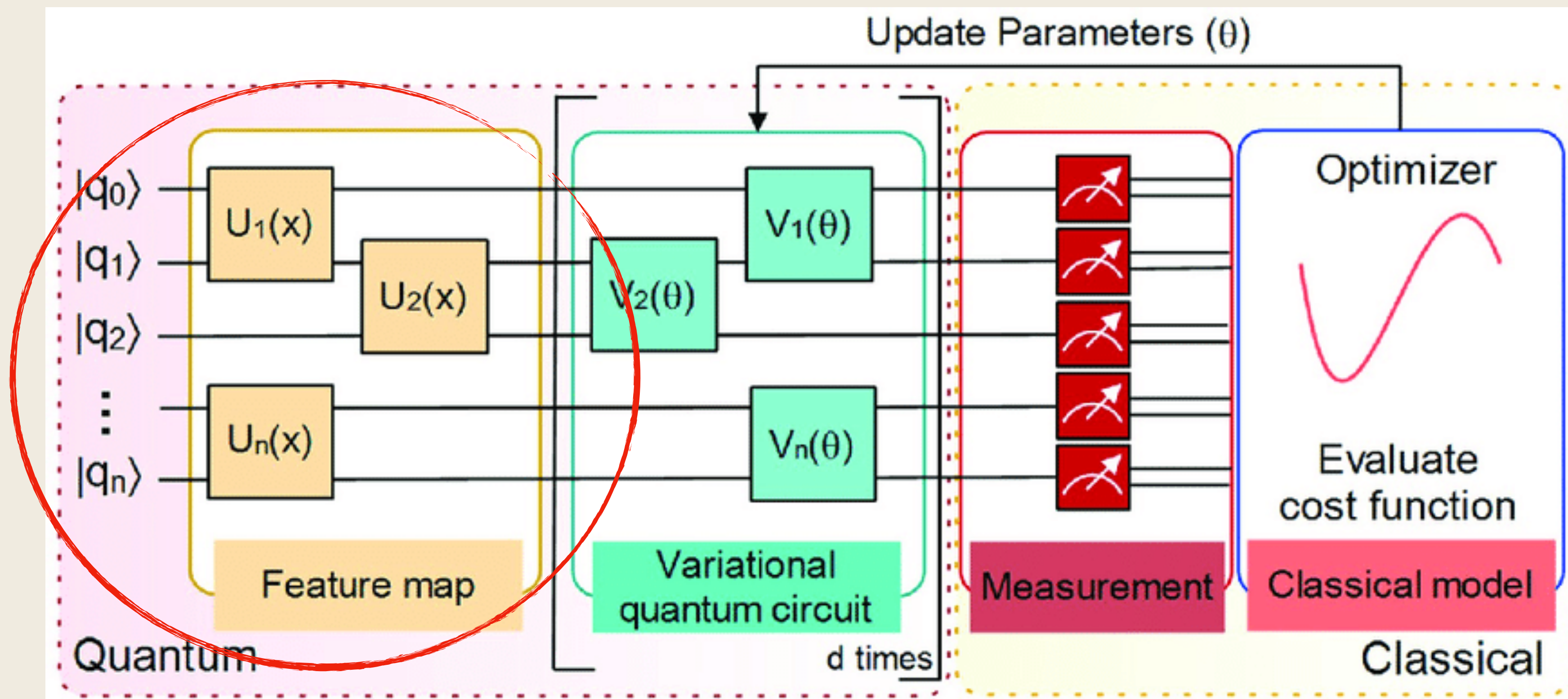
- Have achieved compression with 8 qubits
- Almost demonstrated storage/retrieval of IceCube simulation data with 14 qubits

# Plan with IceCube data

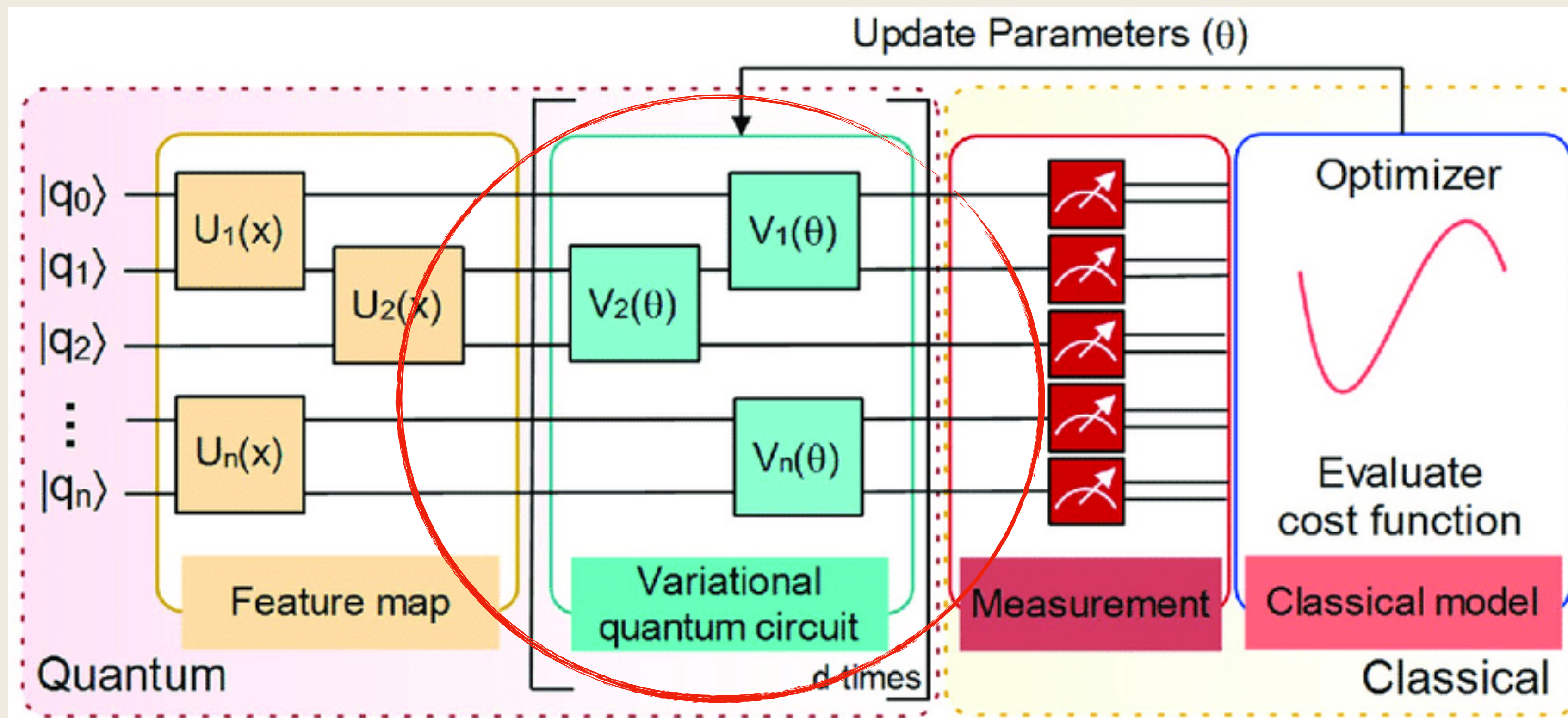
- Take set of IceCube data: tracks/cascades ( $\nu_\mu$  CC/ $\nu_e$  CC)
- Input into our QRAC
- Take states from QRAC, input them to train Variational Quantum Circuit
- First step, will be to investigate if VQC can classify tracks/cascades



# Variational Quantum Circuit

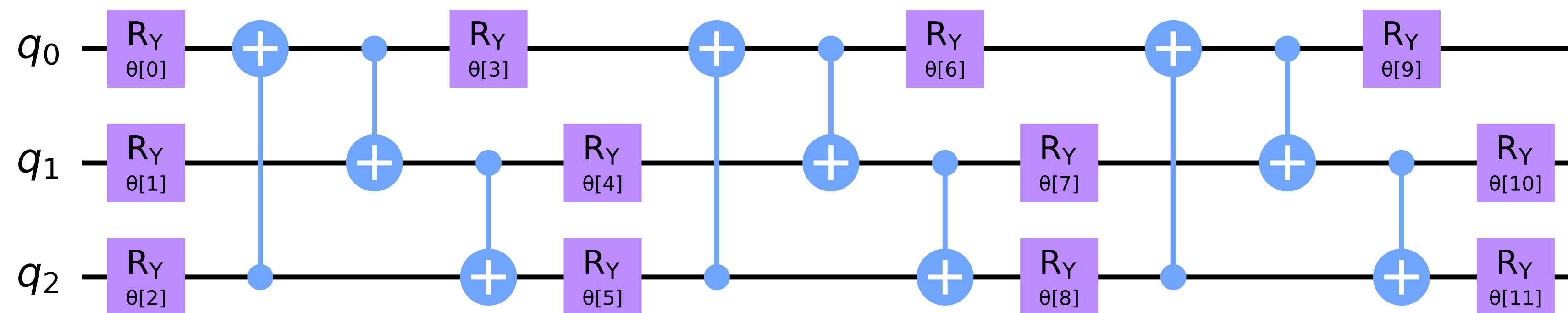


# Variational Quantum Circuit



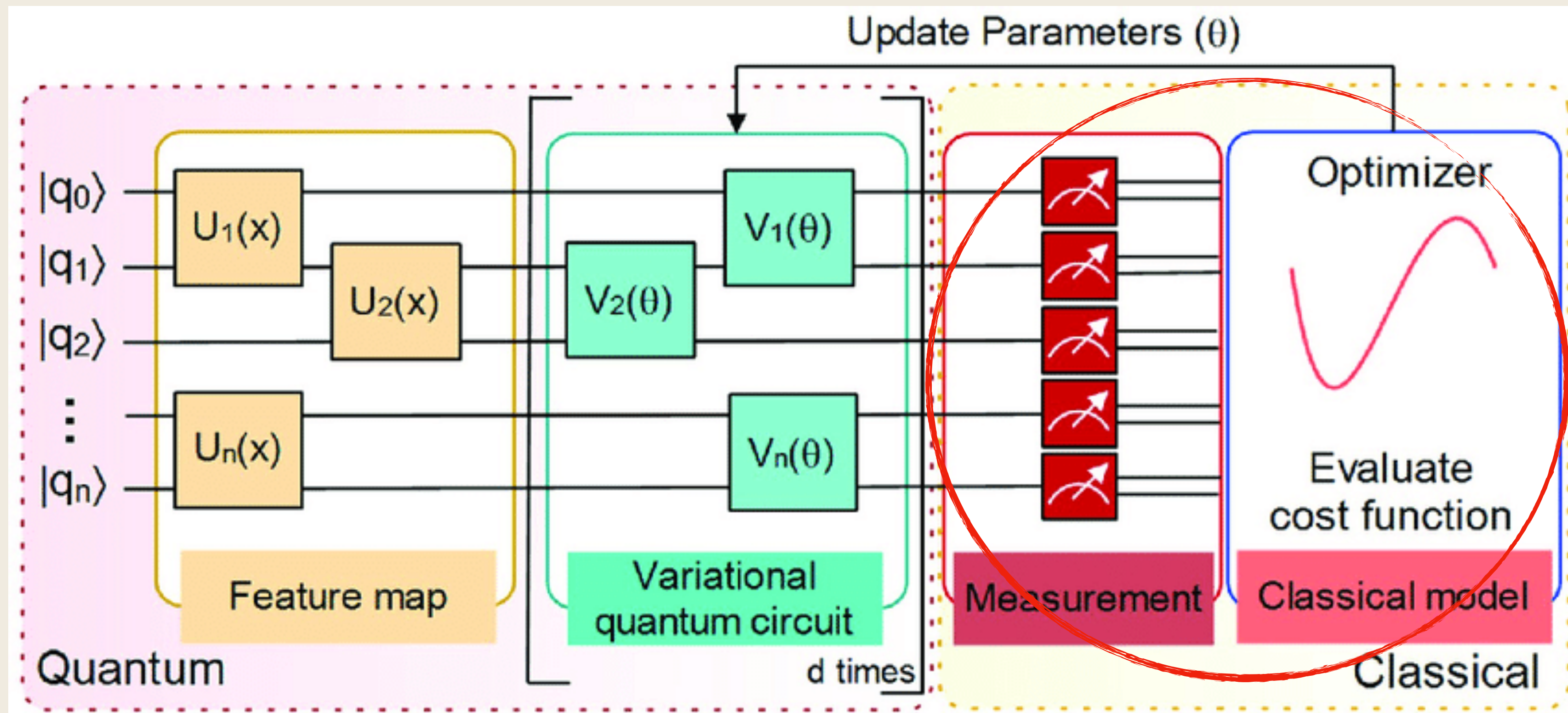
# Model Circuit

- Is the variational part, where machine learning happens, called the “ansatz”
- The ansatz is parameterized by a set of free parameters  $\theta$  that will be updated during training
- The structure of the ansatz, entanglement, type of rotations, number of parameters, number of gates, are all tunable



A 3 qubit example of a “RealAmplitudes” ansatz from qiskit.

# Variational Quantum Circuit



---

# Loss Function

$$\text{CrossEntropyLoss}(\text{predict}, \text{target}) = - \sum_{i=0}^{N_{\text{classes}}} \text{target}_i * \log(\text{predict}_i).$$

- Default: is cross-entropy loss
  - The difference between the ideal distribution (the true labels/target) and the measured distribution
- Calculated across a batch of samples, and the average is taken across the batch to obtain the final loss value for that iteration
- Others like MSE are also used
- Inherently a classical operation



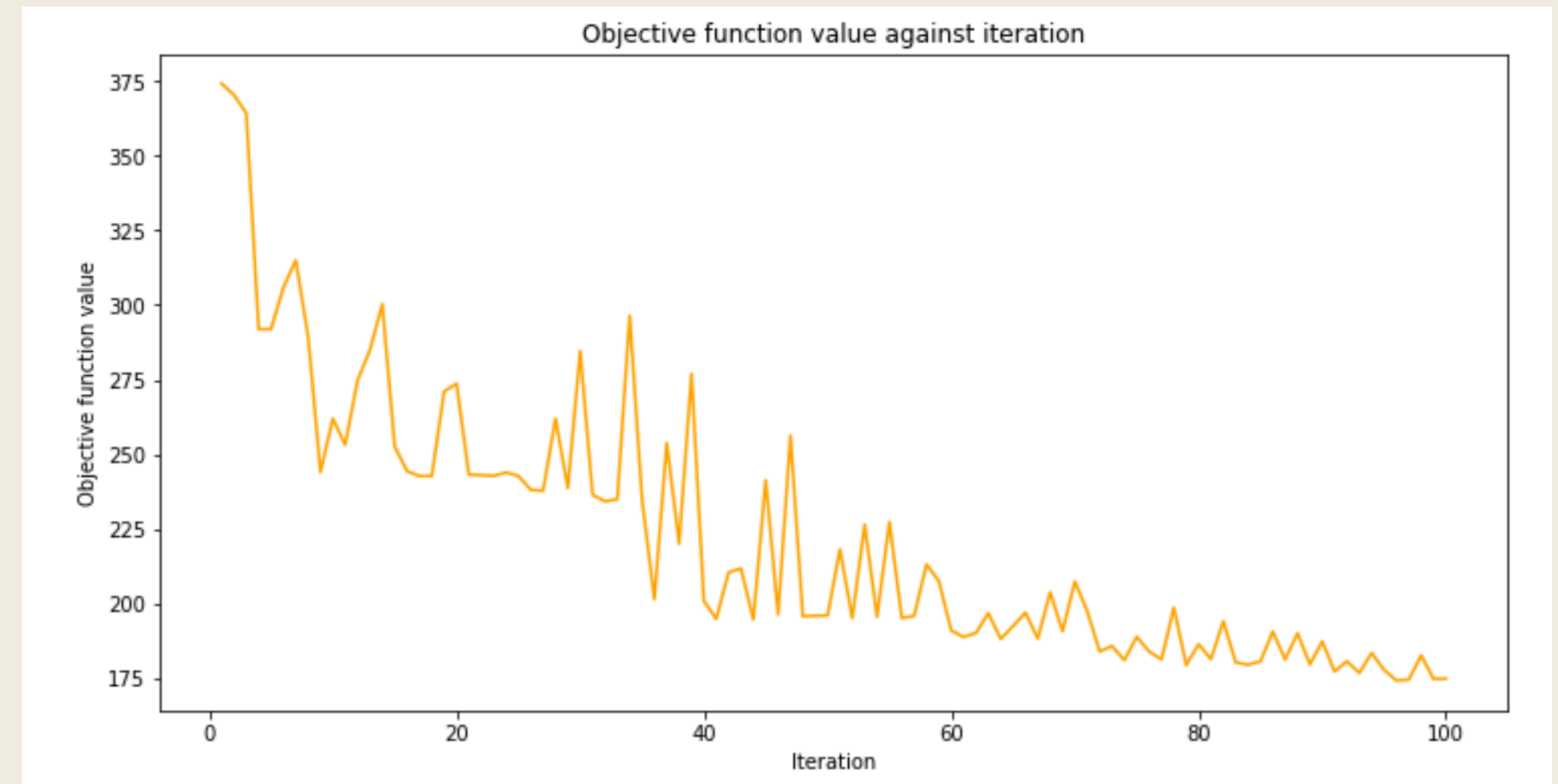
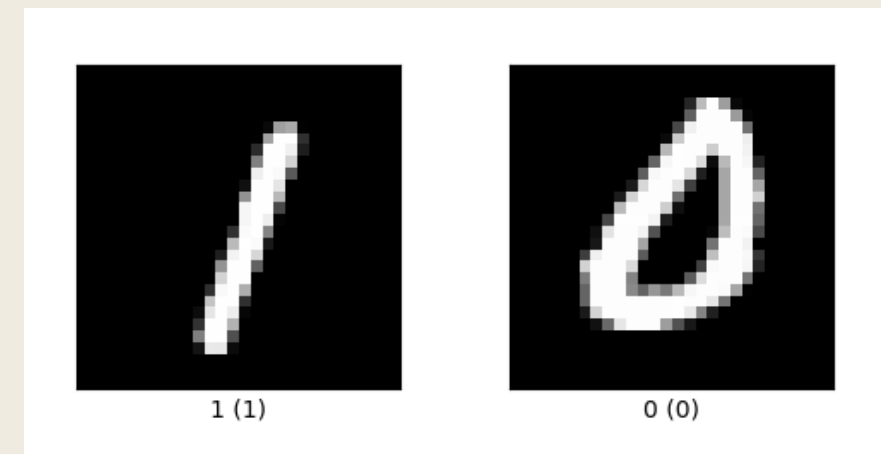
---

# Optimization

- Optimization performed using classical algorithms: COBYLA, ADAM, SPSA, etc.
  - Numerical Differentiation has been the norm for a while using parameter-shift differentiation
  - It's unclear how intermediate derivatives could be stored/reused inside a quantum computation
- Investigating if we can use quantum optimization techniques like quantum annealing
- A lot is still unknown about what is optimal, more of an art to test
- "There is often little theoretical motivation for any given optimizer, although recent work has been to analyze the training dynamics of QVCs (VQCs)"

# Proof of Concept using Qiskit

- MNIST: image classification 0s and 1s
- Input image via amplitude encoding into 8 qubit VQC
- ~about 100 tunable parameters
- Iterations are a hyper parameter
- COBYLA, cross-entropy loss function
- Training score: 97%
- Testing score: 98%



---

# What's next?

- Comparing VQC structure with amplitude encoding of IceCube simulation data with QRAC encoding
- Takes ~16 qubits to encode largest IceCube events via amplitudes
- Need larger quantum computers than publicly available (applied for ORNL grant)
- Now have access to ~100 qubit QCs
- Choose optimal qubits in target QCs, implement custom feature maps, testing ansatz types, optimizations, loss functions etc.
- Transition away from Qiskit to PennyLane (a QML framework)



---

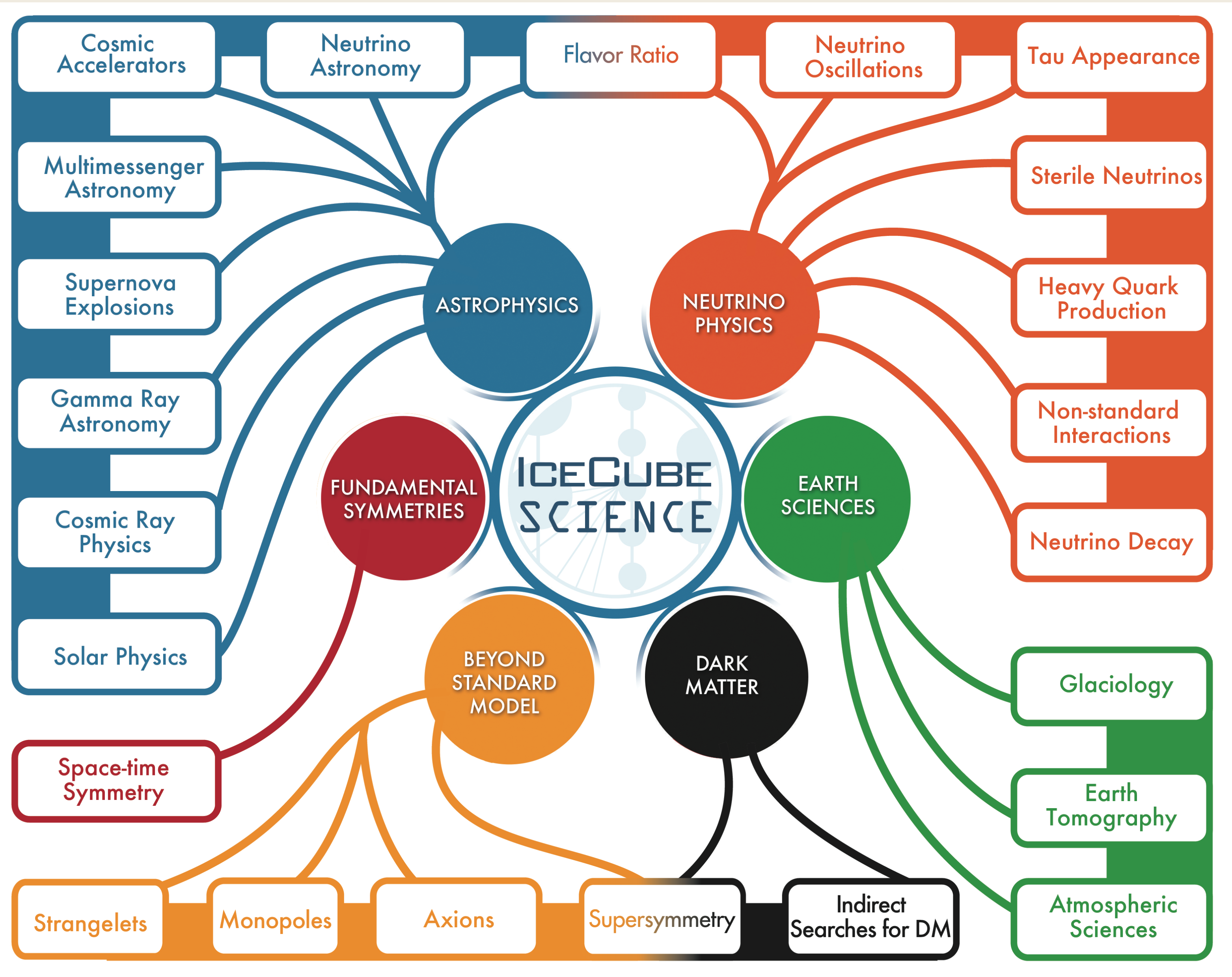
# Conclusion/Takeaway

- ML in neutrino physics has been very successful the past two decades and will continue to contribute
- Yet with concerns about the streetlight effect and increasing data loads in next generation experiments/observatories, we might need new tools
- QC can take advantage of additional degrees of freedom to enhance data compression
- QCML is still a burgeoning field; much is still unknown about optimizations, actual quantum speedups, feasibility on current NISQ computers, etc.
- ***“The question on whether quantum computers can really play a role in identifying practical ML applications is still wide open, and it is unlikely to be decided by theoretical proofs...”***

---

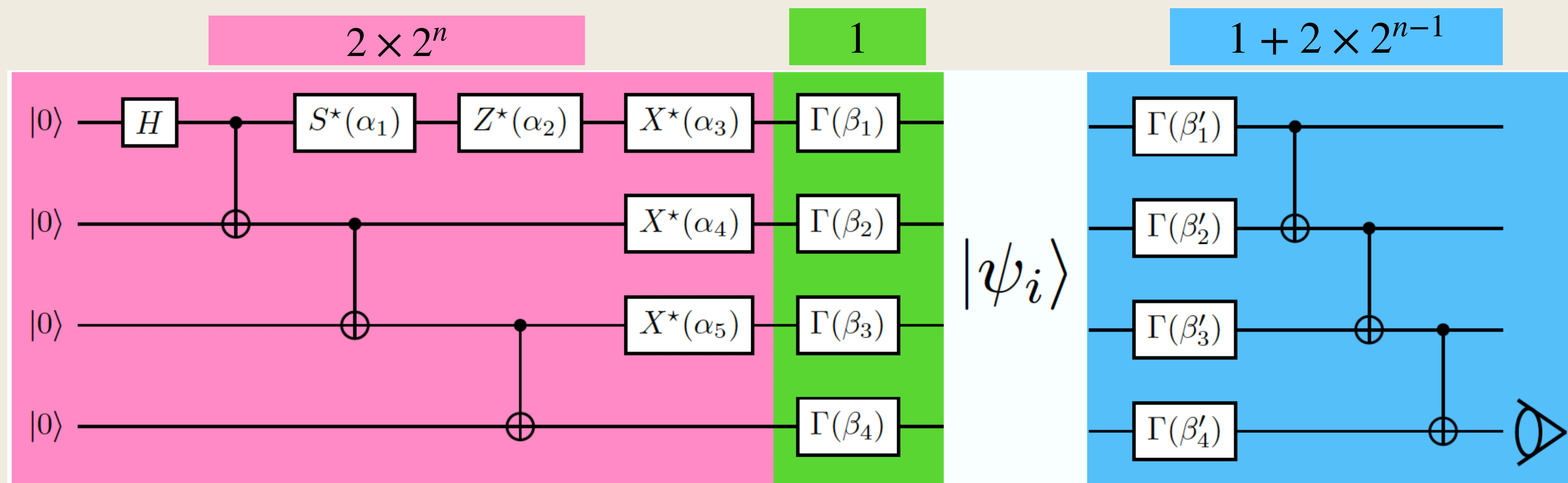
Gracias! Thank you! Preguntas?  
Questions?

New arm of hectopus:  
Quantum Computing!





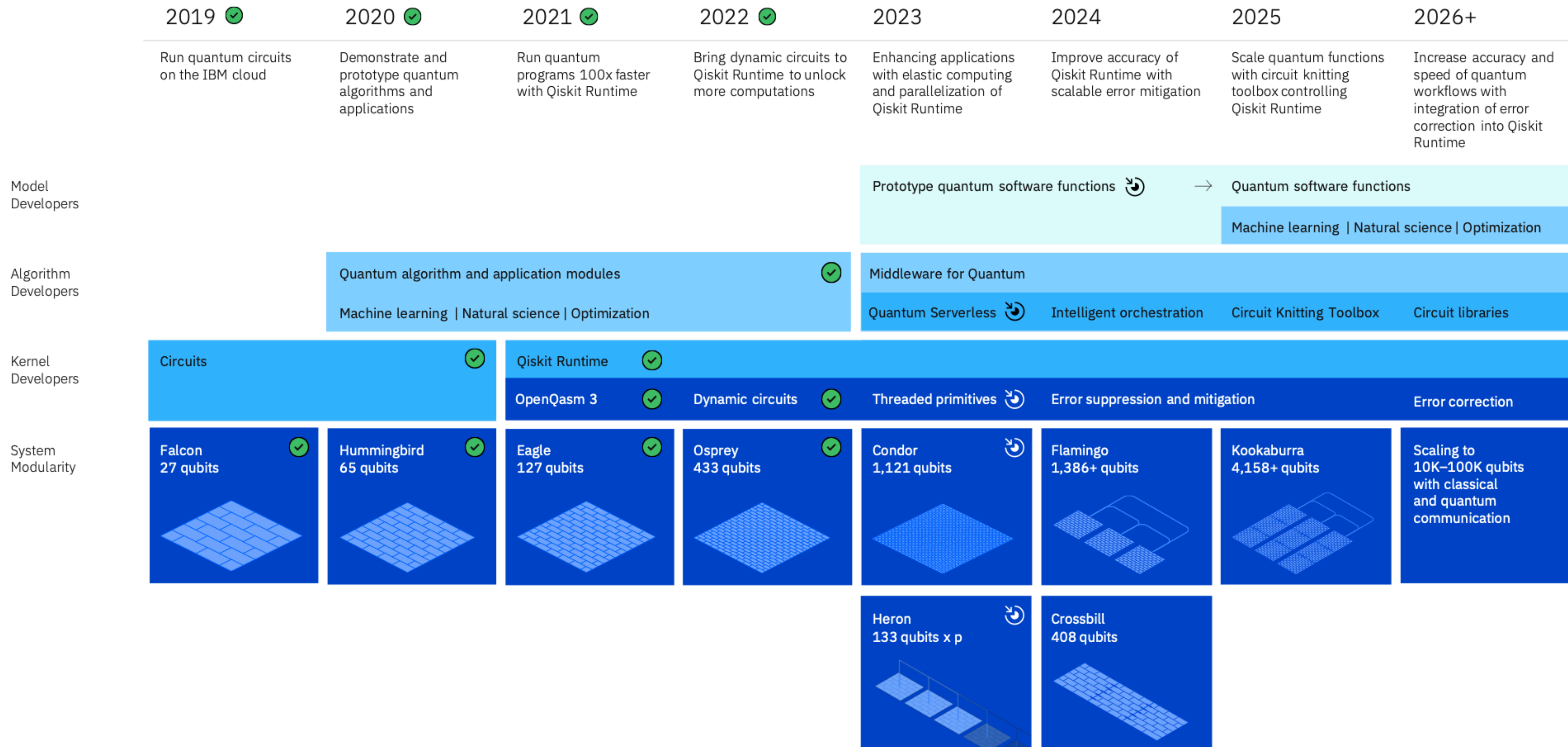
# Measuring Context Eigenstates

- The values for one context are all related. Only need to calculate one "fingerprint" and figure out mapping
- Scales like  $4^N$
- For 12 qubits this can be done in a few hours and stored in 2 MB



# Development Roadmap

Executed by IBM   
On target 





---

# Overview

- Introduction to Quantum Computing
- Application of QC to HEP
- QCML in HEP
- QCML for neutrino astronomy
- Variational Quantum Circuits
- Conclusion