# Hybrid Recurrent Architectures for Quantum-Classical NLP

**Stephen Clark**

OxML 2023

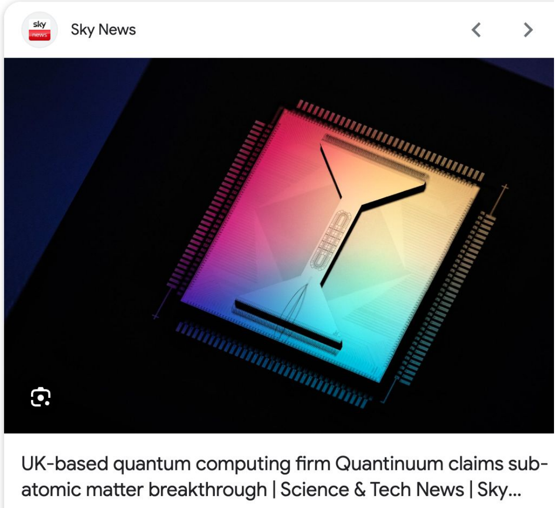
University of Oxford

11 July 2023

QUANTINUUM

# Quantum Computing

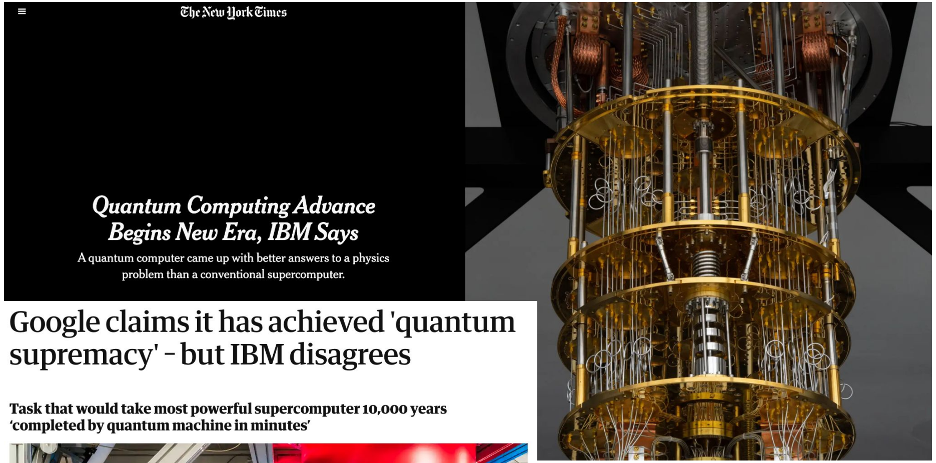






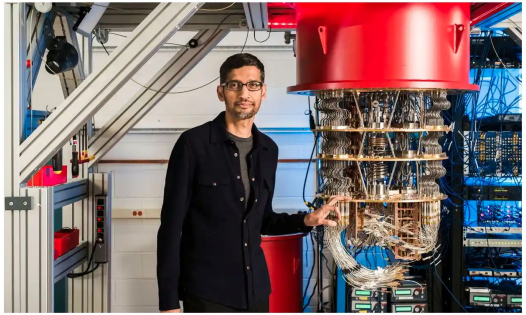

QUANTINUUM

# Quantum Computers

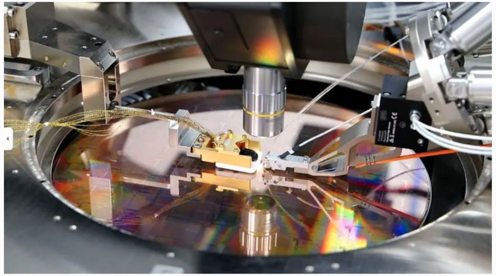



## Google claims it has achieved 'quantum supremacy' - but IBM disagrees

**Task that would take most powerful supercomputer 10,000 years 'completed by quantum machine in minutes'**

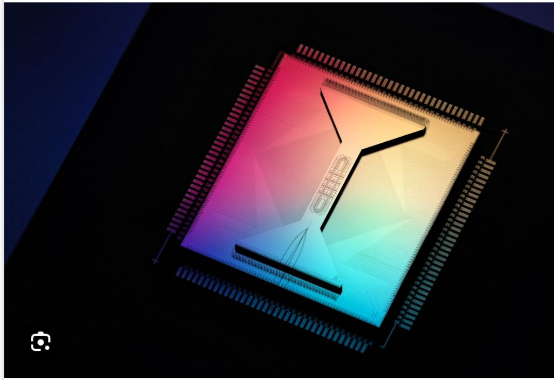


Sundar Pichai, pictured with the Sycamore Quantum processor, compared the feat to building the first rocket to reach space. Photograph: Reuters



PsiQuantum PSIQUANTUM

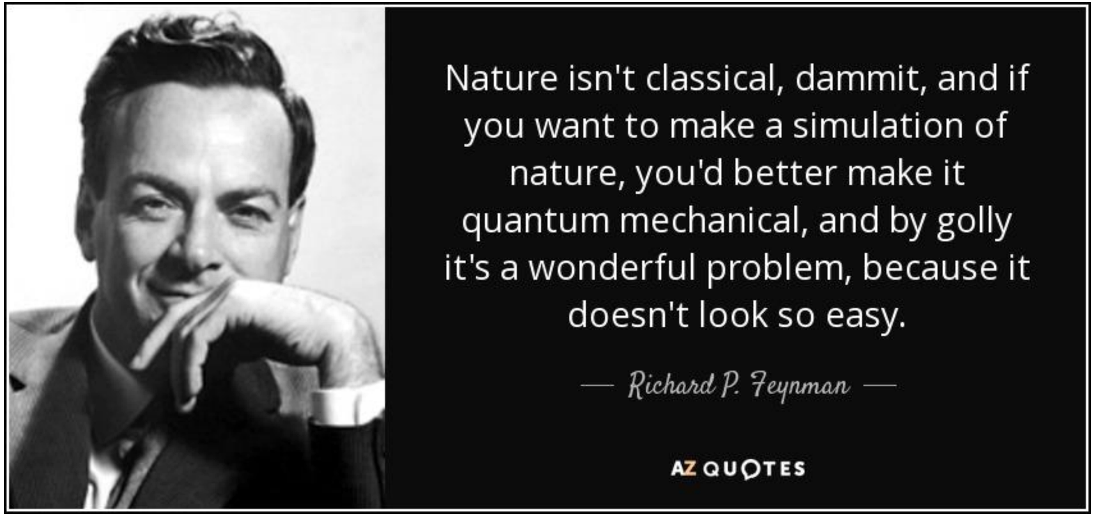
Sky News



UK-based quantum computing firm Quantinuum claims sub-atomic matter breakthrough | Science & Tech News | Sky...

QUANTINUUM

# Quantum Chemistry

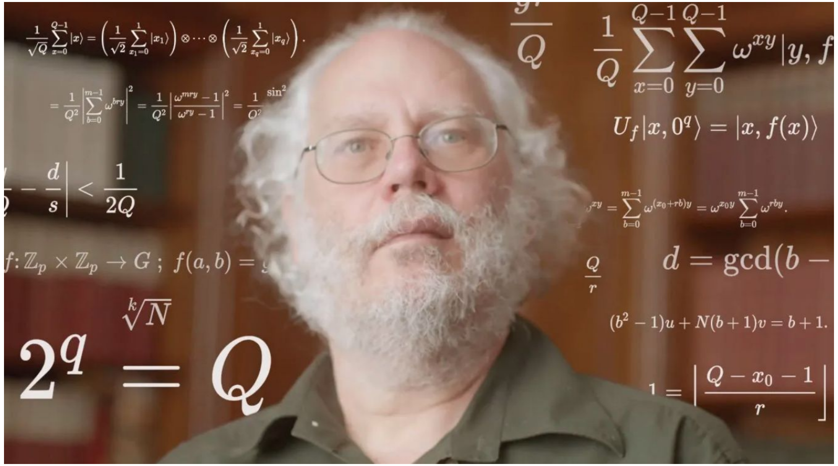



Early 1980s


QUANTINUUM

# Quantum Cryptography

## How Peter Shor's Algorithm Dooms RSA Encryption to Failure




Mid-1990s

(origin) CQ

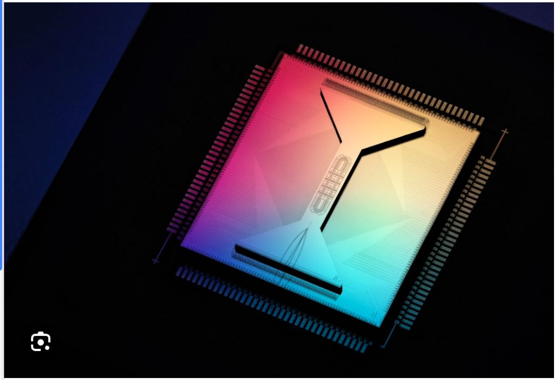
SECURE TODAY,
SECURE TOMORROW

QUANTINUUM

# Quantum ML and NLP



Sky News



UK-based quantum computing firm Quantinuum claims sub-atomic matter breakthrough | Science & Tech News | Sky...



QUANTINUUM

# Lecture Outline

1. Introduction to quantum computing / quantum circuits
2. Application to sequence classification
    a. our hybrid quantum RNN architectures
    b. sentiment analysis experiments

QUANTINUUM

# The State of a Classical Bit

$\psi$

$0$

# The State of a Classical Bit

$\psi$

1

# The State of a Qubit

$$|\psi\rangle$$

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad \alpha, \beta \in \mathbb{C} \qquad |\psi\rangle \in \mathbb{C}^2$$

*superposition*

# The State of a Qubit

$$|\psi\rangle$$

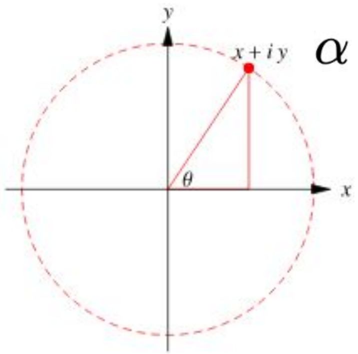
*amplitudes*

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad \alpha, \beta \in \mathbb{C} \qquad \boxed{|\alpha|^2 + |\beta|^2 = 1}$$

*superposition*

QUANTINUUM

# The State of a Qubit




$|\psi\rangle$

*amplitudes*

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad \alpha, \beta \in \mathbb{C} \qquad \boxed{|\alpha|^2 + |\beta|^2 = 1}$$

*superposition*

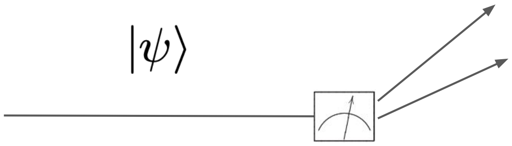
QUANTINUUM

# Measuring a Qubit

$|0\rangle$ with probability $|\alpha|^2$

$|1\rangle$ with probability $|\beta|^2$

$|\psi\rangle$

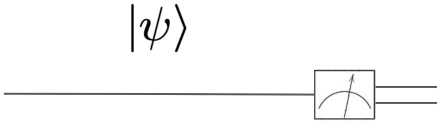


$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

*superposition*

$$|\alpha|^2 + |\beta|^2 = 1$$

# Measuring a Qubit (scalar output)




$$|\psi\rangle = |0\rangle$$

1 with probability $|\alpha|^2$

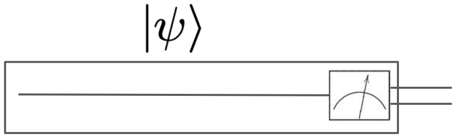
-1 with probability $|\beta|^2$

$$|\psi\rangle = |1\rangle$$

$$\cancel{|\psi\rangle = \alpha|0\rangle + \beta|1\rangle}$$

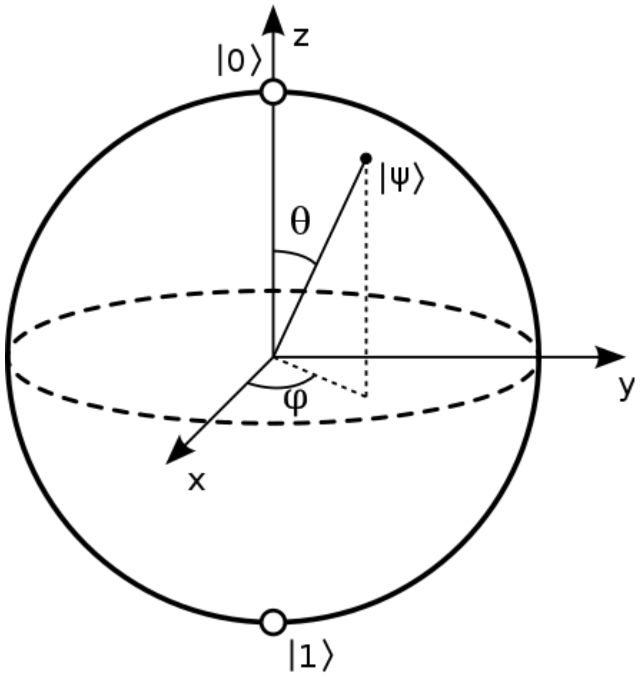
*~~superposition~~*

# Measuring a Qubit (many times)




$$\langle M \rangle \in [-1, 1]$$

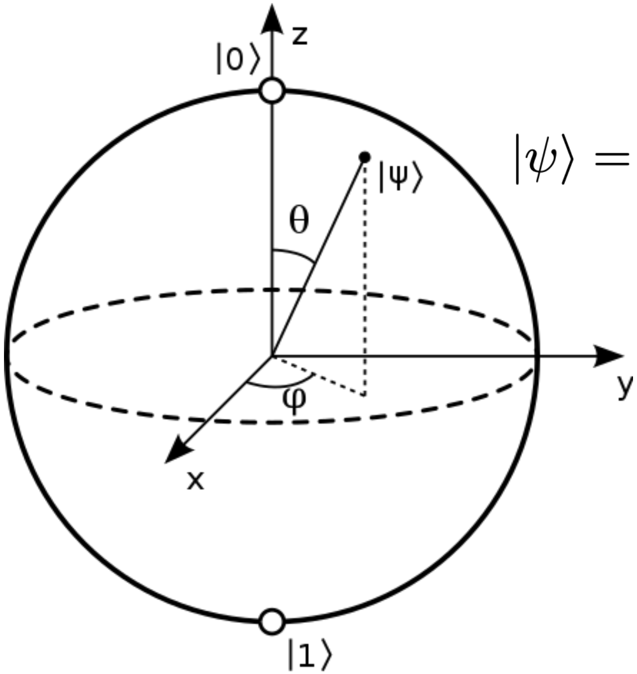
# The Bloch Sphere Representation of a Qubit




$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

https://en.wikipedia.org/wiki/Bloch_sphere

# The Bloch Sphere Representation of a Qubit

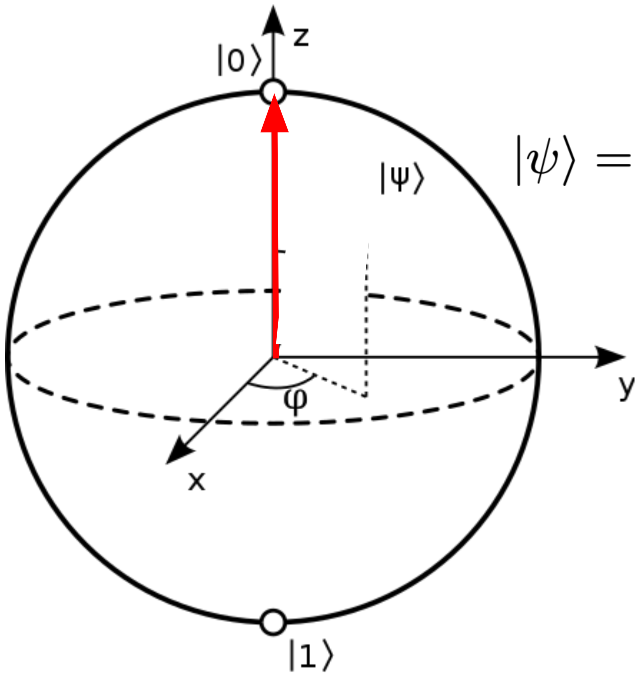



$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

https://en.wikipedia.org/wiki/Bloch_sphere

QUANTINUUM

# "The Collapse of the Wave Function"



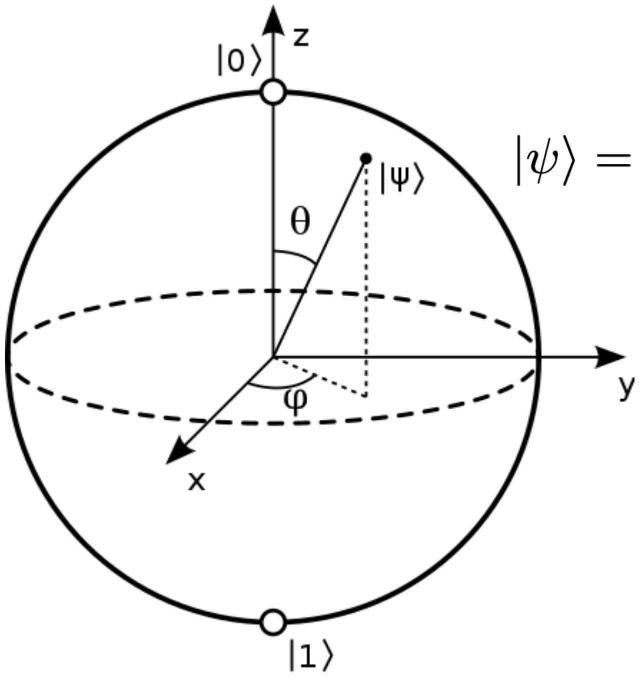

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

$$\boxed{\cos^2(\theta/2)}$$

https://en.wikipedia.org/wiki/Bloch_sphere

QUANTINUUM

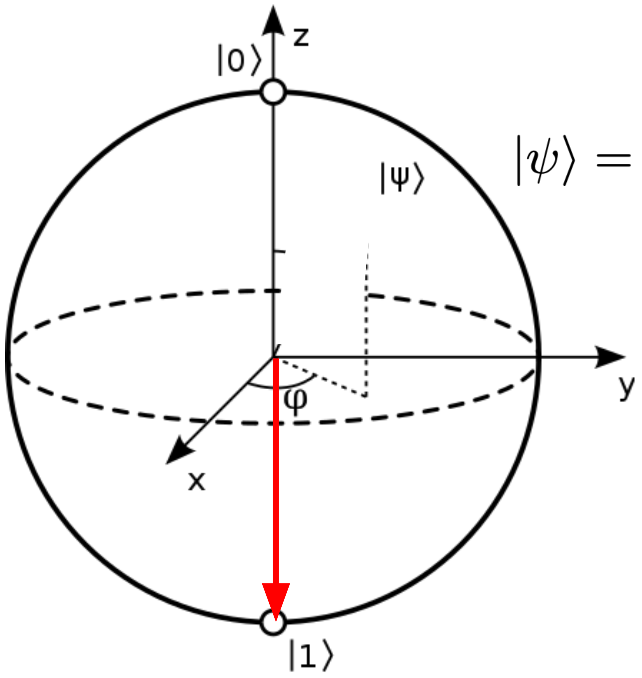
# The Bloch Sphere Representation of a Qubit




$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

https://en.wikipedia.org/wiki/Bloch_sphere

# “The Collapse of the Wave Function”

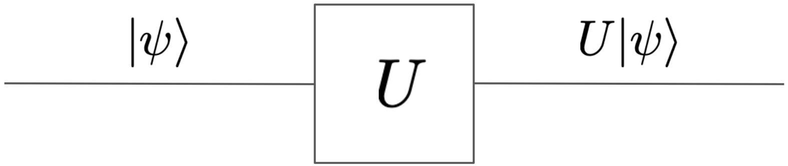



$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

$$\boxed{\sin^2(\theta/2)}$$

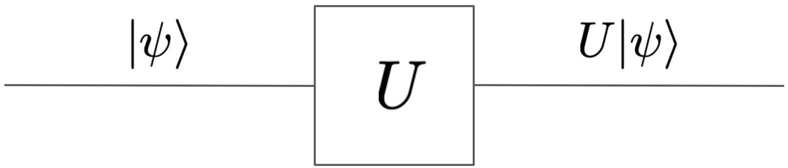
https://en.wikipedia.org/wiki/Bloch_sphere

# Linear Transformations of a Qubit




$$U : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha U|0\rangle + \beta U|1\rangle$$

# Unitary Transformations of a Qubit

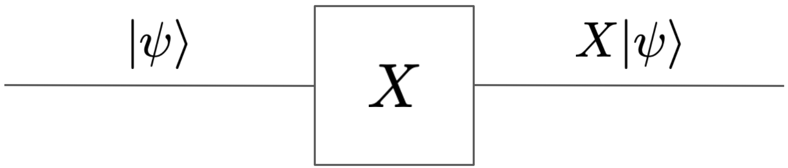



$$U : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha'|0\rangle + \beta'|1\rangle$$

$$\boxed{|\alpha'|^2 + |\beta'|^2 = 1}$$

# 1-Qubit Quantum Gates

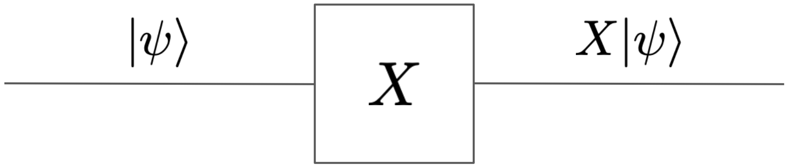
quantum Not gate




$$X : |0\rangle \mapsto |1\rangle$$
$$X : |1\rangle \mapsto |0\rangle$$

# 1-Qubit Quantum Gates

quantum Not gate *acts linearly*

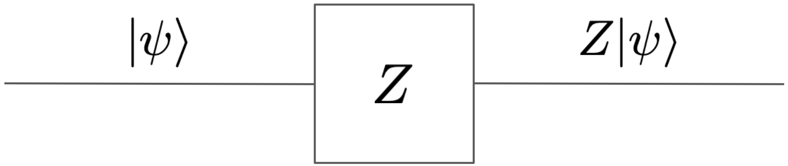



$$X : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|1\rangle + \beta|0\rangle$$

QUANTINUUM

# 1-Qubit Quantum Gates

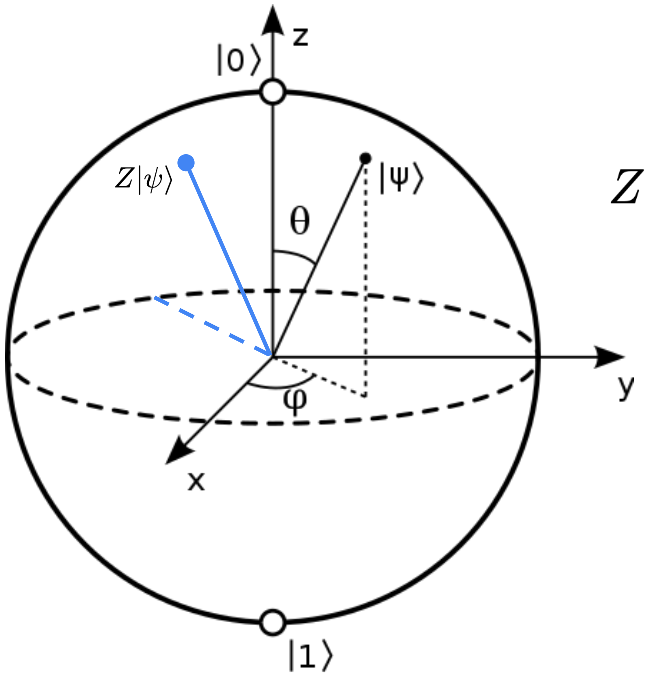
Pauli Z Gate




$$Z : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle - \beta|1\rangle$$

# 1-Qubit Quantum Gates

Pauli Z Gate *rotates about the Z axis*

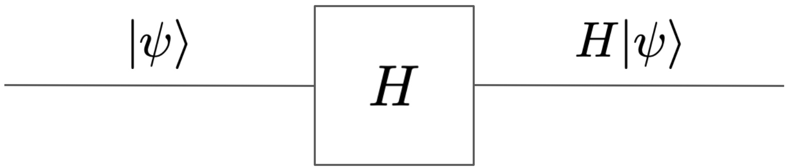



$$Z : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle - \beta|1\rangle$$

QUANTINUUM

# 1-Qubit Quantum Gates

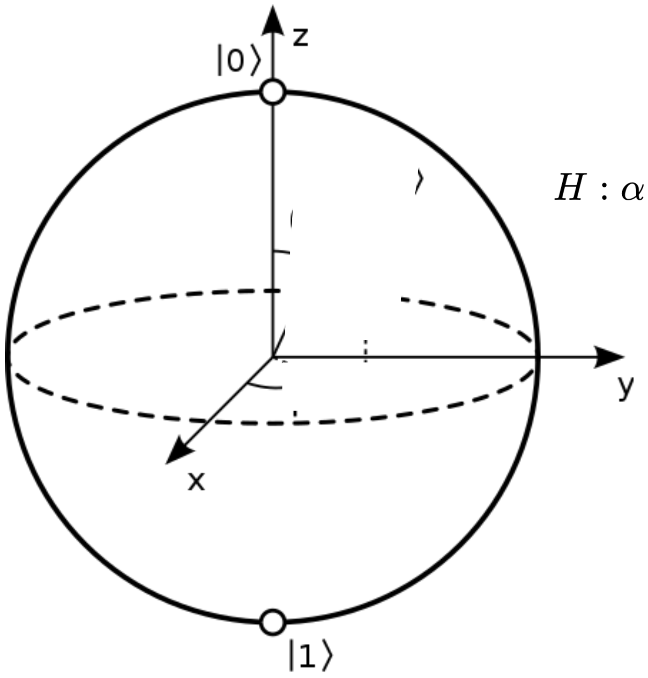
Hadamard Gate




$$H : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

# 1-Qubit Quantum Gates

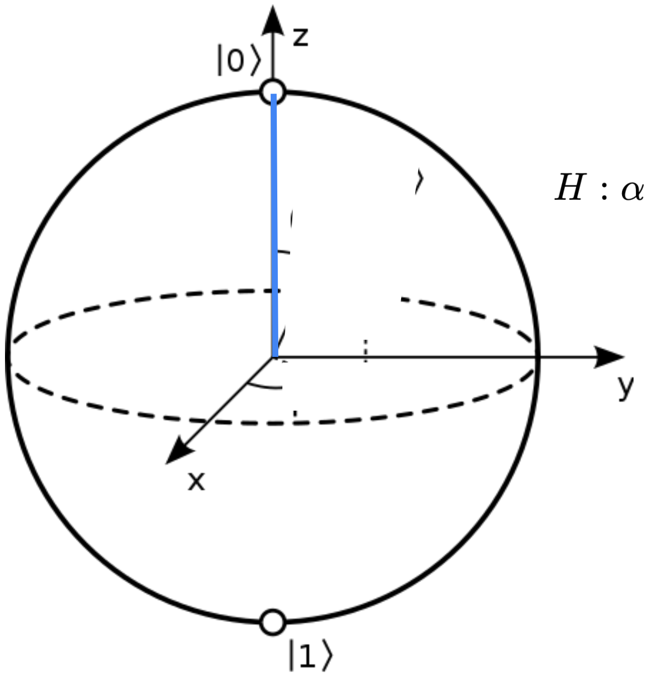
Hadamard Gate *introduces superposition*




$$H : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

# 1-Qubit Quantum Gates

Hadamard Gate *introduces superposition*

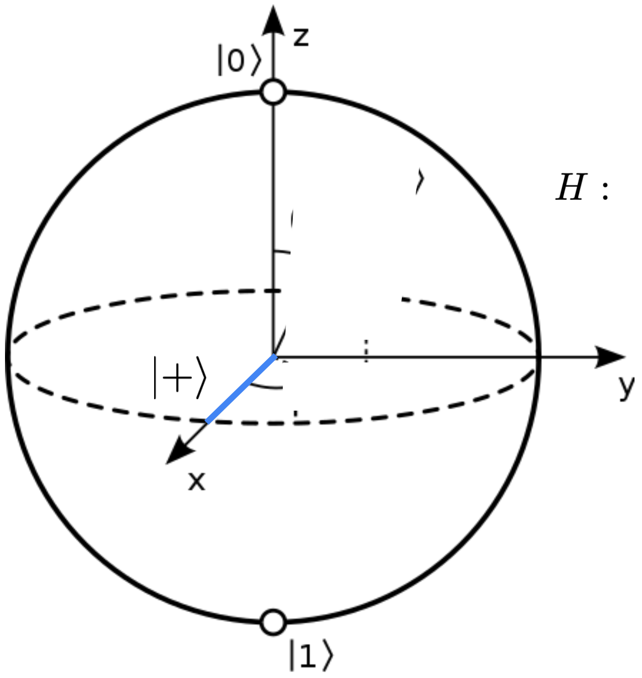



$$H : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

QUANTINUUM

# 1-Qubit Quantum Gates

Hadamard Gate *introduces superposition*

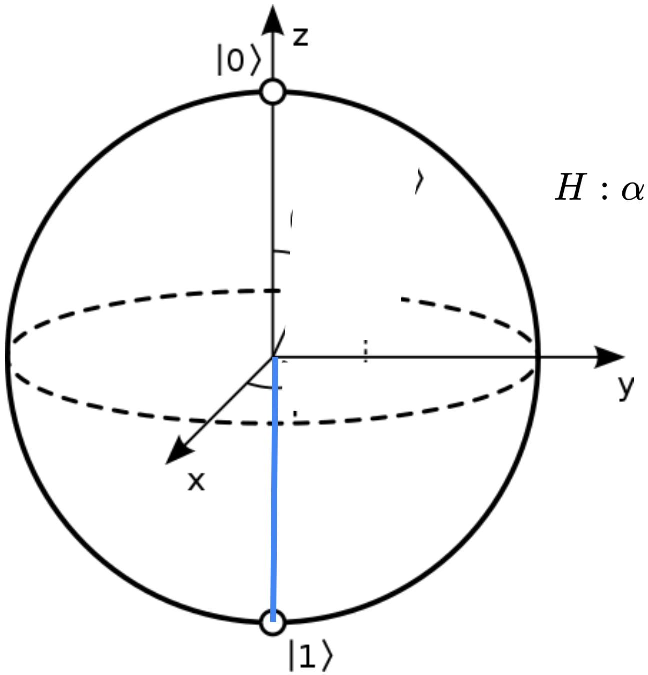



$$H: \quad |0\rangle \mapsto \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

QUANTINUUM

# 1-Qubit Quantum Gates

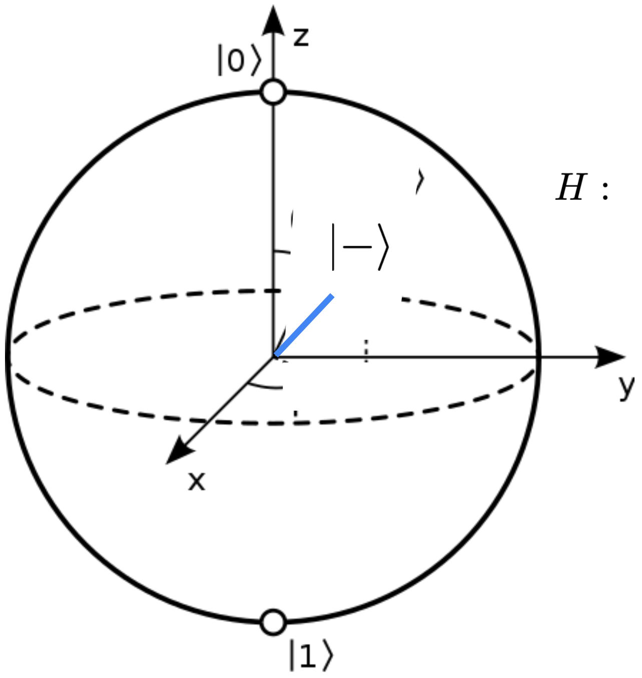
Hadamard Gate *introduces superposition*




$$H : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

# 1-Qubit Quantum Gates

Hadamard Gate *introduces superposition*




$$H : \quad |1\rangle \mapsto \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

QUANTINUUM

# The State of Many Qubits

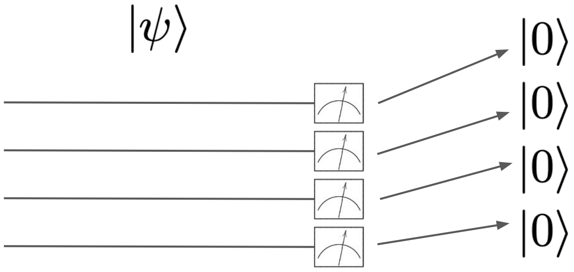
$$|\psi\rangle$$

$$|\psi\rangle \in \mathbb{C}^{2^4}$$

$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

# Measuring Many Qubits

$$|\psi\rangle \in \mathbb{C}^{2^4}$$

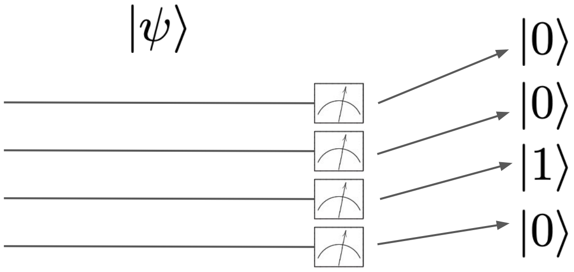



$$|\psi\rangle = \alpha_{0000}\boxed{|0000\rangle} + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

$$\boxed{|\alpha_{0000}|^2}$$

QUANTINUUM

# Measuring Many Qubits

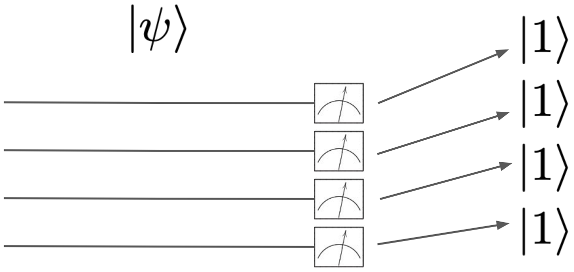



$$|\psi\rangle \in \mathbb{C}^{2^4}$$

$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}\boxed{|0010\rangle} + \ldots \alpha_{1111}|1111\rangle$$

$$\boxed{|\alpha_{0010}|^2}$$

# Measuring Many Qubits



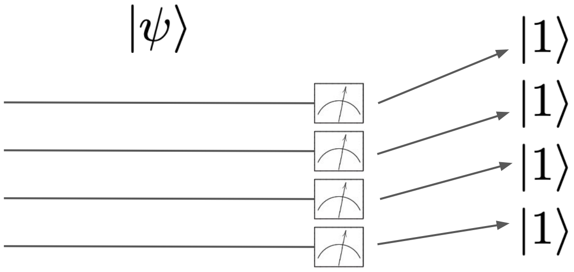

$$|\psi\rangle \in \mathbb{C}^{2^4}$$

$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}\boxed{|1111\rangle}$$

$$\boxed{|\alpha_{1111}|^2}$$

# Measuring Many Qubits

$$|\psi\rangle \in \mathbb{C}^{2^4}$$




$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

$$\boxed{\sum_{b\in\{0,1\}^4} |\alpha_b|^2 = 1}$$

# Product States

$$|\psi\rangle$$

$$|\psi\rangle \in \mathbb{C}^{2^2}$$

$$|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

# Product States

$$|\psi\rangle$$

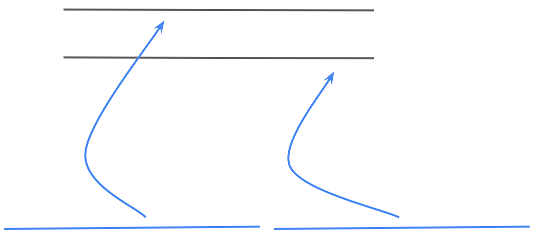
$$|\psi\rangle \in \mathbb{C}^{2^2}$$

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

# Product States

$$|\psi\rangle \in \mathbb{C}^{2^2}$$

$$|\psi\rangle$$



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

# Entangled States

$|\psi\rangle$

$$|\psi\rangle \in \mathbb{C}^{2^2}$$

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

QUANTINUUM

# Entangled States

$$|\psi\rangle \in \mathbb{C}^{2^2}$$

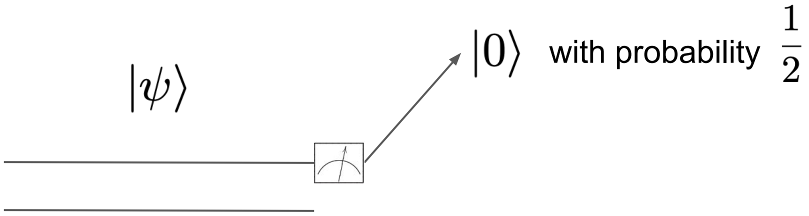
$$|\psi\rangle$$

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \underline{\quad\phi\quad} \; \underline{\quad\mu\quad}$$

?

QUANTINUUM

# (Strong) Correlation




$$|\psi\rangle \in \mathbb{C}^{2^2}$$

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

# (Strong) Correlation

$|\psi\rangle$

$|0\rangle$ with probability $\frac{1}{2}$

$|0\rangle$ with probability $1$

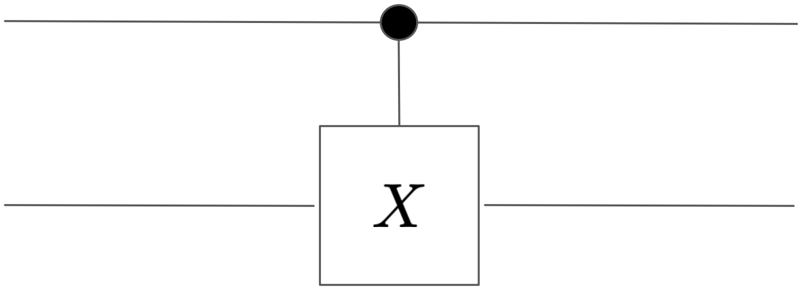
$|\psi\rangle \in \mathbb{C}^{2^2}$

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

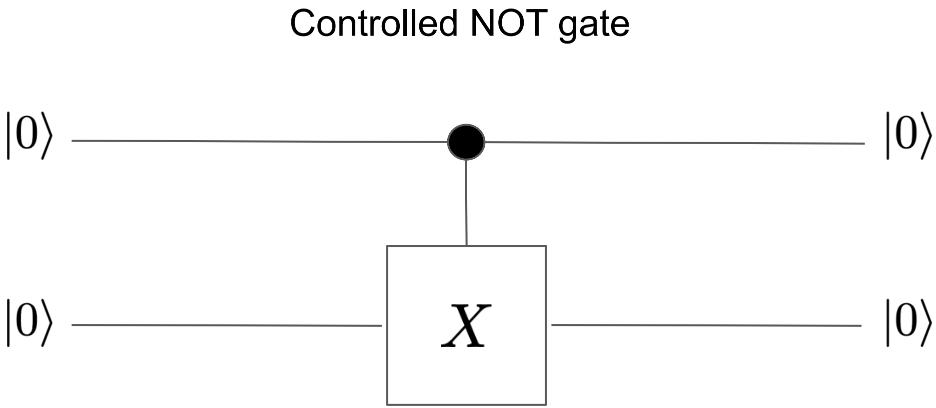
QUANTINUUM

# Entangling Qubits

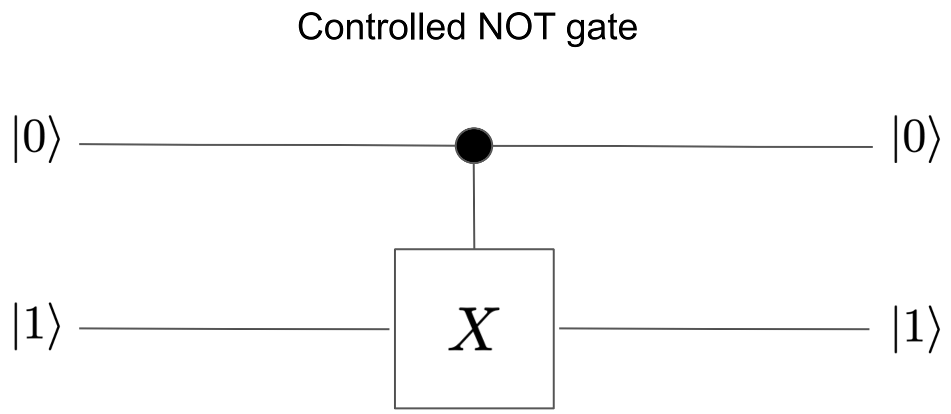
Controlled NOT gate



# Entangling Qubits

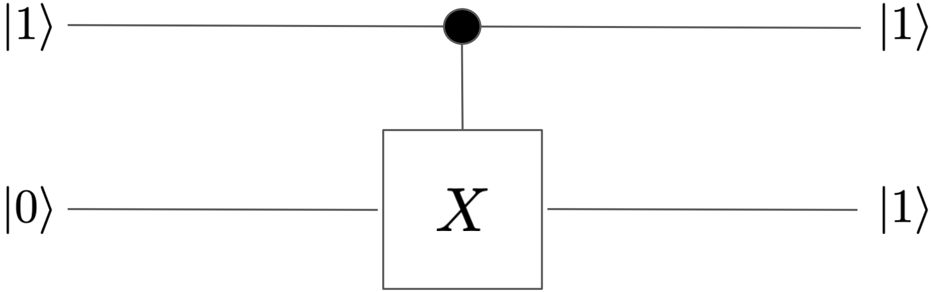


# Entangling Qubits



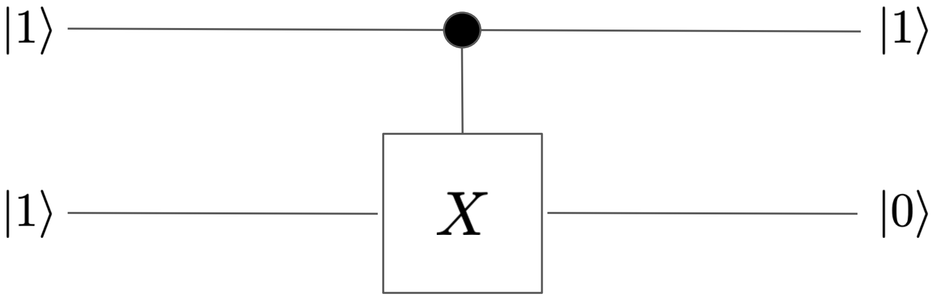
# Entangling Qubits

Controlled NOT gate



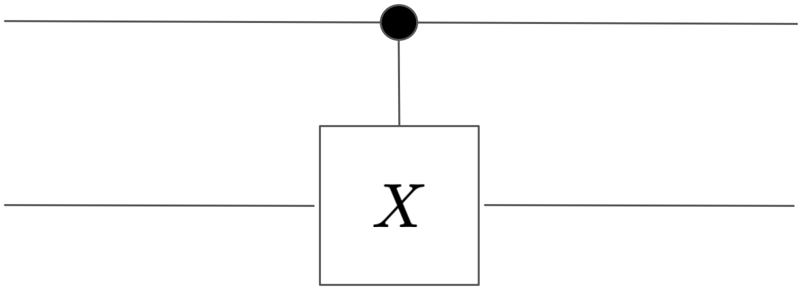
# Entangling Qubits

Controlled NOT gate



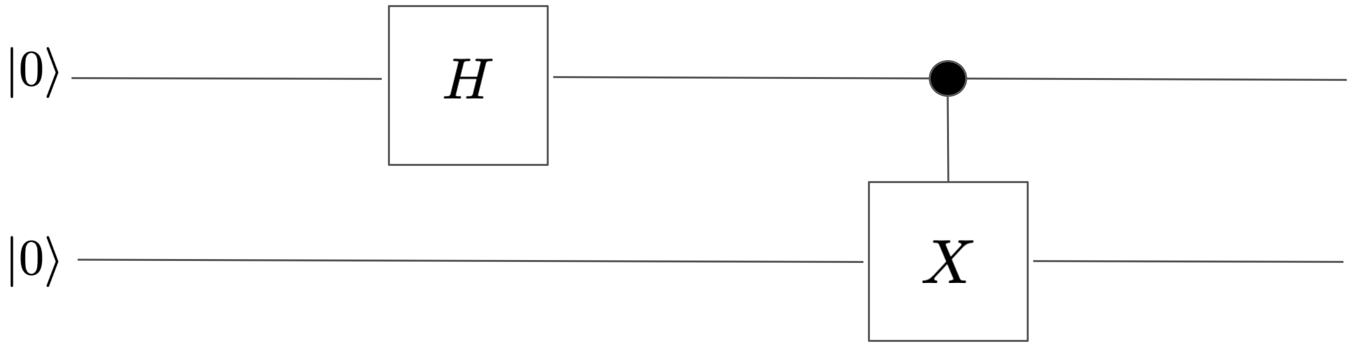
# Entangling Qubits

Controlled NOT gate *acts linearly*

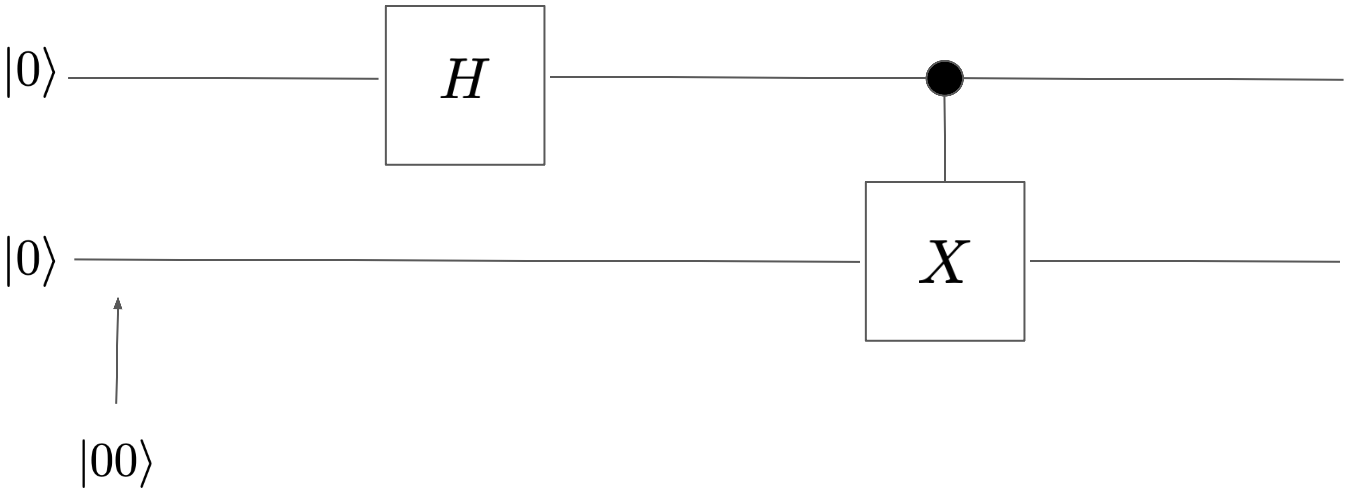



$$CX : \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle \mapsto \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|11\rangle + \alpha_{11}|10\rangle$$

# Creating a Bell State

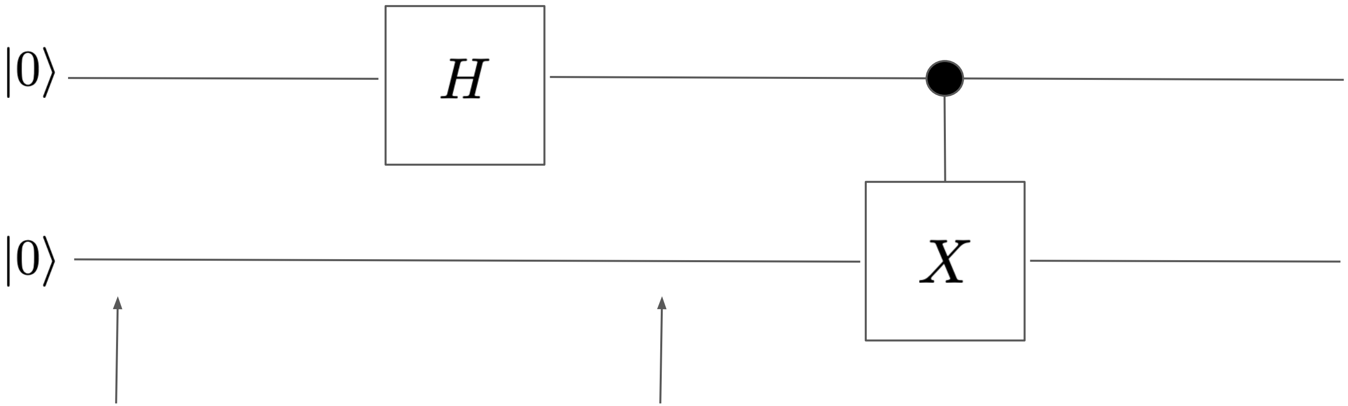


# Creating a Bell State




QUANTINUUM

# Creating a Bell State

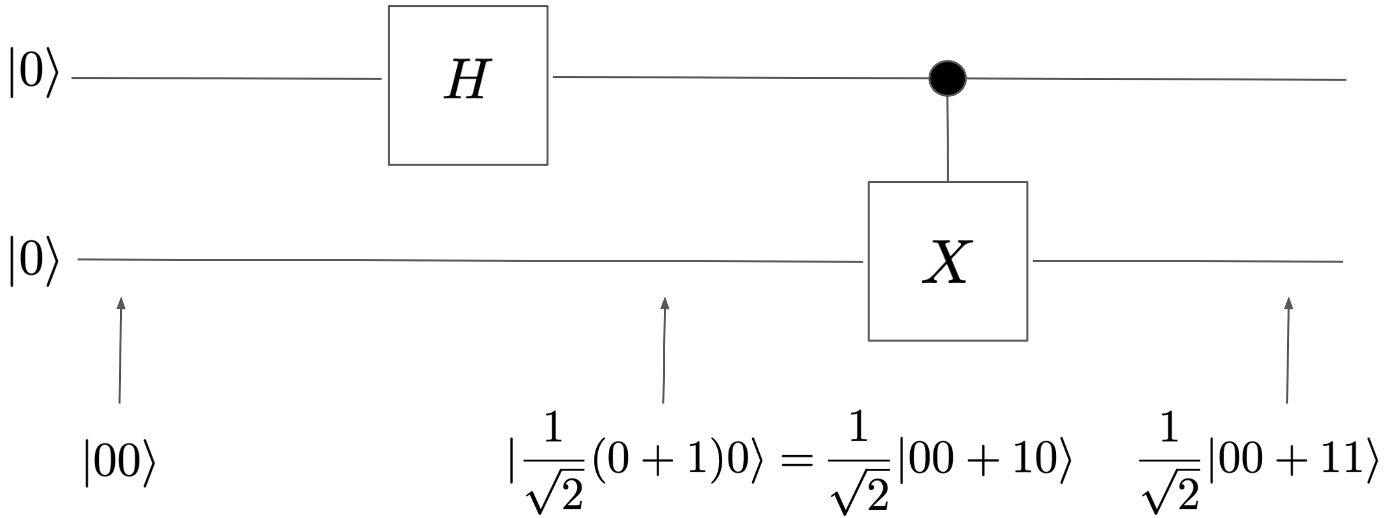



$$|00\rangle$$

$$|\frac{1}{\sqrt{2}}(0+1)0\rangle = \frac{1}{\sqrt{2}}|00+10\rangle$$

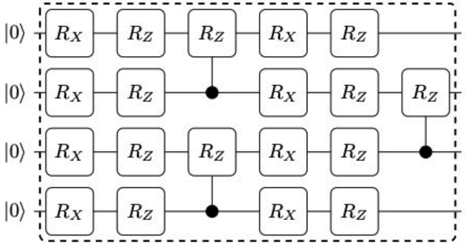
QUANTINUUM

# Creating a Bell State

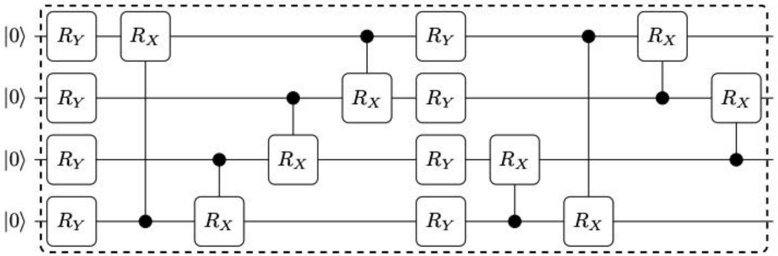



QUANTINUUM

# Quantum Circuits




Circuit 7



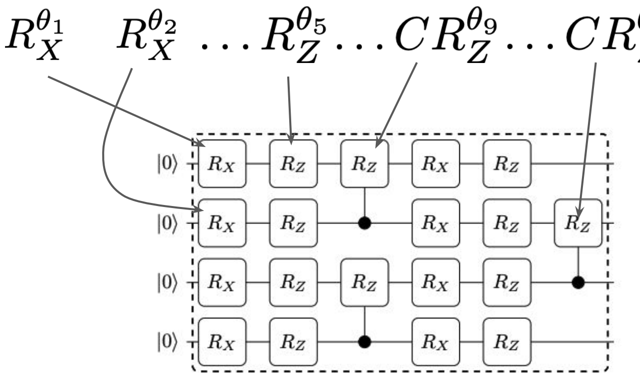

Circuit 14

**Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms**

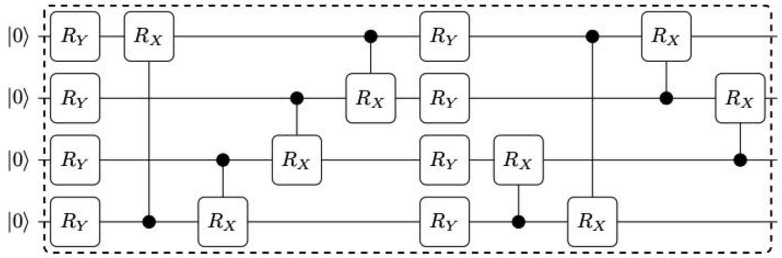
Sukin Sim,[1,2,*] Peter D. Johnson,[2] and Alán Aspuru-Guzik[2,3,4,5,†]

QUANTINUUM

# Parameterised Quantum Circuits (PQCs)




Circuit 7




Circuit 14

**Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms**

Sukin Sim,[1,2,*] Peter D. Johnson,[2] and Alán Aspuru-Guzik[2,3,4,5,†]

# Intermission

# Lecture Outline

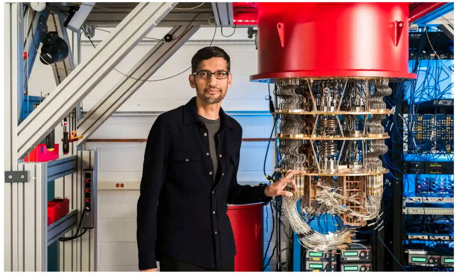
1. Introduction to quantum computing / quantum circuits
2. Application to sequence classification
   a. our hybrid quantum RNN architectures
   b. sentiment analysis experiments

# What's the Goal?

- Quantum Supremacy / Advantage?
  - execute an algorithm on a quantum computer that cannot be simulated efficiently on a classical computer (with an exponential speedup, e.g. Shor's algorithm)
  - solve a *useful* problem faster on a quantum computer than on a classical computer

## Google claims it has achieved 'quantum supremacy' - but IBM disagrees

**Task that would take most powerful supercomputer 10,000 years 'completed by quantum machine in minutes'**



Sundar Pichai, pictured with the Sycamore Quantum processor, compared the feat to building the first rocket to reach space. Photograph: Reuters

# What’s the Goal?

- Getting ready for better quantum hardware
  - so we can run potentially better models
  - many advances in QNLP / QML are likely to be experimental
- Other potential benefits of quantum models
  - perhaps more interpretable / explainable / transparent
  - perhaps more sample efficient / better at generalisation
  - something different to a DNN / Transformer!
- The models today can be simulated efficiently (with small qubit counts)

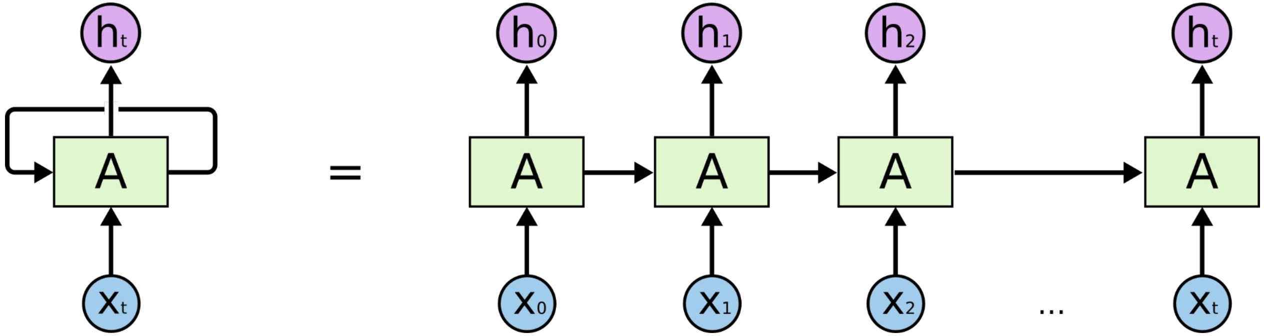
Perspective Open Access

Is Quantum Advantage the Right Goal for Quantum Machine Learning?

Maria Schuld and Nathan Killoran
PRX Quantum **3**, 030101 – Published 14 July 2022

QUANTINUUM

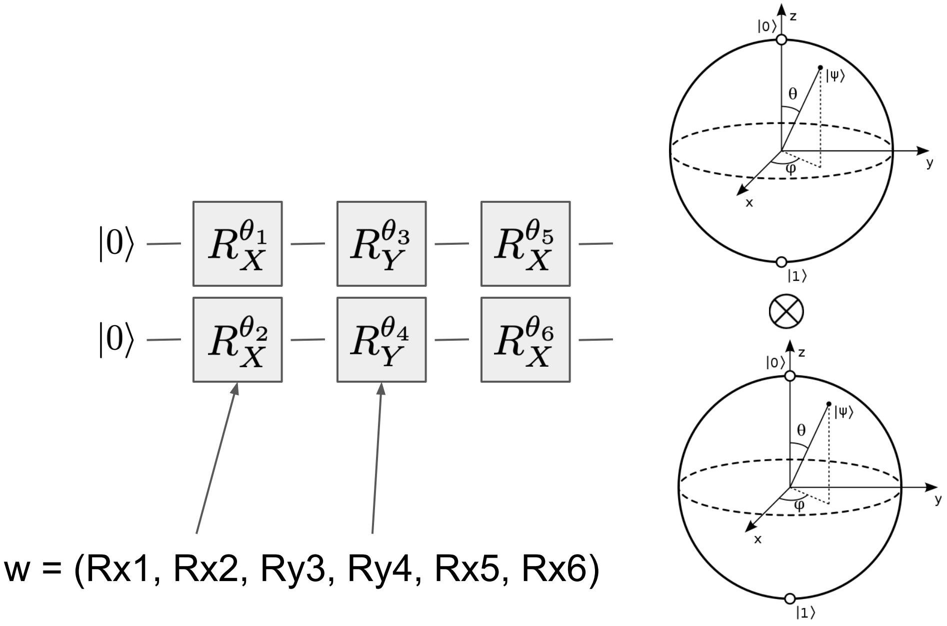
# Recurrent Neural Networks (RNNs)




$$h_t = f(x_t\mathbf{U} + h_{t-1}\mathbf{W})$$

From Colah's blog: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

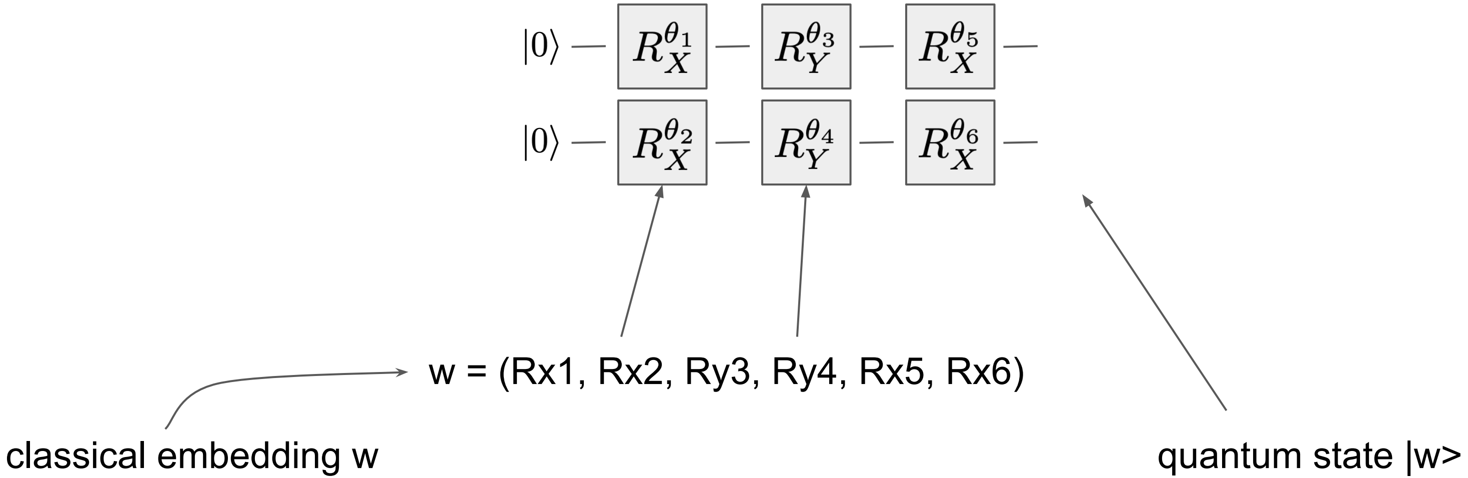
# Angle Encoding for Words



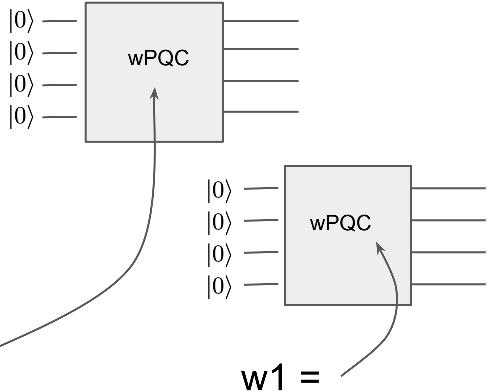

*Angle encoding*

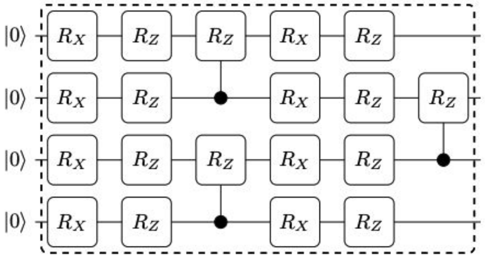
QUANTINUUM

# Angle Encoding for Words



# Word Encoding using PQCs



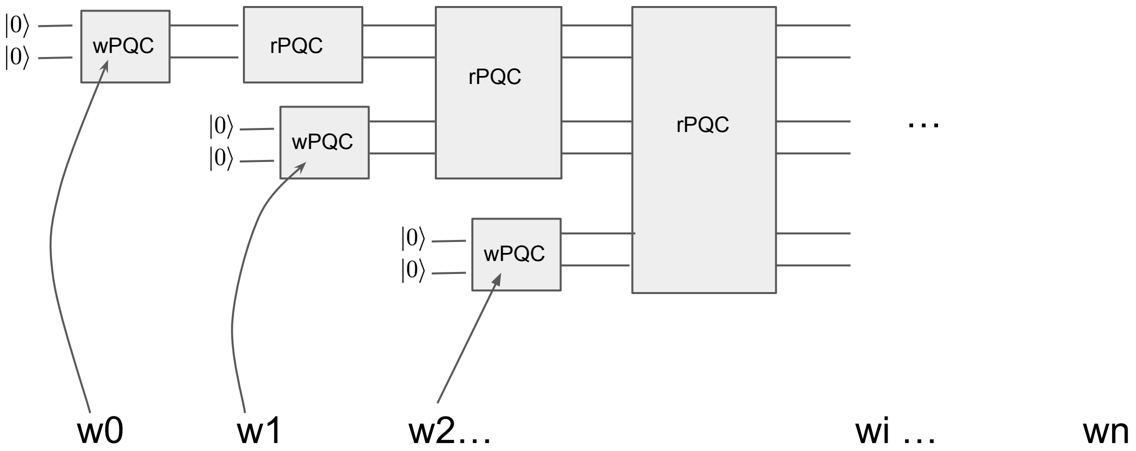



Circuit 7

w0 w1 = … wi … wn

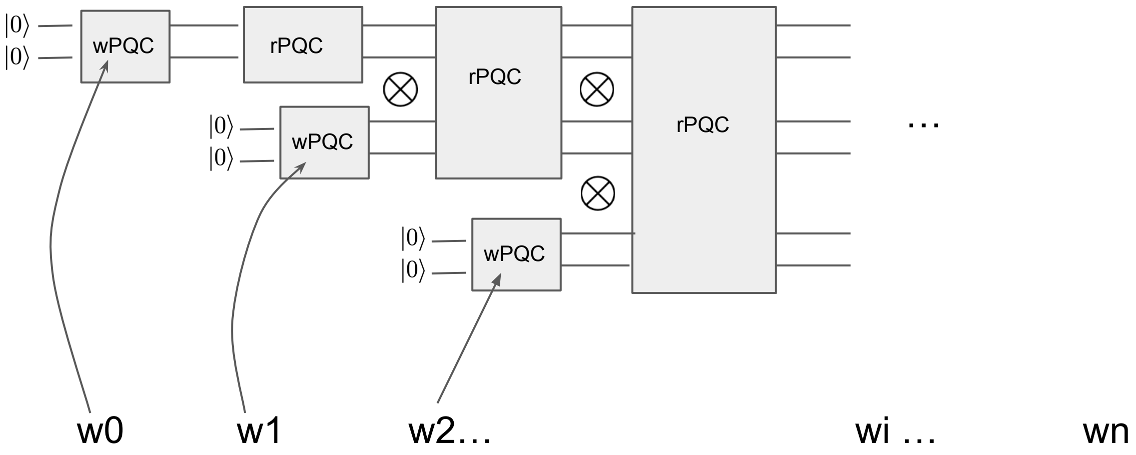
(Rx1, Rx2, …, Rz5, …, CRz9, …)

# qRNN Take One




QUANTINUUM

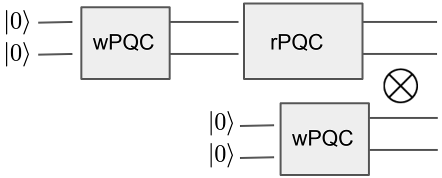
# qRNN Take One




*tensor product*

QUANTINUUM

# Tensor Product

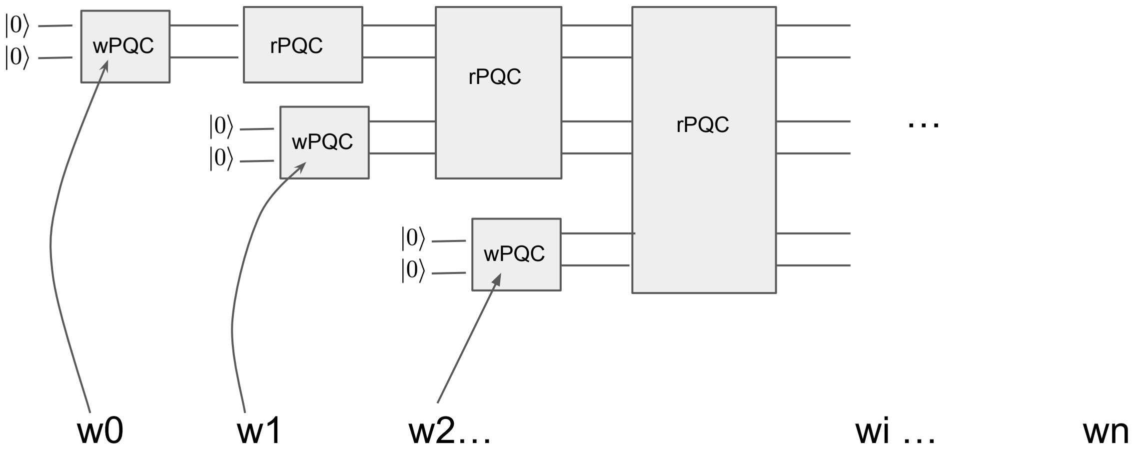



$$(\alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle) \otimes (\beta_{00}|00\rangle + \beta_{01}|01\rangle + \beta_{10}|10\rangle + \beta_{11}|11\rangle)$$

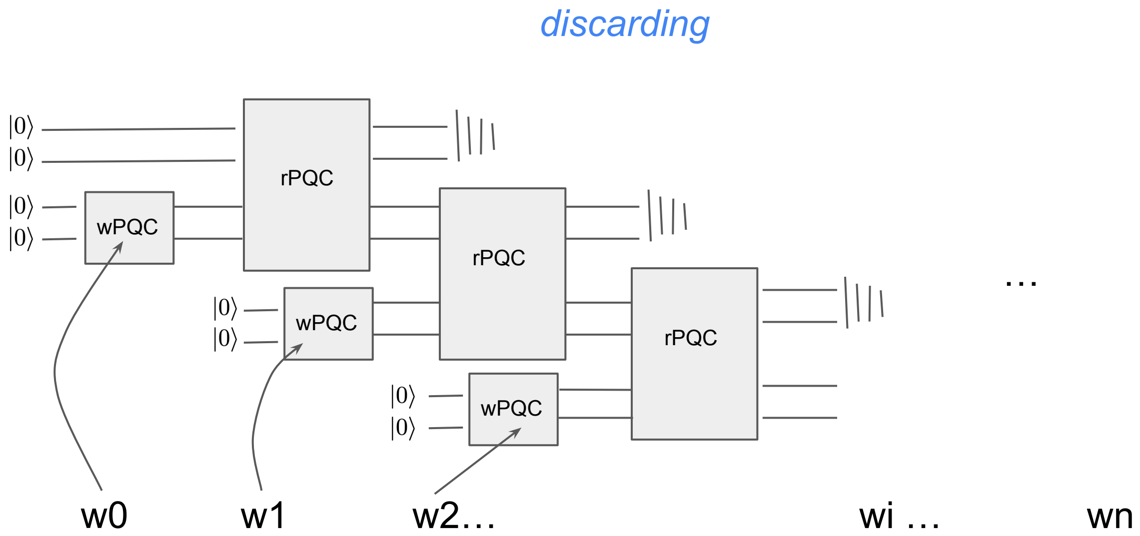
$$= \alpha_{00}\beta_{00}|0000\rangle + \alpha_{00}\beta_{01}|0001\rangle + \ldots \alpha_{01}\beta_{00}|0100\rangle + \alpha_{01}\beta_{01}|0101\rangle \ldots \alpha_{11}\beta_{10}|1110\rangle + \alpha_{11}\beta_{11}|1111\rangle$$

QUANTINUUM

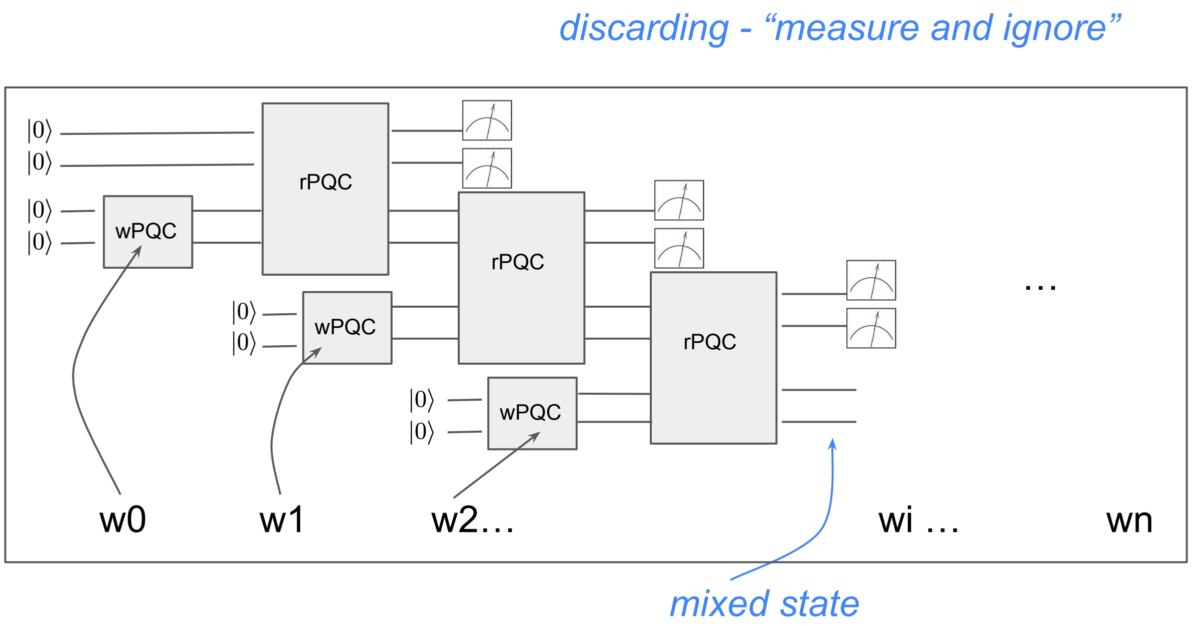
# qRNN Take One



# qRNN Take One

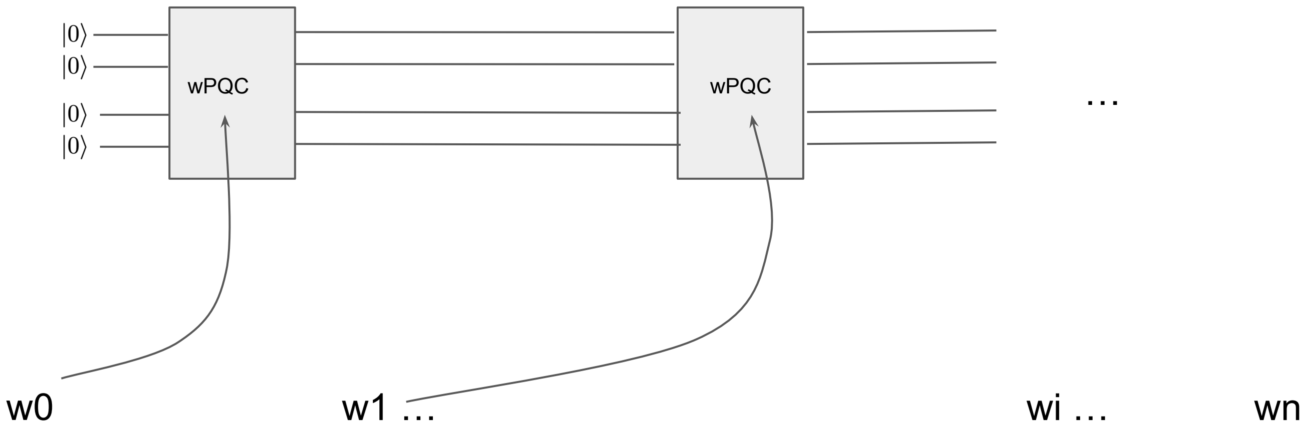


# qRNN Take One

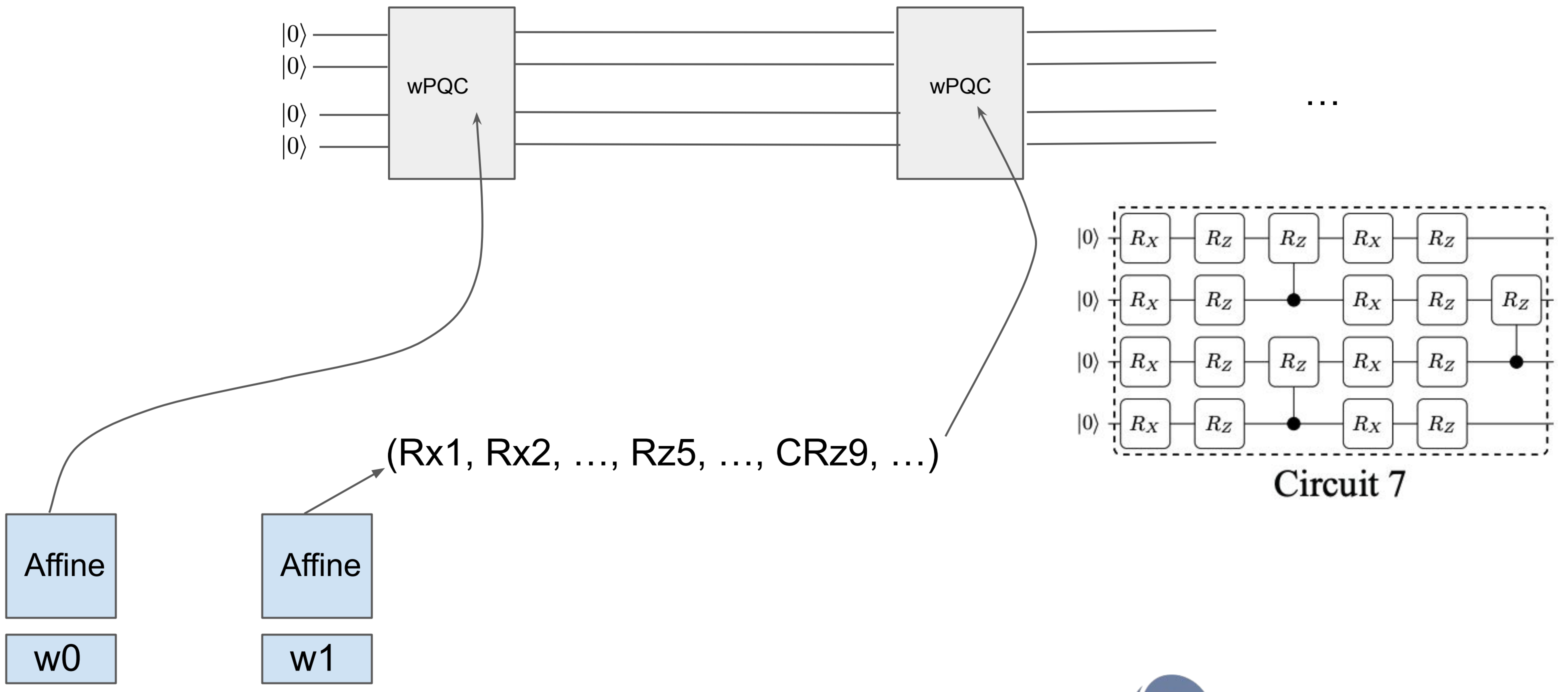



QUANTINUUM

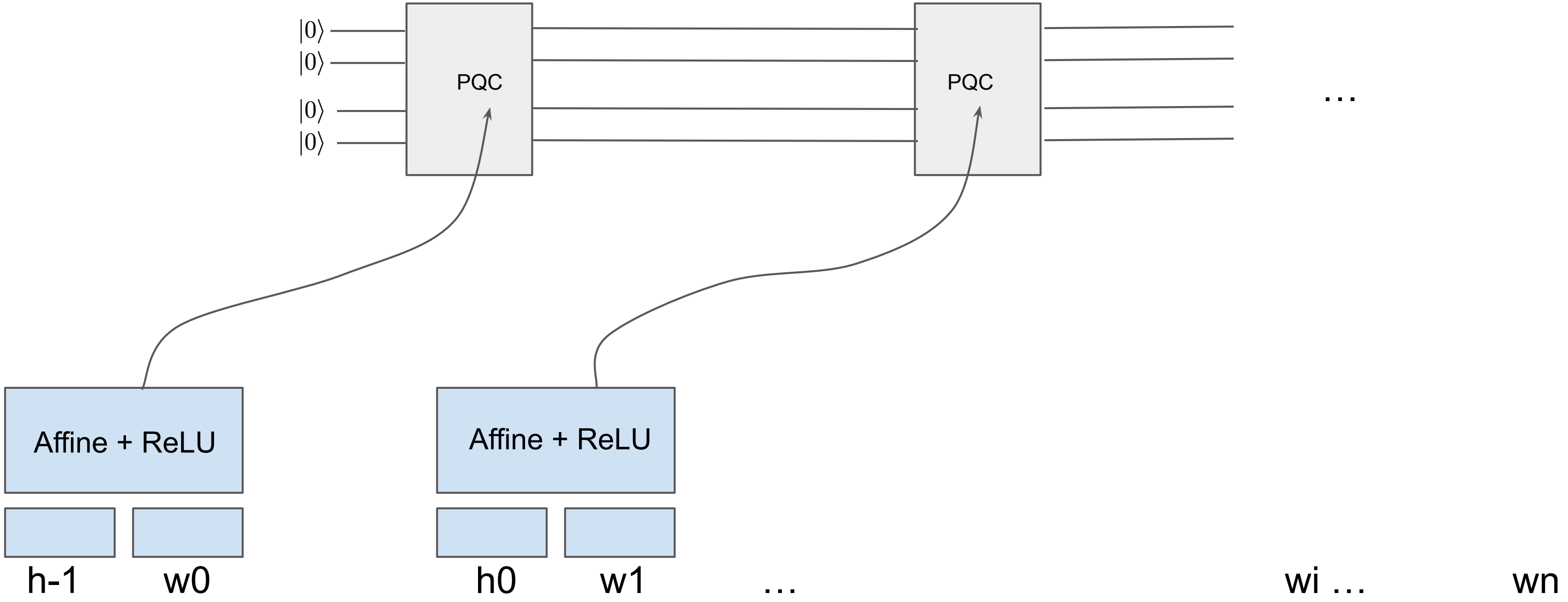
# qRNN Take Two



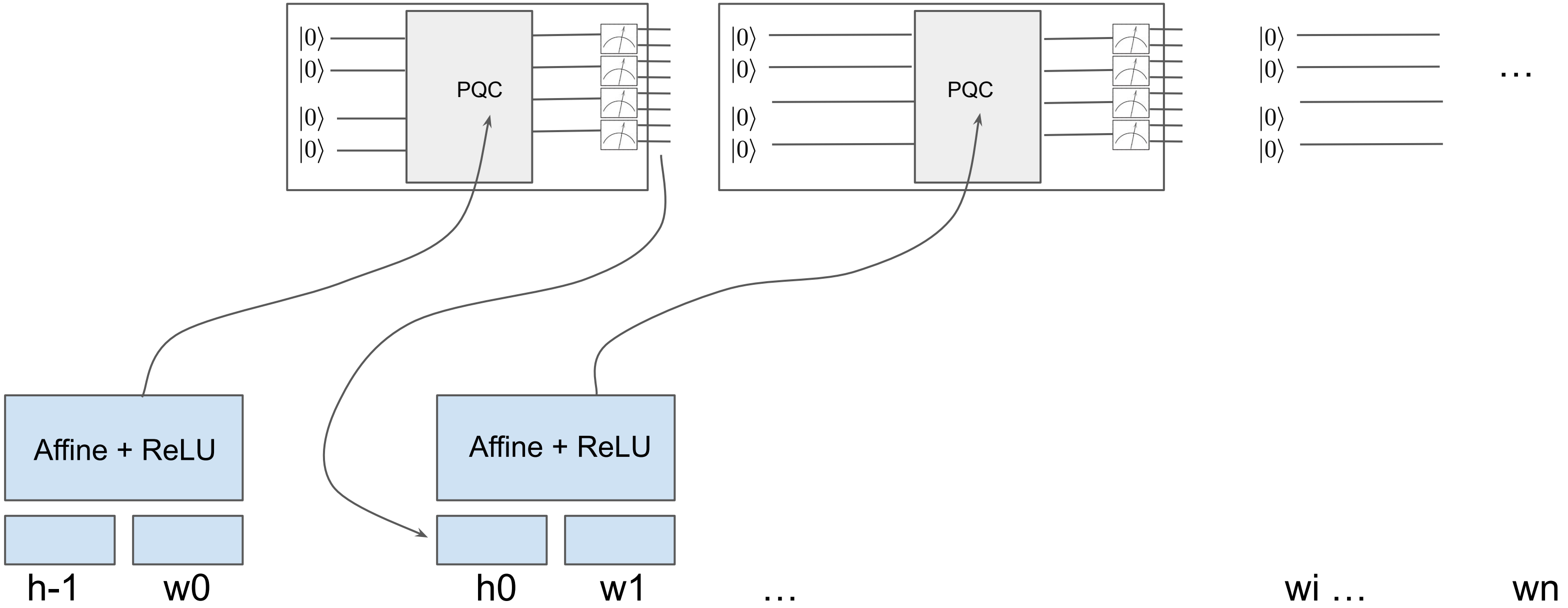
# qRNN Take Two



# qRNN Take Two (Hybrid)

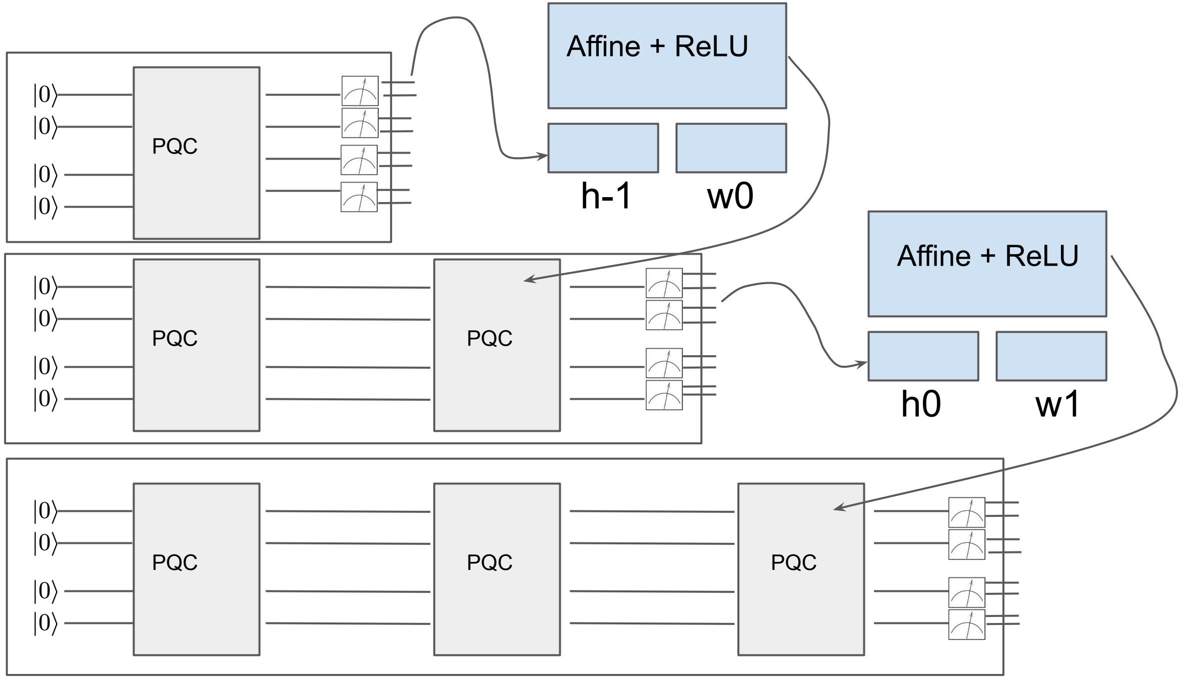


# qRNN Take Two (Hybrid w/Measure)



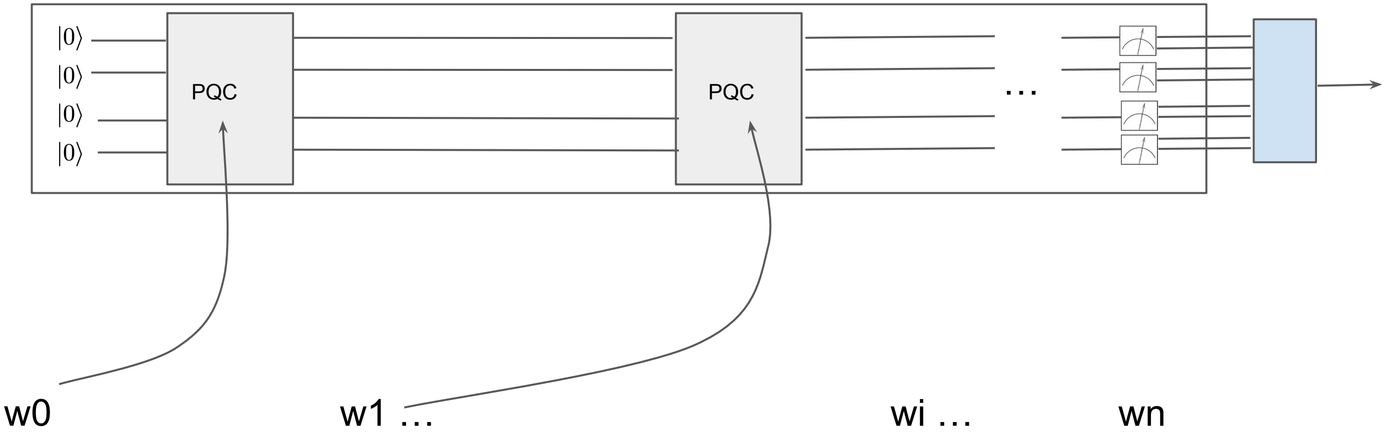

QUANTINUUM

# qRNN Take Two (Coherent Hybrid, “Unrolled”)

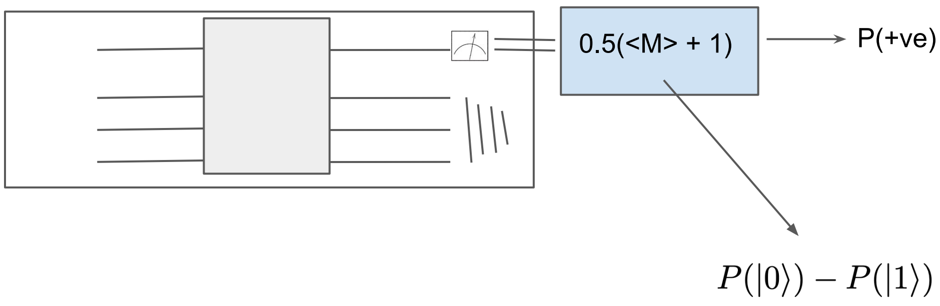



QUANTINUUM

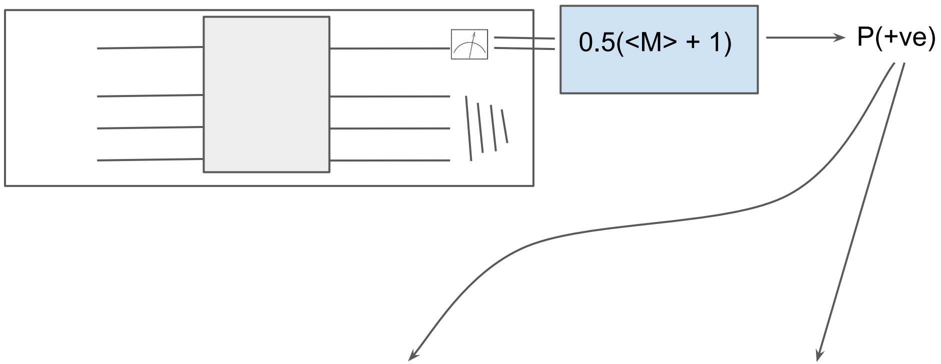
# Output



# Probabilistic Output

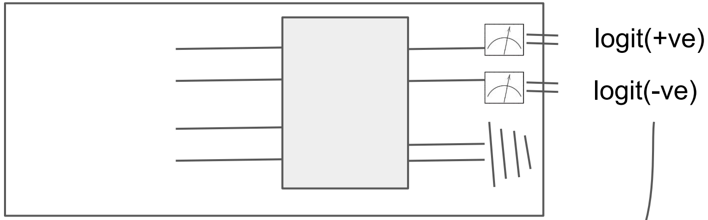


# Probabilistic Output for Training



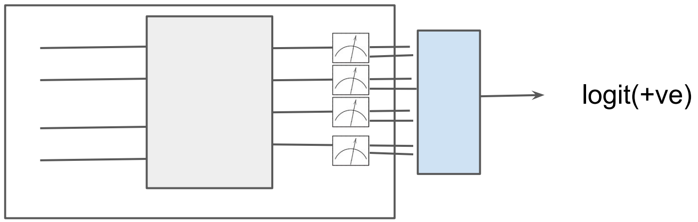

$$-(y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

# Logits Output




$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}} \quad CE = -\sum_i^C t_i log(f(s)_i)$$

# Neural Output



# The Task

- Sentiment analysis (Rotten Tomatoes dataset)
- 8,530 training examples (well balanced); 1,066 dev examples
- Simple binary classification task

```
if you sometimes like to go to the movies to have fun , wasabi is a good place to start .          1
emerges as something rare , an issue movie that's so honest and keenly observed that it doesn't feel like one .          1

simplistic , silly and tedious .          0
it's so laddish and juvenile , only teenage boys could possibly find it funny .          0
```

QUANTINUUM

# Baseline / Goal

- Goal is *not* to beat the s-o-t-a
- Goal (at this stage) is to be competitive with a classical vanilla RNN

# Hybrid Toolkits

PENNYLANE

lambeq

TensorFlow

Quantum

Torch

Quantum

Qiskit

QUANTINUUM

# Hybrid Toolkit

- Requirements for classical simulation:
  - easily interfaces with PyTorch (or TensorFlow, JAX, …)
  - fast to train on real-world datasets
  - accommodates batching
  - **essentially PyTorch ML library with complex number linear algebra**

Torch Quantum

**A PyTorch Library for Quantum Simulation and Quantum Machine Learning**

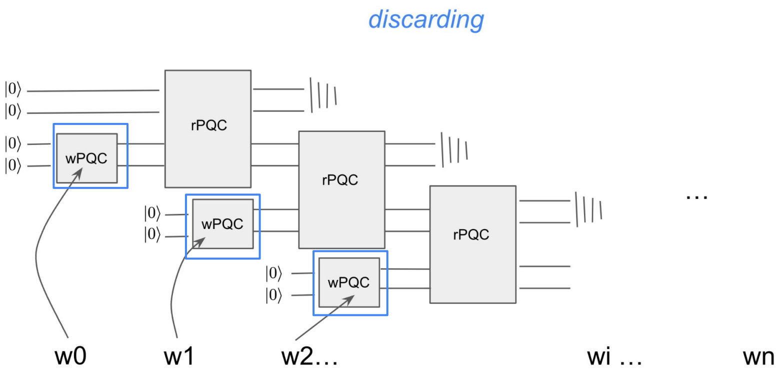
Faster, Scalable, Easy Debugging, Easy Deployment on Real Machine

QUANTINUUM

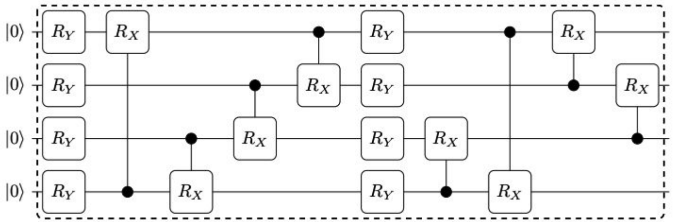
# “Stairs” Architecture in Practice

- We added density matrices to TorchQuantum (for mixed states)
- Choice of PQC:

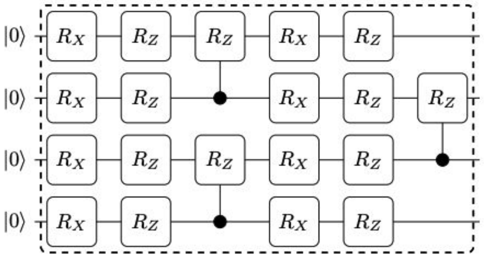
```
'2x4_ryzxy':
    [
        {'input_idx': [0], 'func': 'ry', 'wires': [0]},
        {'input_idx': [1], 'func': 'ry', 'wires': [1]},
        {'input_idx': [2], 'func': 'rz', 'wires': [0]},
        {'input_idx': [3], 'func': 'rz', 'wires': [1]},
        {'input_idx': [4], 'func': 'rx', 'wires': [0]},
        {'input_idx': [5], 'func': 'rx', 'wires': [1]},
        {'input_idx': [6], 'func': 'ry', 'wires': [0]},
        {'input_idx': [7], 'func': 'ry', 'wires': [1]},
    ],
```



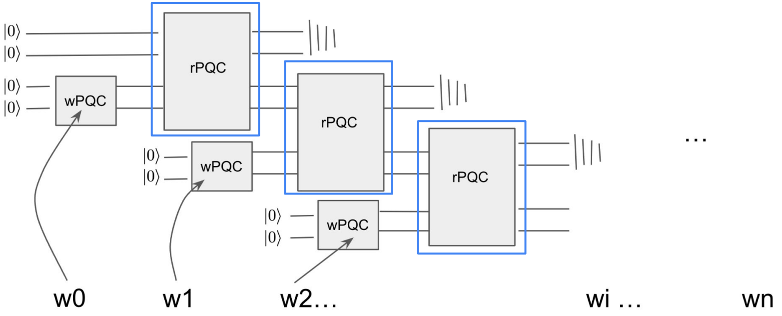
# “Stairs” Architecture in Practice



Circuit 14



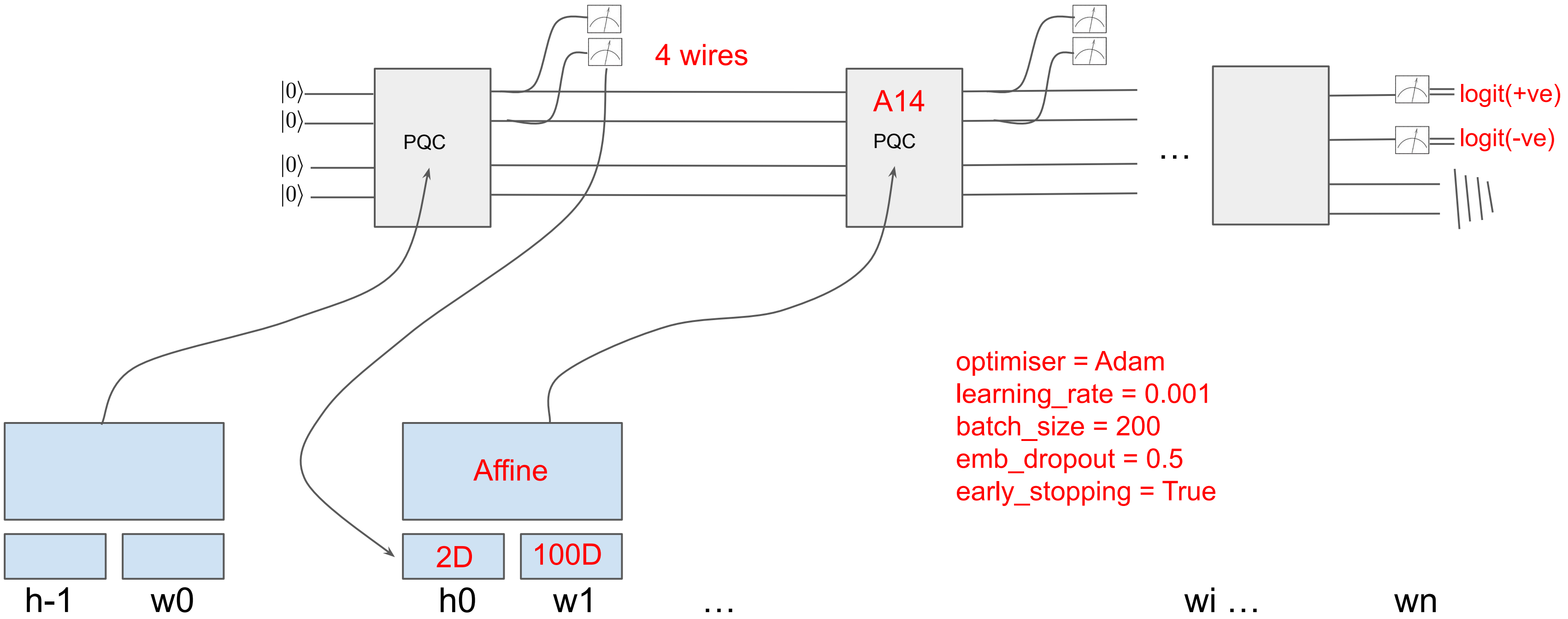
Circuit 7




**Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms**

Sukin Sim,[1,2,*] Peter D. Johnson,[2] and Alán Aspuru-Guzik[2,3,4,5,†]

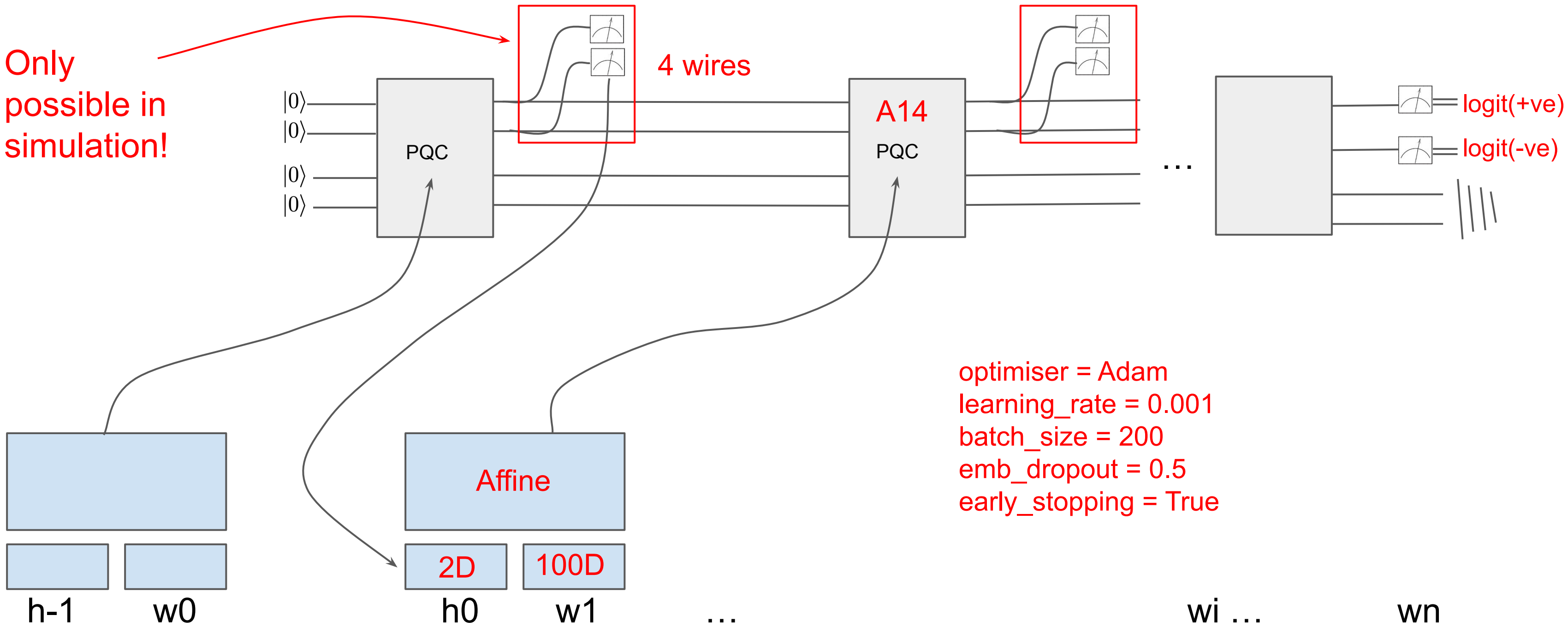
# Experimental Settings




optimiser = Adam
learning_rate = 0.001
batch_size = 200
emb_dropout = 0.5
early_stopping = True

QUANTINUUM

# Experimental Settings




optimiser = Adam
learning_rate = 0.001
batch_size = 200
emb_dropout = 0.5
early_stopping = True

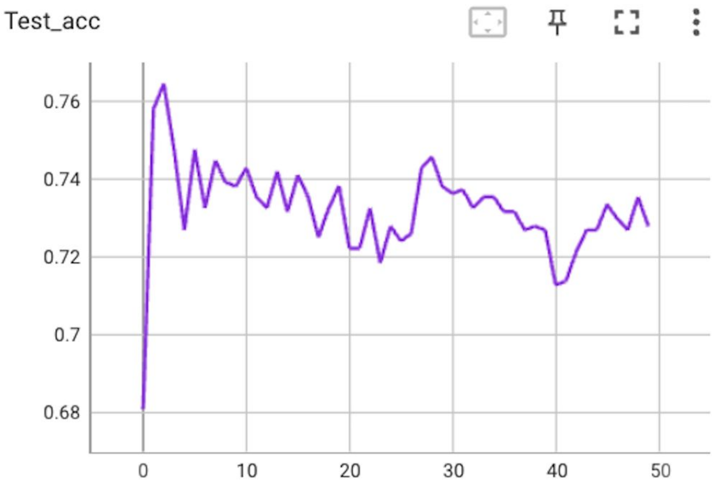
QUANTINUUM

# Results on Dev Set

| 4 wires | Dev acc |
|---|---|
| Classical RNN | 75.6 |
| “Stairs” (w/feedback, measure 2) | 76.5 |
| “Stairs” (wo/feedback, measure 2) | 75.1 |
| “Cups” (w/feedback, measure all + affine) | 76.6 |
| “Cups” (wo/feedback, measure all + affine) | 76.2 |

QUANTINUUM

# Learning Curve




NVidia A30 GPU, PyTorch 1.12:

~5 secs / epoch for 1 wire (pure state)
~11 secs / epoch for 2 wires (pure state)
~14 secs / epoch for 4 wires (pure state)
~26 secs / epoch for 8 wires (pure state)

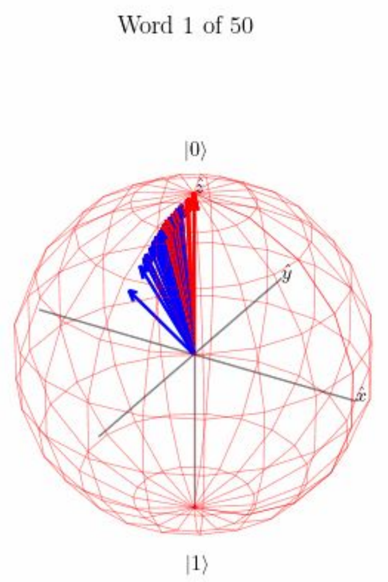
QUANTINUUM

# Results (# Wires)

| # wires | Dev acc<br>“cups” w/A14 + feedback |
|---|---|
| Classical RNN | 75.6 |
| 1 | 72.1 (rxzx) |
| 2 | 75.8 |
| 4 | 76.6 |
| 10 | 77.3 (A7, +ReLU on input) |

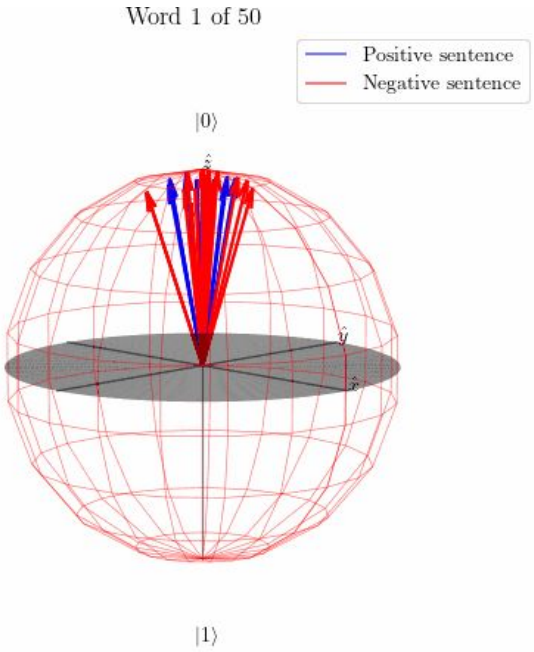
QUANTINUUM

# Visualisation on One Qubit



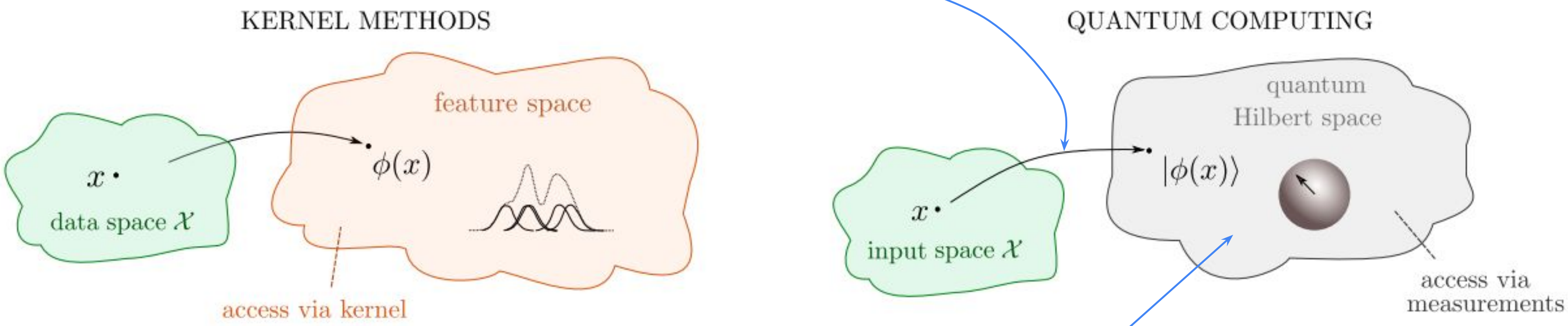

QUANTINUUM

# Visualisation on One Qubit



# Where’s the (Potential) Advantage?

Non-linear mapping




Linear (unitary) transformations happen in this (potentially very large) space

**Supervised quantum machine learning models are kernel methods**

Maria Schuld
*Xanadu, Toronto, ON, M5G 2C8, Canada*

# So What's the Current State of Quantum Hardware?


June 30, 2023

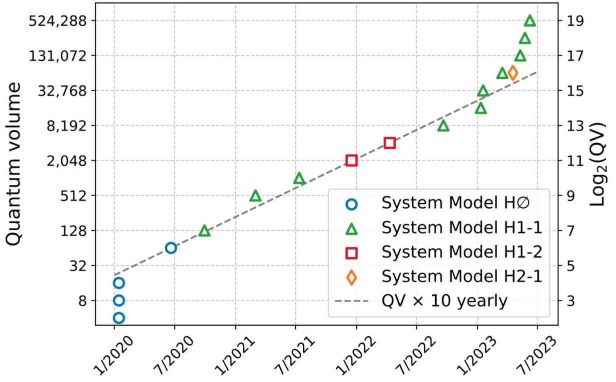


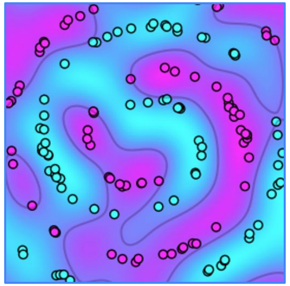

*Figure 1: H-series progress quantum volume improvement trajectory*

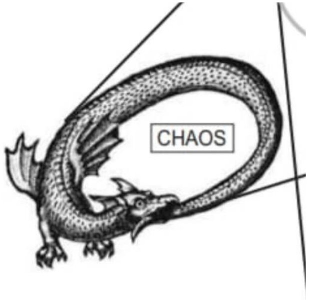
QUANTINUUM

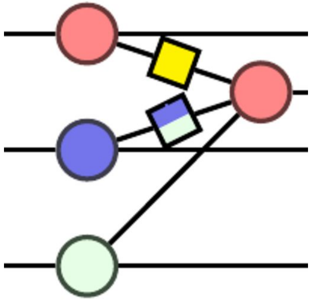
# Future Work

- Apply the models to more tasks
  - sequence labelling, language modelling, translation, …
- Apply pre-training / fine-tuning paradigm
- Develop more hybrid architectures
  - based on CNNs (e.g. MERA-like), transformers, …
- Run on quantum hardware

# The Oxford Hybrid NLP Team





Wenduan Xu, Konstantinos Meichanetzidis, Douglas Brown, Gabriel Matos,
Charlie London, Richie Yeung, Carys Harvey, Nikhil Khatri, Stephen Clark

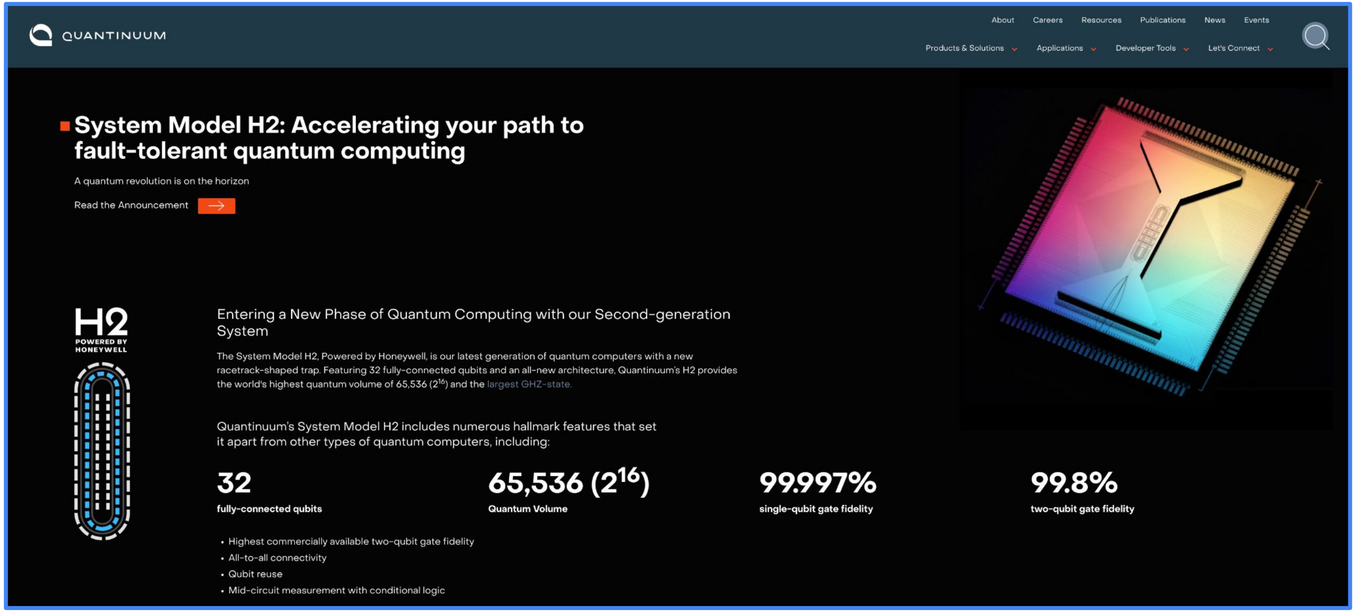
QUANTINUUM

# Readings

- Quantum Computing
  - *Quantum Computation and Quantum Information,* Nielsen and Chuang, 2010
  - *Quantum Computing since Democritus,* Aaronson, 2013
- Quantum Mechanics
  - *Quantum Mechanics: The Theoretical Minimum,* Susskind and Friedman, 2015
- Quantum ML
  - *Machine Learning with Quantum Computers,* Schuld and Petruccione, 2021
- Quantum NLP
  - *Today's research* - keep an eye out on arXiv and for blog post!
  - *QNLP in Practice,* Lorenz et al., JAIR, 2023
- Diagrammatic Reasoning
  - *Picturing Quantum Processes,* Coecke and Kissinger, 2017

QUANTINUUM

# The Future is (Almost) Here




QUANTINUUM