Quantum Recurrent Architectures for Natural Language Processing

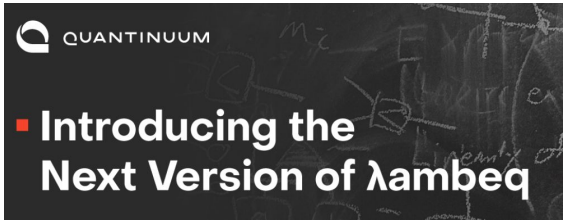
Stephen Clark

CLASP, University of Gothenburg

6 September 2023

QUANTINUUM

Quantum Computing



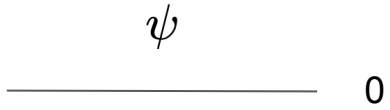
UK-based quantum computing firm Quantinuum claims sub-atomic matter breakthrough | Science & Tech News | Sky...

QUANTINUUM

Talk Outline

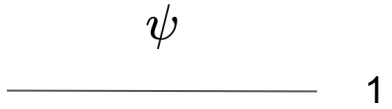
1. Introduction to quantum computing / quantum circuits
2. Application to sequence classification
    a. our quantum RNN architectures
    b. sentiment analysis experiments

The State of a Classical Bit



QUANTINUUM

The State of a Classical Bit



QUANTINUUM

The State of a Qubit

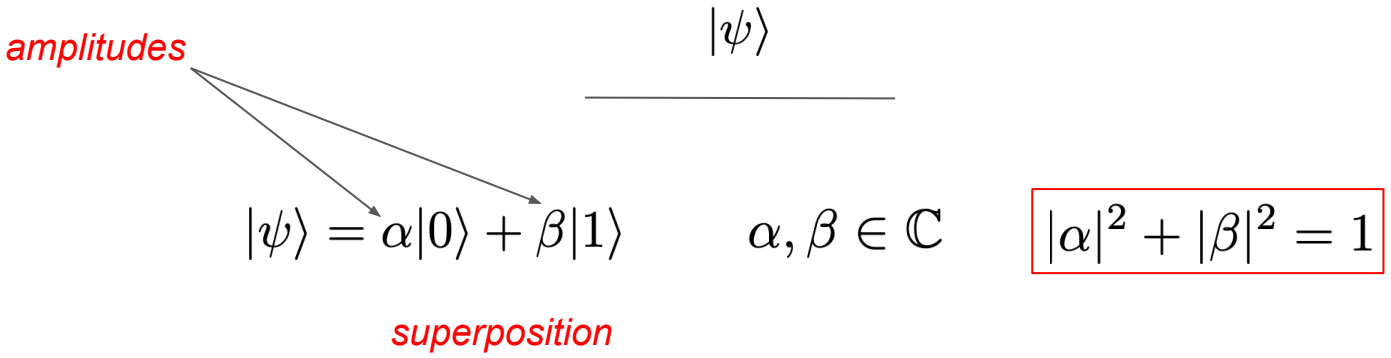
$$|\psi\rangle$$

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

superposition

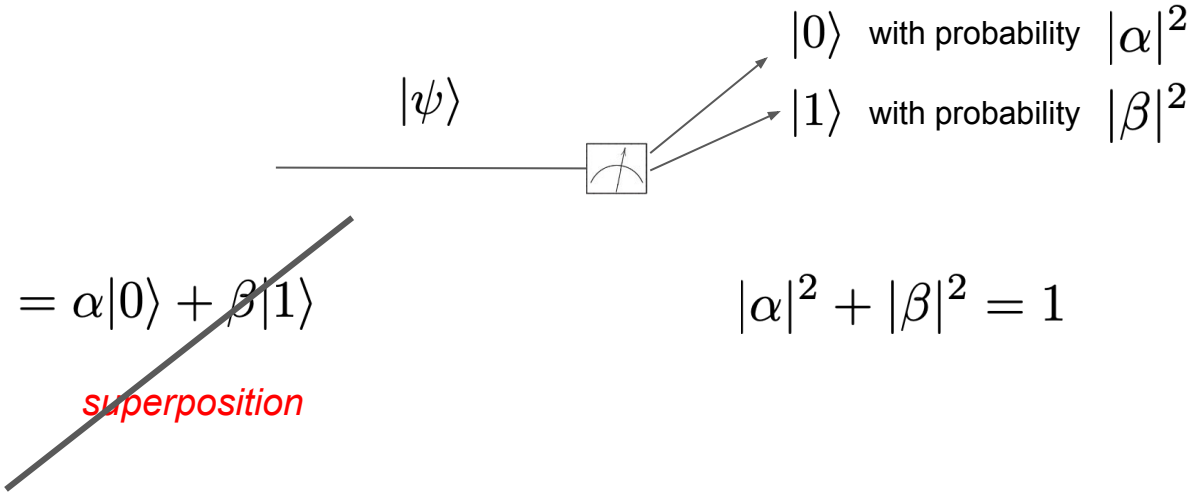
QUANTINUUM

The State of a Qubit



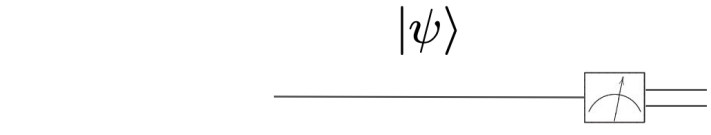
QUANTINUUM

Measuring a Qubit



QUANTINUUM

Measuring a Qubit (scalar output)



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

superposition

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

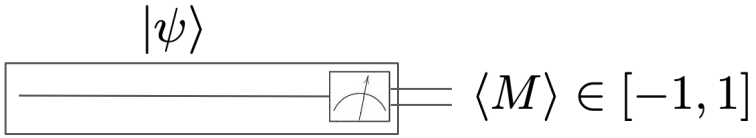
1 with probability $|\alpha|^2$

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

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

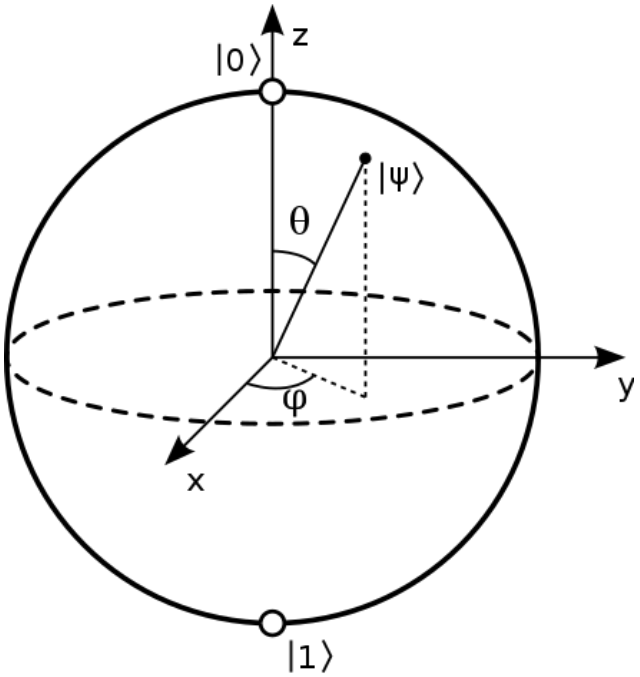
QUANTINUUM

Measuring a Qubit (many times)



QUANTINUUM

The Bloch Sphere Representation of a Qubit

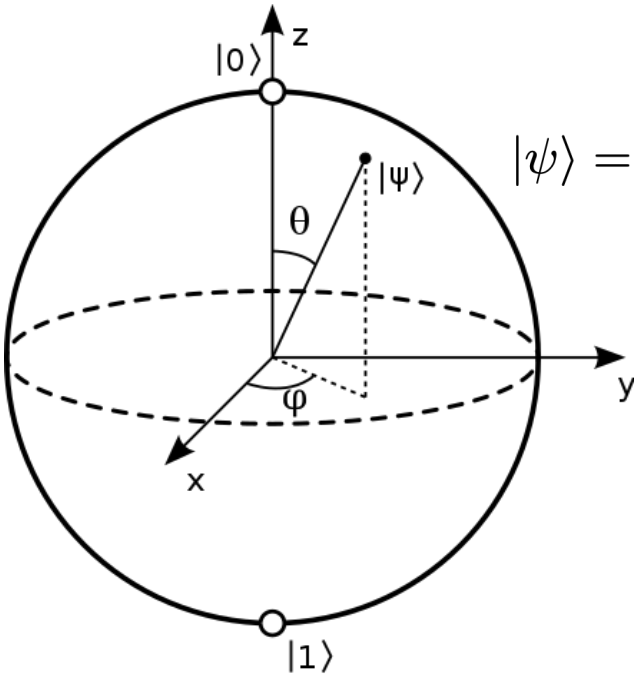


$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

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

QUANTINUUM

The Bloch Sphere Representation of a Qubit

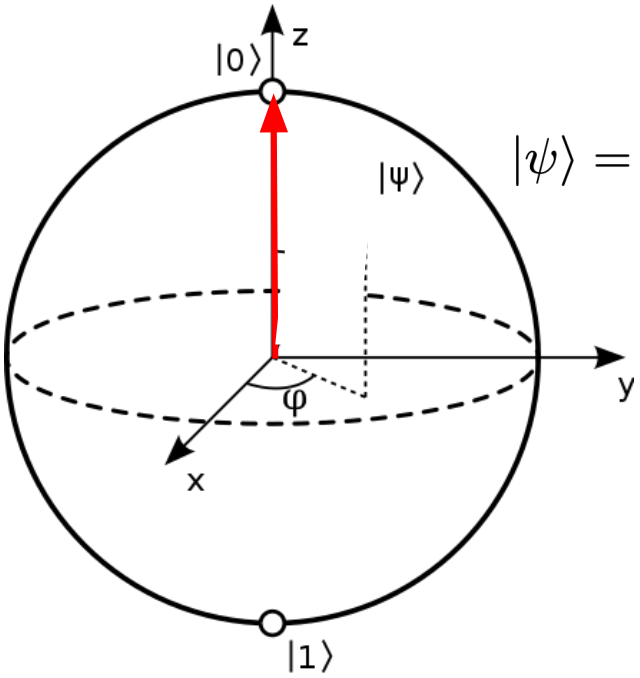


$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

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

QUANTINUUM

“The Collapse of the Wave Function”



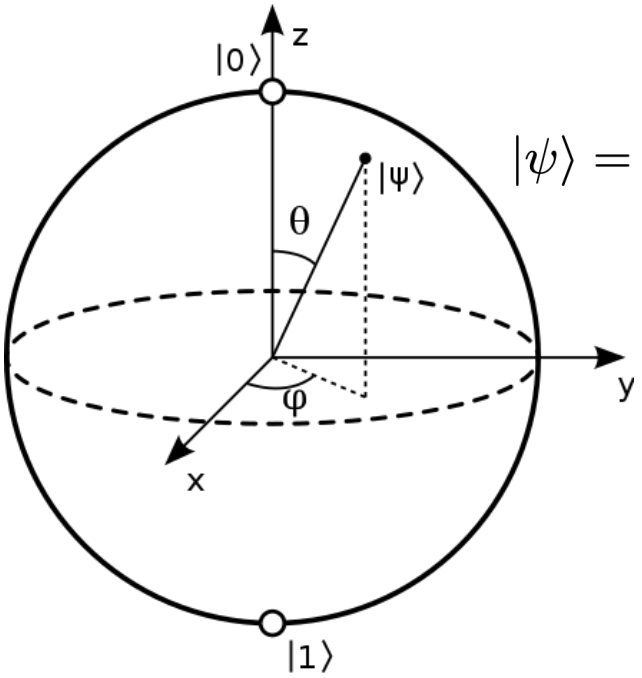
$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

$$\cos^2(\theta/2)$$

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

QUANTINUUM

The Bloch Sphere Representation of a Qubit

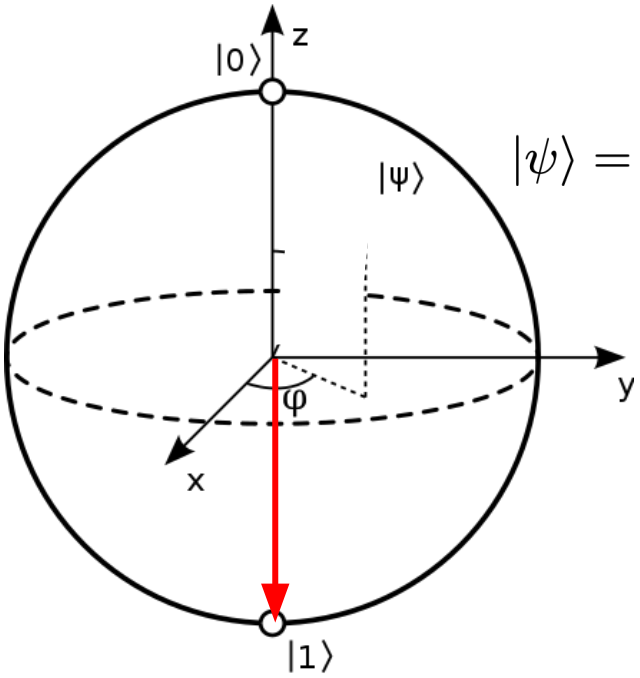


$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

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

QUANTINUUM

“The Collapse of the Wave Function”



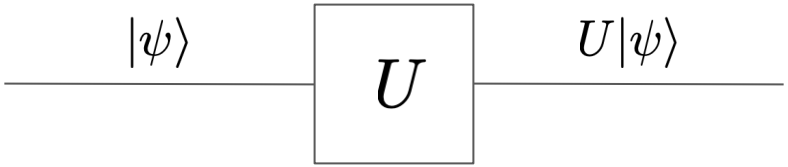
$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

$$\sin^2(\theta/2)$$

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

QUANTINUUM

Unitary Transformations of a Qubit



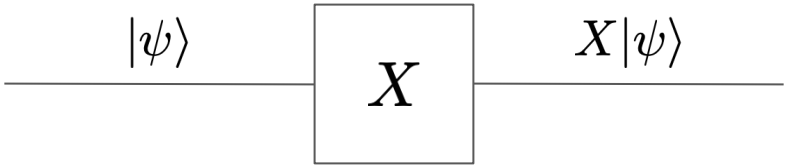
$$U : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha'|0\rangle + \beta'|1\rangle$$

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

QUANTINUUM

1-Qubit Quantum Gates

quantum Not gate

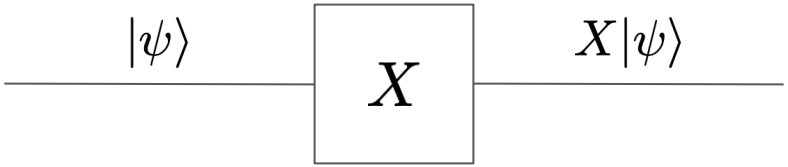


$$X : |0\rangle \mapsto |1\rangle$$
$$X : |1\rangle \mapsto |0\rangle$$

QUANTINUUM

1-Qubit Quantum Gates

quantum Not gate *acts linearly*

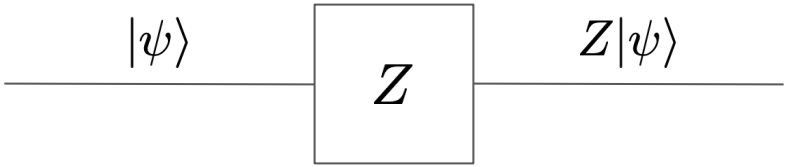


$$X : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|1\rangle + \beta|0\rangle$$

QUANTINUUM

1-Qubit Quantum Gates

Pauli Z Gate

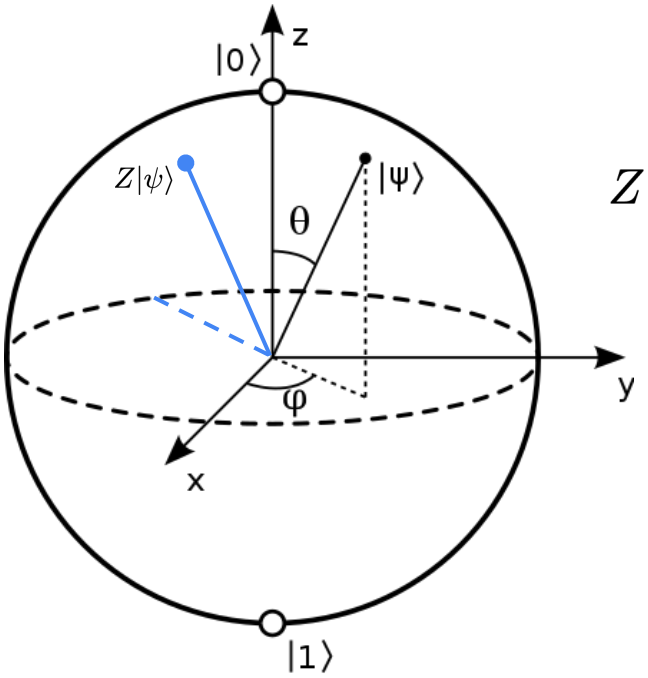


$$Z : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle - \beta|1\rangle$$

QUANTINUUM

1-Qubit Quantum Gates

Pauli Z Gate *rotates about the Z axis*

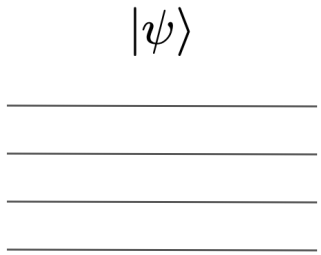


$$Z : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle - \beta|1\rangle$$

QUANTINUUM

The State of Many Qubits

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

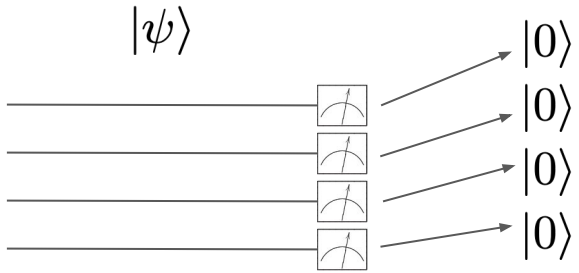


$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

QUANTINUUM

Measuring Many Qubits

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



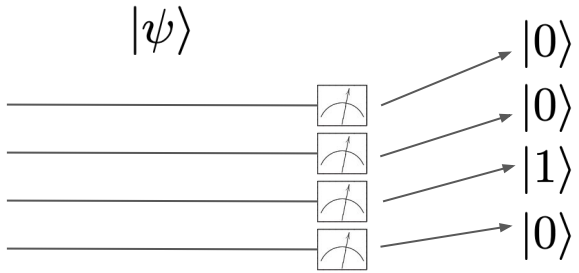
$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

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

QUANTINUUM

Measuring Many Qubits

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



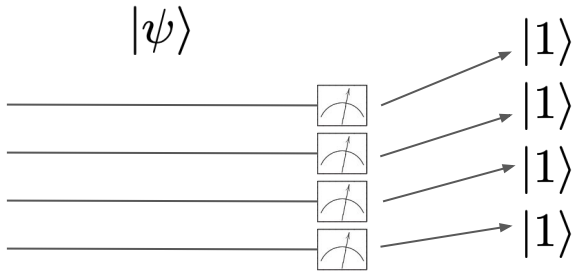
$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

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

QUANTINUUM

Measuring Many Qubits

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



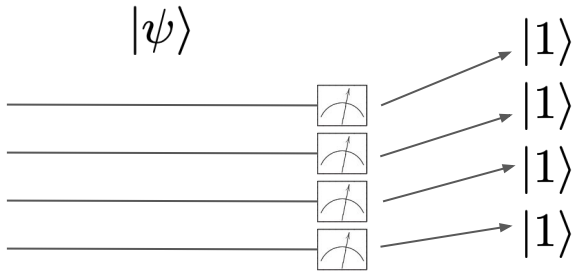
$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

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

QUANTINUUM

Measuring Many Qubits

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



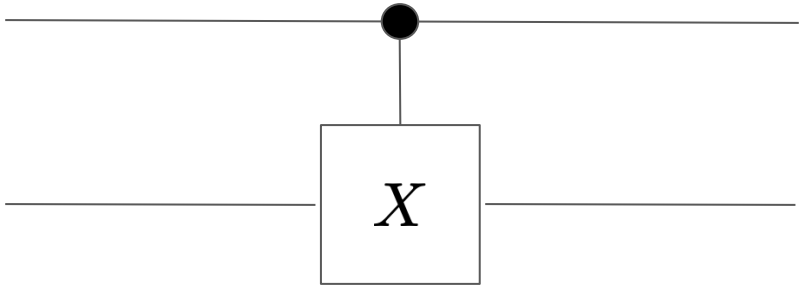
$$|\psi\rangle = \alpha_{0000}|0000\rangle + \alpha_{0001}|0001\rangle + \alpha_{0010}|0010\rangle + \ldots \alpha_{1111}|1111\rangle$$

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

QUANTINUUM

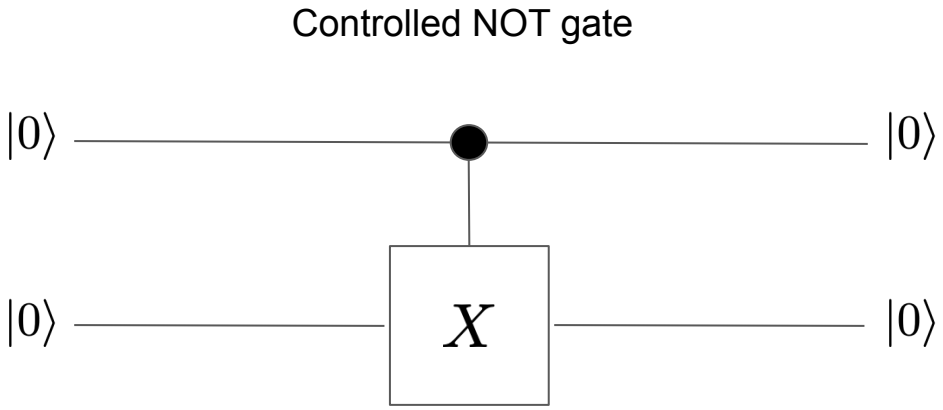
Entangling Qubits

Controlled NOT gate



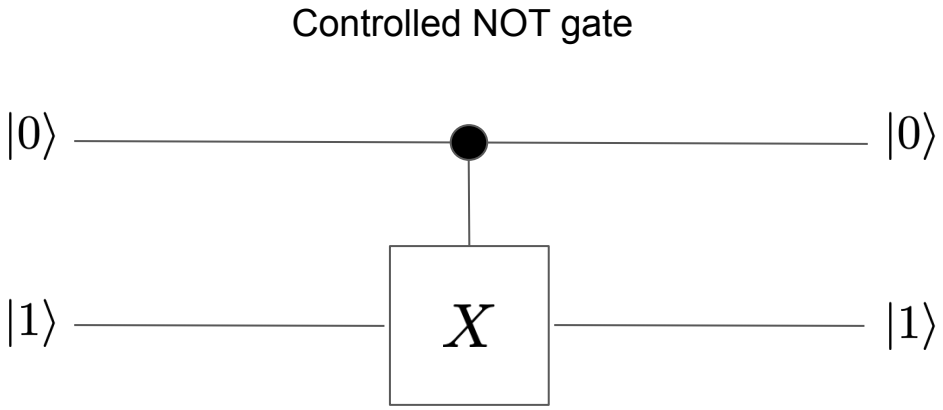
QUANTINUUM

Entangling Qubits



QUANTINUUM

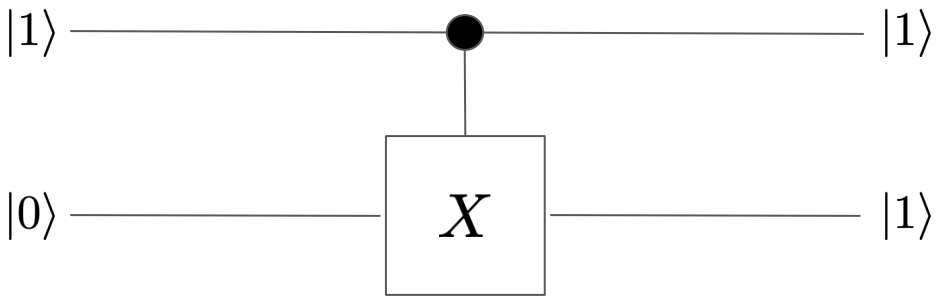
Entangling Qubits



QUANTINUUM

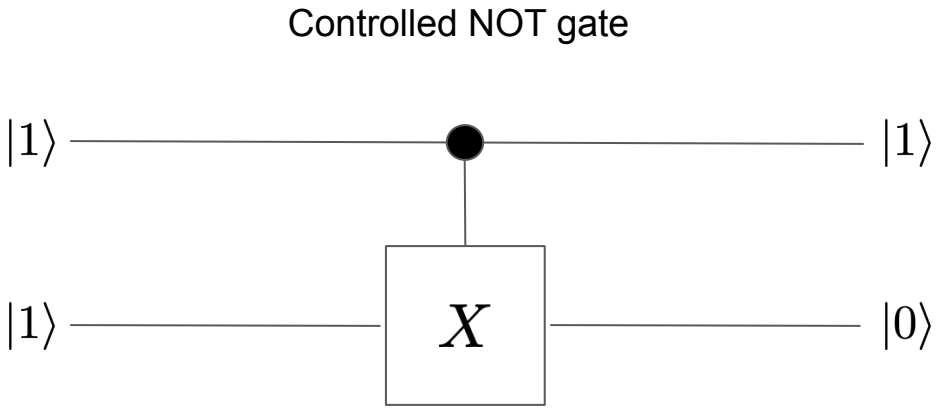
Entangling Qubits

Controlled NOT gate



QUANTINUUM

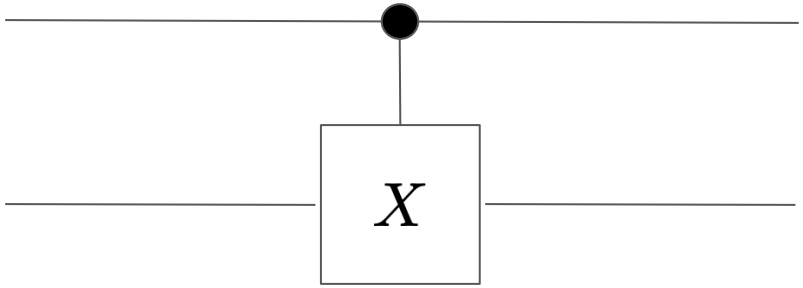
Entangling Qubits



QUANTINUUM

Entangling Qubits

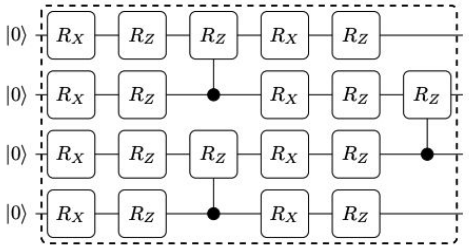
Controlled NOT gate *acts linearly*



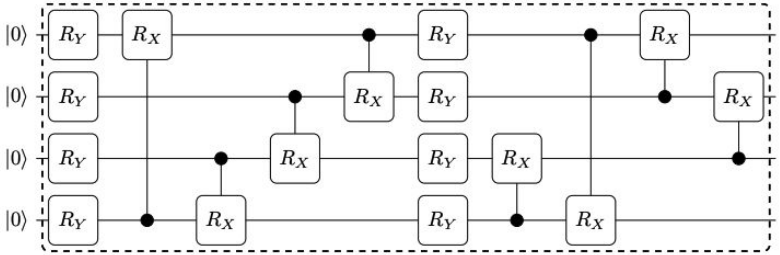
$$CX : \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle \mapsto \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|11\rangle + \alpha_{11}|10\rangle$$

QUANTINUUM

Quantum Circuits



Circuit 7



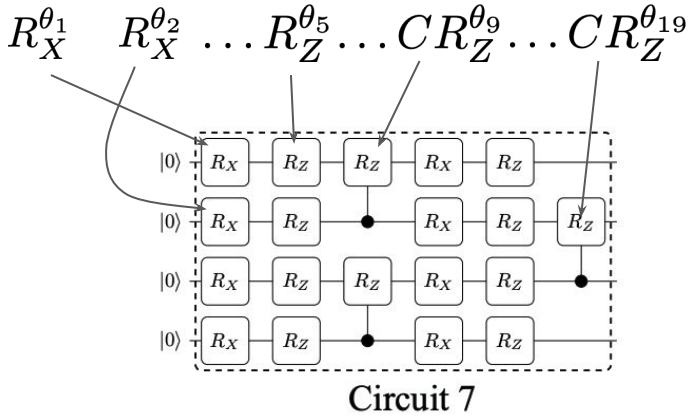
Circuit 14

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

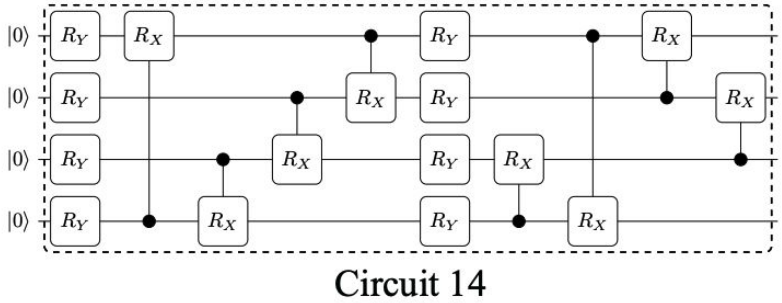
Sukin Sim,[1,2,*] Peter D. Johnson,[2] and Alán Aspuru-Guzik[2,3,4,5,†]

QUANTINUUM

Parameterised Quantum Circuits (PQCs)



Circuit 7



Circuit 14

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

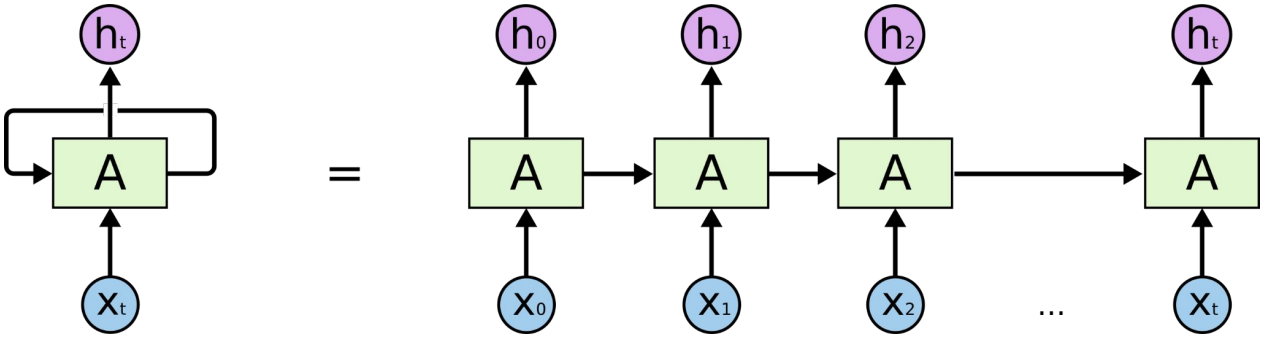
Sukin Sim,[1,2,*] Peter D. Johnson,[2] and Alán Aspuru-Guzik[2,3,4,5,†]

QUANTINUUM

Talk Outline

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

Recurrent Neural Networks (RNNs)

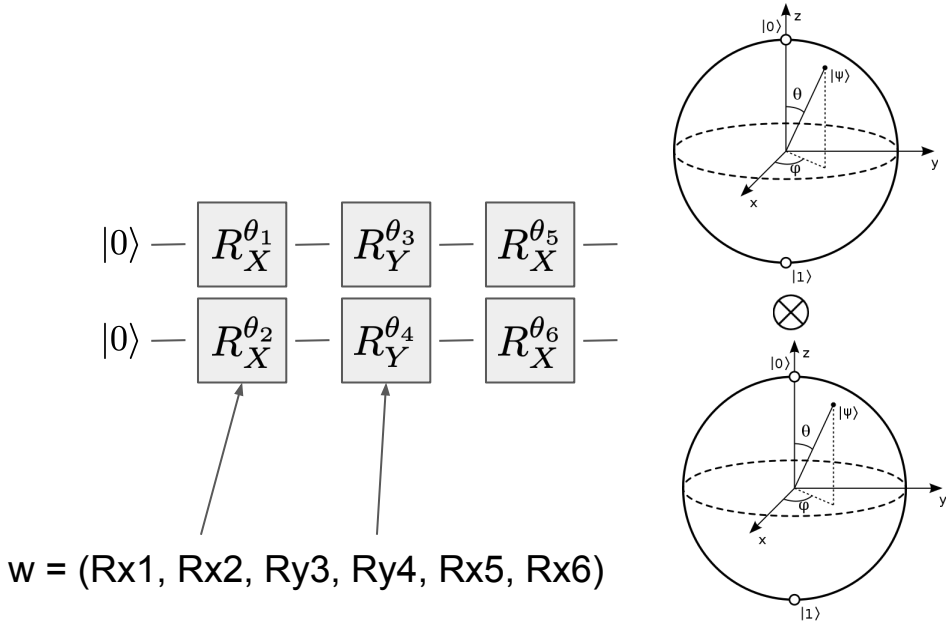


$$h_t = f(x_t \mathbf{U} + h_{t-1} \mathbf{W})$$

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

QUANTINUUM

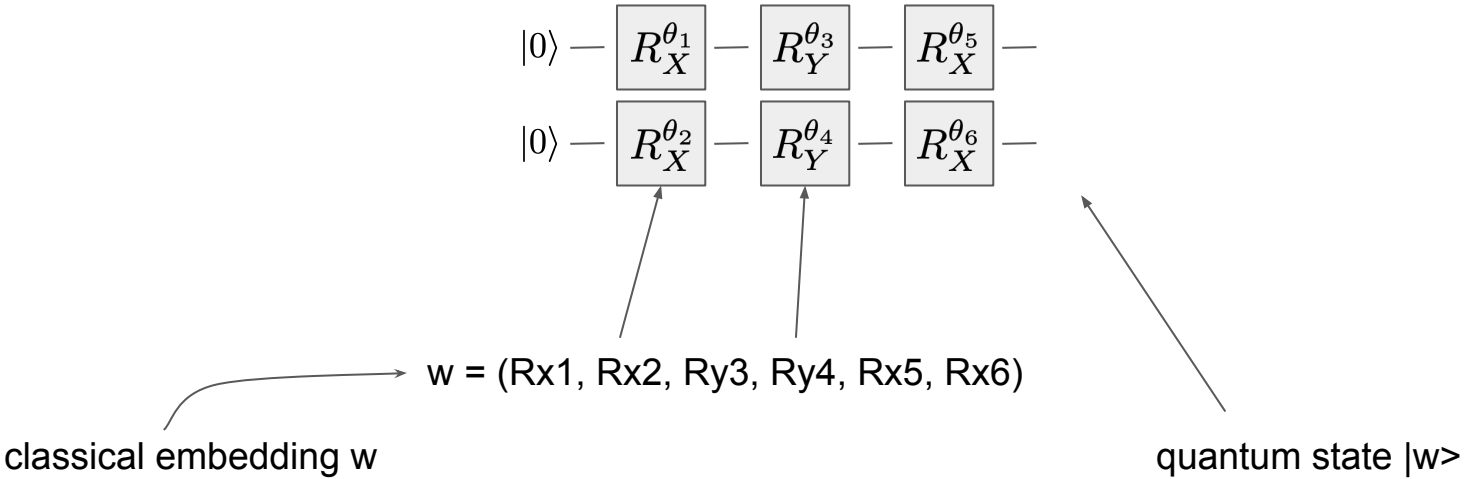
Angle Encoding for Words



Angle encoding

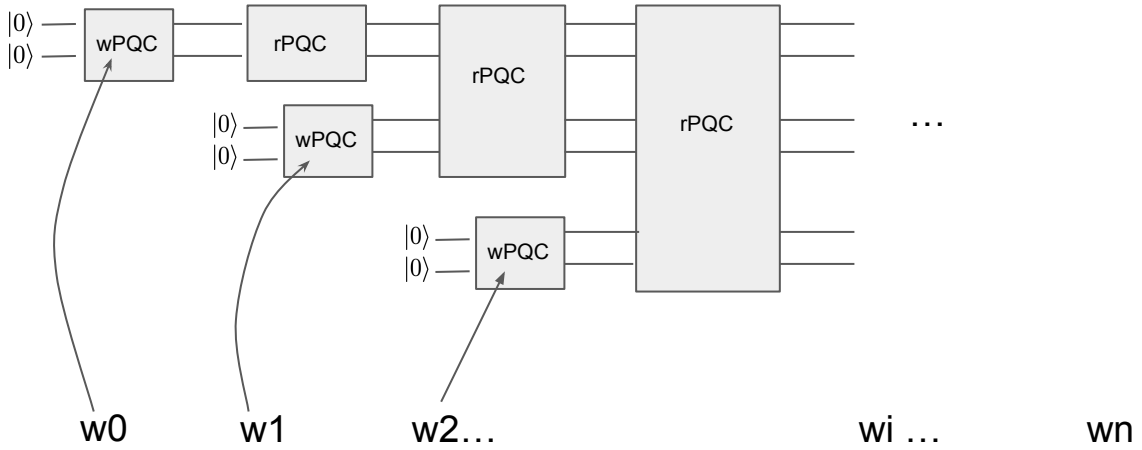
QUANTINUUM

Angle Encoding for Words



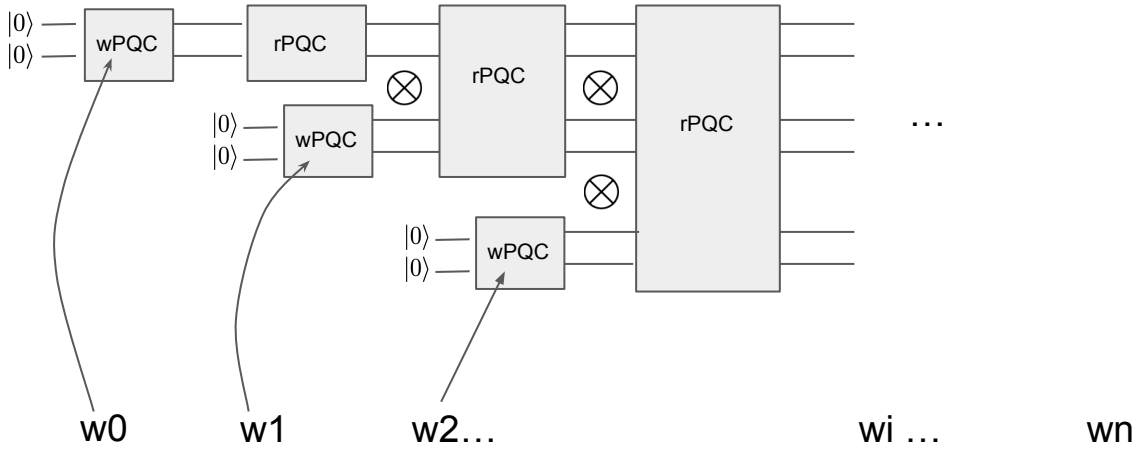
QUANTINUUM

qRNN Take One



QUANTINUUM

qRNN Take One

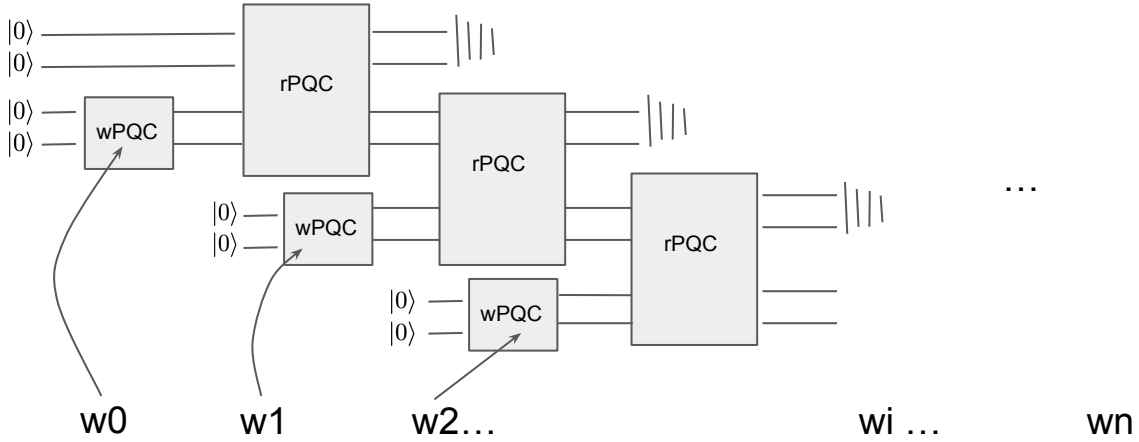


tensor product

QUANTINUUM

qRNN Take One

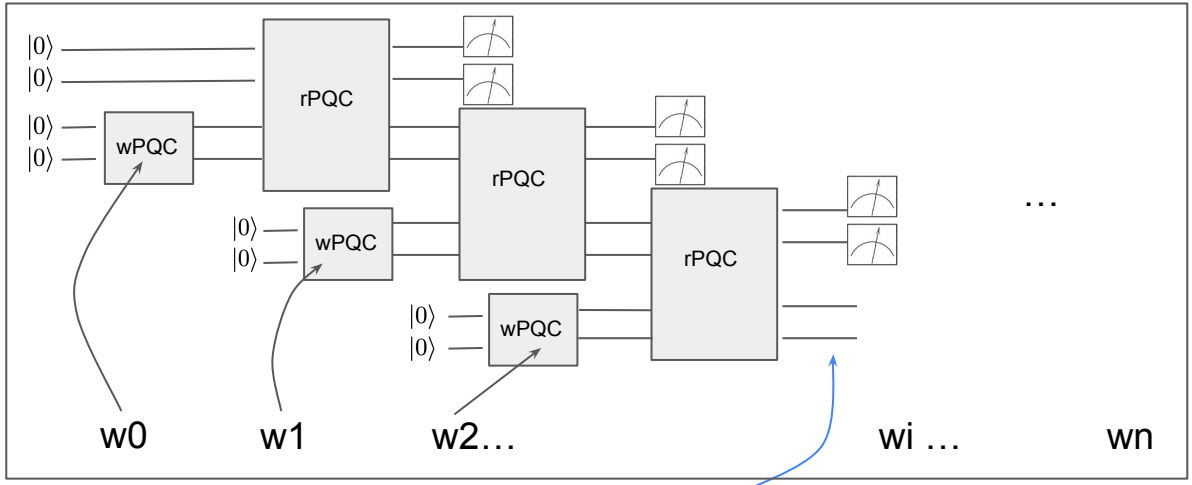
discarding



QUANTINUUM

qRNN Take One

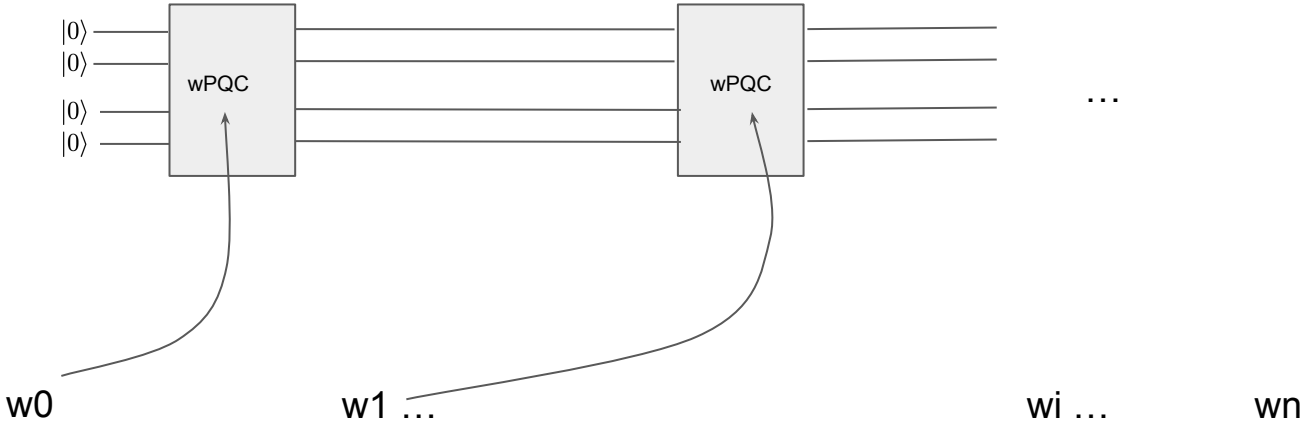
discarding - “measure and ignore”



mixed state

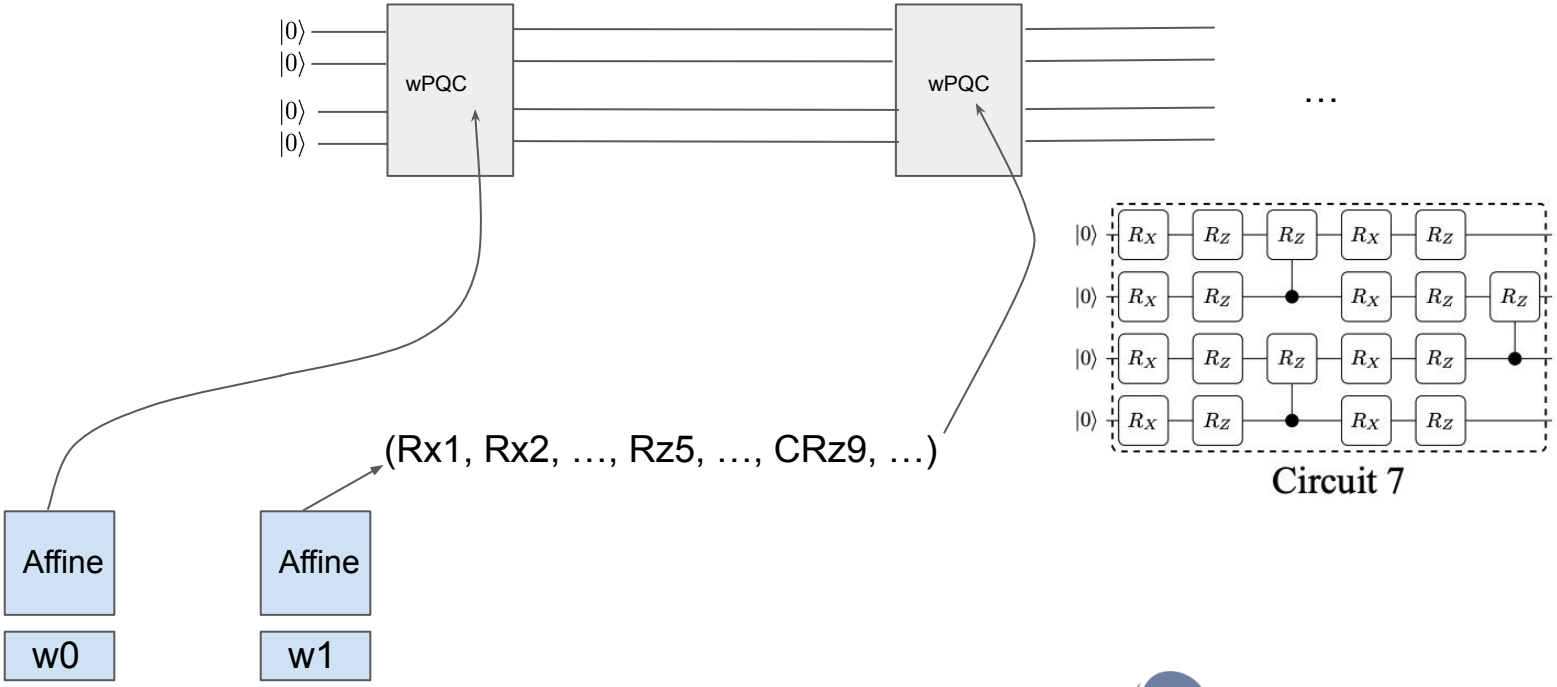
QUANTINUUM

qRNN Take Two



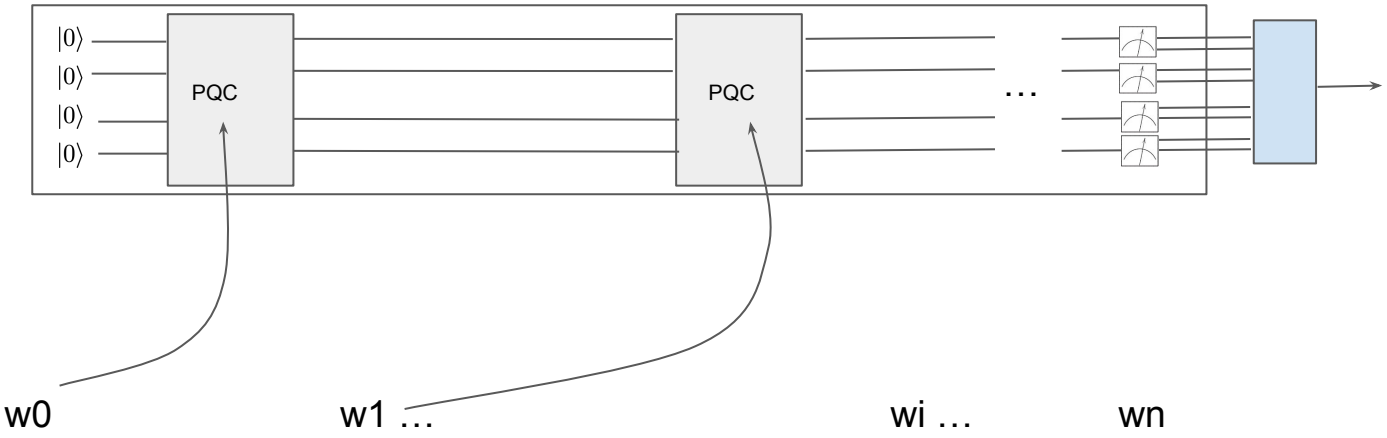
QUANTINUUM

qRNN Take Two



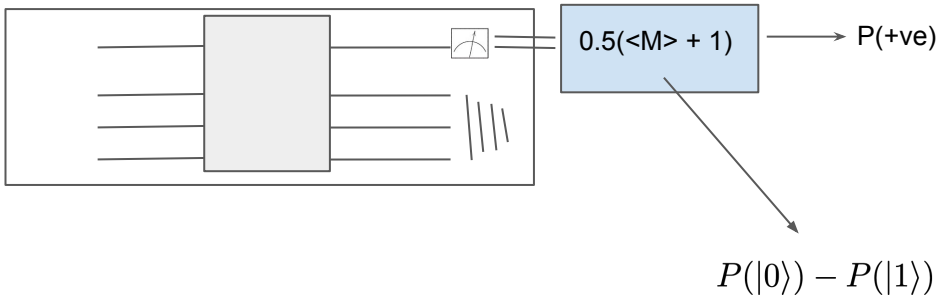
QUANTINUUM

Output



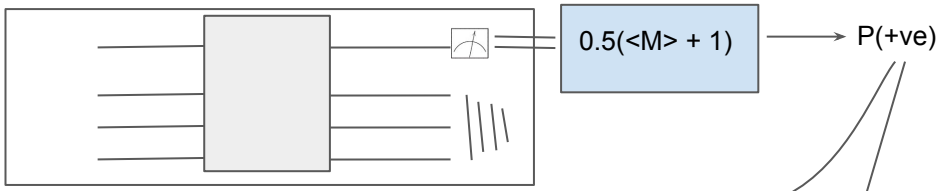
QUANTINUUM

Probabilistic Output



QUANTINUUM

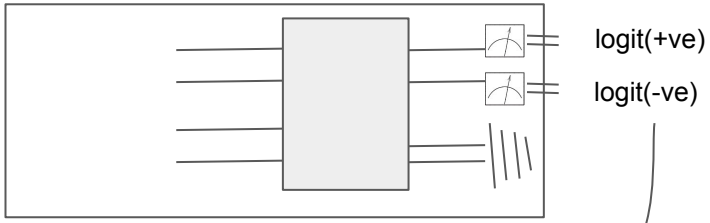
Probabilistic Output for Training



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

QUANTINUUM

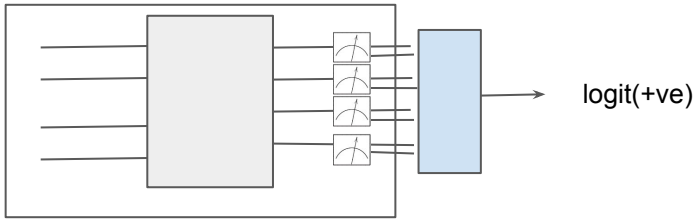
Logits Output



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

QUANTINUUM

Neural Output



QUANTINUUM

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
```

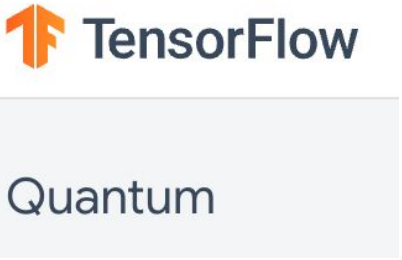
QUANTINUUM

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

lambeq



Hybrid Toolkit

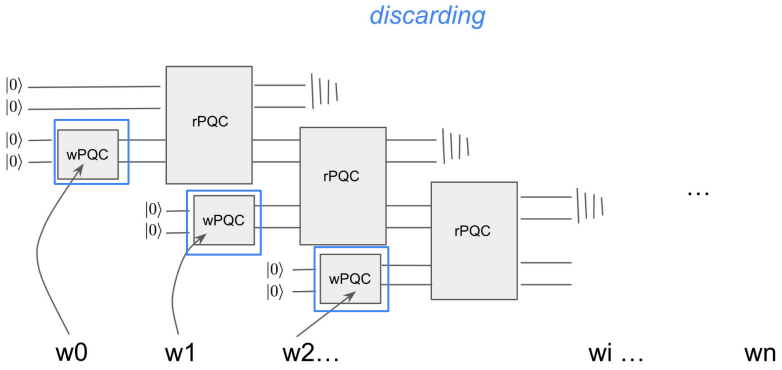
- Requirements for classical simulation:
  - easily interfaces with PyTorch (or TensorFlow, JAX, …)
  - fast to train on real-world datasets
  - accommodates batching
  - **essentially PyTorch ML library with complex number linear algebra**

QUANTINUUM

“Stairs” Architecture in Practice

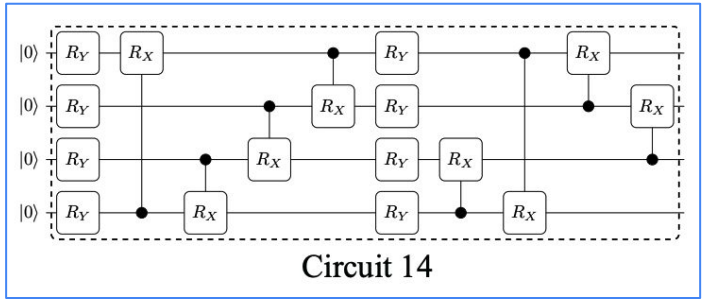
- We added density matrices to TorchQuantum (for mixed states)
- Choice of PQC:

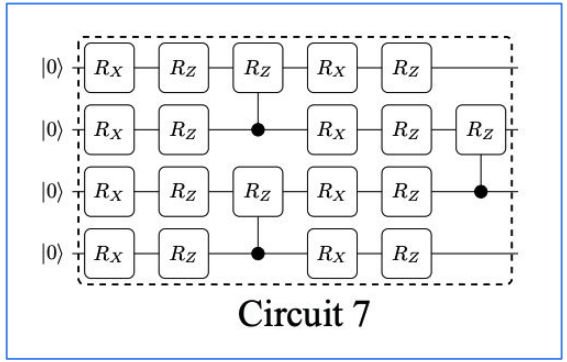
```
'2x4_ryzxy':
    [
        {'input_idx': [0], 'func': 'ry', 'wires': [0]},
        {'input_idx': [1], 'func': 'ry', 'wires': [1]},
        {'input_idx': [2], 'func': 'rz', 'wires': [0]},
        {'input_idx': [3], 'func': 'rz', 'wires': [1]},
        {'input_idx': [4], 'func': 'rx', 'wires': [0]},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

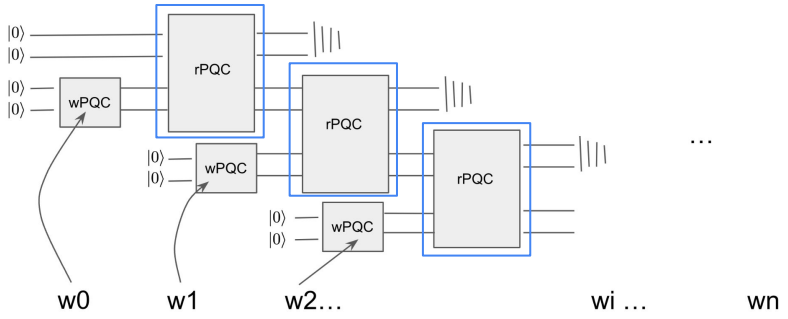


QUANTINUUM

“Stairs” Architecture in Practice





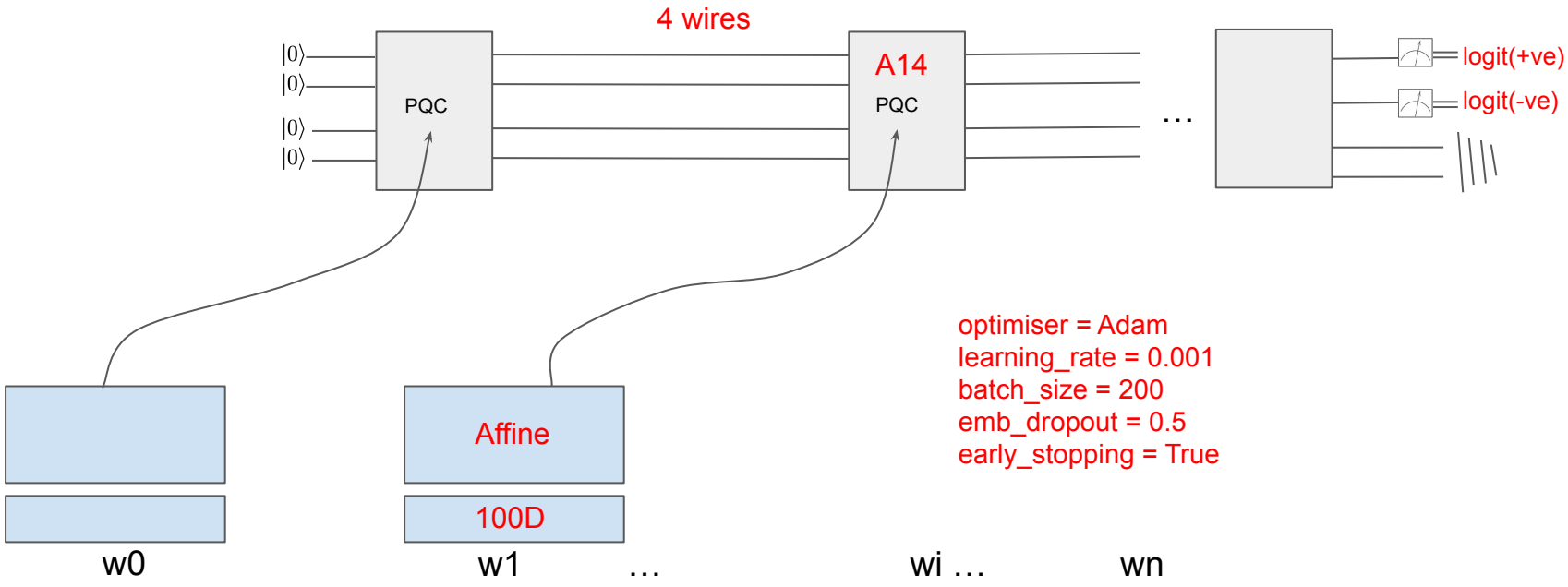


Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms

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

QUANTINUUM

Example Experimental Settings



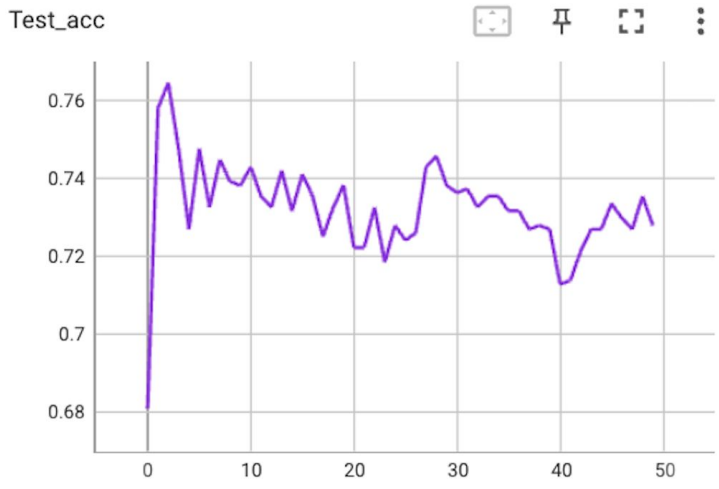
optimiser = Adam
learning_rate = 0.001
batch_size = 200
emb_dropout = 0.5
early_stopping = True

QUANTINUUM

Results on RT Test Set

4 wires	Acc
Classical RNN	77.2
“Flat”	77.9
“Stairs”	79.5
Classical GRU	78.7
Classical LSTM	79.4
LSTM (Dai and Le, 2015)	79.7

Learning Curve

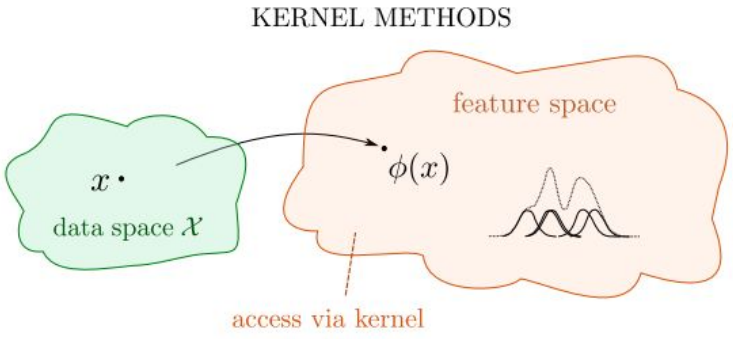


NVidia A30 GPU, PyTorch 1.12:

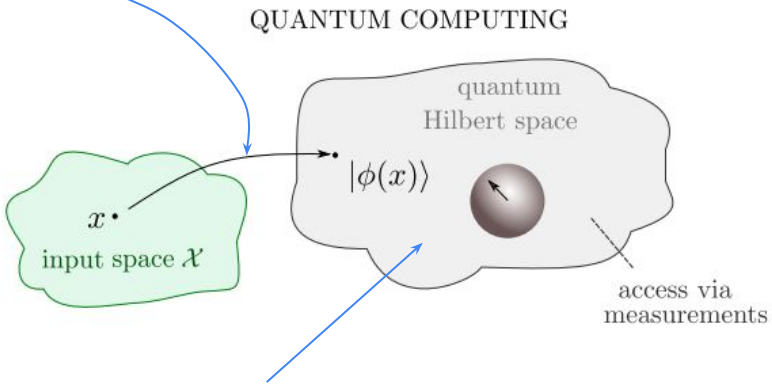
~5 secs / epoch for 1 wire (pure state)
~11 secs / epoch for 2 wires (pure state)
~14 secs / epoch for 4 wires (pure state)
~26 secs / epoch for 8 wires (pure state)

QUANTINUUM

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

QUANTINUUM

So What's the Current State of Quantum Hardware?

June 30, 2023

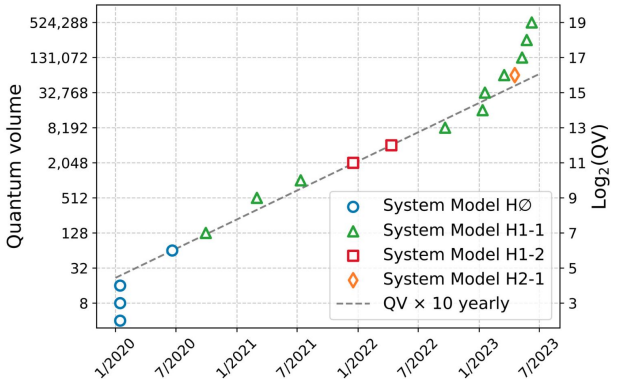


Figure 1: H-series progress quantum volume improvement trajectory

QUANTINUUM

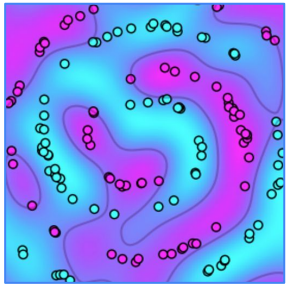
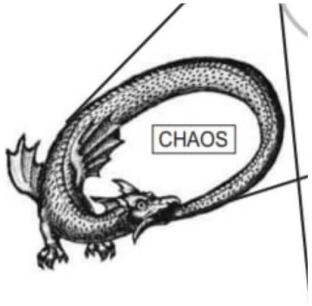
Future Work

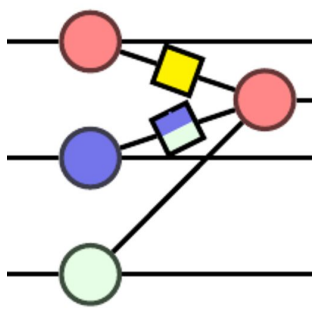
- Apply the models to more tasks
  - sequence labelling, language modelling, translation, …
- Apply pre-training / fine-tuning paradigm
- Develop more hybrid architectures
  - based on CNNs (e.g. MERA-like), transformers, …
- Run on quantum hardware

The Oxford Hybrid NLP Team





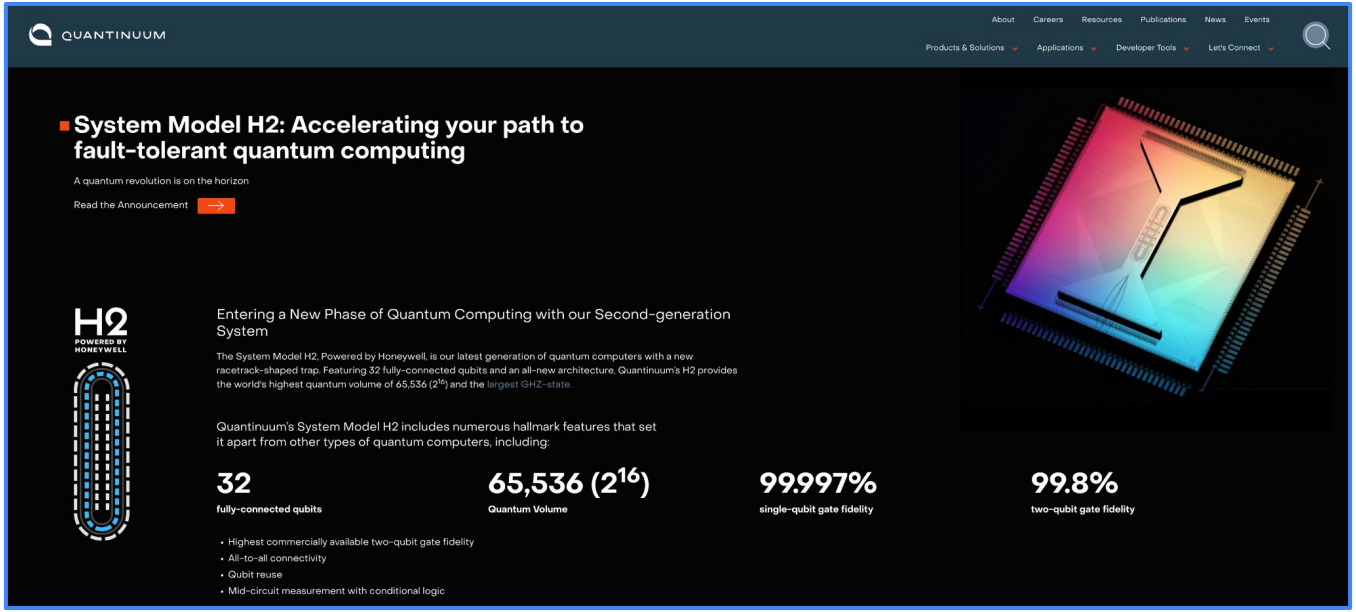







Wenduan Xu, Konstantinos Meichanetzidis, Douglas Brown, Gabriel Matos,
Charlie London, Richie Yeung, Carys Harvey, Nikhil Khatri, Stephen Clark

QUANTINUUM

The Future is (Almost) Here



QUANTINUUM