Quantum Machine Learning

(Slightly) More Than Just a Buzzword

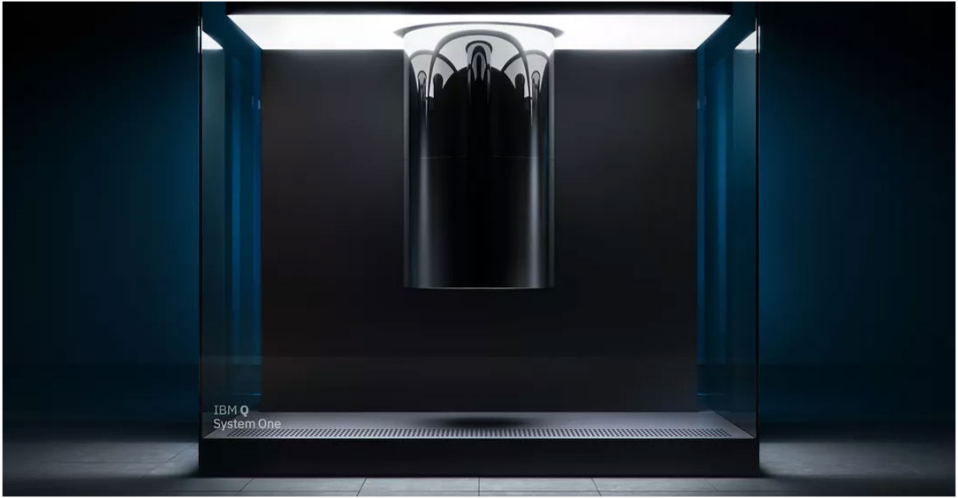
Wilfrid Somogyi

University College London

Table of contents

Variational Quantum Circuits

Awaiting the Quantum Singularity



NISQ Computers

VQAs are well-suited to the constraints imposed by noisy intermediate-scale quantum (NISQ) computers.

They use classical optimization methods and borrow techniques from machine learning.

This allows useful problems to be solved with much shorter circuits.

Quantum Algorithms

Variational quantum algorithms (VQAs) encompass a range of different techniques for solving both quantum and classical problems

Classical Problems

- Classifiers
- Autoencoders
- Generative models
- Solving linear systems
- Factoring

Quantum problems

- Eigensolvers
- Quantum simulation
- Cryptography
- Quantum annealing

Quantum Neural Networks

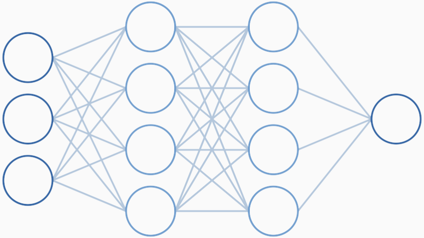
Classical Neural Networks

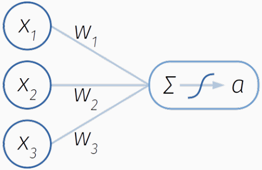
Fancy way of learning a parameterised function

$$f(\vec{x}; \vec{w})$$

Parameters are refined using various optimization methods

- Batch gradient descent
- Stochastic gradient descent
- Particle swarm
- Genetic algorithms





Quantum Neural Networks

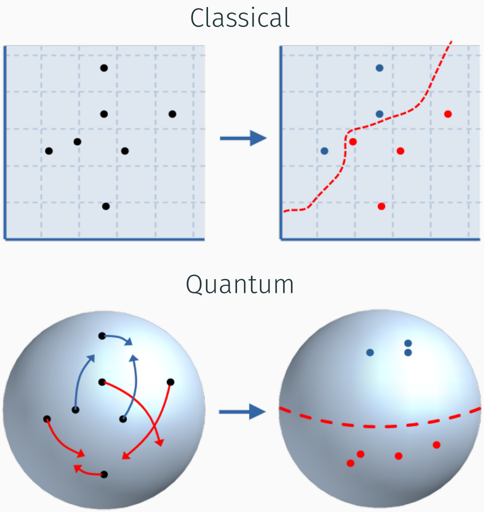
Fourier series approximator using gates with trainable parameters

$$f(x;\vec{w}) = \sum_{\omega} c_{\omega} e^{i\omega x}$$

Repeated encodings increase the number of frequencies

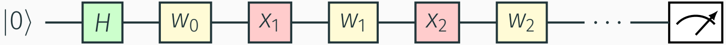
QNNs learn an arbitrary function to encode data in a quantum state

A simple measurement linearly separates the data



Feature Encoding

Today we have used alternating trainable layers with feature encoding rotations to build a single qubit QNN



To really leverage the power of quantum computers we need multi-qubit models...

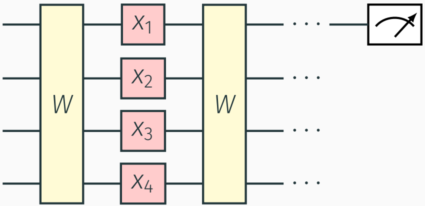
Scalable Multi-Feature Encoding

Multi-qubit Models

Quantum computers have just tens of qubits

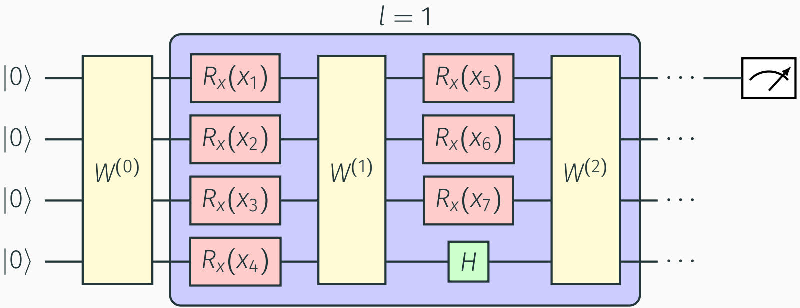
Current multi-qubit QNNs encode one feature per qubit, or less

Limits the type of problem you can solve



Scalable Ansatz

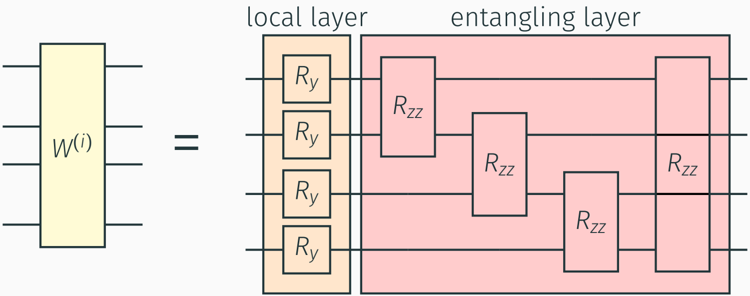
Develop a scalable multi-feature encoding ansatz to embed features $n > n_{\text{qubits}}$



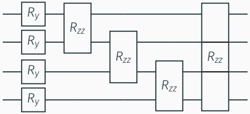
Pad additional qubits with Hadamard gates to increase the dimension of the Hilbert space

Trainable Layer

The trainable layer consists of a local rotation layer and an entangling layer

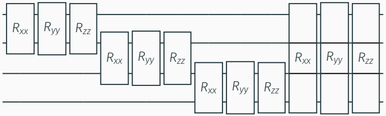


Universality



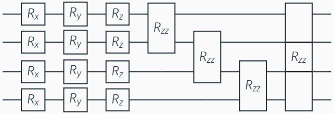
1-axis local rotation + n.n. Ising gates

Total gates: $2n$



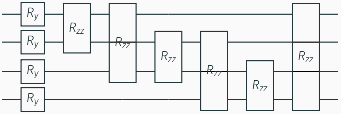
No local rotation + 3-axis n.n. Ising gates

Total gates: $3n$



3-axis local rotation + n.n. Ising gates

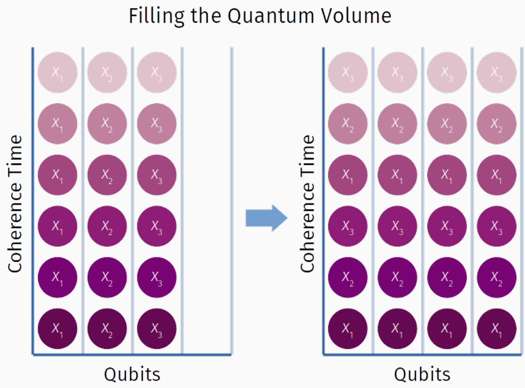
Total gates: $4n$



1 axis local rotation + fully connected Ising gates

Total gates: $\frac{n}{2}(n+1)$

Coherence vs. Qubits



Utilise all available qubits *and* deepest possible circuit

Multi-feature encoding results in deeper circuits

Understand how noise on real NISQ computers effects QNN training

Quantum Noise

Noise Channels

1. Gate Noise: Finite probability of failing to execute a given gate

$$U\,|\psi\rangle\langle\psi| \to \sqrt{1-p}U\,|\psi\rangle\langle\psi| + \sqrt{p}\mathbb{I}\,|\psi\rangle\langle\psi|$$

2. Amplitude Damping: Energy loss to the environment

$$|1\rangle_S\,|0\rangle_E \to \sqrt{1-p}\,|1\rangle_S\,|0\rangle_E + \sqrt{p}\,|0\rangle_S\,|1\rangle_E$$

3. Phase Damping: Quantum decoherence due to adiabatic interactions with the environment

$$|0\rangle_S\,|0\rangle_E \to \sqrt{1-p}\,|0\rangle_S\,|0\rangle_E + \sqrt{p}\,|0\rangle_S\,|1\rangle_E$$
$$|0\rangle_S\,|0\rangle_E \to \sqrt{1-p}\,|0\rangle_S\,|0\rangle_E + \sqrt{p}\,|0\rangle_S\,|1\rangle_E$$

Data Collection

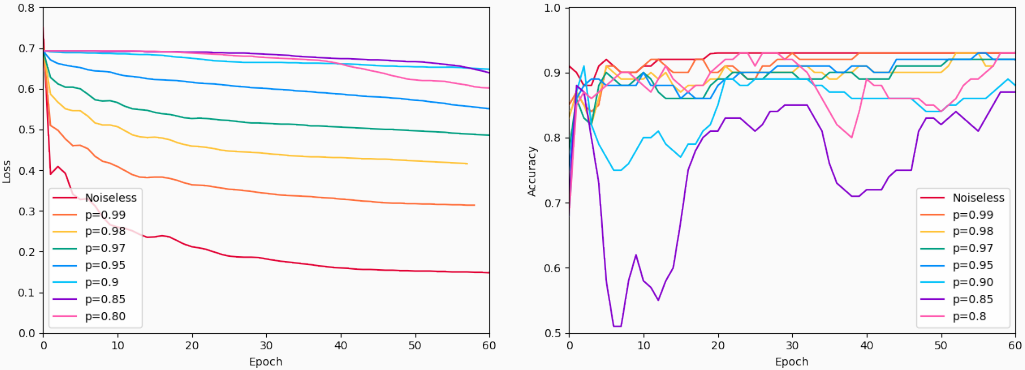
Studying the effect of noise is challenging:

- 10 samples of different levels of noise for each channel
- 50+ training instances of each sample for error bars
- 100 epochs for convergence of each QNN instance
- Each epoch trains using 100 samples from HTRU2
- Each sample uses 1000 circuit shots to obtain an expectation value

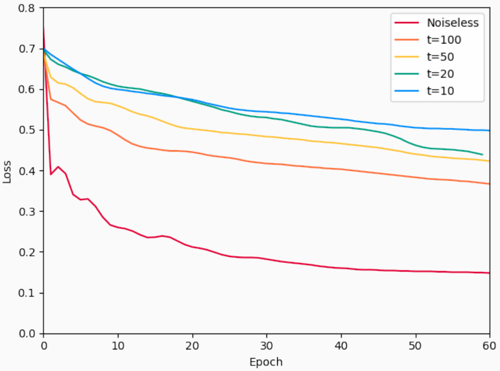
This means $10 \times 3 \times 50 \times 100 \times 100 \times 1000 = 15$ billion circuits shots in total!

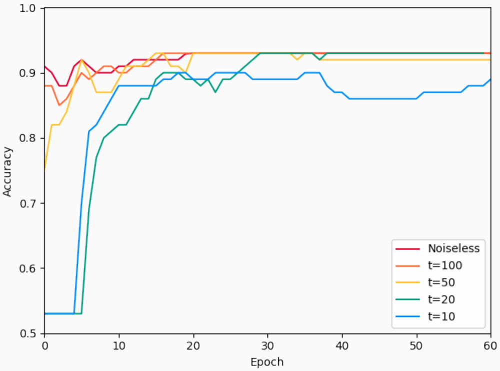
Noisy simulations using myQLM on the Atos Quantum Learning Machine

Gate Execution Noise

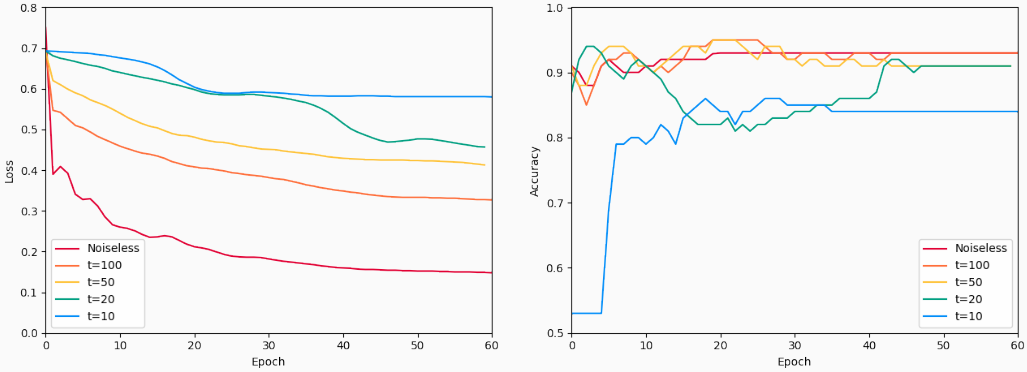


Amplitude Damping Noise





Dephasing Noise



Conclusions

Box classification preserves accuracy as loss increases due to noise

Step-like reductions in loss may imply the model learns how to defend against noise
- e.g repeated data encodings can protect against gate execution loss

High stochasticity makes it difficult to draw concrete conclusions without error bars...

But initial results suggest QNNs are highly resistant to the effect of noise