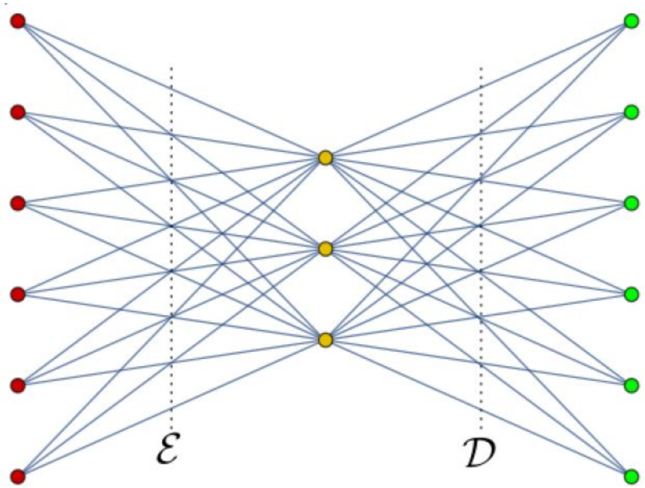
Quantum AutoEncoder

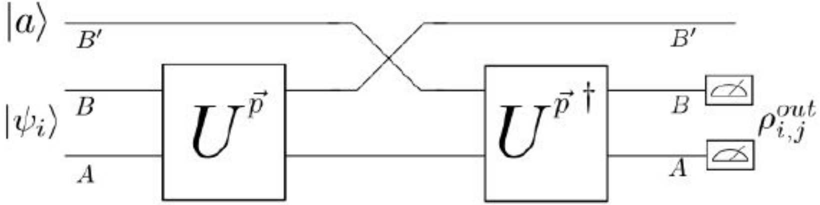
QHAAT Team

Main Reference: arXiv:1612.02806v2

Compression of State Space

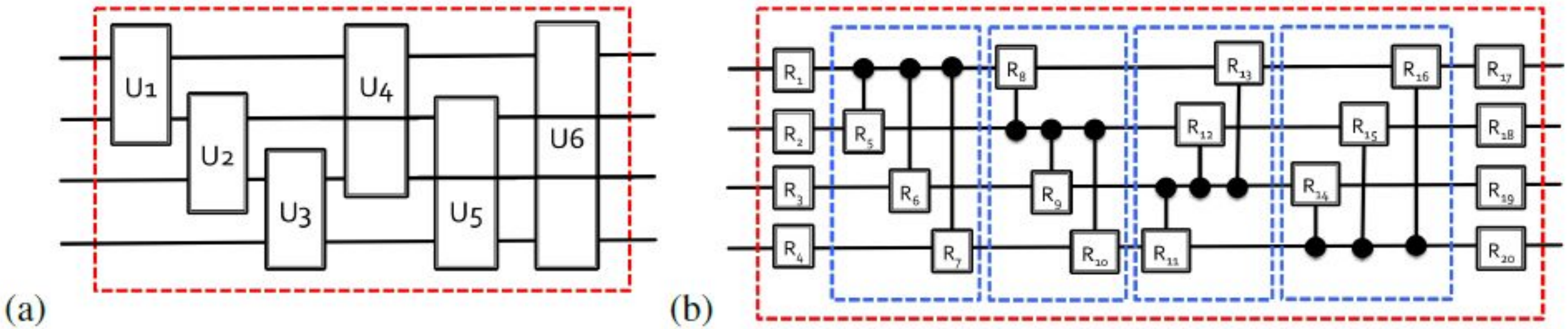


Classical

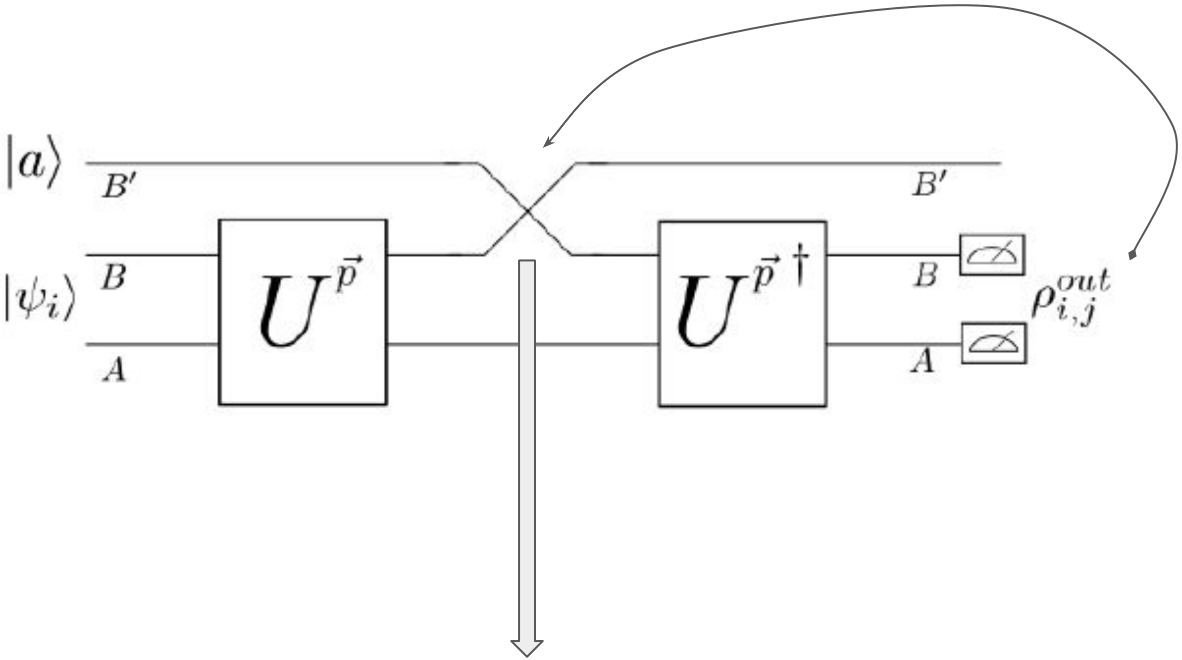


Quantum

Generic Operators:



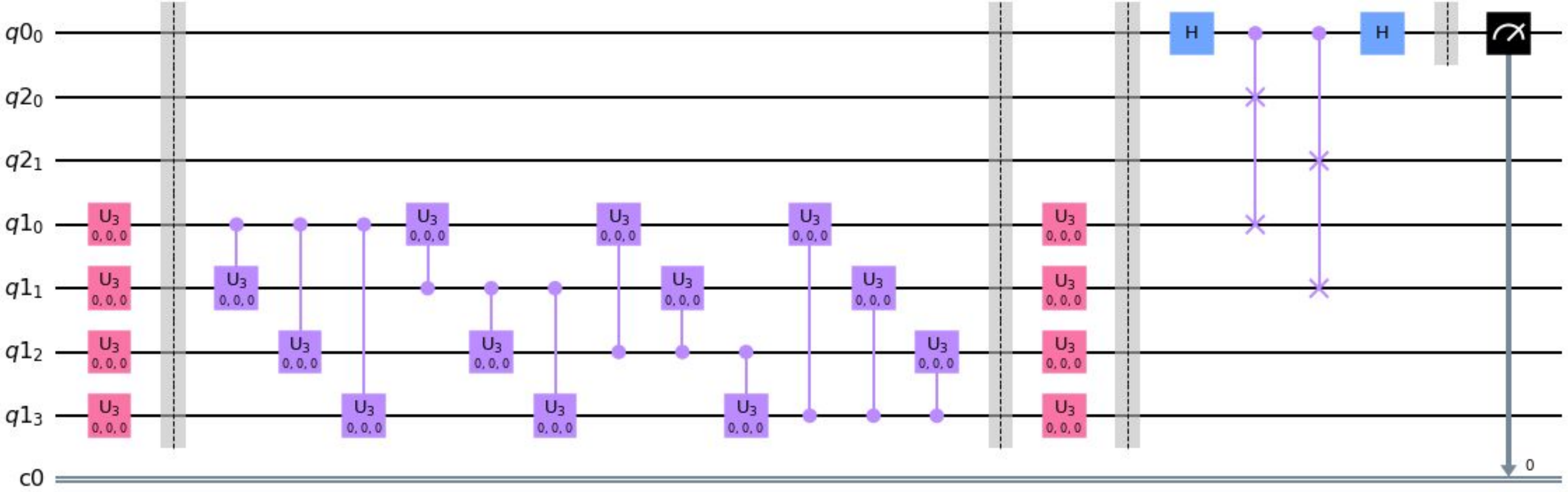
Optimization Metric



Evaluate as Swap Gate

Take advantage of the unitarity!

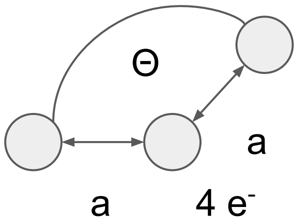
Qiskit Circuit Diagram:



Applied test cases:

A) Trivial sanity check: leave two of the qubits almost 0

Converged to fidelity of $1\text{-}10^{-11}$



B) Toy molecular orbitals.

Converged to testing fidelity of 99.389%, verification fidelity of 99.051%

C) MNIST handwritten digits, reduced to 14x14

Still running, currently fidelity 0.823.

Discussion

In terms of the parameter space, the number of parameter scales as the number of qubits squared. Although compared to the state space scaling with the exponential of the number of qubits, this is smaller. But for the current scale of quantum computers, this still poses some limitations.

Classically, we can train it by calculating the fidelity analytically and perform optimization using standard algorithms. In principle, we can calculate the gradient of the parameters analytically, but we did not have time to explore that. We evaluated the fidelity function with perturbations to approximate the gradient, which is limiting the training speed. We would explore this option in the future for a model trained classically but evaluated quantumly.

Discussion

For QPU based training, we have no analytical formula for the fidelity. Evaluating high fidelity requires high amount of sampling for the swap test. We will need to prepare multiple copies of input state and reference state, where preparing the latter is apparently easier than the former. We also need many trail perturbations of the parameters to approximate the gradient for training. Moving forward, we might need a better training scheme that calculate the gradient in a quantum fashion?

Furthermore, we believe our encoder circuit used in this model need to be improved. Particularly, we want to design a more powerful circuit, for example, consisting of general d-qubit gates, where d2. Note that, as long as d is small, the number of trainable parameters is polynomial in terms of input dimension. We also want to try higher dimensional data and higher compression rate in our future work.

Thank you for your attention!