

Assessing the Effect of Overparameterization in Quantum Neural Networks

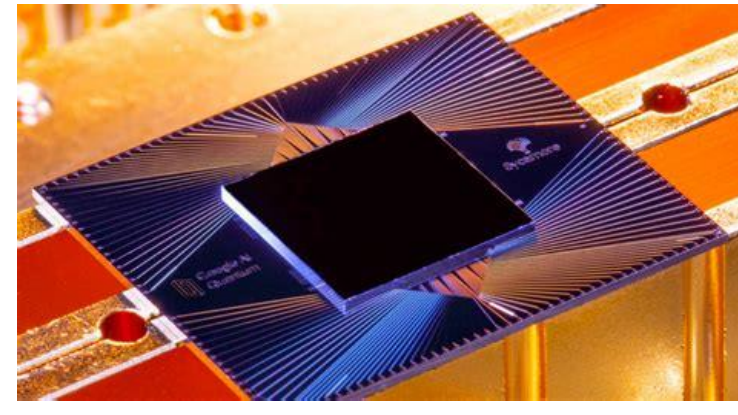
Ryan Schwartz, Computer Science
Mentor: Christian Arenz, PhD

To what extent does increasing the number of parameters in a quantum machine learning algorithm increase the rate of convergence when solving chemical problems?

1. Motivation

Quantum computing [1]

The field of Quantum Computation aims to use the properties of quantum mechanical systems to solve classically intractable problems.

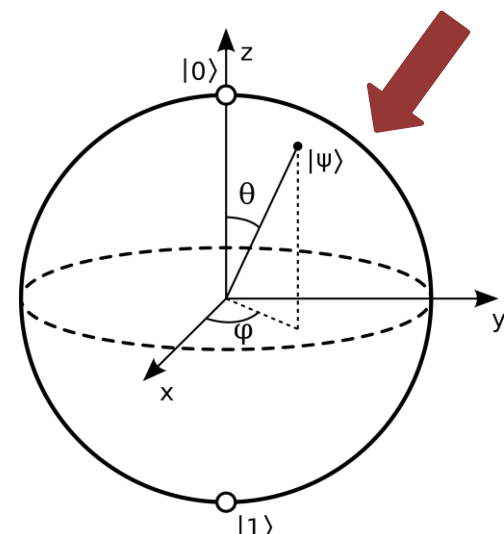


Picture of Sycamore chip taken from Google Quantum AI

Variational quantum algorithms (VQA) [2]

Today's quantum devices can be used in tandem with classical methods to solve high-value problems in an iterative loop, which can be understood as training a quantum neural network.

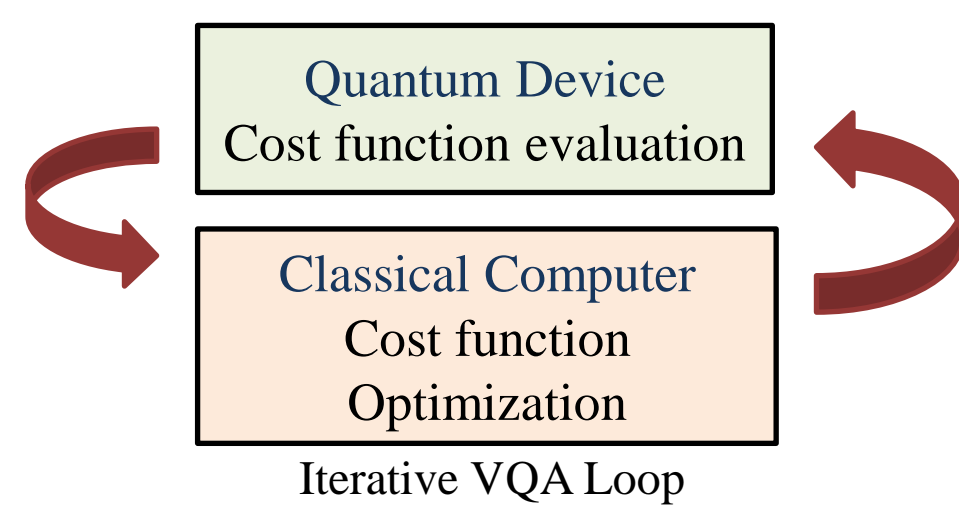
In quantum computing a unit of computation is a qubit which can be represented by the Bloch sphere



Picture taken from Wikipedia

VQA applications

Companies such as Google, IBM, and Honeywell are investigating these algorithms for their applications in chemistry, combinatorial optimization, finance, and more.



Problem: poor convergence to correct solutions

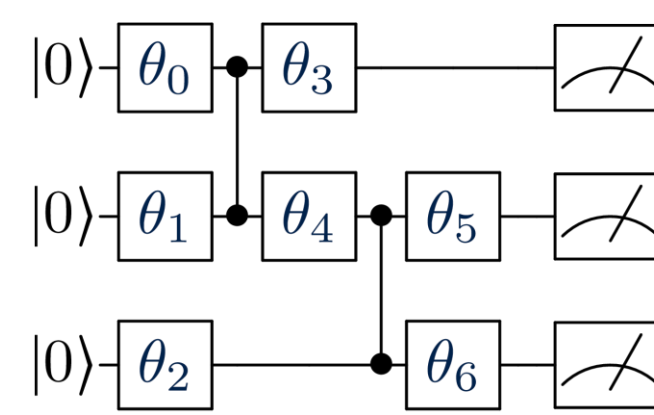
2. Theory of VQA's

First, we formulate the problem of interest as an optimization problem

$$\min_{\vec{\theta} \in \mathbb{R}^M} J(\vec{\theta})$$

Quantum Device

- Evaluates cost function by measuring qubits



Parametrized quantum circuit
Picture taken from entropicalabs.io

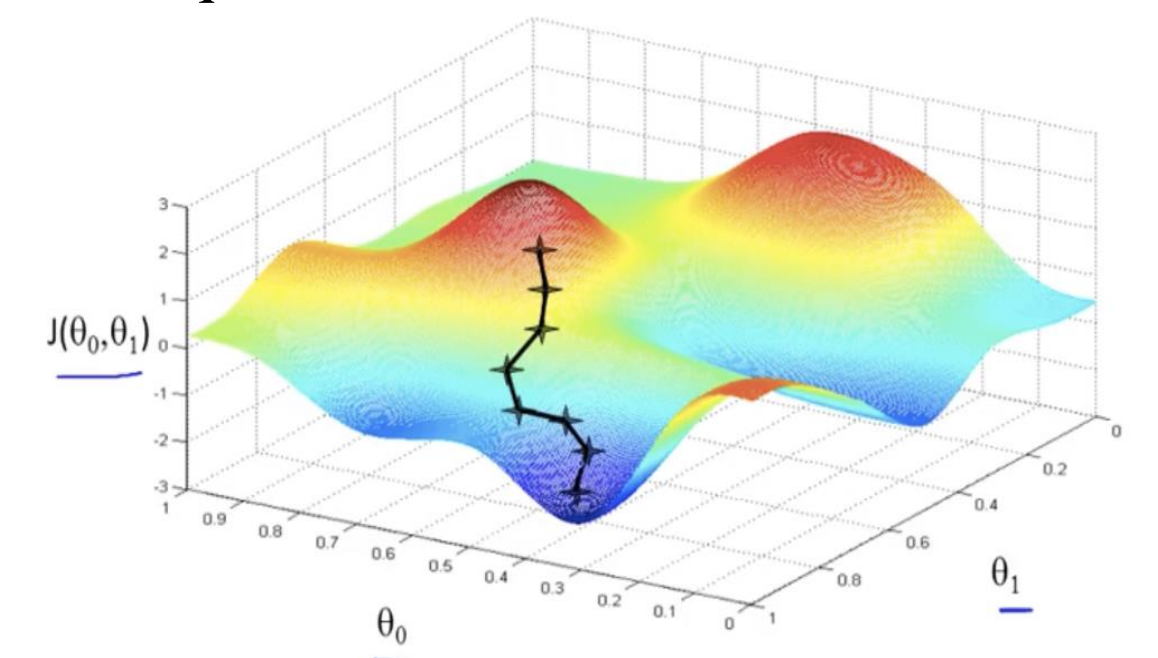
$$J(\vec{\theta}) = \langle 0|U^\dagger(\vec{\theta})H_p U(\vec{\theta})|0\rangle$$

Problem Hamiltonian encodes the solution to the problem of interest, which correspond to the lowest energy of the system qubits

Classical Computer

- Updates the parameters by e.g., using the gradient descent algorithm to descend the landscape of J

$$\vec{\theta}_{n+1} = \vec{\theta}_n - \gamma \nabla J(\vec{\theta}_n)$$



Landscape of J taken from Medium

Can we make the landscape easier to descend by adding more parameters to the quantum circuit, thereby improving convergence[3][4]?

3. Results

1 Qubit System Case Study with 3 variational parameters

Comparing different types of overparameterization for a fixed 1 qubit problem Hamiltonian (2 x 2 matrix):

Overparameterization 1: U_1

a sequence of rotations about 3 axes by $\theta_1, \theta_2, \theta_3$

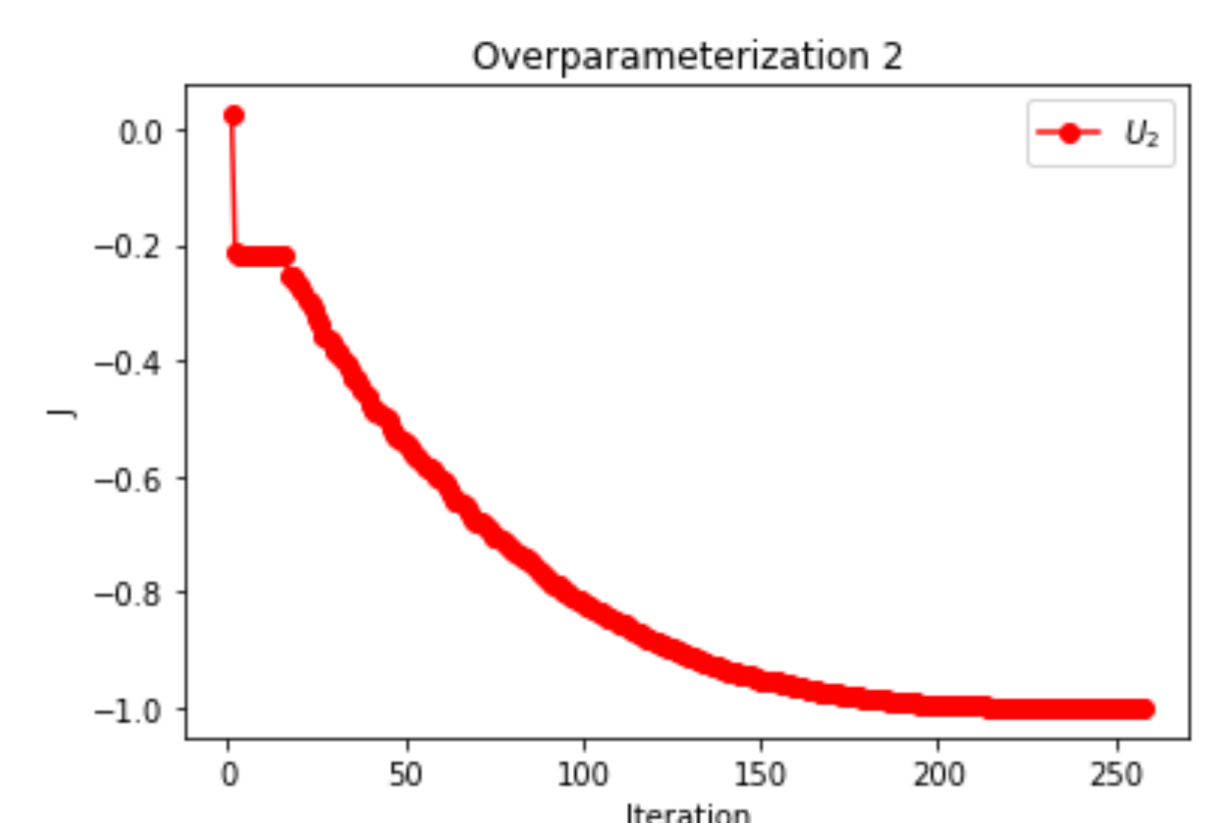
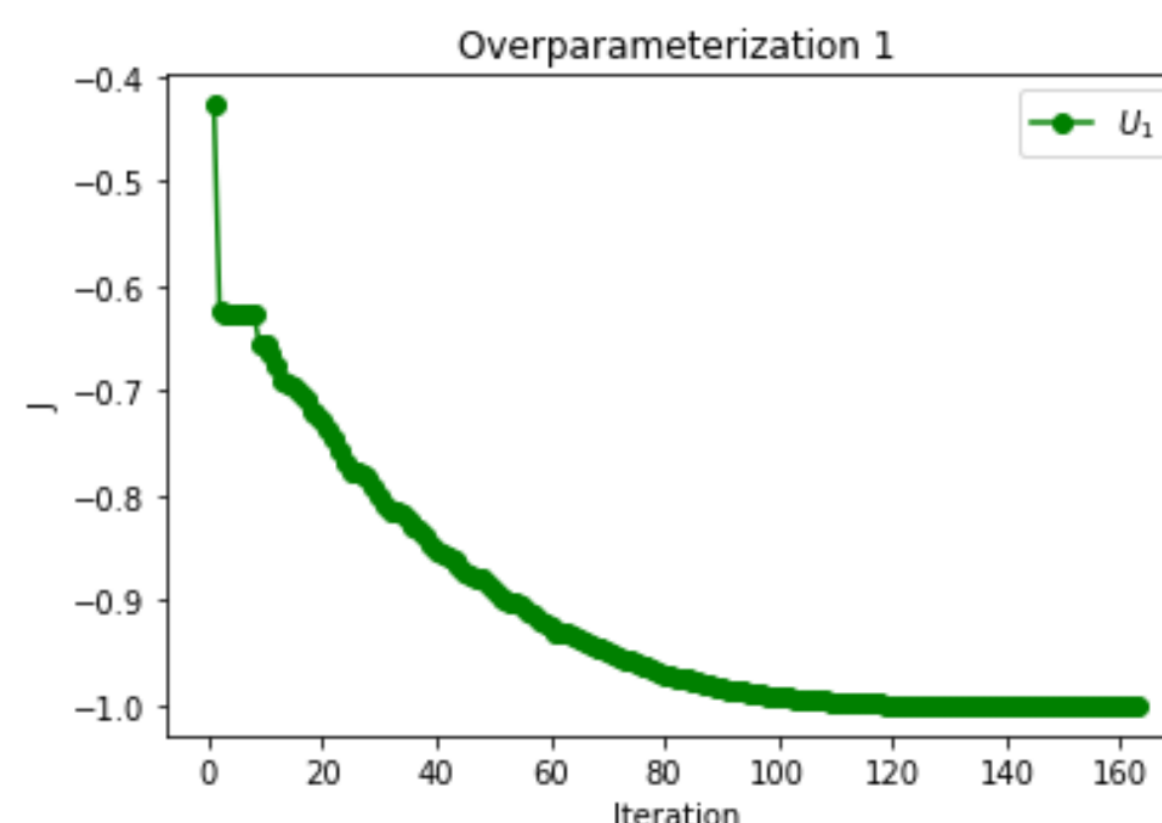
Overparameterization 2: U_2

1 rotation around an arbitrary direction defined by $\theta_1, \theta_2, \theta_3$

Numerical results

We found that for randomly chosen initial values of, $\theta_1, \theta_2, \theta_3$ the value of J consistently converges to -1. (The lowest possible value of J).

This suggests that U_1 and U_2 are sufficiently similar in helping to descend the optimization landscape. Additionally, it shows parameterizing the quantum circuit with 3 parameters, is sufficient to consistently reach convergence for a single-qubit system.



4. Conclusion and Outlook

Summary

We found that in the case of both types of overparameterization, similar convergence to the global minimum energy state was reached with 3 parameters using the gradient descent algorithm on the cost function J. This suggests that both types of overparameterization allows for favorable (local-optima-free) optimization landscapes

Open questions

How can we overparameterize to obtain similar convergence behavior for multi qubit systems?

How many parameters are needed to obtain favorable optimization landscapes in multi qubit systems?

5. References

- [1] Nielsen, M. and Chuang, I. Quantum Computing and Quantum Information. Cambridge University Press, Cambridge (2003)
- [2] M. Cerezo et al., Variational Quantum Algorithms, Nature Reviews Physics 3, 625 - 644 (2021)
- [3] A. Magann et al., From Pulses to Circuits and back again: A quantum optimal control perspective on Variational Quantum Algorithms. PRX Quantum 2, 010101 (2021)
- [4] J. Lee et al., Progress toward favorable landscapes in quantum combinatorial optimization, Phys. Rev. A 104, 032401 (2021)