# Investigation of Barren Plateaus in Quantum Neural Network Development

Quantum neural networks (QNN) are machine learning models that are inspired by the workings of classical artificial neural networks, but which utilise quantum circuits for their representation and a mix of quantum and classical training algorithms.

A common approach to implementing QNNs are variational quantum algorithms (VQA), which take advantage of classical computation for the optimisation of a parametrised quantum circuit to train QNN, while using quantum machines to model the landscape of the loss function and efficiently estimate its gradient. VQA allows training of well-designed QNNs to rapidly converge to a solution while avoiding many problems present in training classical neural networks. QNNs, however, have their own trainability issues, which can manifest in large variational QNN circuits, either due to their depth, a large number of qubits, or poor initialisation of their parameters. One such problem is the formation of large flat areas in the cost function landscape, called barren plateaus, which impede effective circuit optimisation.

The methods of eliminating barren plateaus or mitigating their presence in circuit optimisation have been proposed. However, the efficacy of each method in the context of a given QNN architecture and the selected training data remains an open question. This project, therefore, aims to investigate the effectiveness of different approaches to mitigate the emergence of barren plateaus in various QNN developmental circumstances. To this end, three different approaches were selected to deal with barren plateaus and were then evaluated against different VQA quantum circuit structures (depth, qubits and initialisation), initially using a random gradient landscape and then on a sample classification problem.

## Authors


**Thanh Nguyen**
**Institute of Research and Development,**
**Duy Tan University**

**Prof. Jacob Cylbulski**
**School of IT,**
**Deakin University**


## Objective

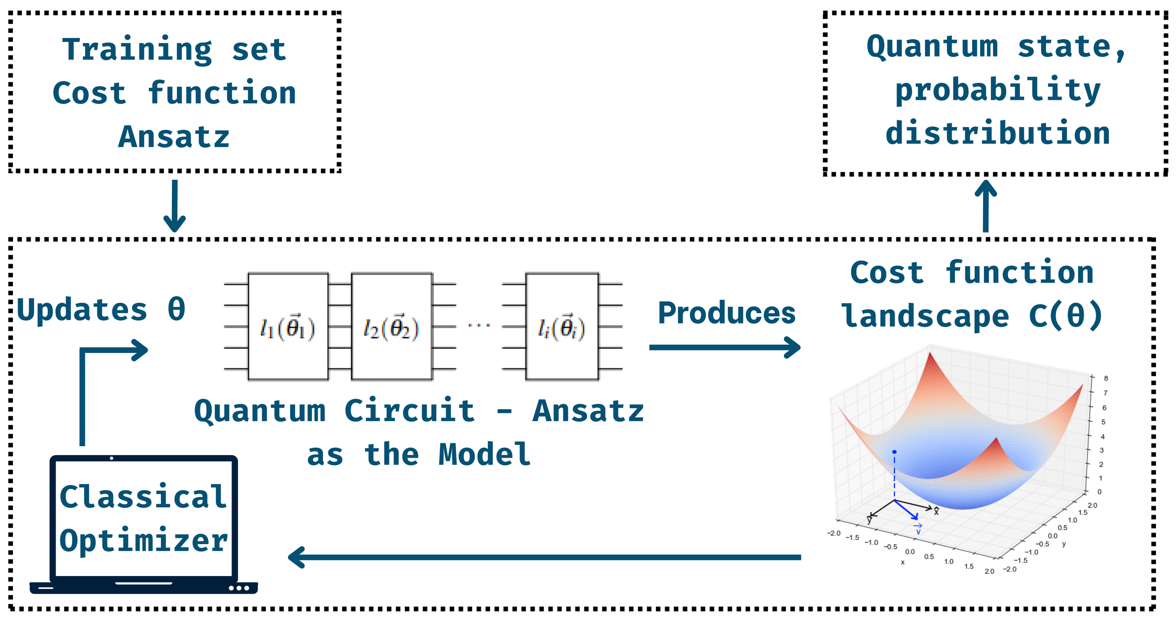
- Provide the background information for quantum computing, VQA and QNN, which enables the definition of the research question.
- Define the BP problem and the causes that would lead to this phenomenon.
- Investigate several methods to mitigate or avoid BP.
- Compare advantages and disadvantages of the identified methods to identify the most appropriate approaches in different circumstances

## Introduction

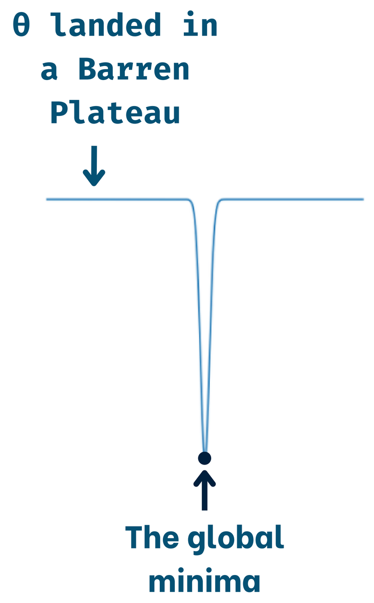
Barren Plateaus is the training difficulty that impedes Variational Quantum Algorithm and Quantum Neural Networks training. If the gradient based optimizer runs into Barren Plateaus, the Objective Function landscape becomes flat, causing failure to converge.

We investigate and compare three different methods that counter barren plateaus, which address the factors lead to barren plateaus in VQA and QNN training:

- **The Parameterized Circuit design (number of qubits and circuit depth);**
- **Randomized Parameters Initialization.**




**Typical Structure of Variational Quantum algorithm. The Ansatz acts as the neural network that accept trainable parameters to provide predictions on data**

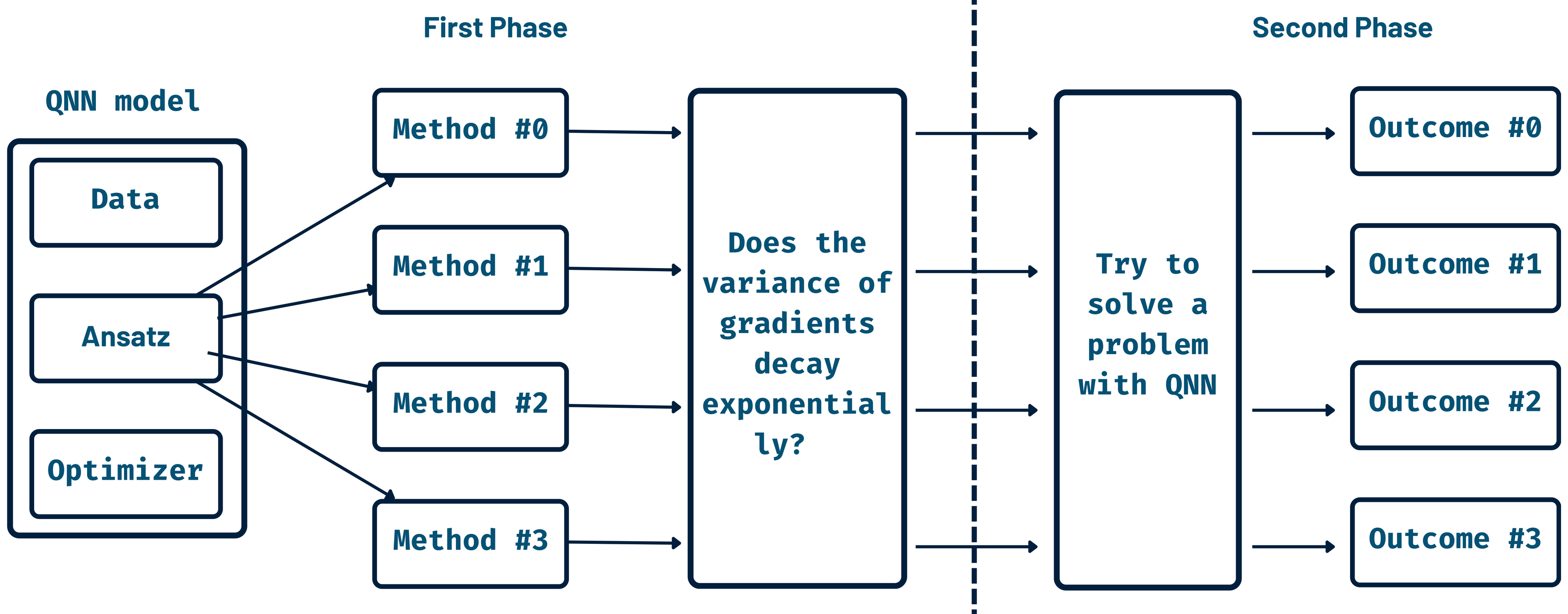



**Visualise the Cost Function landscape produced from the QNN model**

## Methodology

We employ three distinct methods for barren plateaus mitigation, and compare to a base QNN (method 0). The three designs are compared against each other in their gradient decay rate per qubits, and performance in a classification problem.

- We used 100 uniformly randomized parameters to scan the gradient and calculate the gradient variance.
- Use the ansatzes as the neural network to classify a dataset, and compare their performance in terms of accuracy and objective function value. Using COBYLA optimizer and MSE loss.




**(Method 0) Base configuration**
A general multilayered perception network with randomized initial parameters and measurement for all qubits.

**(Method 1) Altering Circuit design:** Shallow depth, Local Cost Function
Shortening the circuit and only use certain qubits instead of a full measurement.

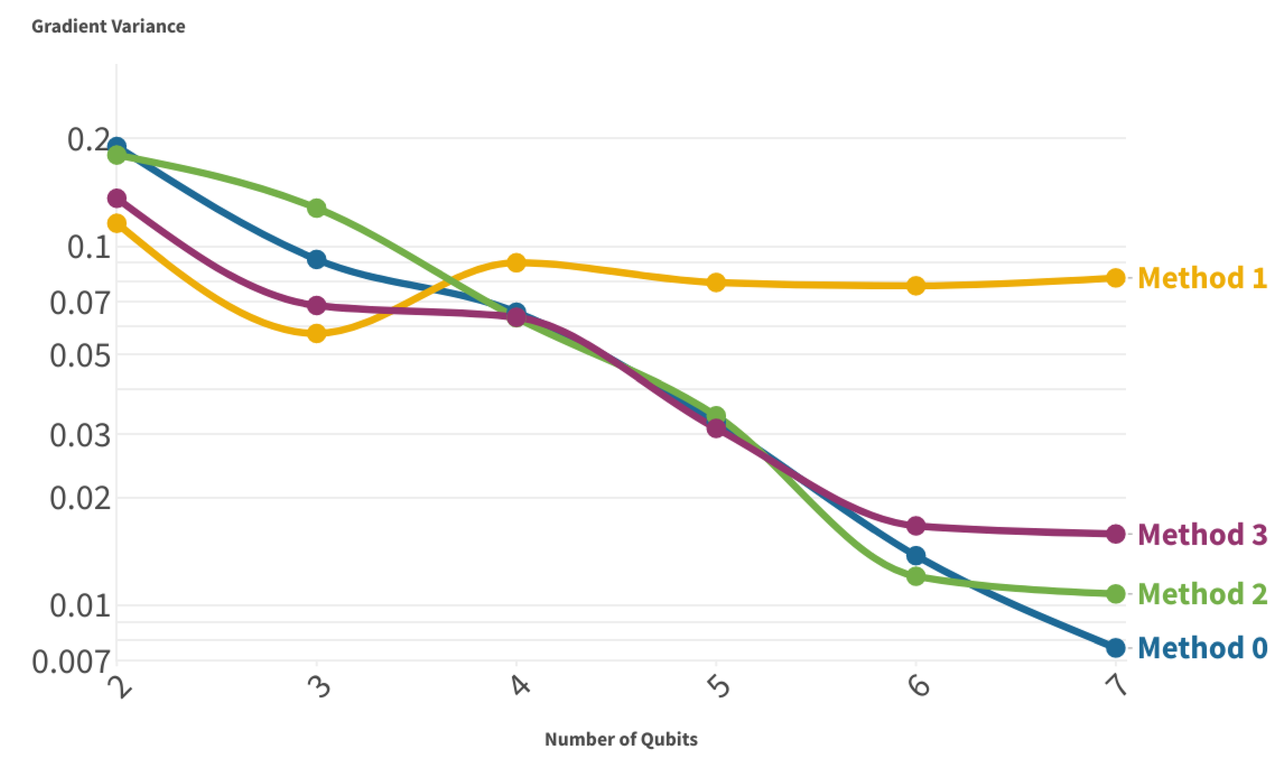
**(Method 2) Altering Parameter Initialization:** Layer-wise Learning
Add one layer and train that specific layer to form the ansatz and the initial parameters until we obtained the desired ansatz length

**(Method 3) Altering Circuit design and Parameters:** Identity Blocks
Forming an identity block takes two layers that inverts each other. We obtain the ansatz and the parameters with this strategy.

## Sampling Gradients Result

Overall, the ansatzes with the local cost function and restriction on circuit depth have their variance values remaining higher and being more consistent for higher qubit count. The ansatz with this treatment, therefore would not possess a barren plateau. On the other hand, the values for the other cases shrink exponentially and eventually, the near-zero gradient around the initial point will expand to a large plateau.

| Method | Vaiance Exponential fit |
|---|---|
| (Method 0) Base Configuration | -0.62 |
| (Method 1) Shallow depth, Local Cost Function | -0.03 |
| (Method 2) Layer-wise Learning | -0.53 |
| (Method 3) Identity Blocks | -0.56 |




Variance decay rates of four ansatzes

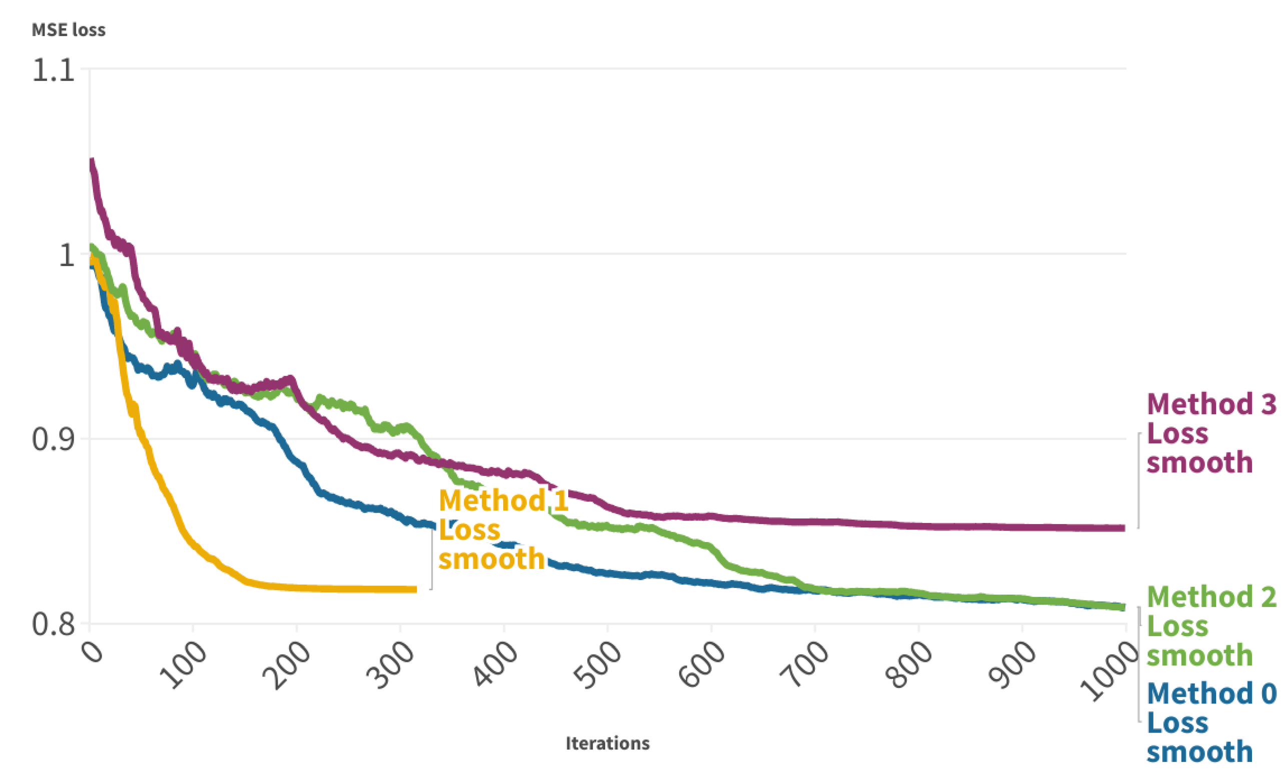
## Data Classification Result

While the unrestricted ansatz (Method 0) has produced better result at 1000 iterations, the other methods accuracy remain unchanged after 150 seconds. The unrestricted method is likely to run into barren plateaus, therefore require more optimization steps to converge.

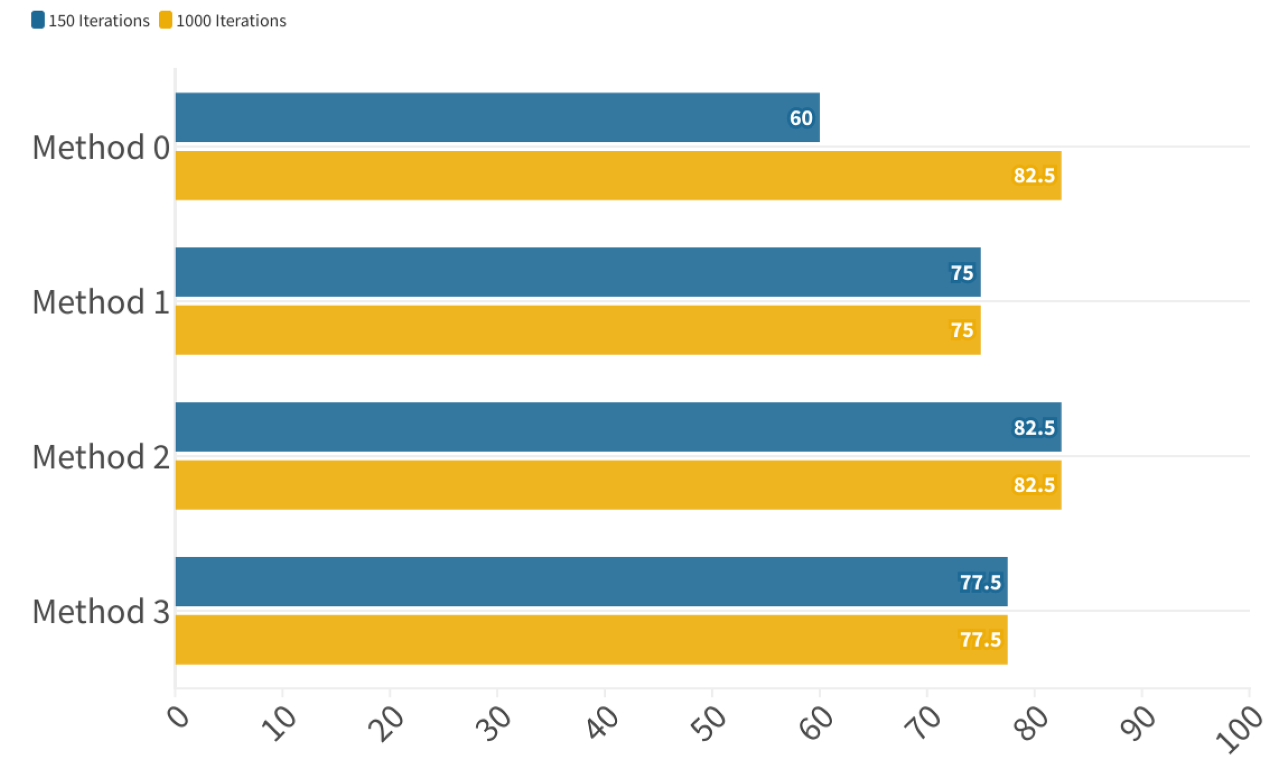
The local cost function - shallow depth ansatz seems to be the best in overcoming barren plateau as expected, and converge early. However it has high error rate due to the shortened circuit.

The last two methods focused on the initial parameters while allowing longer circuits, resulting in better capacity to learn. We can see that with the optimized initial parameters, the ansatz has better accuracy compared to identity blocks initialization. Thus they are suitable for designing ansatzes of higher layer count to learn more complex functions. However, it would take more time to obtain the optimal initial parameter due to the training process (see Methodology section).

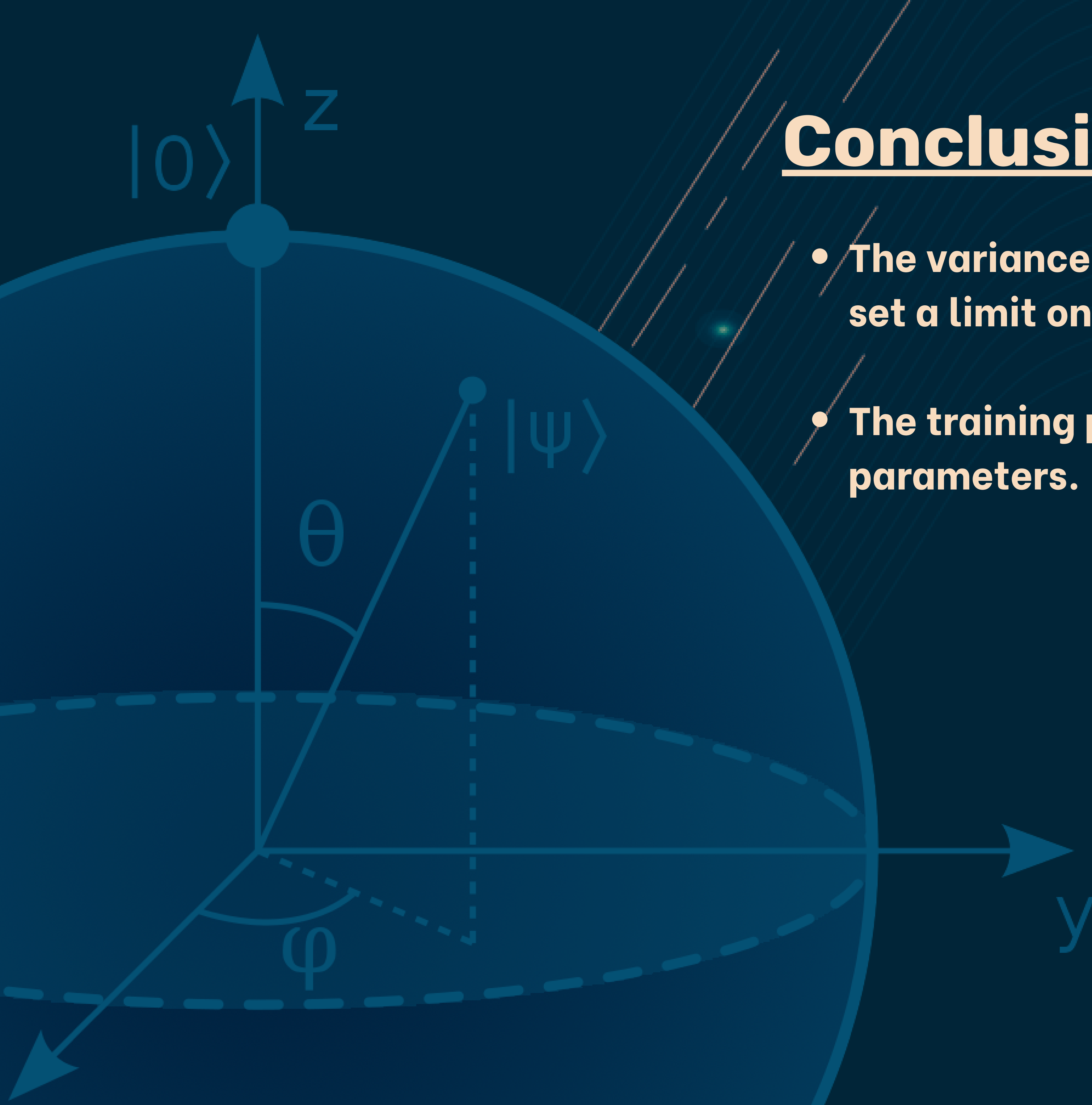
| Method | Circuit depth (layers) | Parameters Count | Accuracy at 150 Iterations | Accuracy at 1000 Iterations |
|---|---|---|---|---|
| (Method 0) Base Configuration | 123 | 105 | 60% | 82.5 |
| (Method 1) Shallow depth, Local Cost Function | 21 | 20 | 75% | 75% |
| (Method 2) Layer-wise Learning | 123 | 105 | 82.5% | 82.5% |
| (Method 3) Identity Blocks | 100 | 100 | 77.5% | 77.5% |




Loss function values of four classifiers




Accuracy score of four classifiers




## Conclusion

- **The variance of gradients of parameterized quantum circuit can be stable if we set a limit on length and the cost function.**
- **The training performance can increase if we carefully select the initial parameters.**

## Related Literature


[1] M. Cerezo et al. Variational quantum algorithms. 3(9):625–644.
[2] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles. Cost function de-pendent barren plateaus in shallow parametrized quantum circuits. 12(1):1–12.
[3] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti. An initialization strategy for addressing barren plateaus in parametrized quantum circuits. 3:214.
[4] A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib. Layerwise learning for quantum neural networks. 3(1):1–11.
[5] R. S. Sutor. Dancing with Qubits: How Quantum Computing Works and How It Can Change the World. Packt Publishing Ltd.