Quantum Circuit Design using Complex-Valued Neural Network in Stiefel Manifold

Sayan Manna, Mahesh Mohan M R
SPAI Group, Department of AI, IIT Kharagpur
sayan2003@kgpian.iitkgp.ac.in

1. Introduction

Quantum computing uses quantum gates to perform operations on qubits via unitary transformations [1]. Designing efficient, hardware-implementable quantum circuits from algorithm descriptions remains a challenge.

- **Goal**: Automate quantum circuit synthesis using machine learning.
- **Challenge**: Maintain unitarity during learning—a fundamental physical constraint.
- **Why ML?** Neural networks can learn complex unitary mappings from input–output quantum state pairs [2].
- **Innovation**: We propose a complex-valued neural network trained on the Stiefel manifold to preserve unitarity at every step.

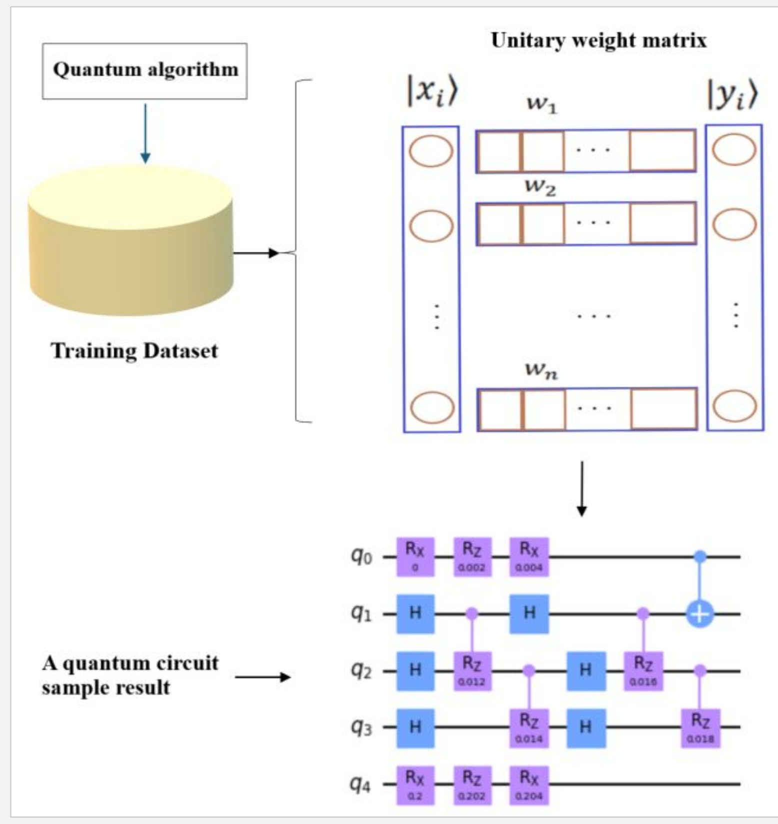


Fig 1: General protocol for quantum circuit synthesis including neural network training and the resulting quantum circuit after the transpilation in Qiskit. [2]

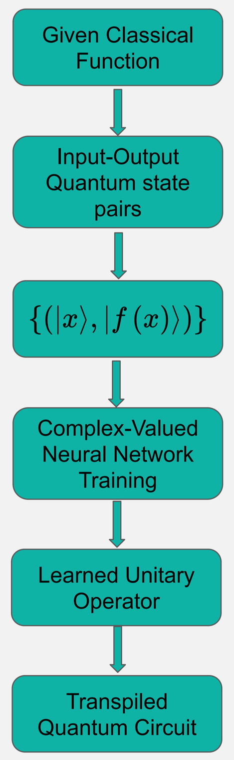


Fig 2: Flow-chart for creating quantum circuit to implement a classical function for a quantum algorithm

2. Existing Solution

Consider training on a n qubit circuit, so total number of basis states is, $N = 2^n$. The methodology is as following [2]:

- Single-layer feed-forward neural network, input state: $|\psi\rangle \in \mathbb{C}^N$, target states: $|\psi_{\text{target}}\rangle \in \mathbb{C}^N$
- Weight matrix $\left(W \in \mathbb{C}^{N\times N}\right)$ represents the unitary transformation.
- Linear activation and no bias. Forward pass: $|\psi'\rangle = W|\psi\rangle$
- Loss function, $\mathcal{L} : \mathbb{C}^{N\times N} \to \mathbb{R}$, $\mathcal{L} = \left|\left||\psi_{\text{target}}\rangle - |\psi'\rangle\right|\right|^2$
- **Unitarity enforcement**: Periodic Gram-Schmidt orthogonalization [2]

$$W^{(k+1)} = \text{GramSchmidt}\left(W^{(k)} - \lambda\nabla_W\mathcal{L}\right)$$

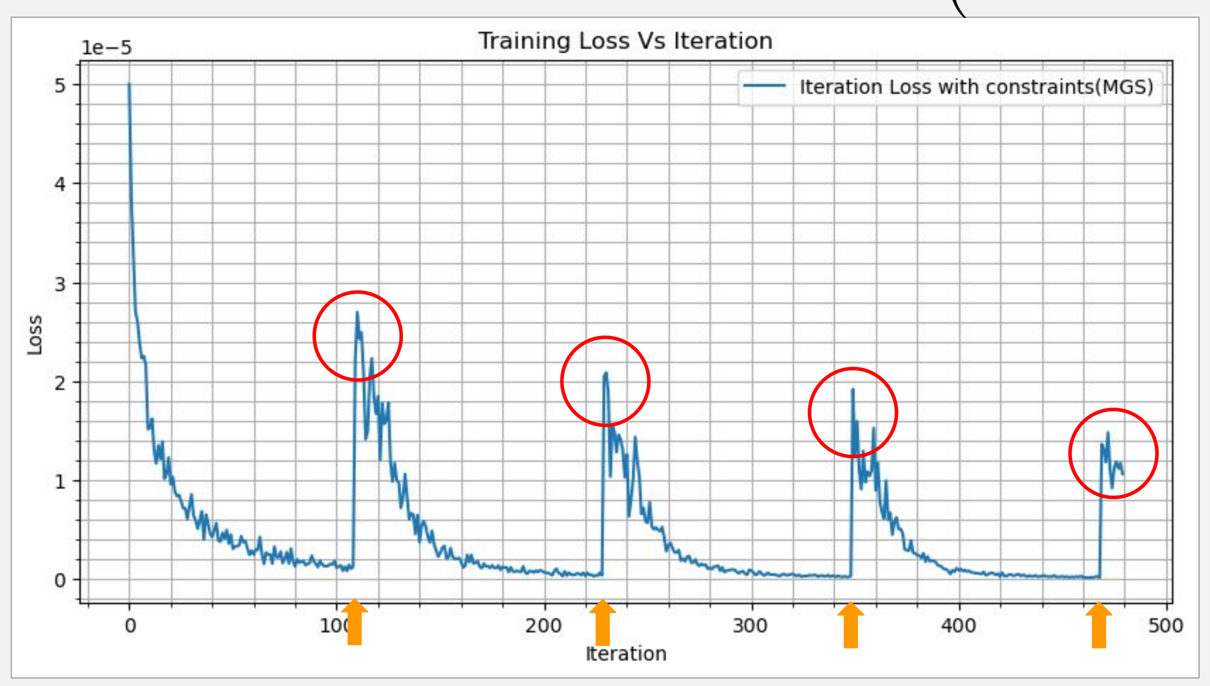


Fig 3: Loss vs Iteration curve of training on a 2-qubit entanglement quantum circuit when Gram-Schmidt orthogonalization is applied to the weight matrix at a fixed interval of epochs to enforce unitarity

Limitations of this approach:

- **Intermittent projection** → loss spikes, unstable training.
- **Numerical instability** and off-manifold weight updates.
- **No smooth optimization** — disrupts gradient-based learning.

3. Methodology: Optimization in Stiefel Manifold

The optimization problem is:

$$\min_W \mathcal{L}(W) \text{ subject to } W^\dagger W = I$$

We constrain weights to the complex Stiefel manifold [3]:

$$M\left(\mathbb{C}^N\right) = \left\{W \in C^{N\times N} | W^\dagger W = I\right\}$$

Algorithm:

- Compute Euclidean gradient: $G = \nabla_W\mathcal{L}\left(W^{(k)}\right)$ using Writinger Calculus [5].
- Compute the skew-Hermitian operator:

$$A = G\left(W^{(k)}\right)^\dagger - W^{(k)}G^\dagger$$

- Update via Cayley transform [4] (smooth & unitary-preserving):

$$\boxed{W^{(k+1)} = \left(I + \frac{\lambda}{2}A\right)^{-1}\left(I - \frac{\lambda}{2}A\right)W^{(k)}}$$

Where $\lambda \in \mathbb{R}$ is the learning rate.

Theoretical Guarantees:

- **Unitarity Preservation:** $W^{(k+1)} \in M\left(\mathbb{C}^N\right)$ for all k
- **Descent Property**: $\mathcal{L}\left(W^{(k+1)}\right) \le \mathcal{L}\left(W^{(k)}\right)$ for all $\lambda > 0$
- **Smooth Optimization**: No jumps or discontinuities.

4. Results

Trained on 2-, 3-, 4-, 5-, 7-, 8- and 10-qubit circuits.

Metric	Gram–Schmidt	Stiefel Manifold
Unitarity Preservation	Intermittent	Continuous
Training Stability	Discontinuous	Smooth
Final Unitary Error	$\sim 10^{-5}$	$\sim 10^{-12}$
Convergence Speed	Slower	Faster
Numerical Stability	Issues	Robust

Here we show example of training on a 7-qubit circuit:

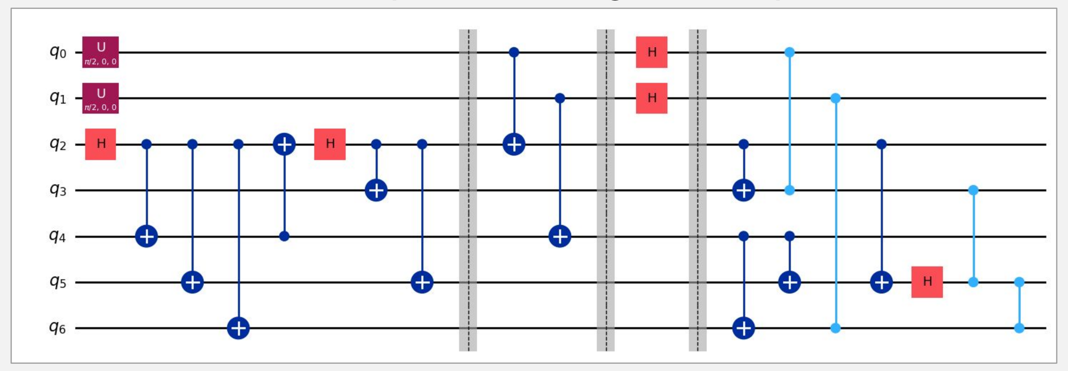
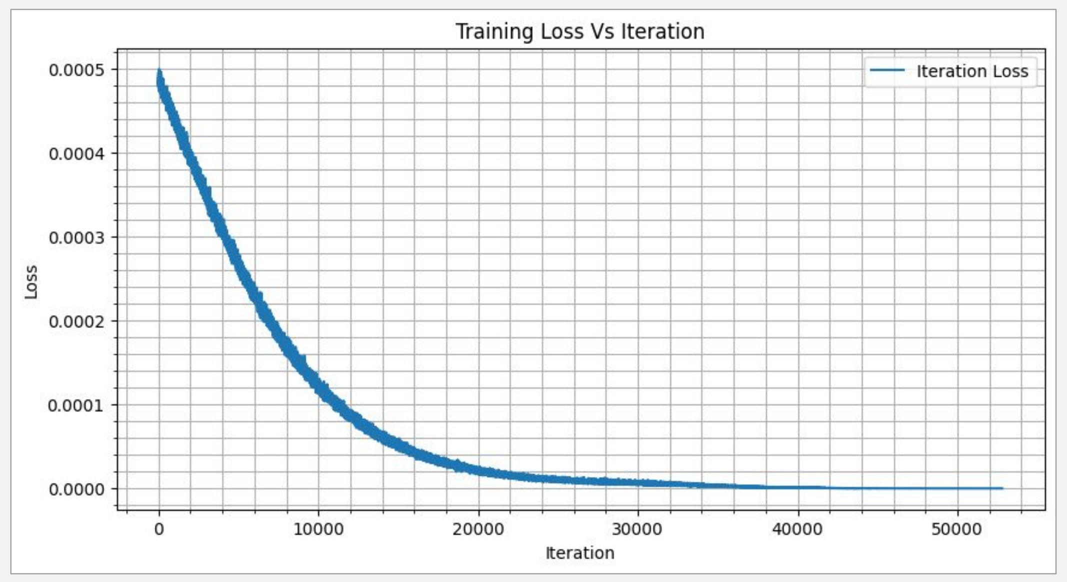


Fig 4: A 7-qubit random quantum circuit from where we collect the input-output quantum state pairs. This is a representation of a quantum algorithm.



- *Number of datapoints used: 15000, number of epochs = 150, learning rate = 0.8*
- *No hidden layer, no bias, linear activation*
- *128 neurons in both input and output layer.*
- *We observe a smoothly decreasing loss curve.*

Fig 5: Training loss vs Iteration

To analyse the training performance we use two metrics - Fidelity and Unitary Error.

Fidelity:

- Fidelity measures how close two quantum states or operations are.
- Ranges from 0 (completely different) to 1 (identical). We want the fidelity to be 1 at the end of training.
- We calculate fidelity between learned unitary matrix W_{final} and target unitary U_{target} as follows:

$$F = \frac{1}{2^n}\left|Tr\left(W_{\text{final}}^\dagger U_{\text{target}}\right)\right|$$

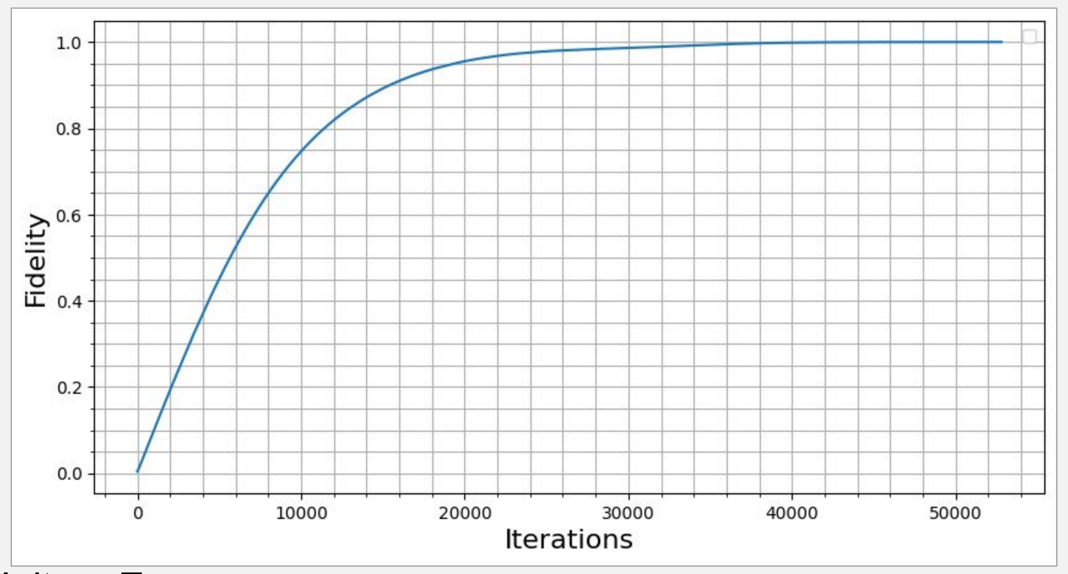


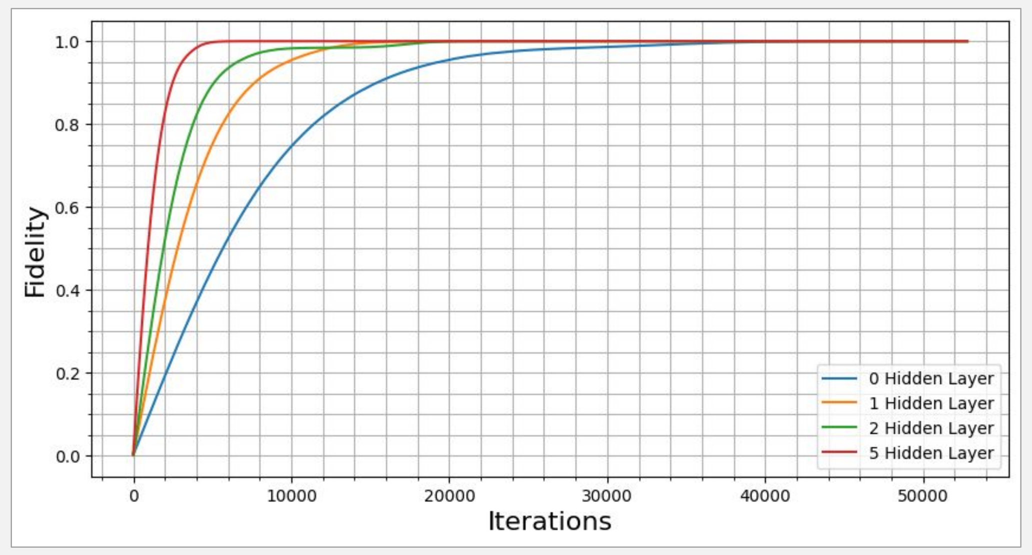
Fig 6: Fidelity vs Iteration

Here, we can observe that the fidelity is reaching 1 as the training progresses (desired).

Unitary Error:
It is calculated at each update to make sure that the weight matrix remains in the Stiefel Manifold after each iteration. We observe that the unitary error is in order of 10^{-11}

$$\left|\left|W^\dagger W - I\right|\right|_F^2 \sim 10^{-11}$$

5. Impact of Multiple Hidden Layers



- More hidden layers → faster convergence.
- Better representation capacity & hierarchical learning.
- Overall unitary: $W_{\text{net}} = \Pi W_i$ stays unitary

Fig 7: 7-qubit circuit training fidelity for: 0HL, 1HL, 2HL, 5HL

5. Future Work

- Optimal hidden layer count vs. saturation.
- Minimum training data requirements for unit fidelity.
- Scaling to 100+ qubit systems.
- Trade-off between depth and training time.

6. References

[1] M. A. Nielsen & I. L. Chuang, Quantum Computation and Quantum Information. Cambridge Univ. Press.
[2] M. Zomorodi et al., "Optimal Quantum Circuit Design via Unitary Neural Networks," arXiv:2408.13211 (2024).
[3] P.-A. Absil, R. Mahony, & R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton Univ. Press (2009).
[4] Wisdom et al., Full-Capacity Unitary RNNs, NeurIPS 2016.
[5] Kreutz-Delgado, K. The Complex Gradient Operator and the CR-Calculus. arXiv:0906.4835, 2009.