# The state-of-the-art quantum technology

*The Transformer & its Applications to High Energy Physics Problems*

Abdualazem Fadol Mohammed

October 21, 2024

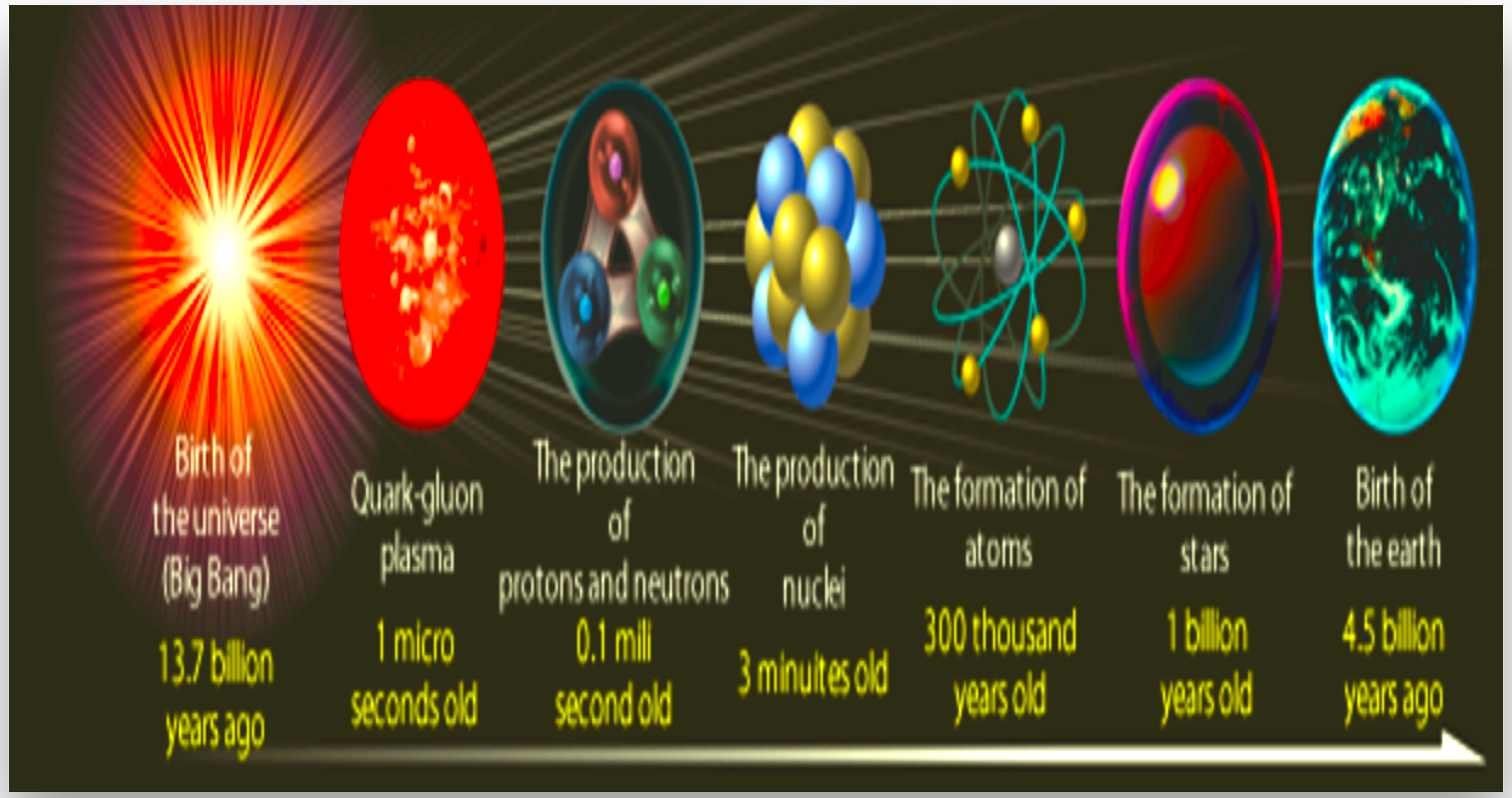
# Overview

- ☐ Introduction
- ☐ Transformer Architecture
- ☐ Hybrid-Quantum Transformer
- ☐ Quantum Transformer
- ☐ Summary

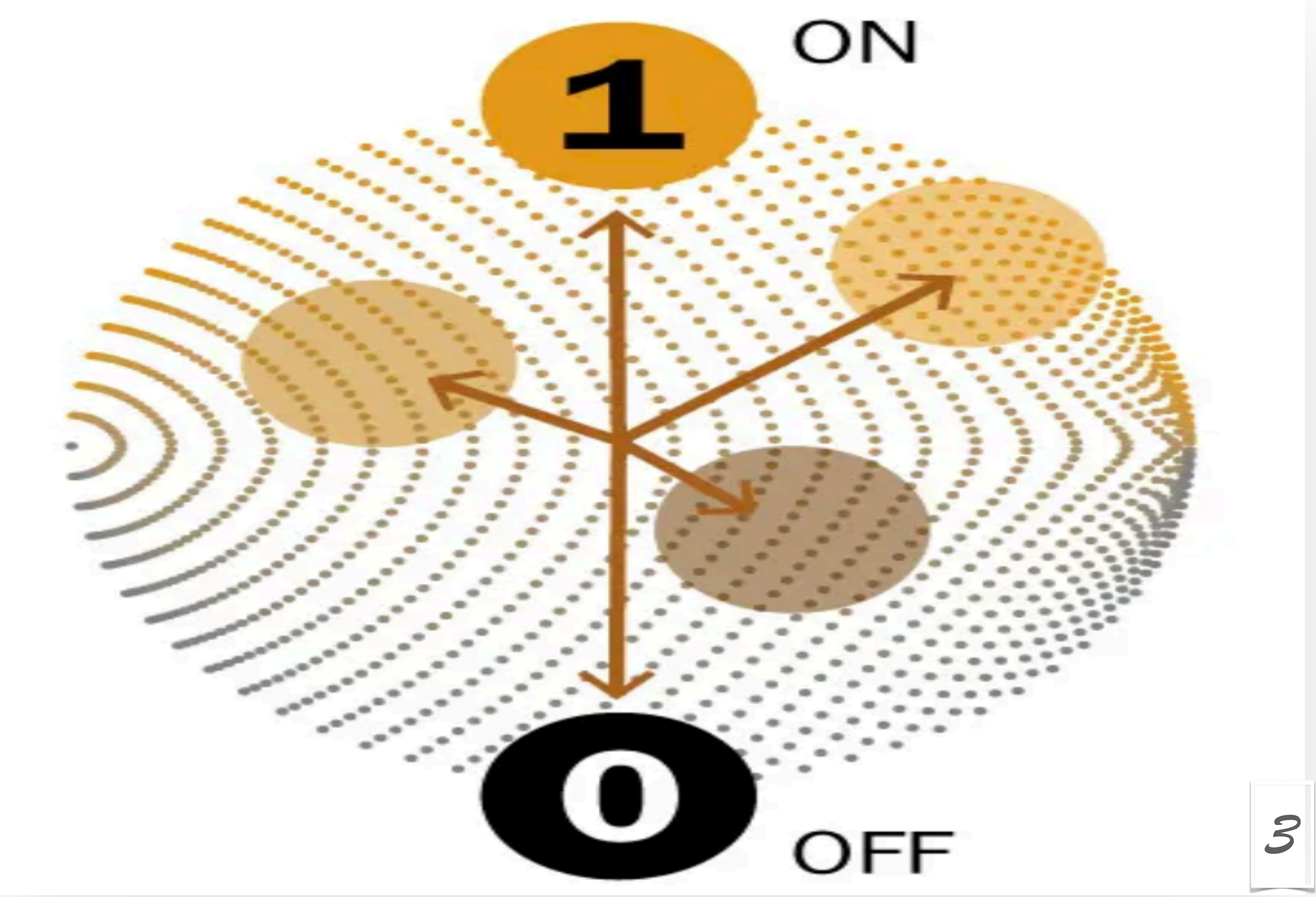

Credited to Thomas Prior for TIME

# Introduction

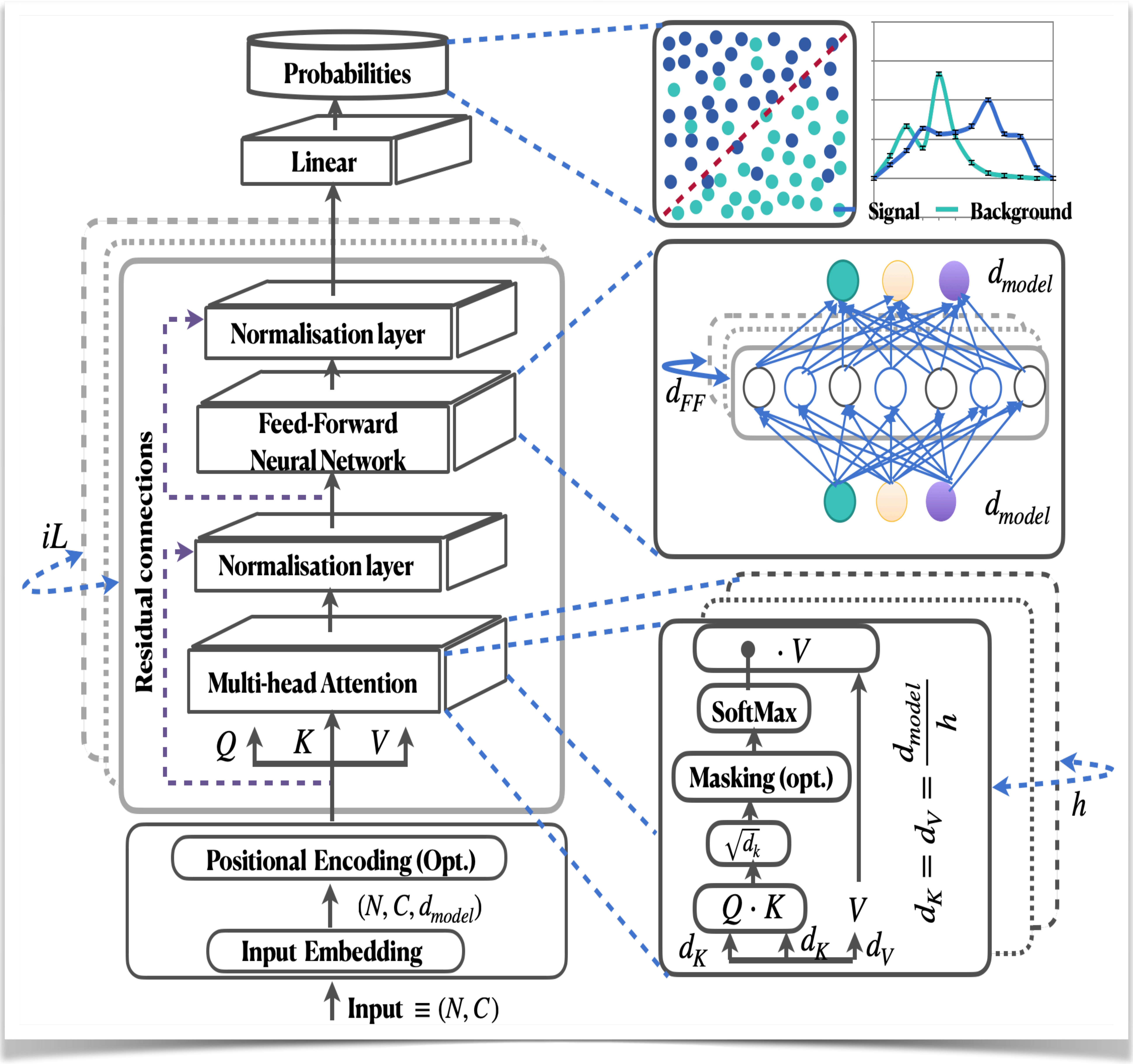
- ☐ **Background & Motivation:**
  - HEP: Understanding fundamental particles and forces.
  - Large datasets, complex computations, advanced simulations.
  - Transformers in HEP: event classification and pattern recognition.




- ☐ **Why Integrate Quantum Computing with Transformers in HEP?**
  - Massive amounts of data require substantial computational power.
  - Quantum-enhanced transformers may offer better performance in processing and analysing HEP data.
  - Integrating cutting-edge quantum technologies.
- ☐ **Goals and Objectives:**
  - Leveraging quantum computing for improved models.
  - Combining classical and quantum computations.
  - Design and evaluate hybrid models & develop quantum algorithms.



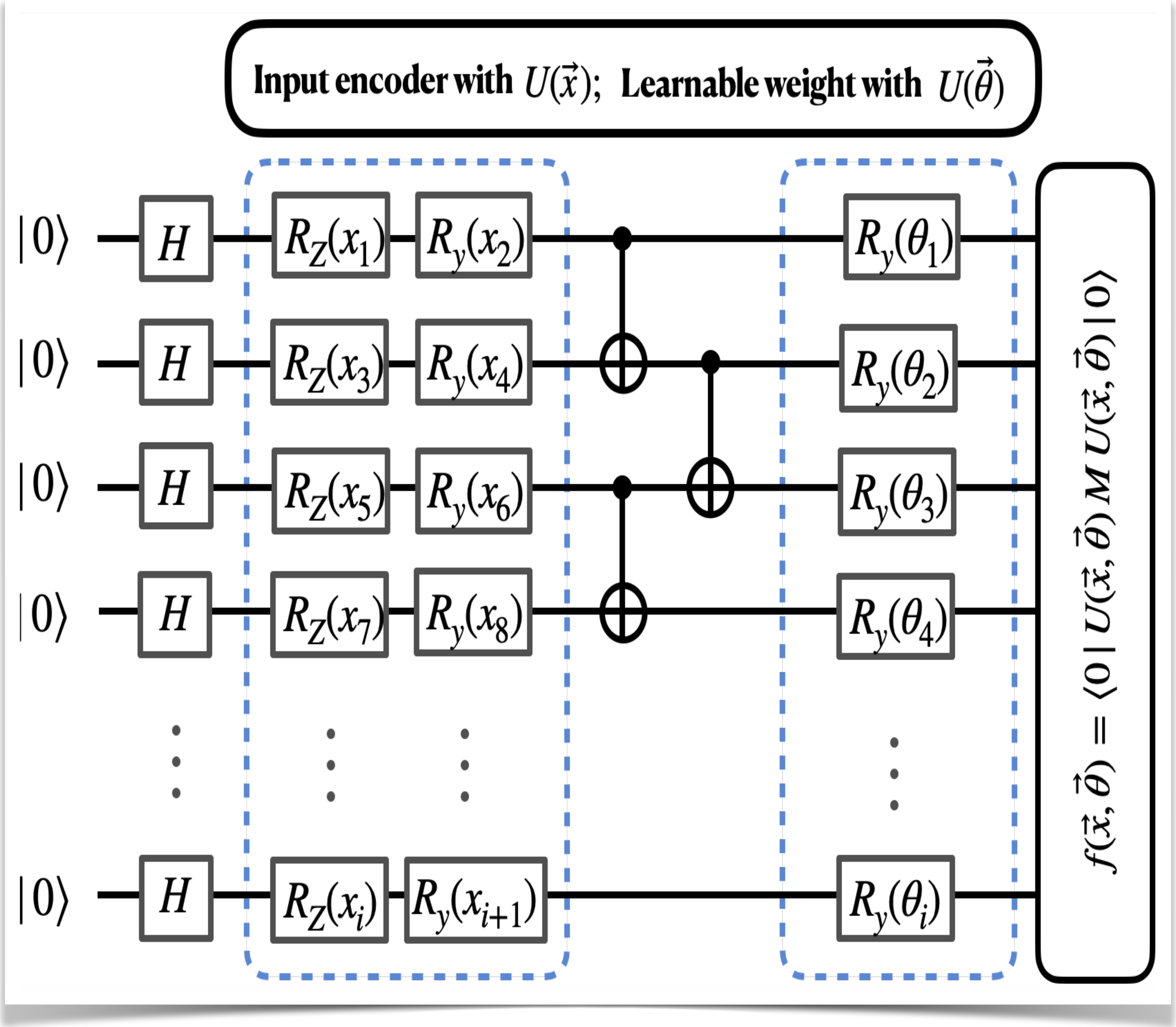
# Classical Transformer




- ☐ Unlike the original transformer,
- ☐ it consists of an encoder only.

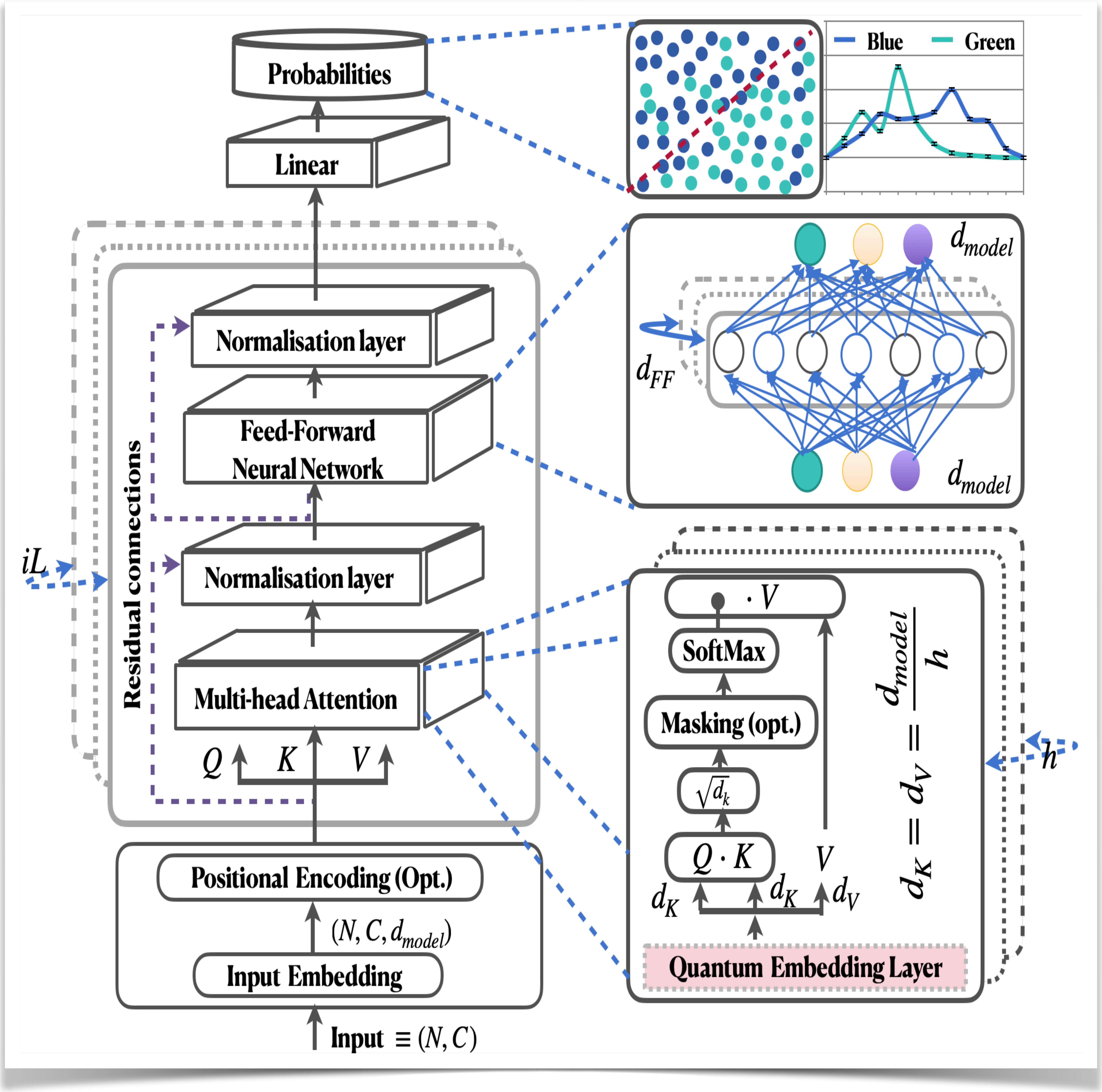
# Hybrid-Quantum Transformer

- **Quantum Embedding Layer:**
  - Variational quantum circuit
  - Inputs are encoded as rotational angles
  - Randmoly initiated weights ($\vec{\theta}$)
- **Measurements:**
  - Counts which is normalised—$2^n$–qubits.
  - Measuring Pauli bases for each qubit.
- The quantum circuit is computed $N$ times.



# Hybrid-Quantum Transformer

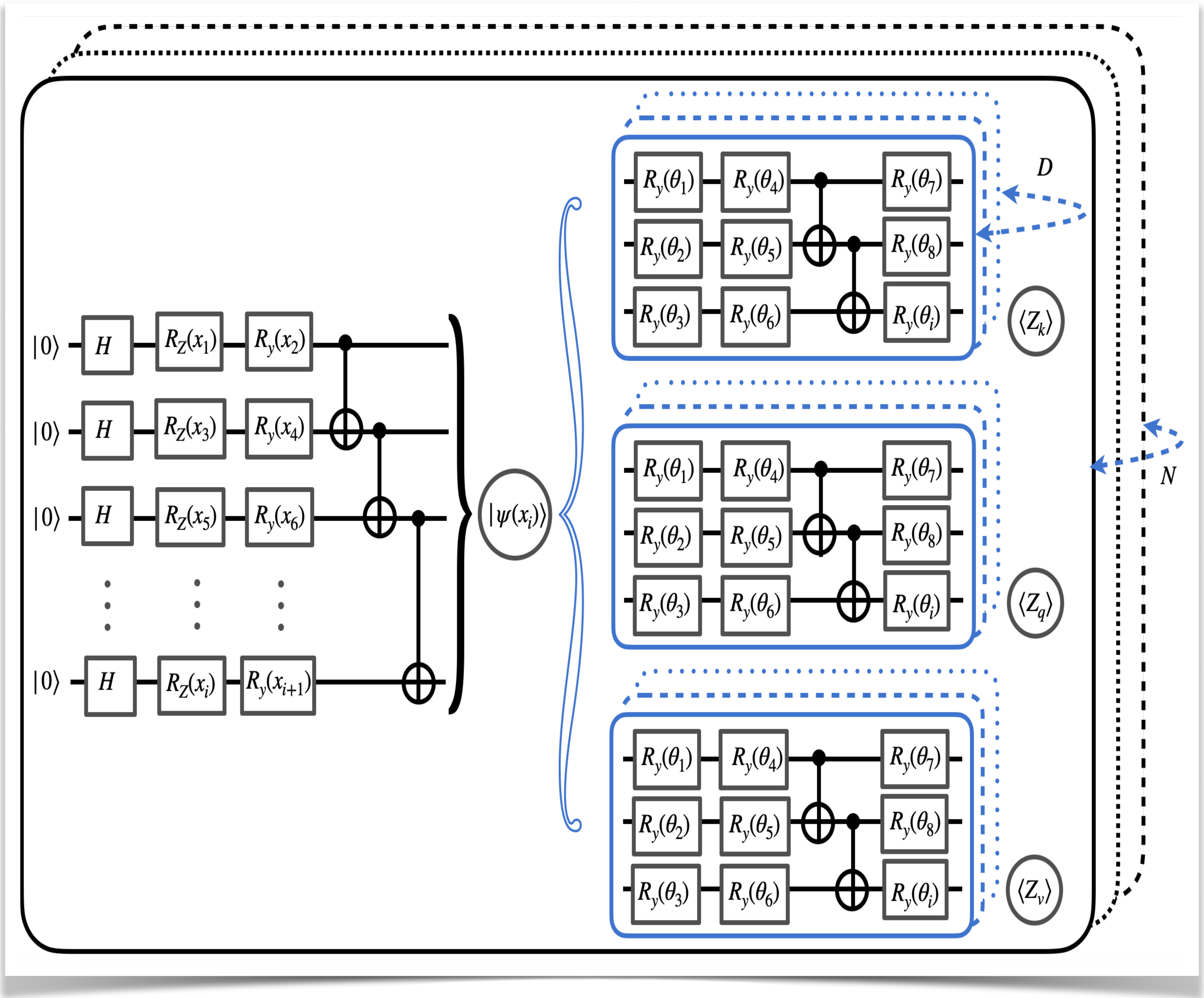
- ☐ It is similar to the classical transformer.
- ☐ We add a quantum embedding layer.
- ☐ It encodes the input dataset.
- ☐ Then, pass them to a linear layer.



# Quantum Transformer

☐ **Quantum self-attention**:

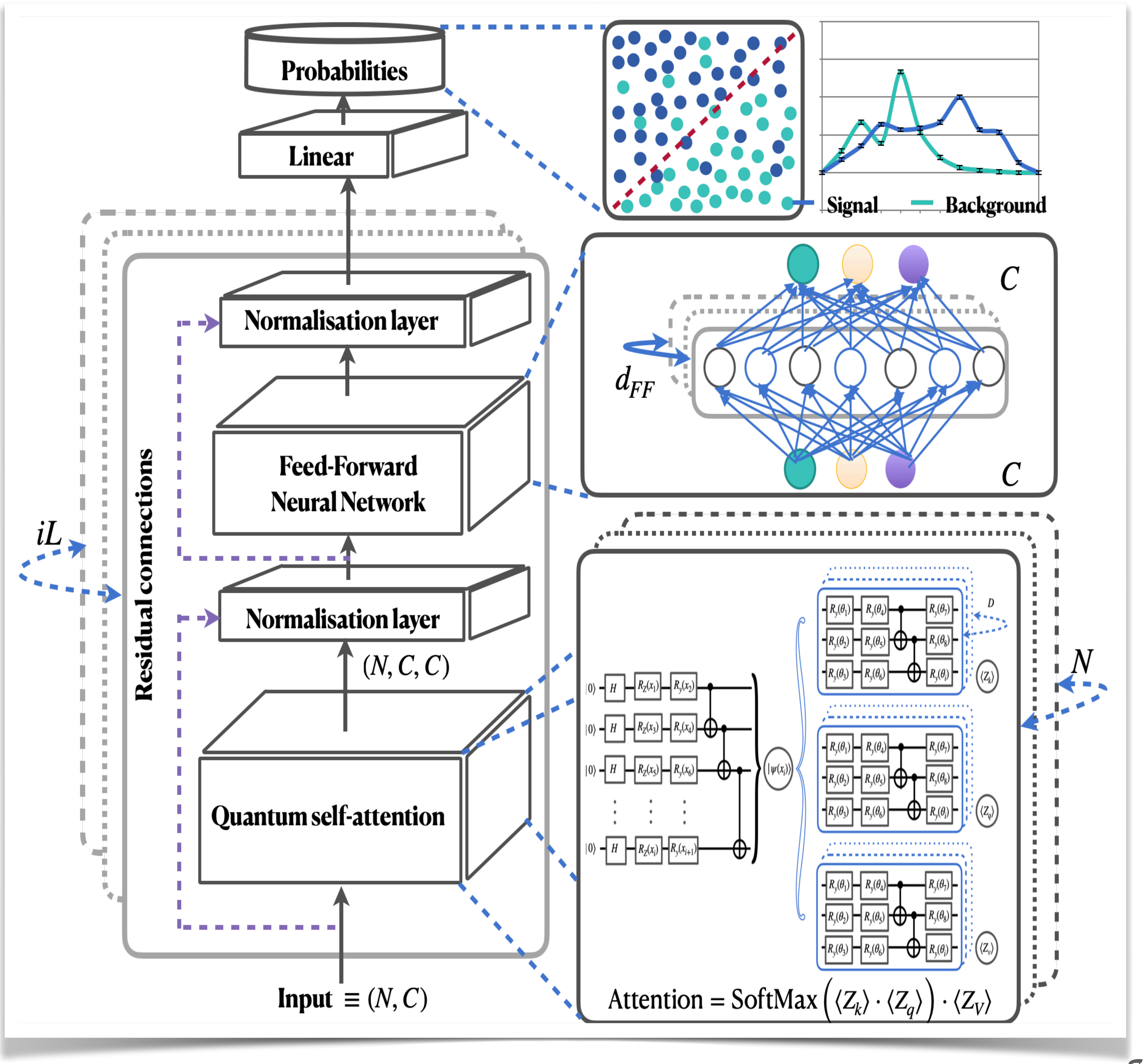
- Initial quantum circuit to encode the input
- Three Ansatzes circuits to get the K, Q, and V
- Each circuit is initialised with a different weight



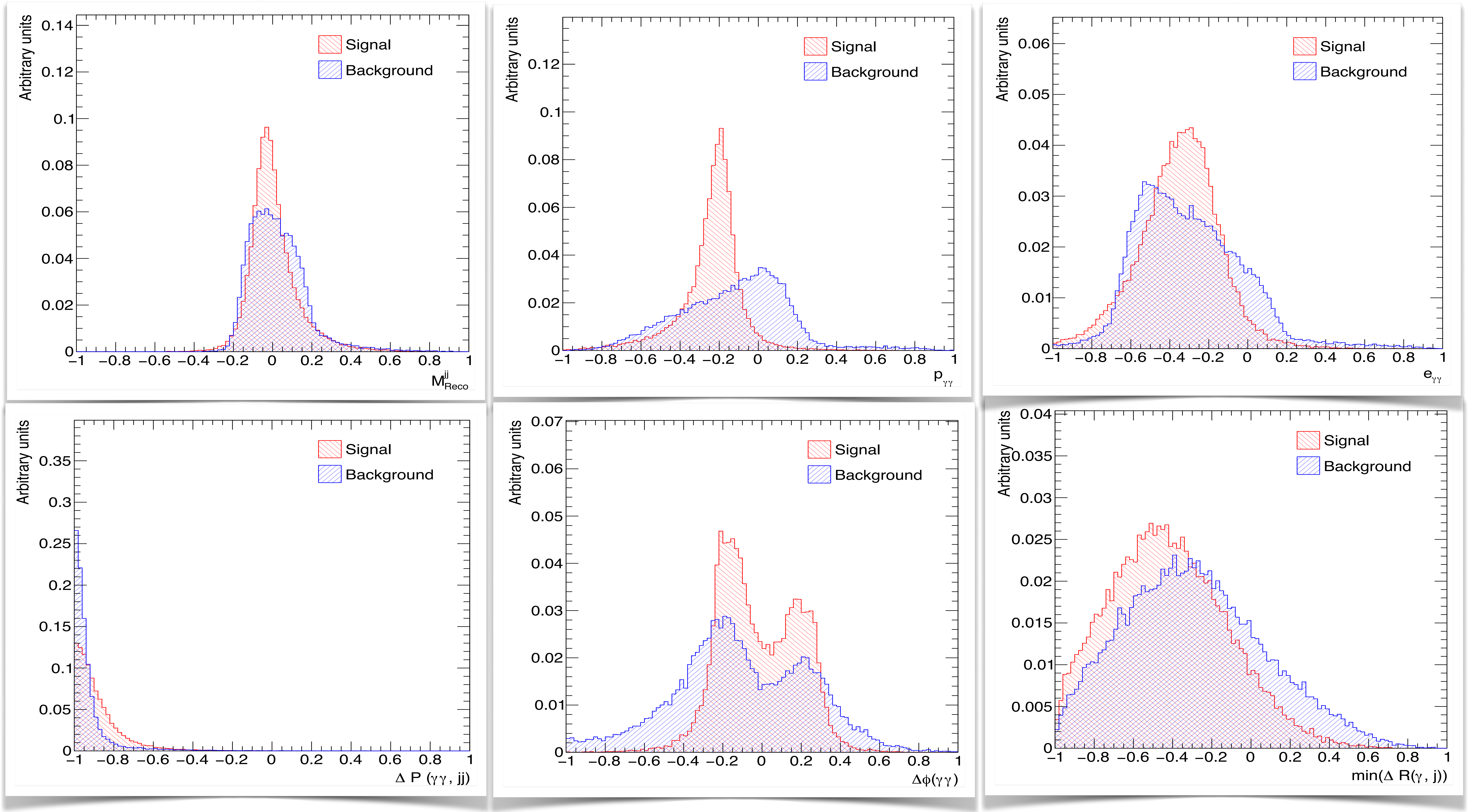

☐ A quantum self-attention inspired by ArXiv: 2205.05625.

# Quantum Transformer

- [ ] Using a quantum self-attention instead of the multi-head attention.



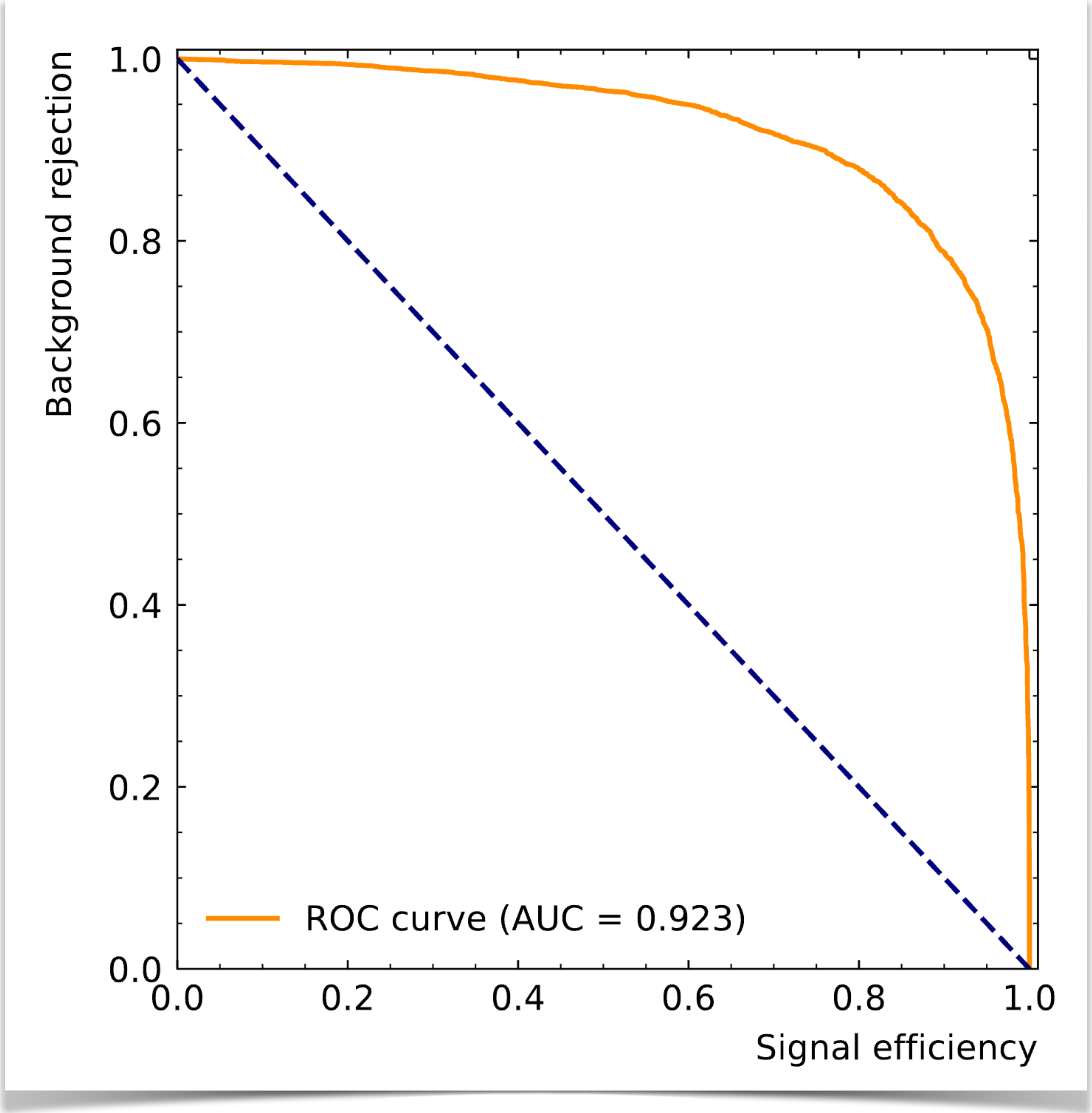
# Datasets: CEPC



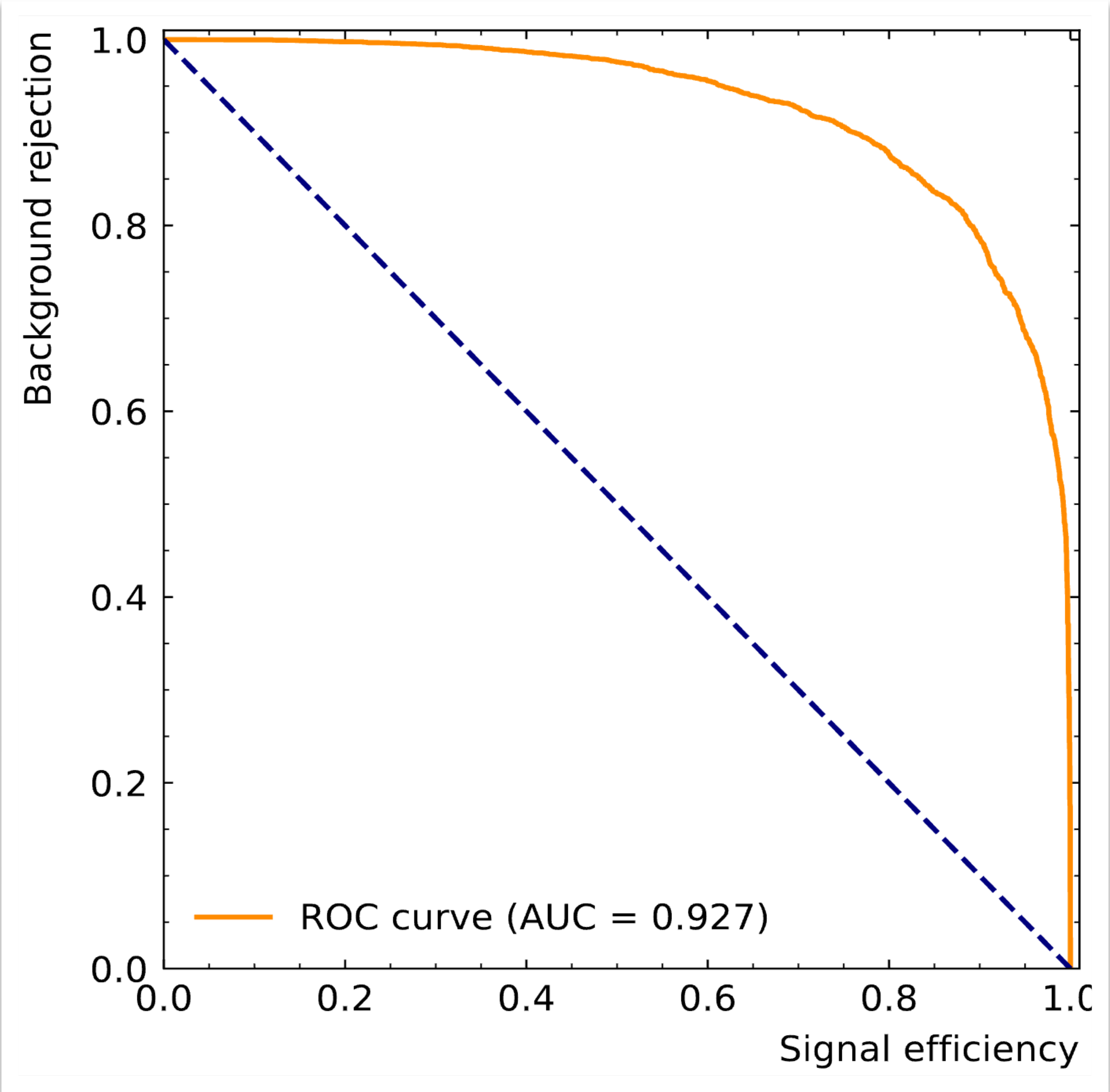

☐ The signal ($e^+e^- \rightarrow ZH \rightarrow \gamma\gamma jj$ )& background ($e^+e^- \rightarrow (Z/\gamma^*)\gamma\gamma$) with 50k events

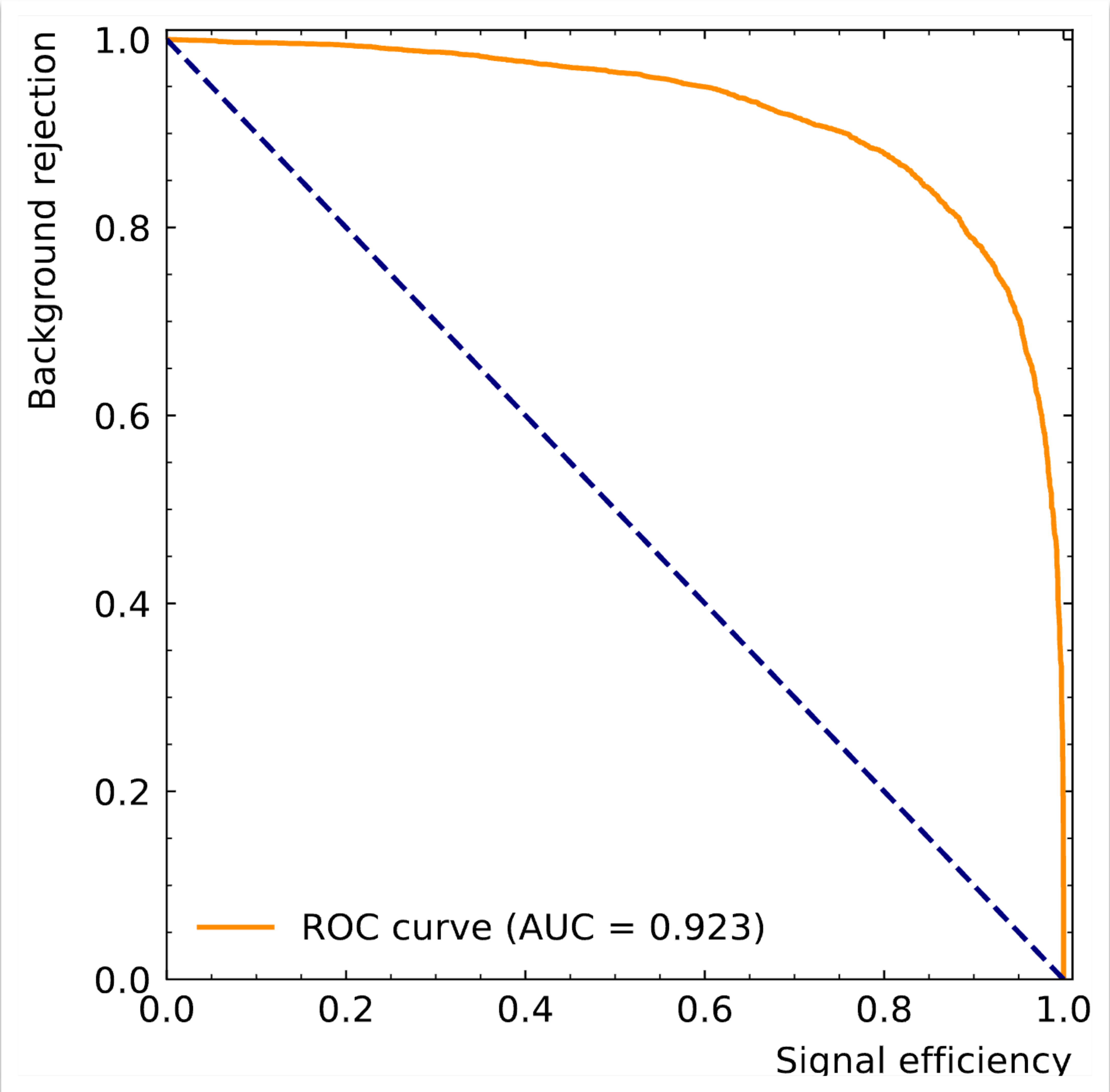
# Hybrid-Quantum Transformer

- Total number of events: 50k.
- Training, validation, and testing: 28k, 12k, and 10k
- Number of variables: six
- L rate & batch: 0.001 & 100
- **Architecture:**
  - $d_{FF} = 10$
  - Dropout = 0.1
  - $iL = 2$
  - $h = 8$
  - Embedded dimension: 512
- Total time for the training and validation: 71h:43m:57s

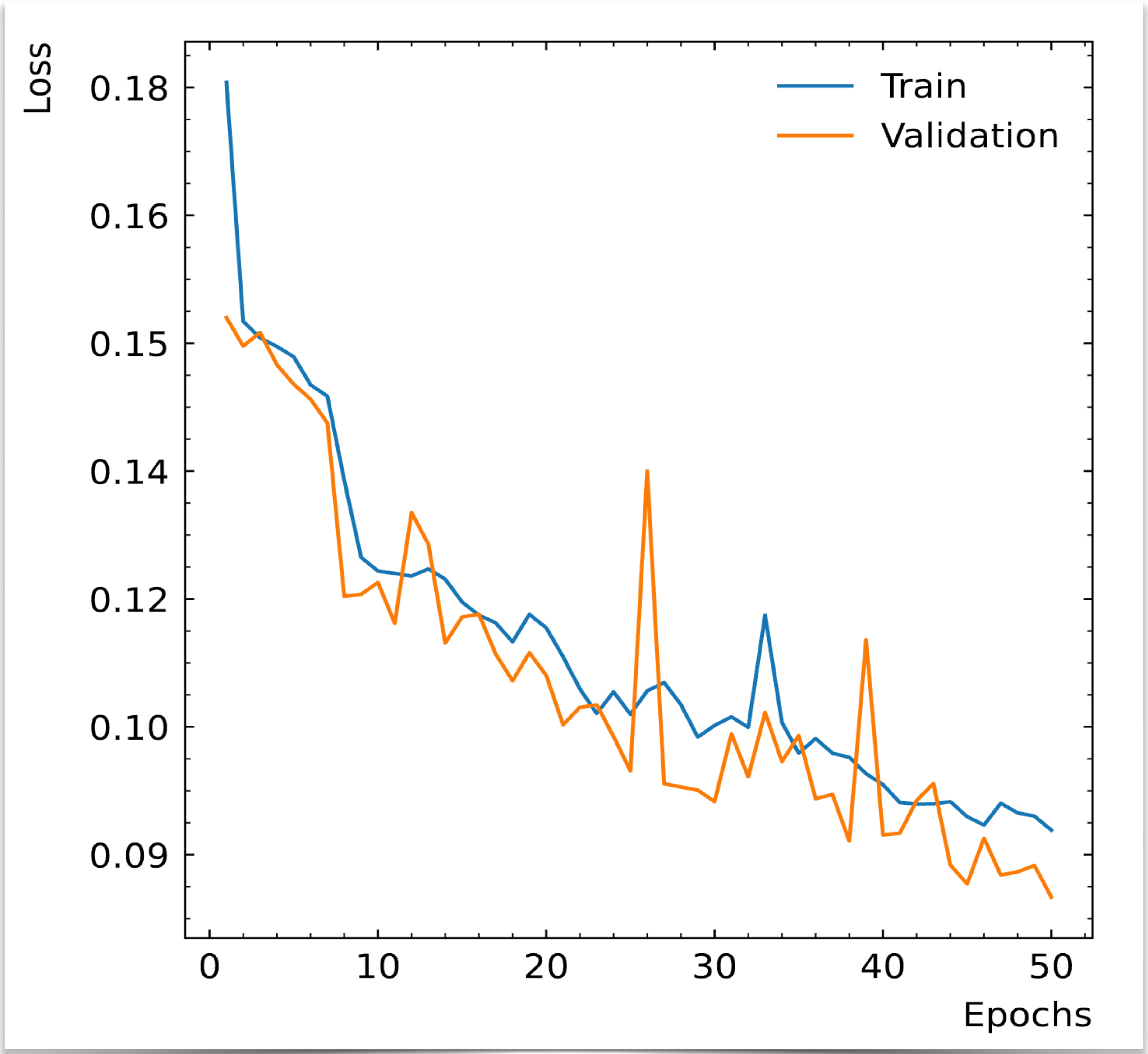


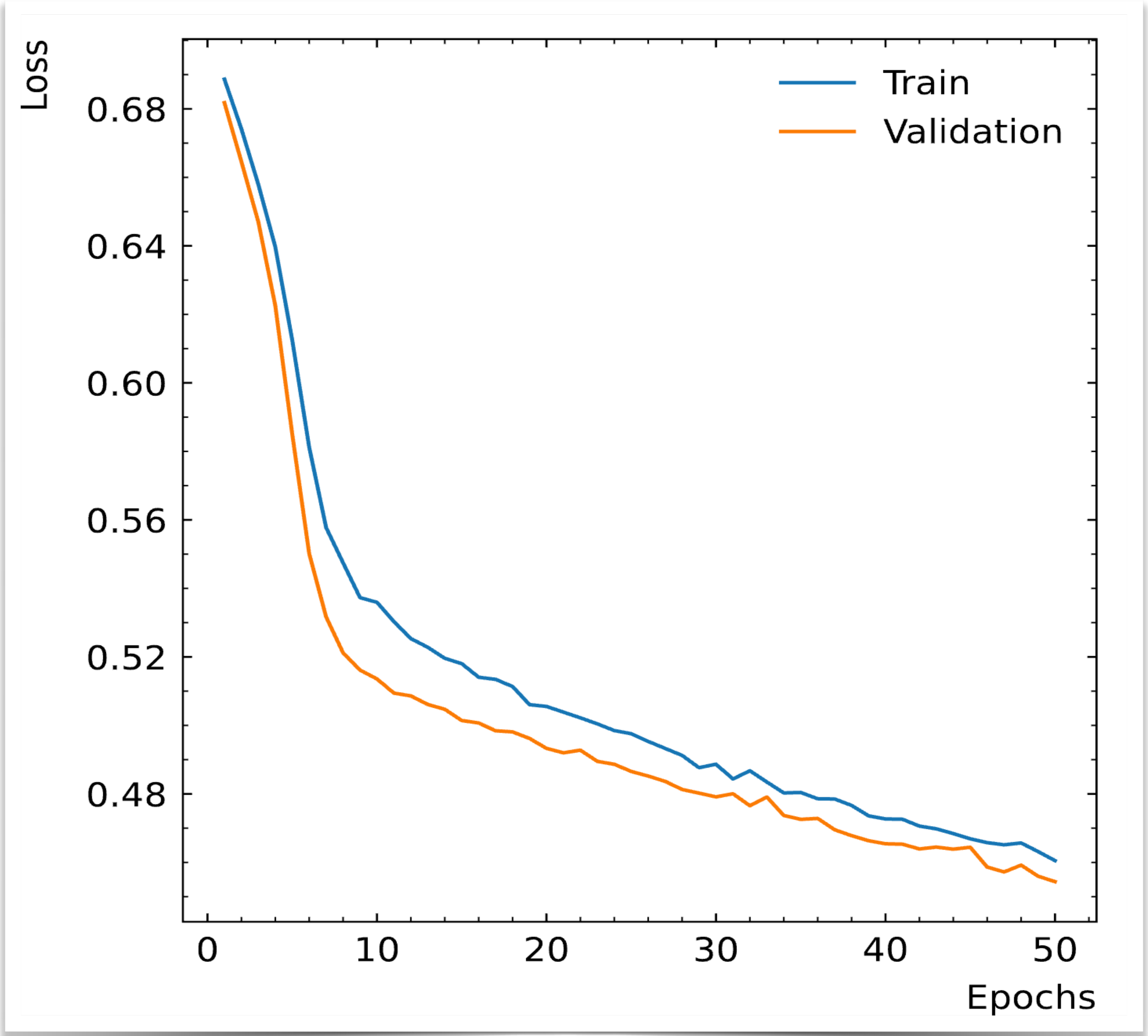
# Transformer & Hybrid-Quantum Transformer



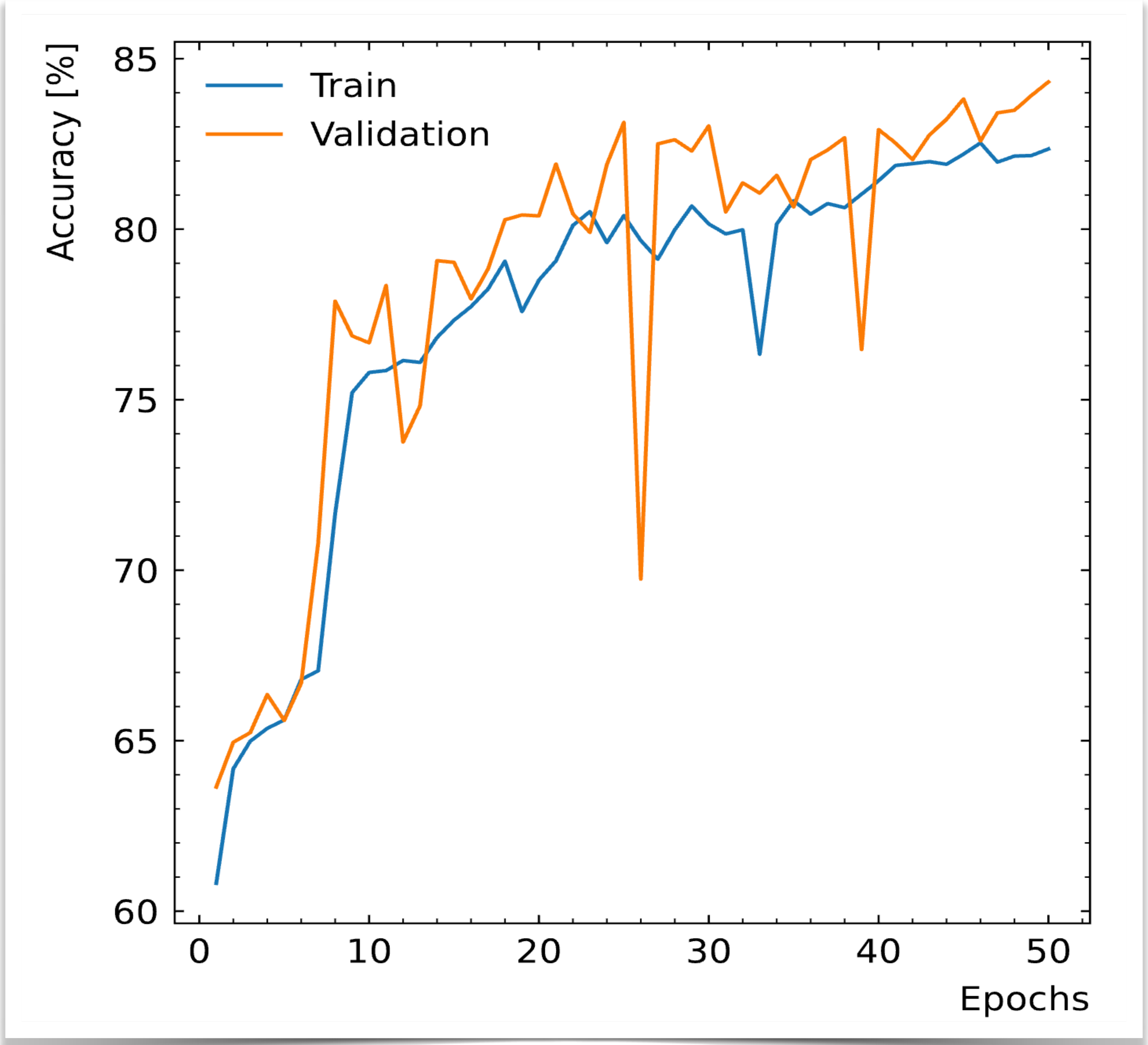



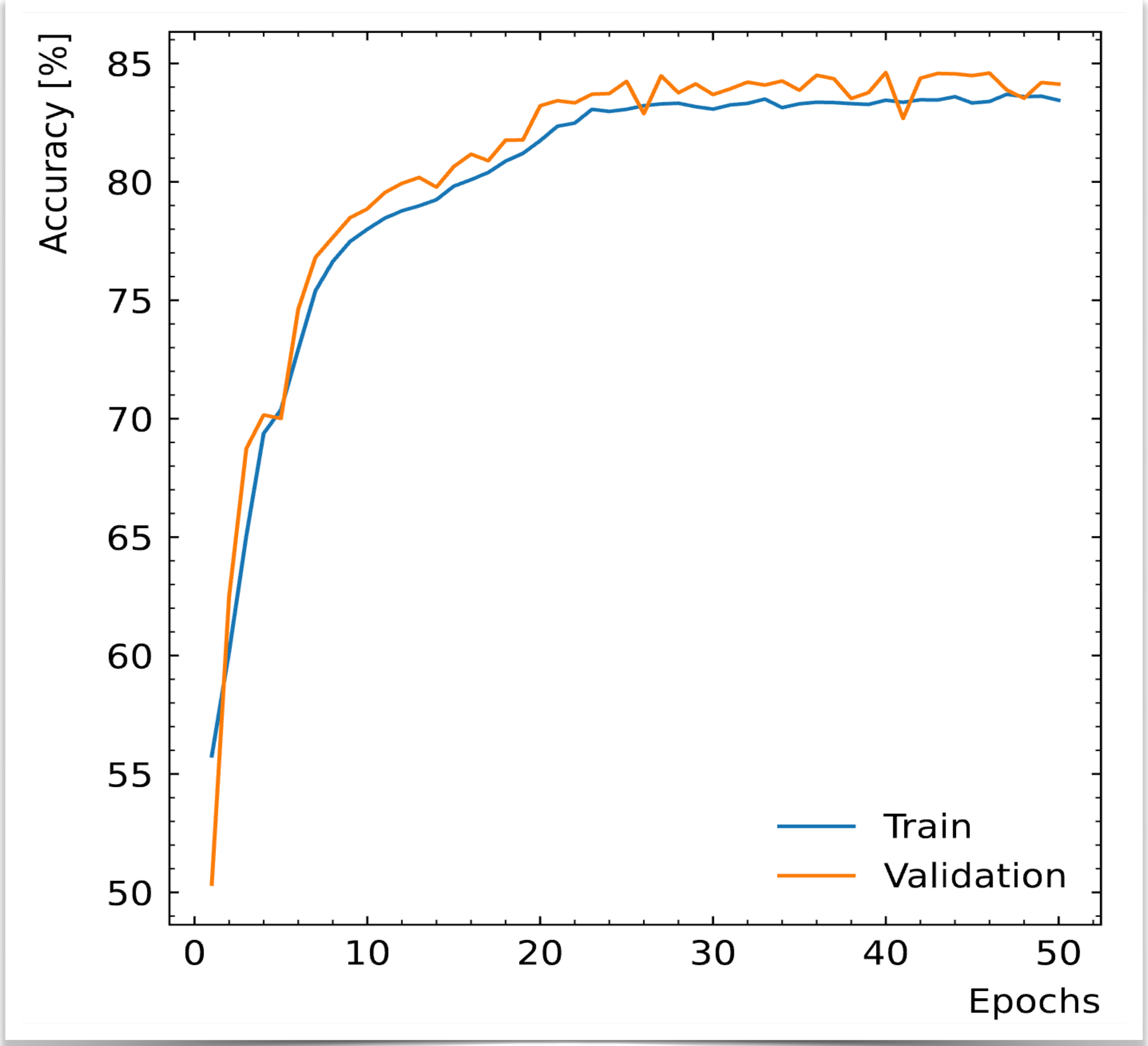
# Transformer & Hybrid-Quantum Transformer



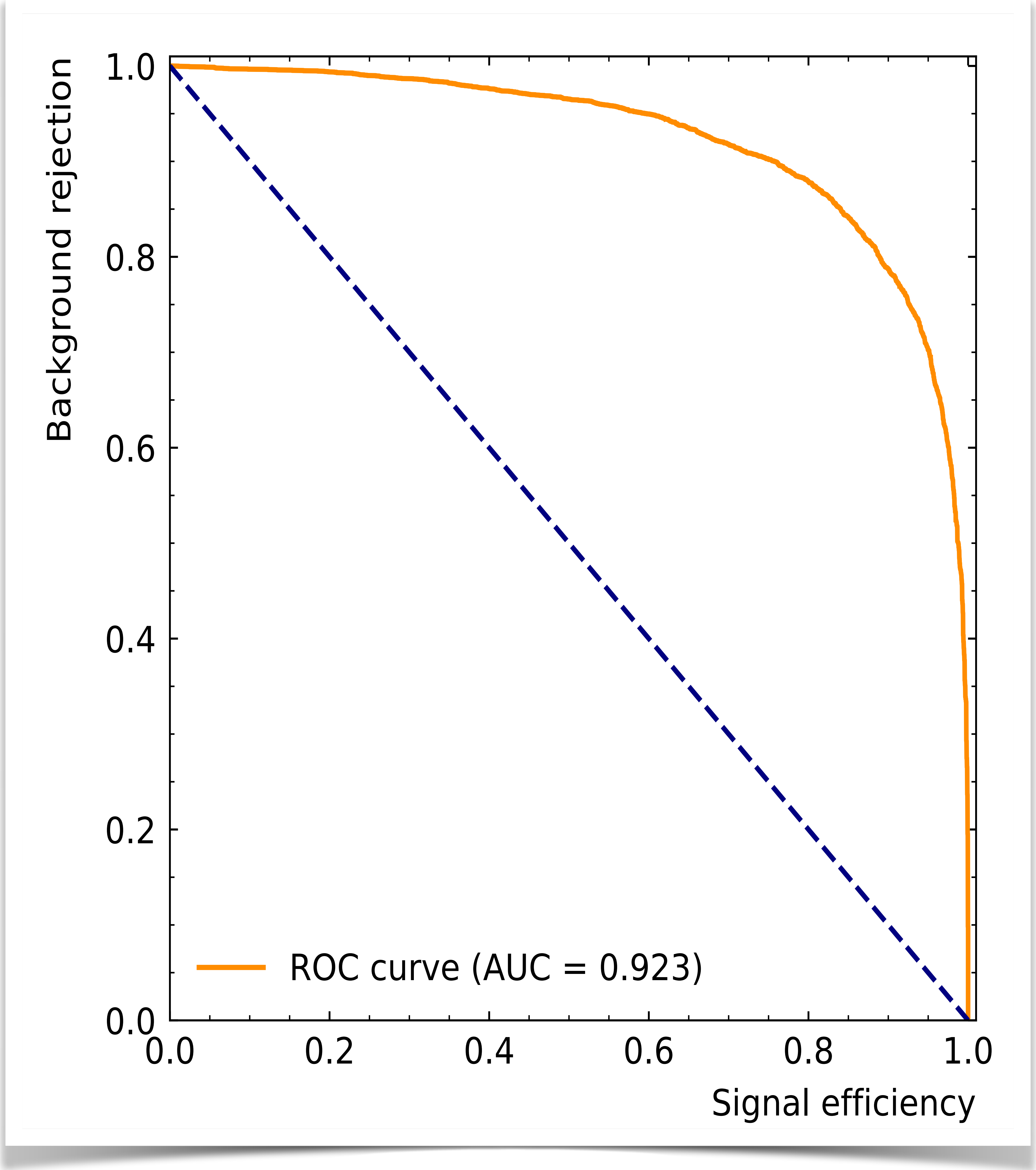
# Transformer & Hybrid-Quantum Transformer



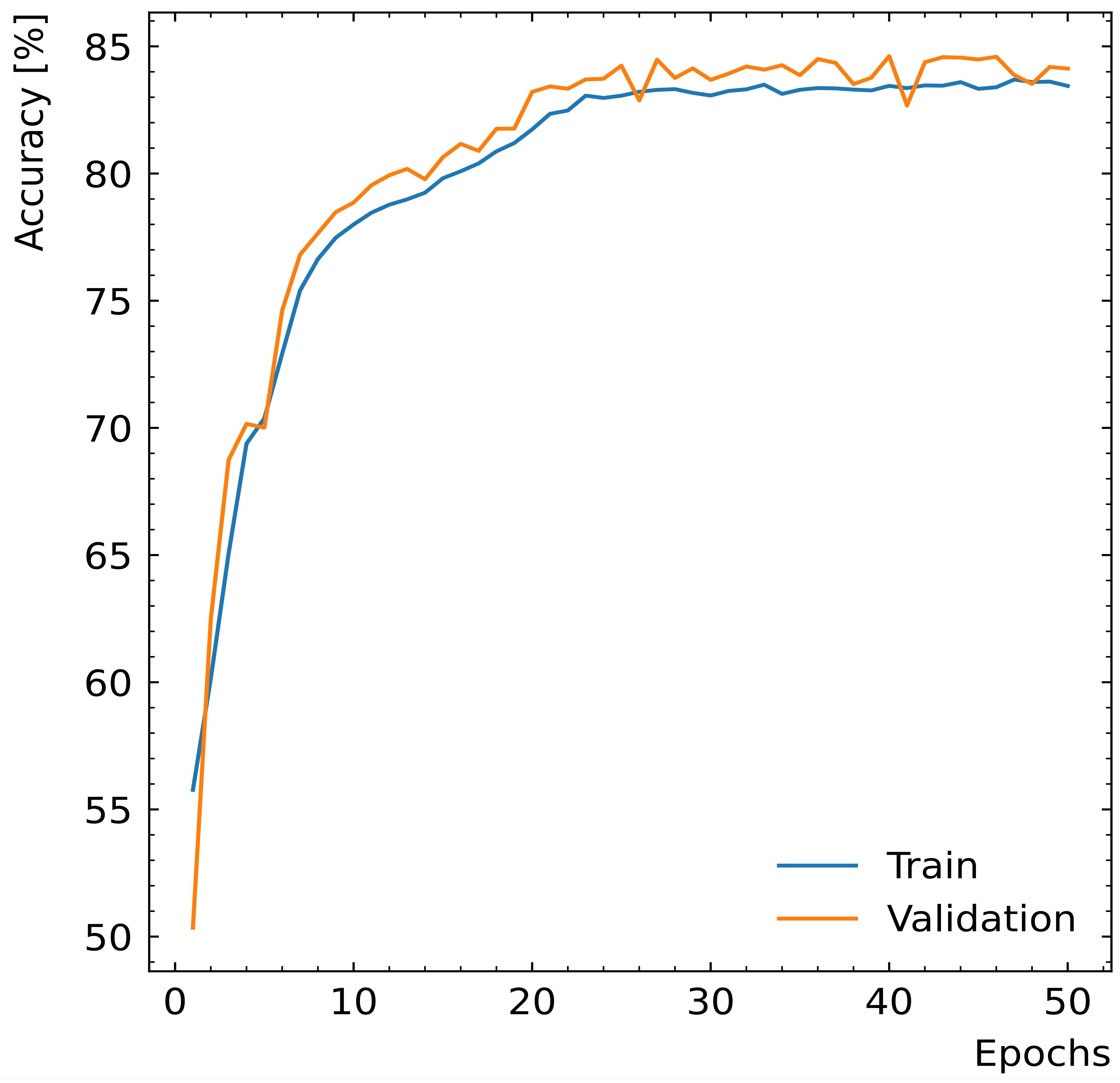



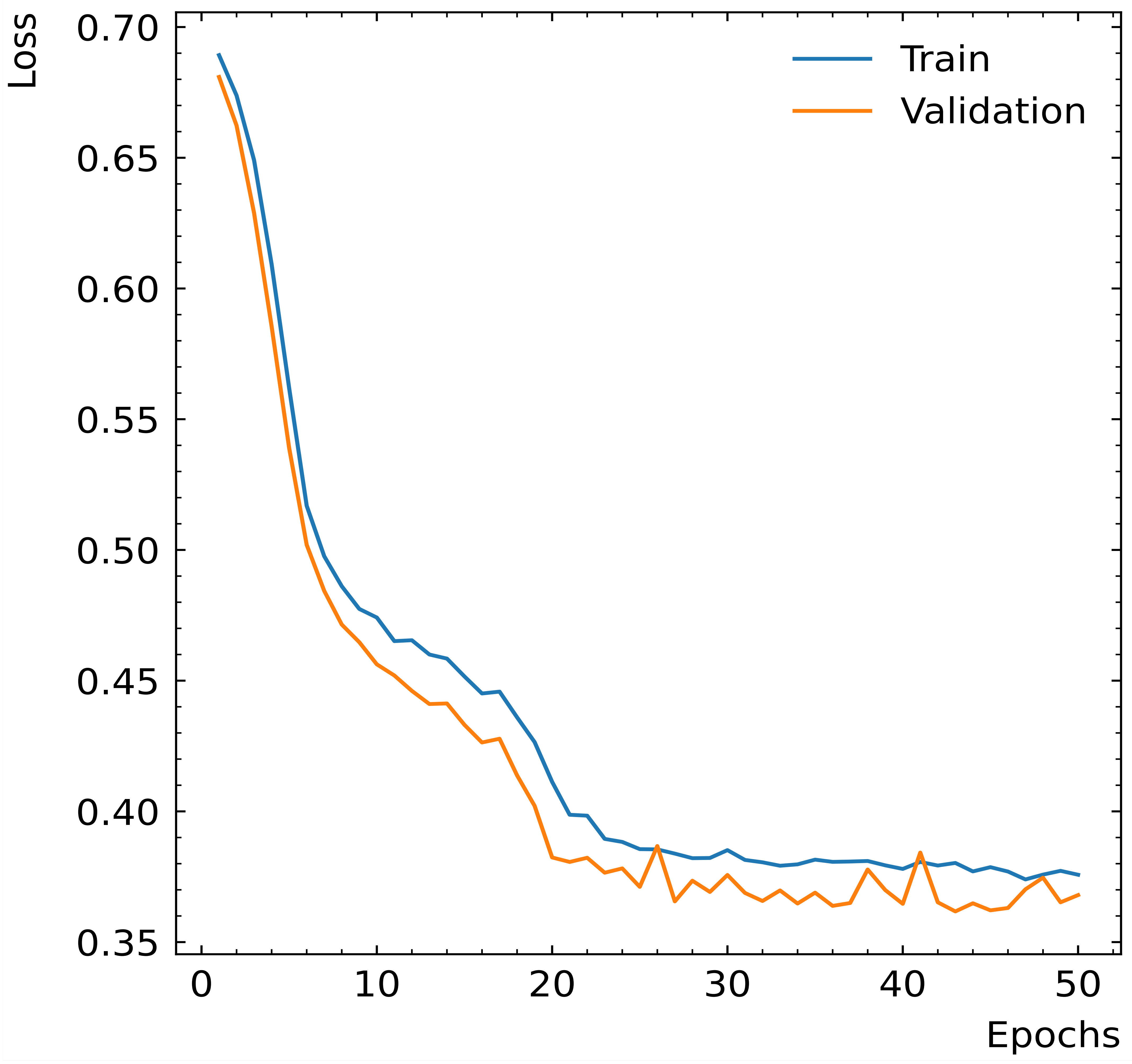
# Quantum Transformer

- Total number of events: 50k.
- Training, validation, and testing: 28k, 12k, and 10k
- Number of variables: six
- L rate & batch: 0.001 & 100
- **Architecture:**
  - $d_{FF} = 10$
  - Dropout = 0.005
  - $iL = 2$
  - $D = 3$
- Total time for the training and validation: 22h:5m:5s



# Quantum Transformer



# Summary

- ☐ Discussed transformers and different types of quantum transformers.
- ☐ Explained the integration of a quantum layer into the transformer architecture.
- ☐ Described the use of variational quantum circuits for enhancing attention mechanisms.
- ☐ Presented the performance comparison between hybrid-quantum and classical transformers.