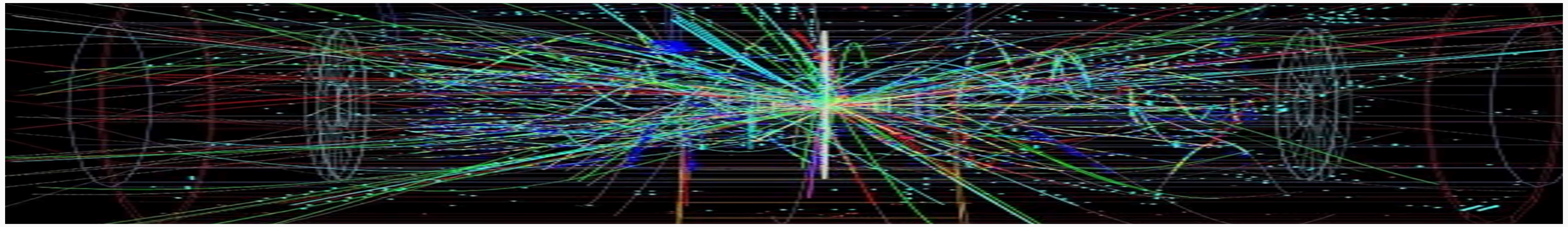
Quantum Computing and Machine Learning Workshop 2024

Integration of quantum computing into transformer architectures for High Energy Physics

Abdualazem Fadol Mohammed
IHEP Quantum Computing Group



August 8, 2024

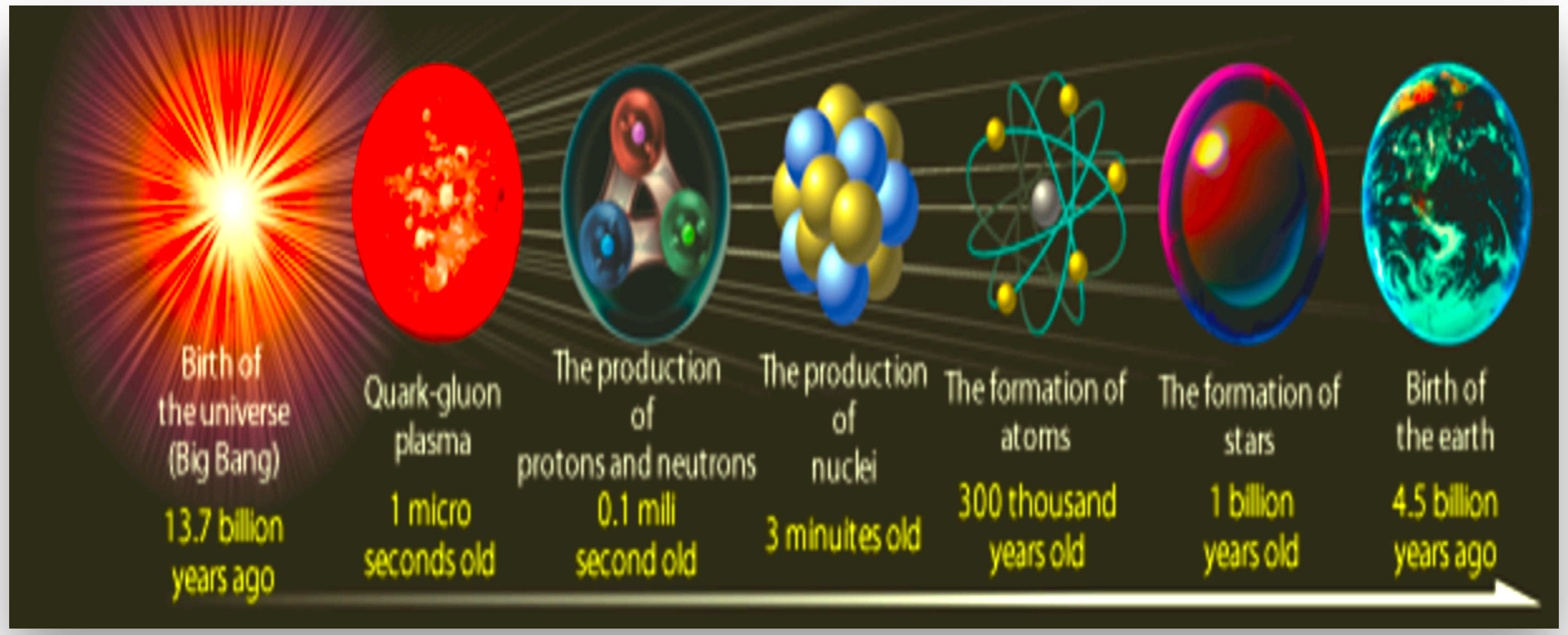
Overview

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

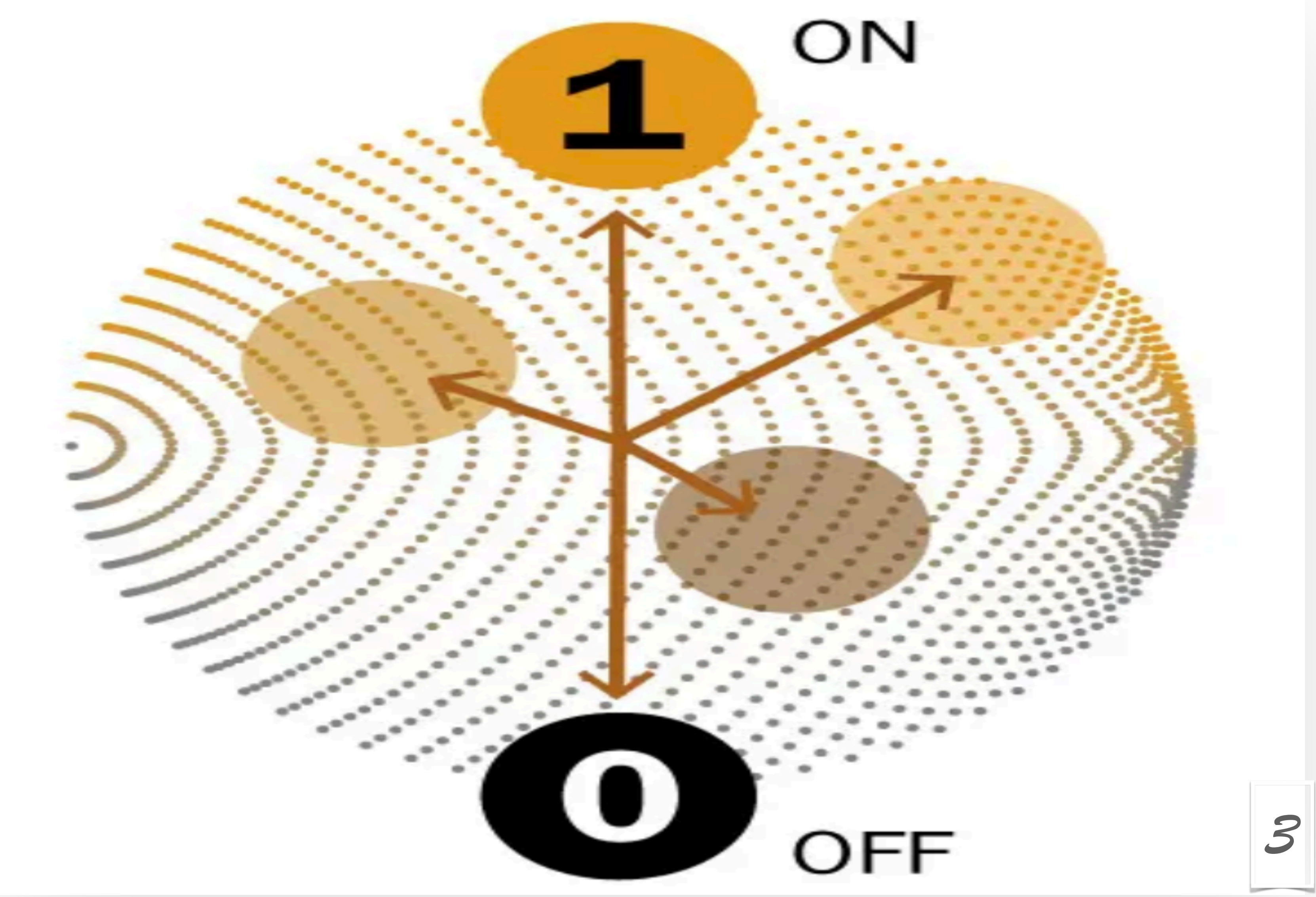
Credited to Thomas Prior for TIME

Introduction

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

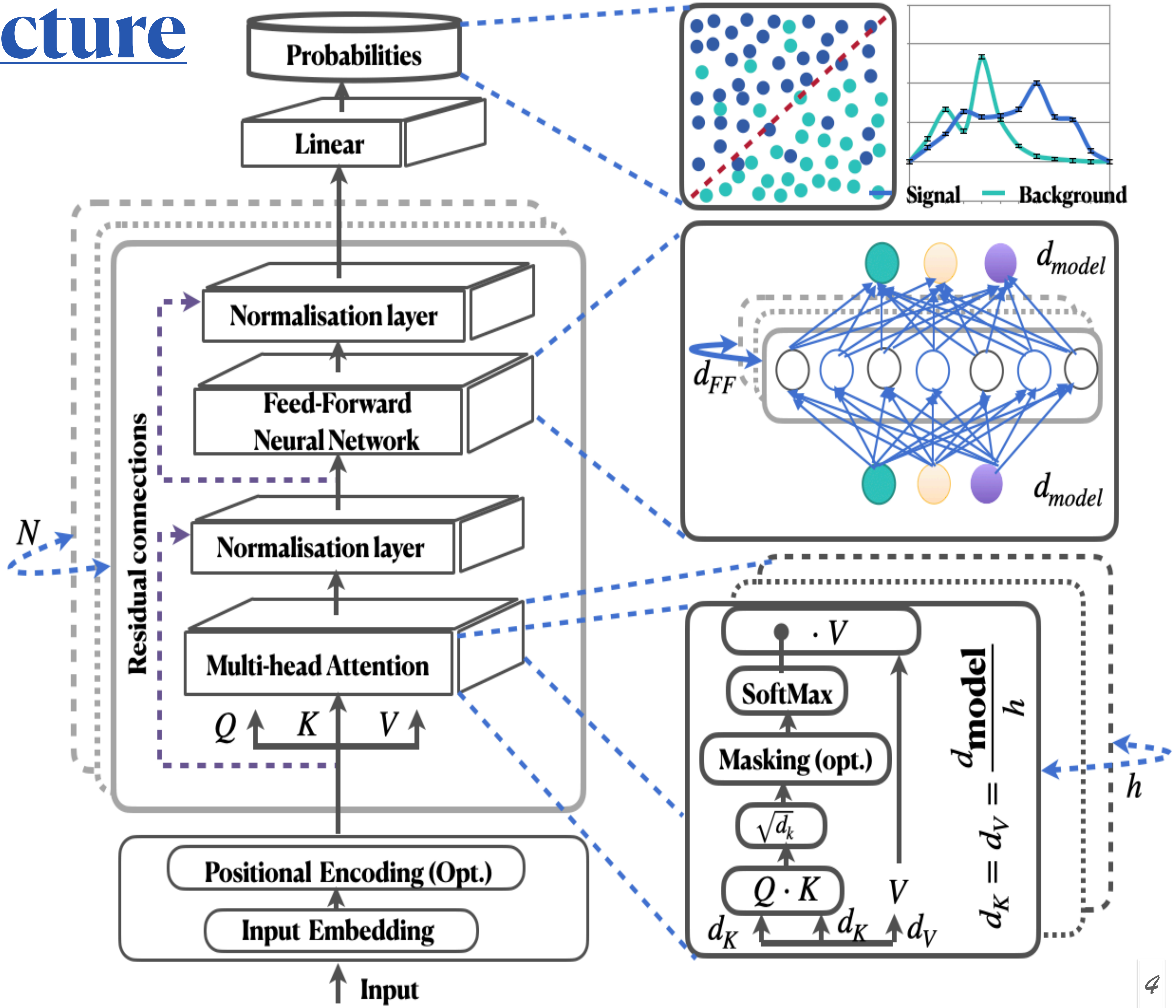


- Why Integrate Quantum Computing with Transformers in HEP?
  - Massive amounts of data require substantial computational power.
  - Quantum-enhanced transformers may offer better performance in processing and analysing HEP data.
  - Integrating cutting-edge quantum technologies.
- Goals and Objectives: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



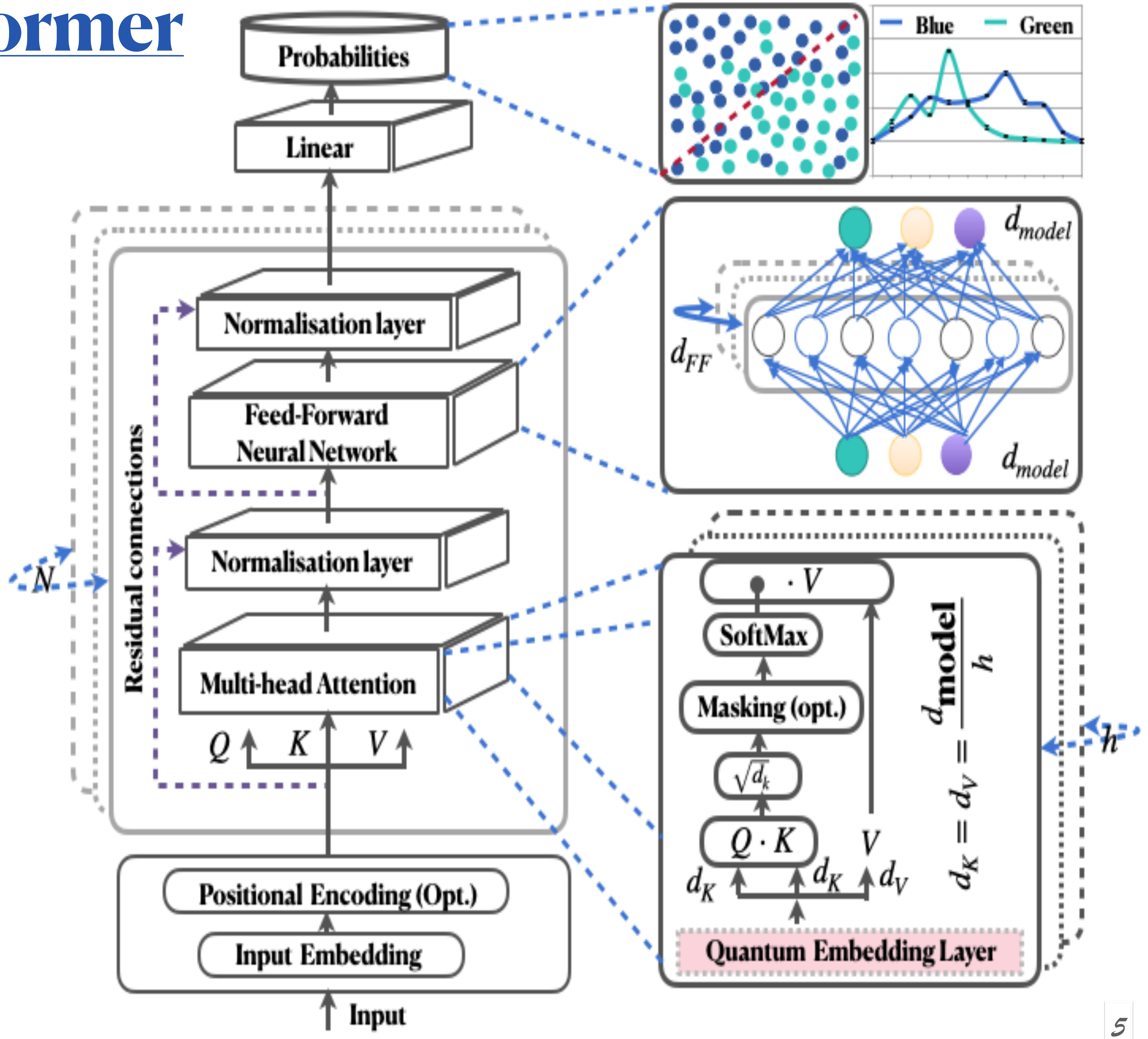
Transformer Architecture

- ☐ It consists of the encoder only.



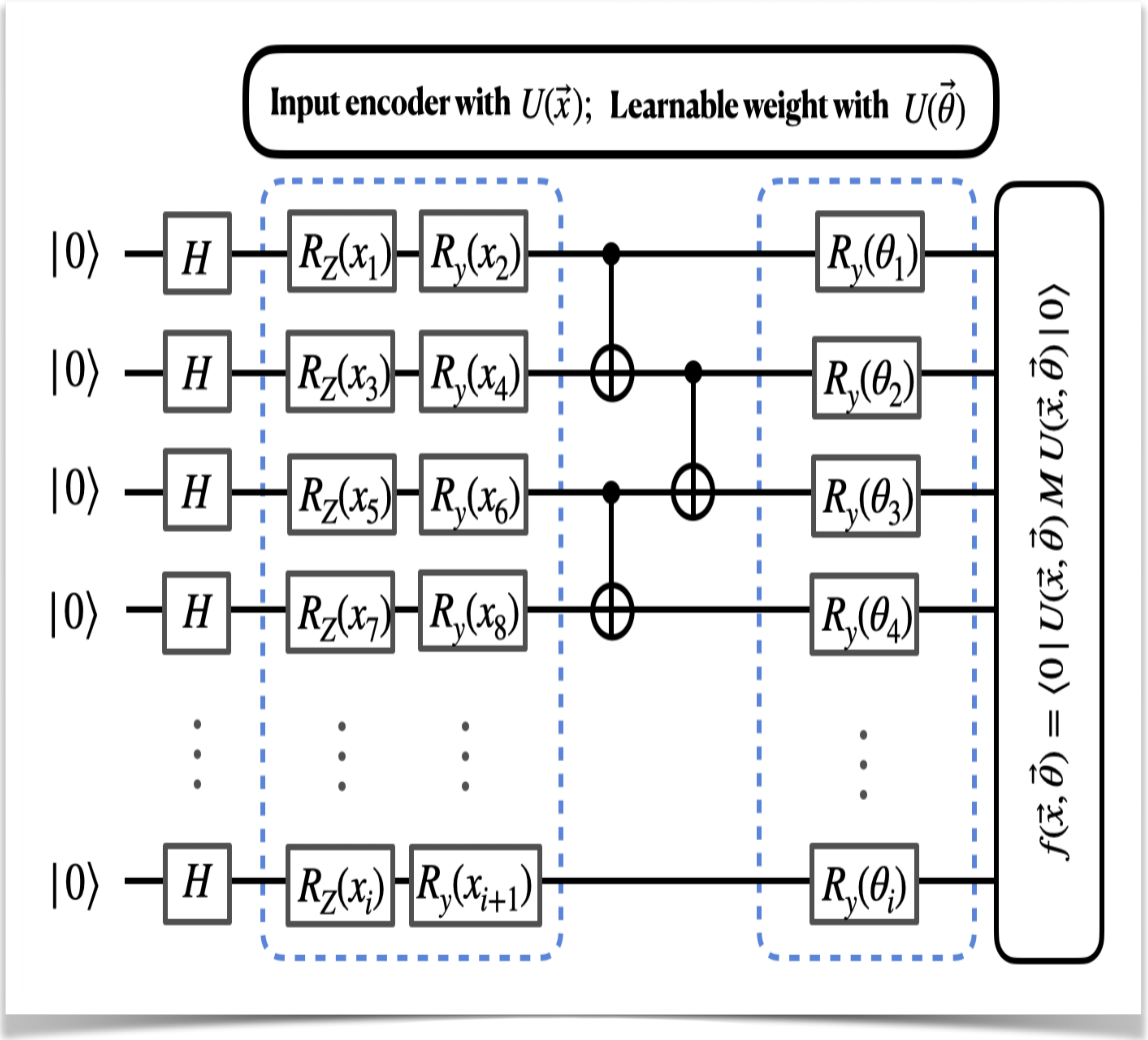
Hybrid-Quantum Transformer

- ☐ We add a quantum layer to the attention block.
- ☐ It encodes the data and then passes them to a linear layer to produce the attention components.

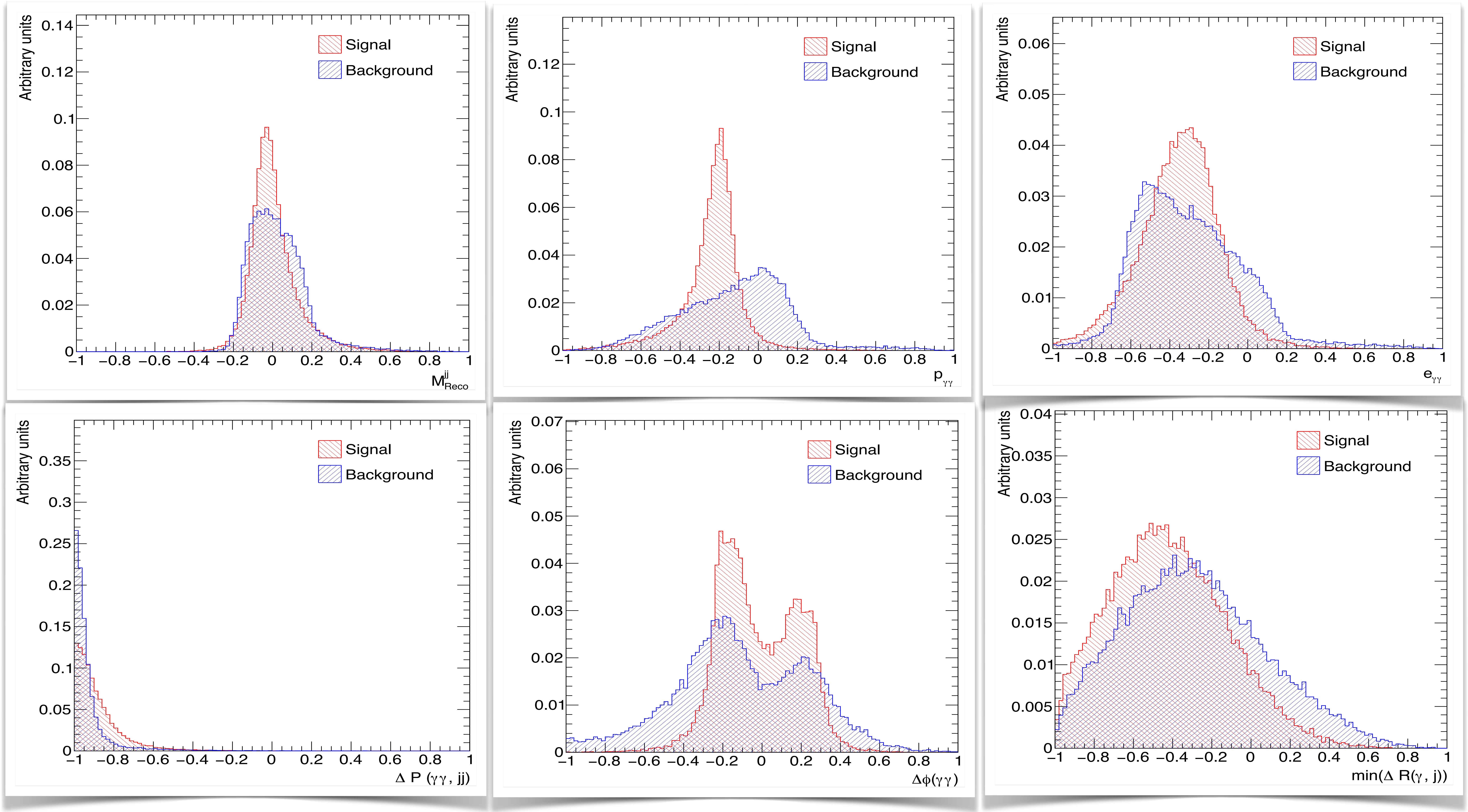


Quantum embedding layer

- ☐ Use a variational quantum circuit:
  - Qubit numbers depend on the variables ($\vec{x}$)
  - Randomly initiated weights of $\vec{\theta}$
  - Weighted average of measurement results
- ☐ Assume that we have a quantum circuit with three qubits:
  - (N=100, C=6) → (N=100, $2^{\text{n-qubits}}$= 8)
  - Use PCA to reduce the dimension to (N=100, C=6)
  - Normalise the outcomes by the measurements
- ☐ The quantum circuit is computed N times.



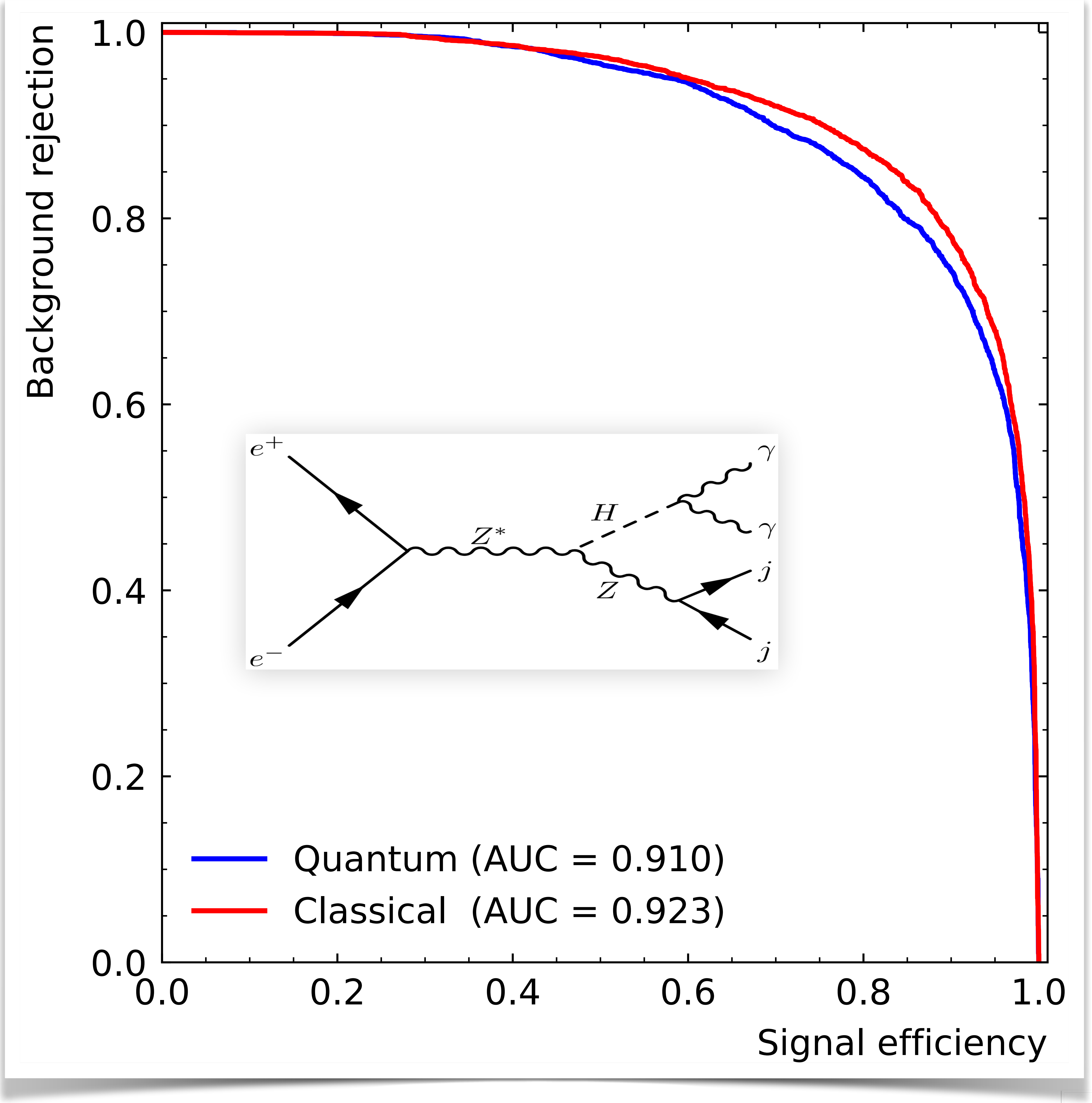
Datasets



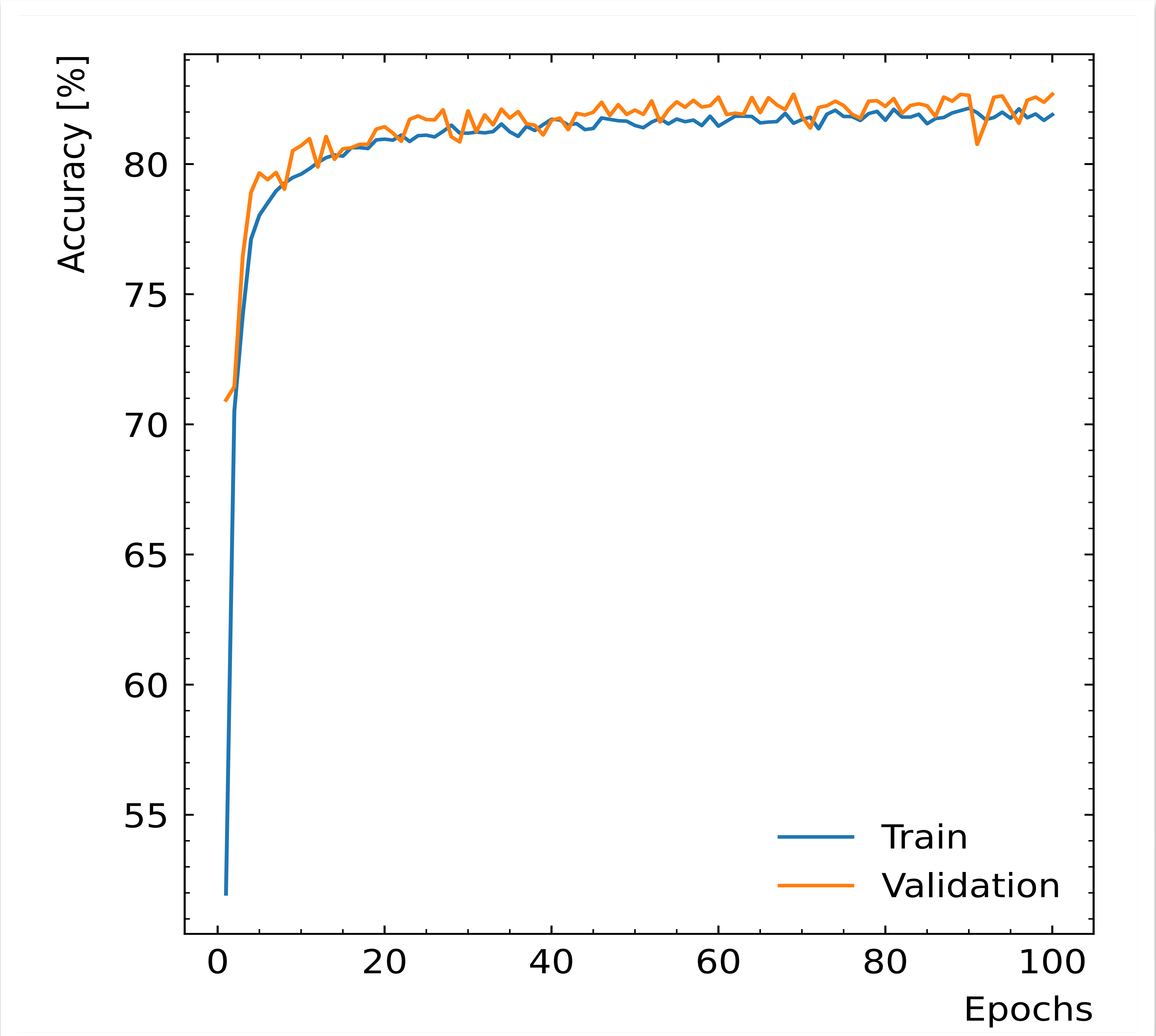
☐ The signal ($e^+e^- \rightarrow ZH \rightarrow \gamma\gamma jj$)& background ($e^+e^- \rightarrow (Z/\gamma^*)\gamma\gamma$) with 50k events

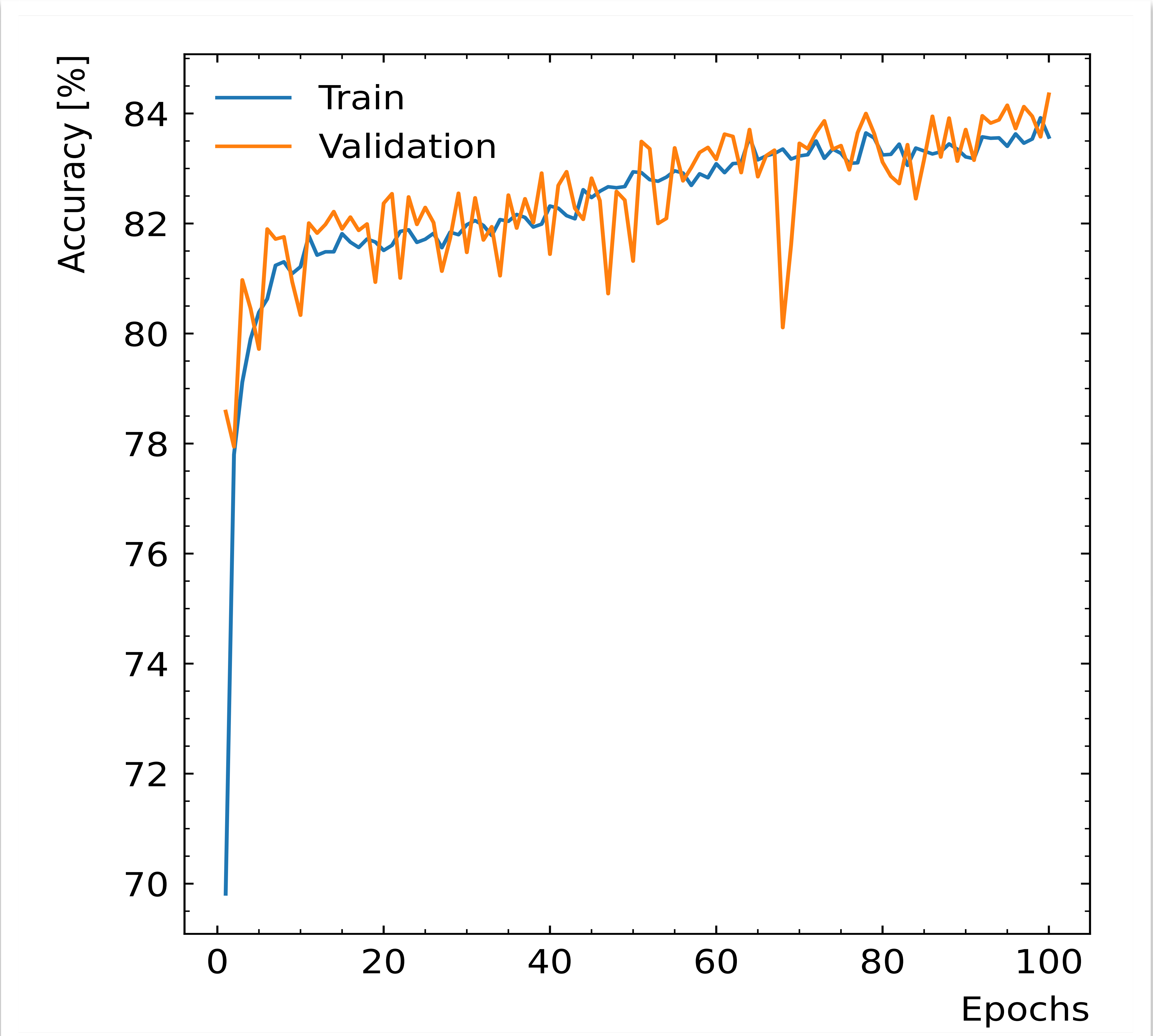
Comparison: ROC curve

- The hybrid-quantum transformer vs the transformer.
- The hybrid-quantum transformer is not fully optimised.
- Possible optimisation would be:
  - Quantum circuit: SPSA and COBYLA
  - Transformer parameters: h, N, d_{model} ...
- A quick scan of the hyperparameters of the transformer model was used.



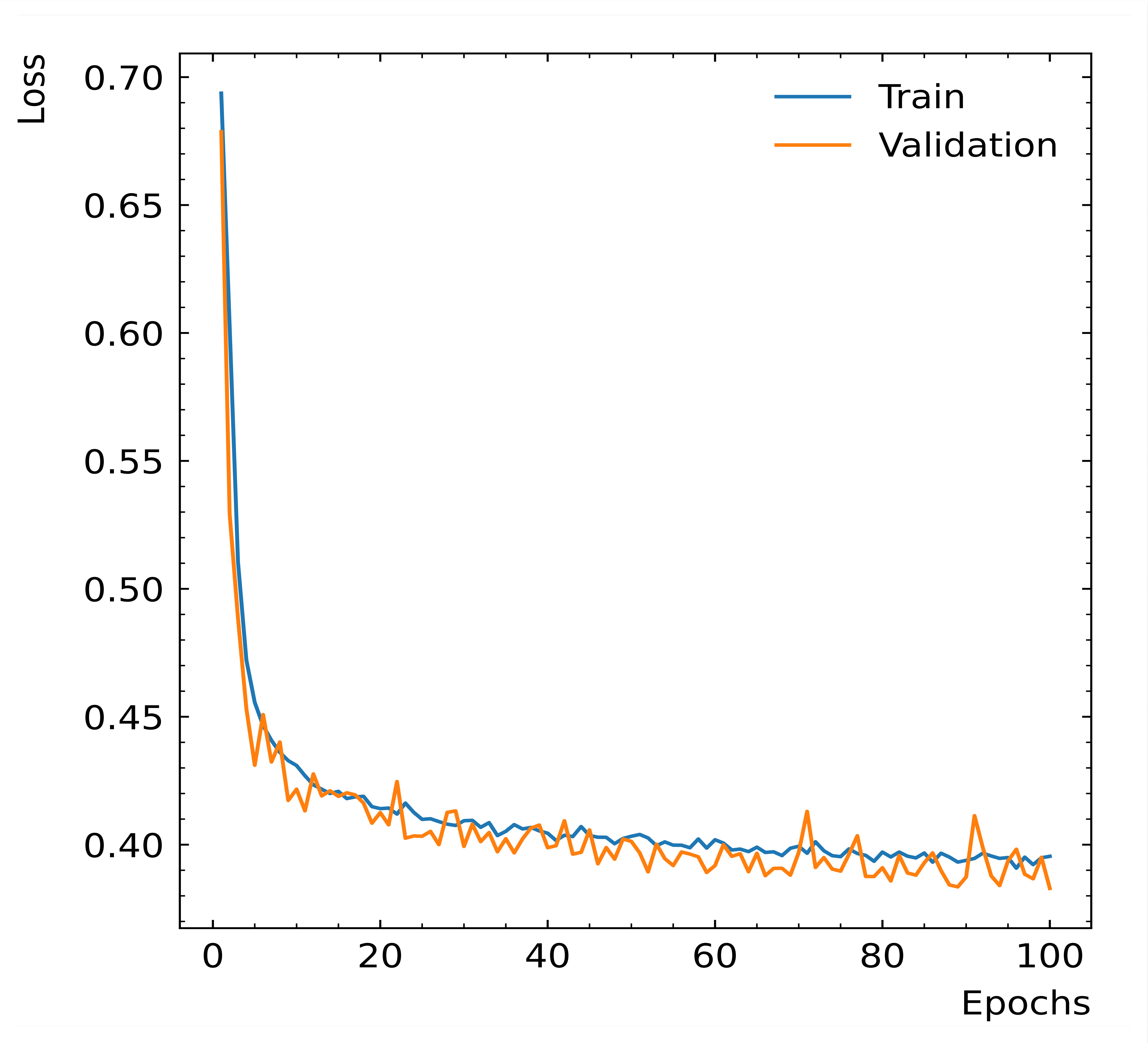
Comparison: Accuracy vs Epochs

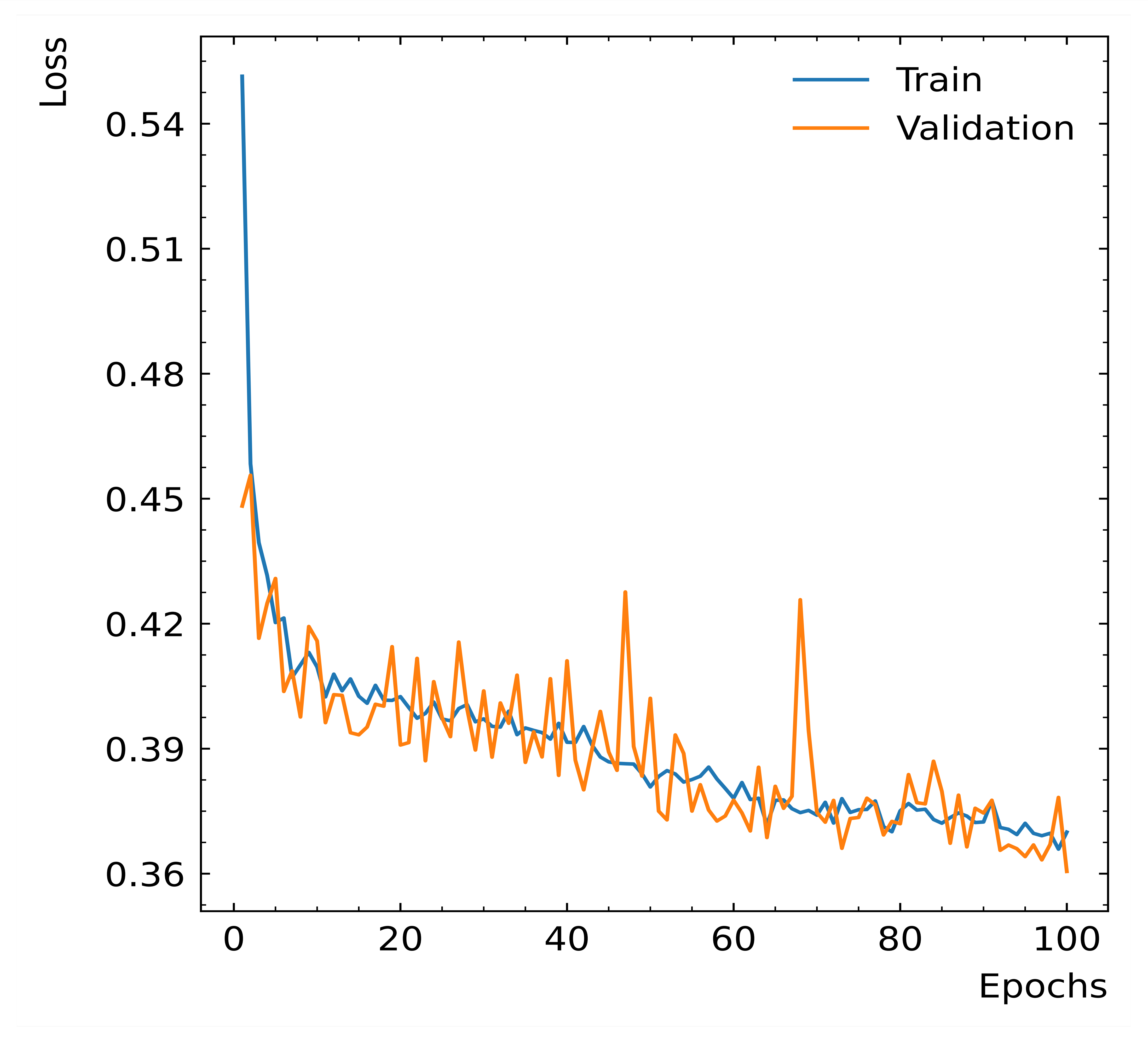




☐ The hybrid-quantum Transformer (left) and the Transformer (right).

Comparison: Loss vs Epochs

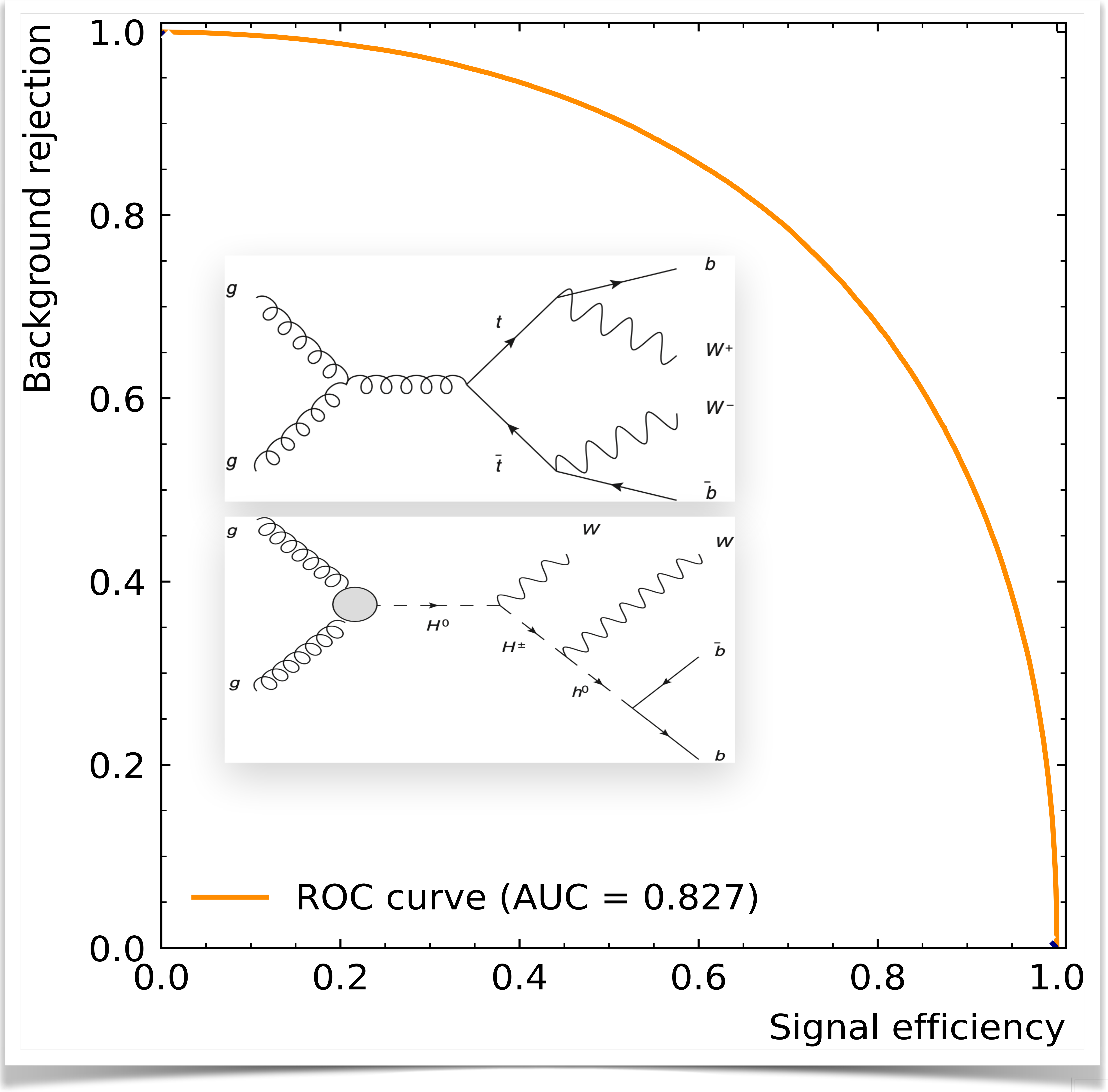




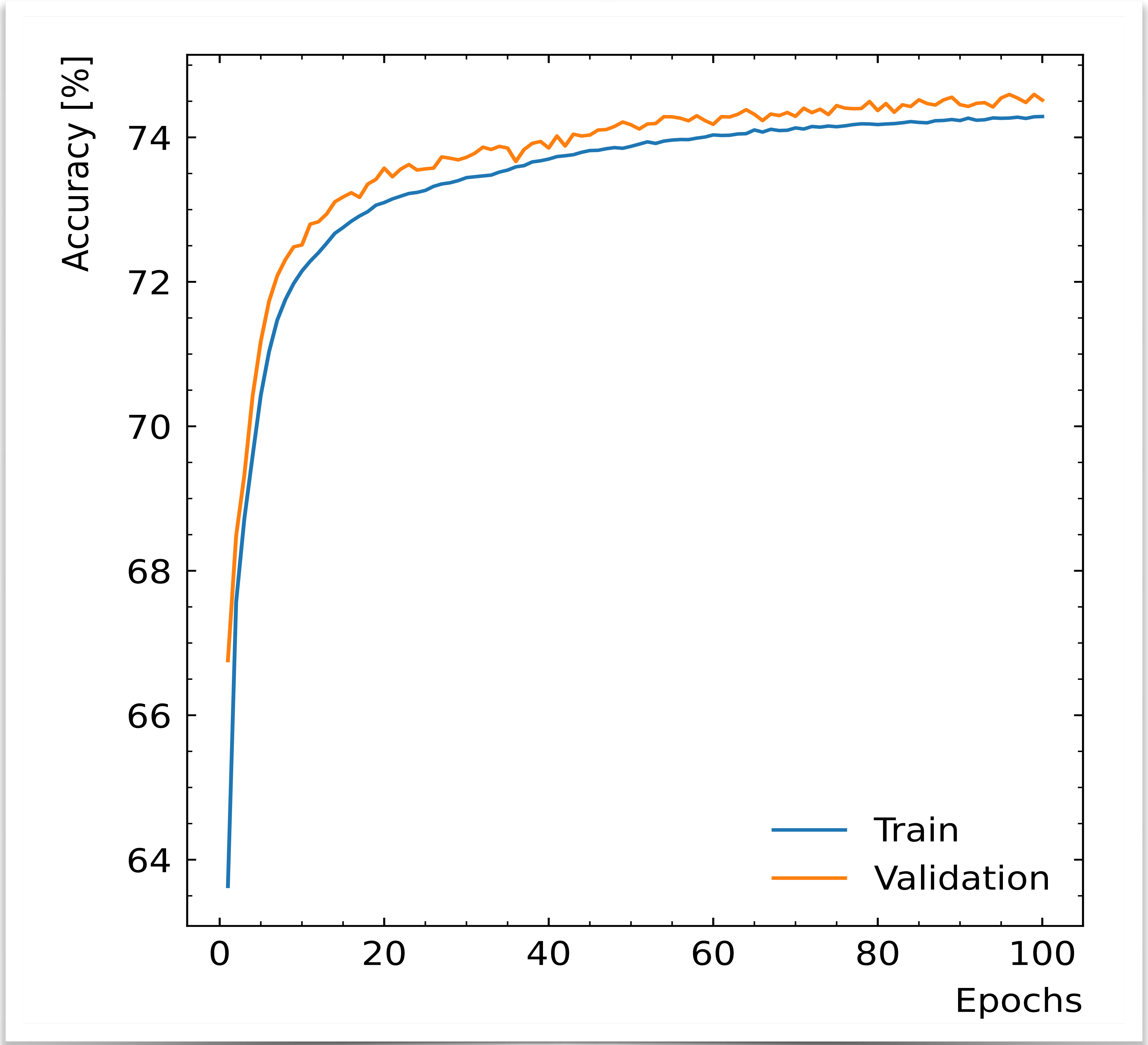
☐ The hybrid-quantum Transformer (left) and the Transformer (right).

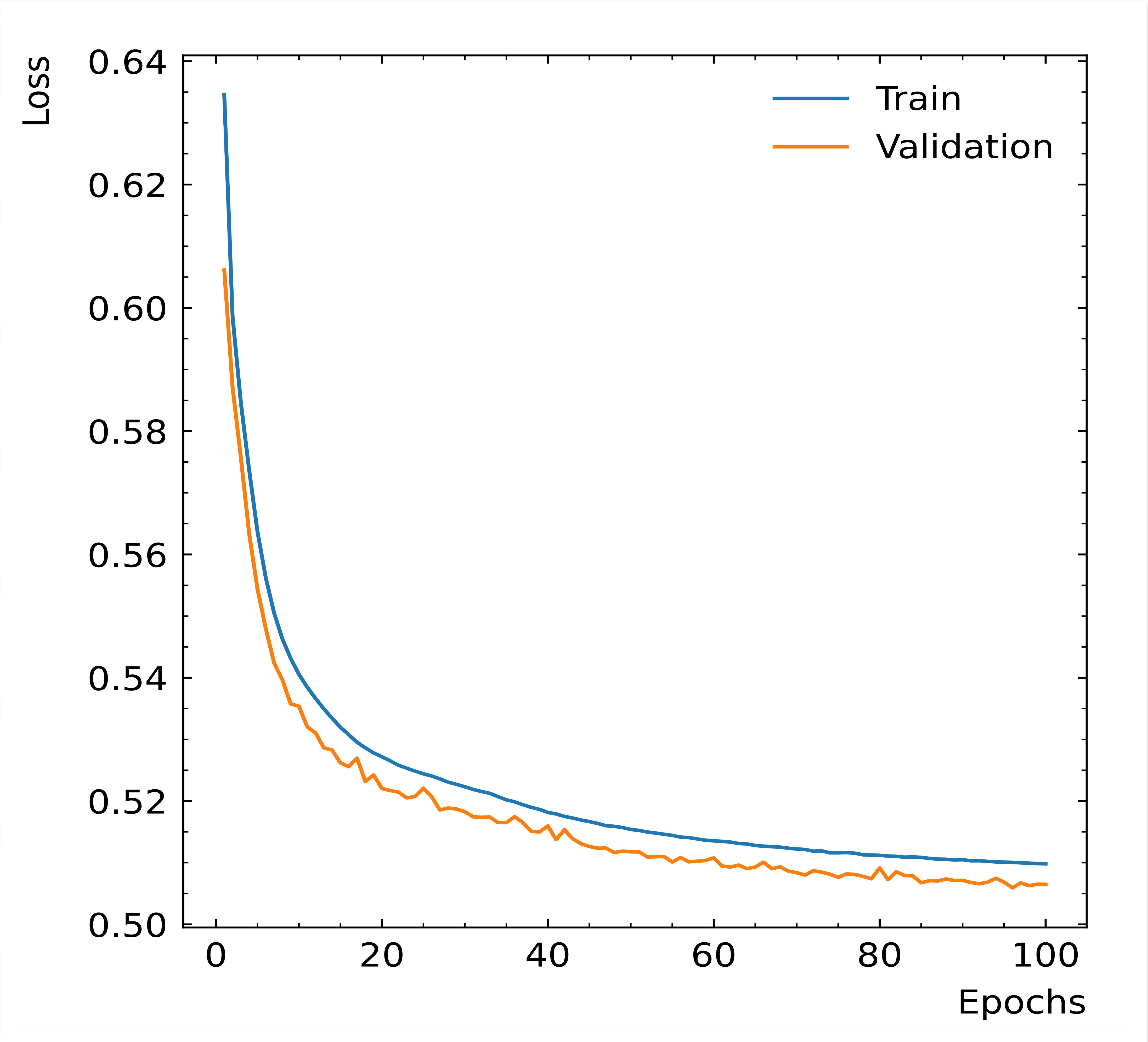
Testing the transformer architecture using LHC data

- ☐ Twenty-one low-level kinematic properties were measured by the particle detectors.
- ☐ Seven high-level features were derived from the twenty-one features.
- ☐ The train, validation and testing: 2.6M, 100k and 500k.
- ☐ A quick scan of the hyperparameters of the transformer model was used.



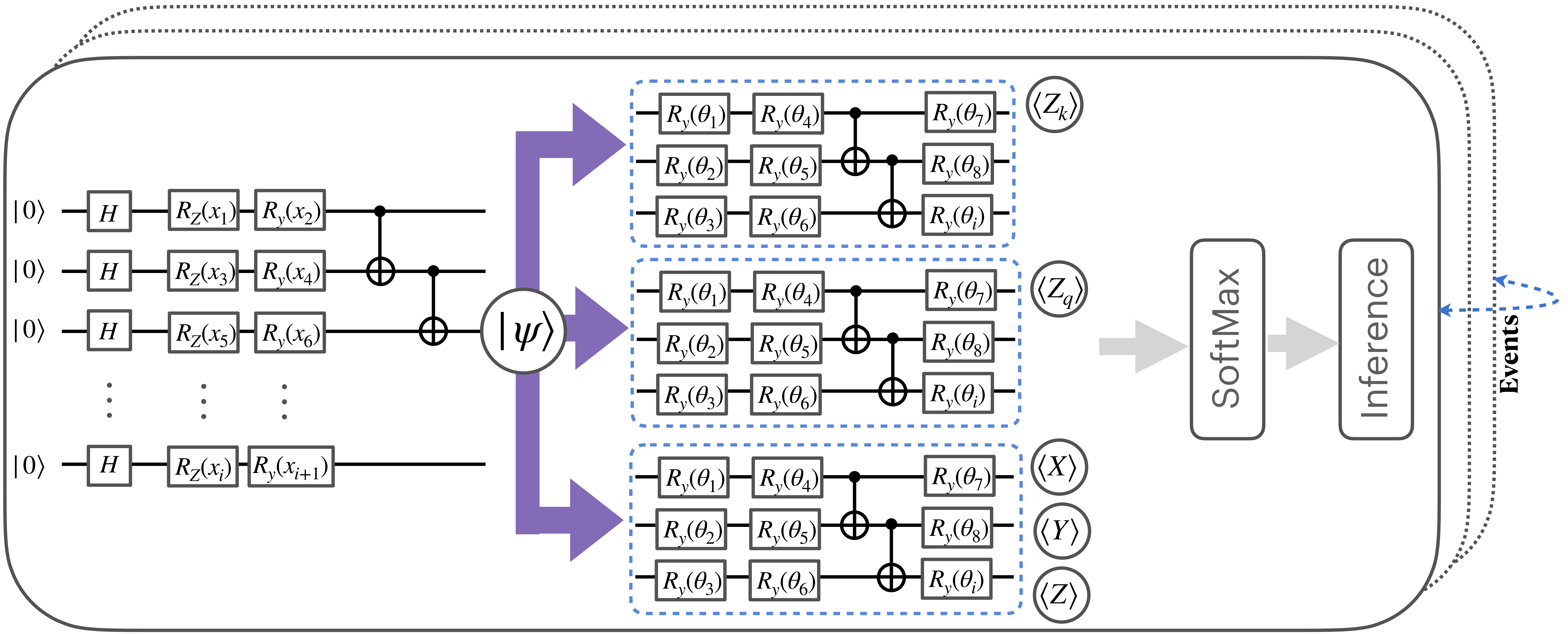
Testing the transformer architecture using LHC data





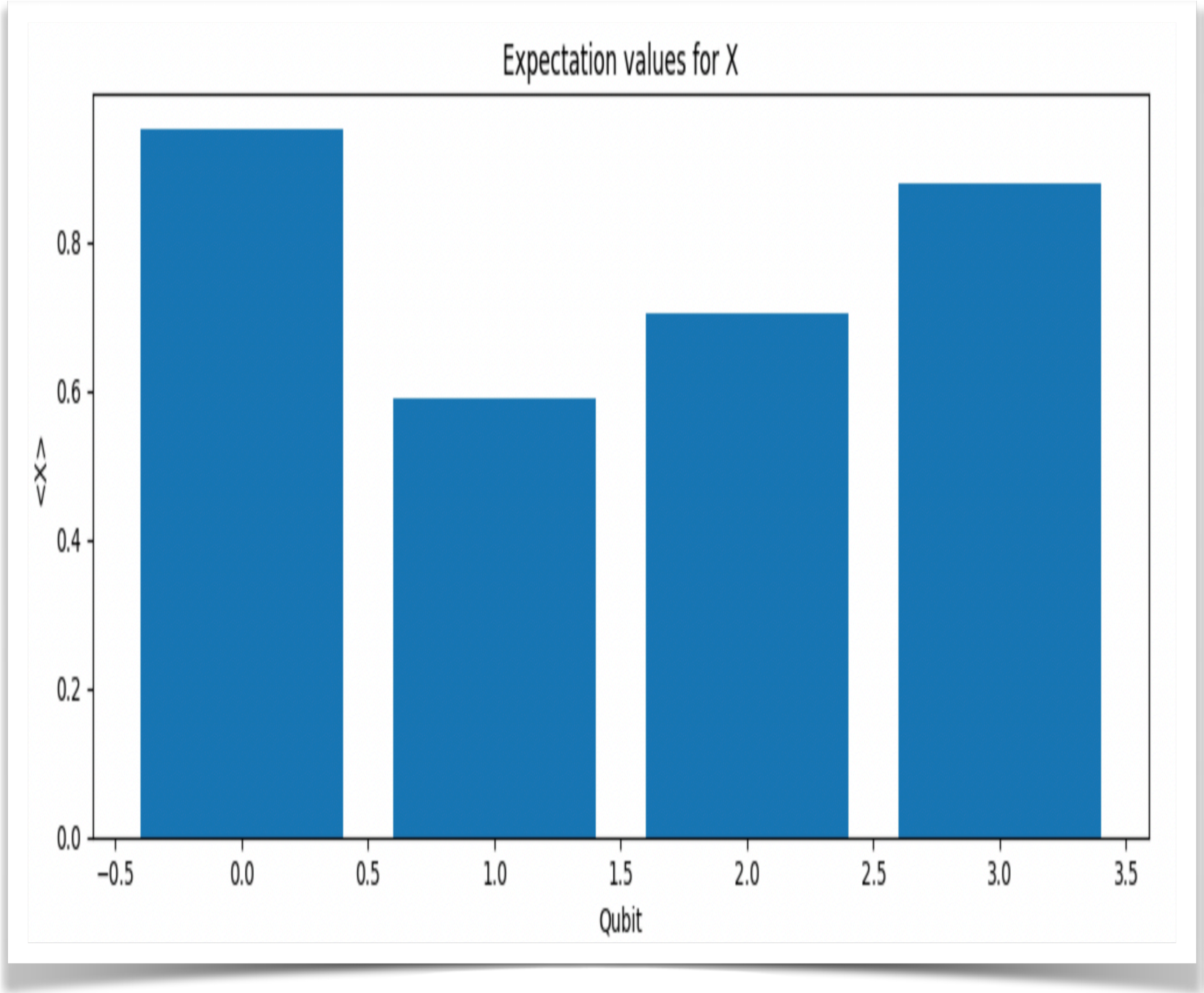
- [ ] The accuracy (left) and the loss (right) against the number of epochs.

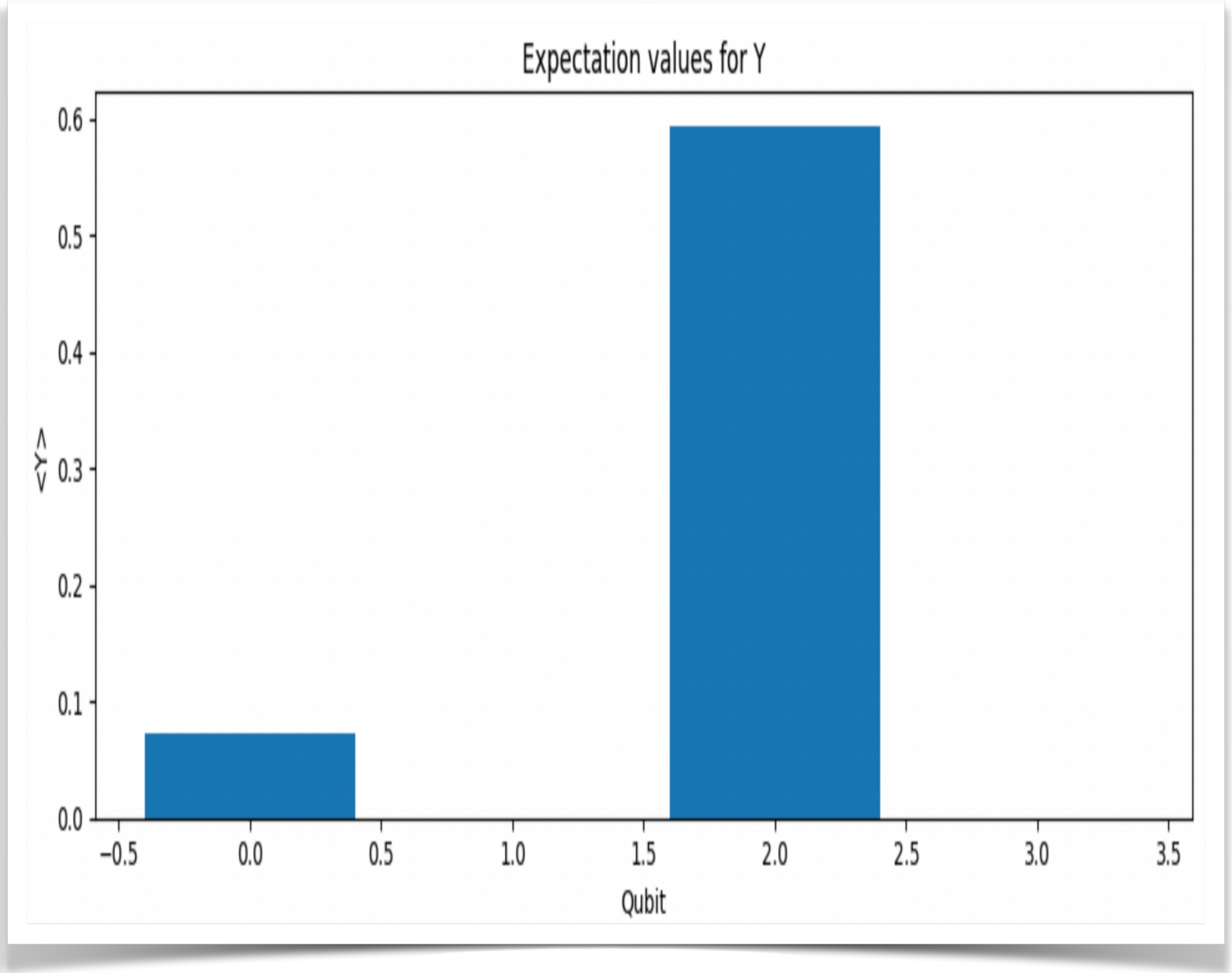
Quantum Transformer

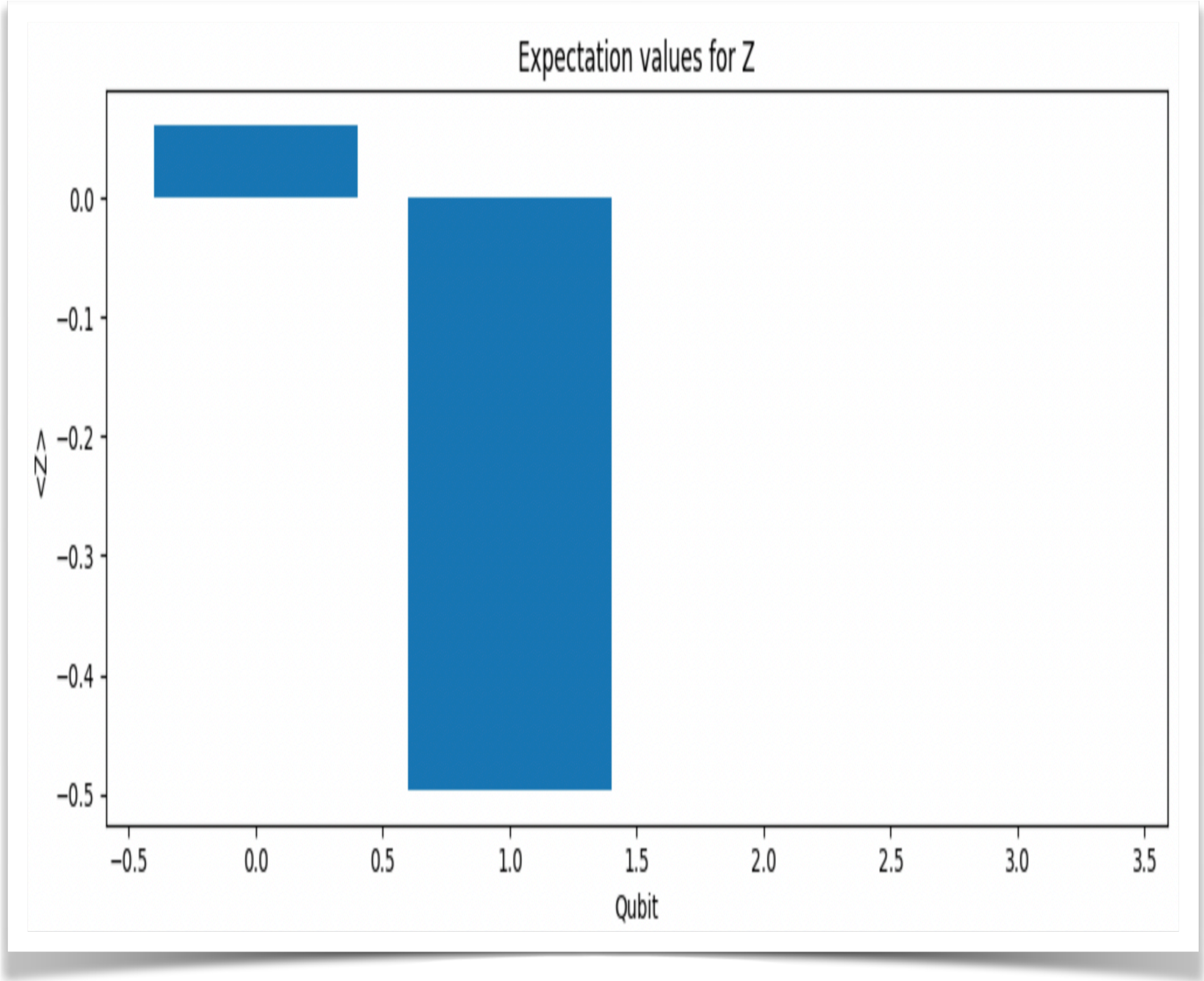


☐ This is still experimental work towards building a pure-quantum transformer inspired by [ArXiv: 2205.05625](https://arxiv.org/abs/2205.05625).

Quantum Transformer



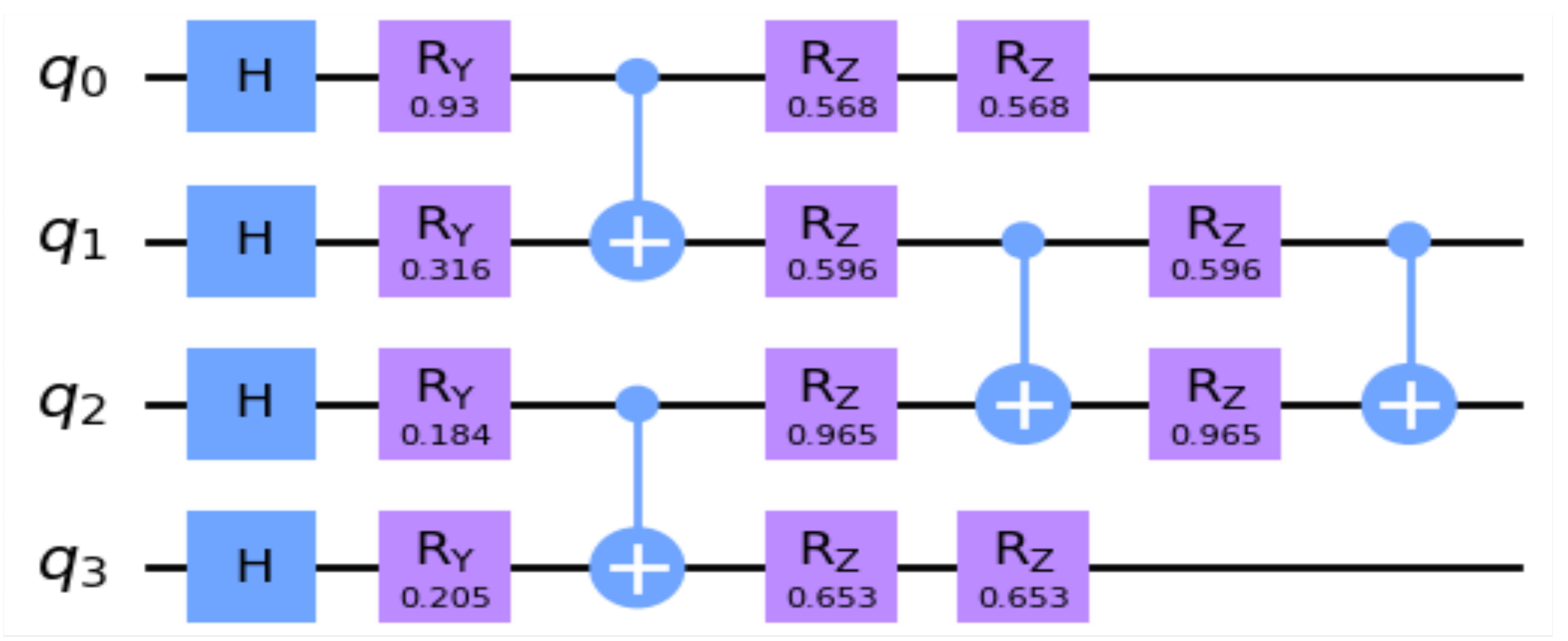


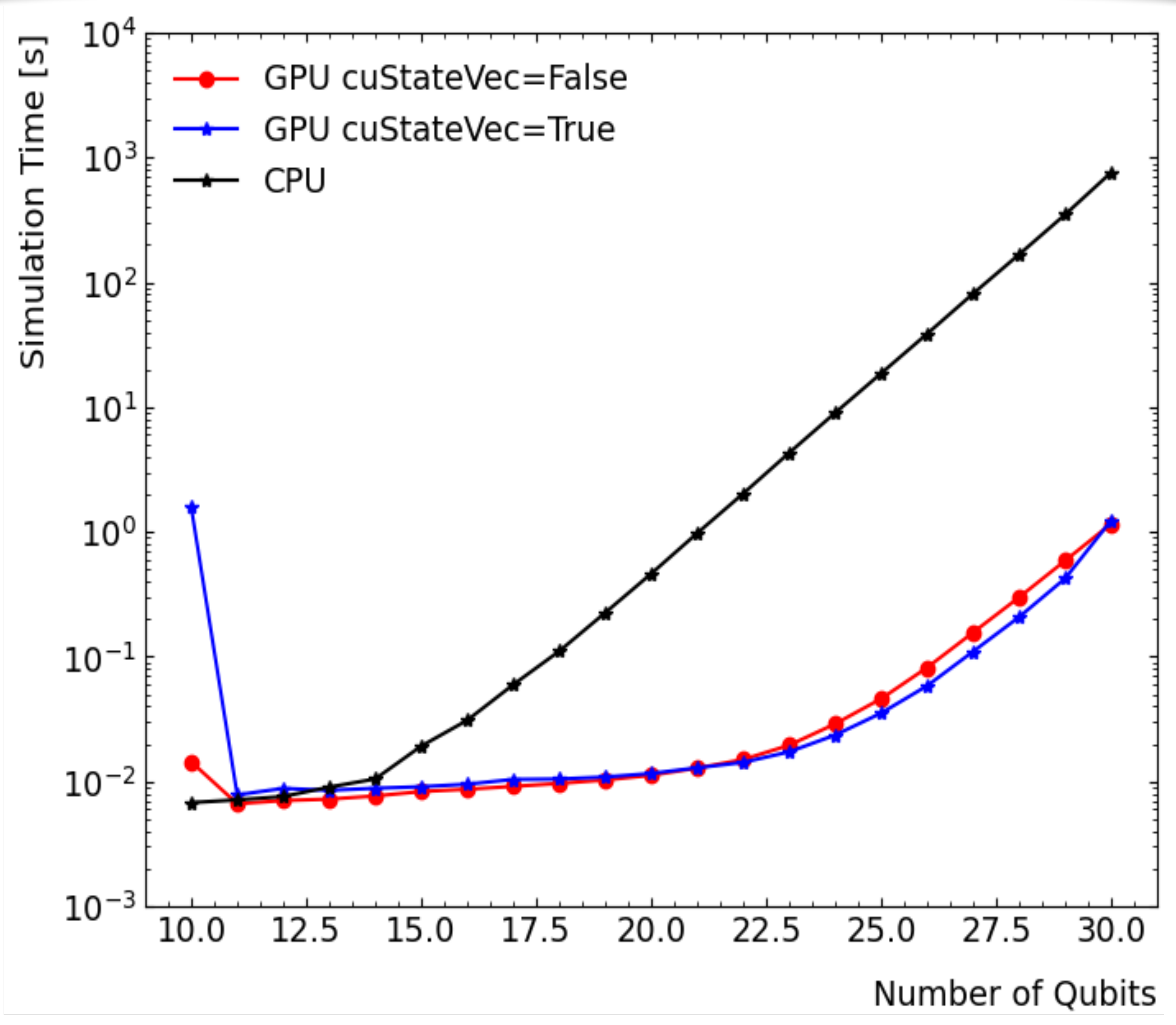


- ☐ Measuring the expectation value for Pauli operators: $\{X, Y, Z\}$ for each qubit. This is taken as “values”.
- ☐ In addition, we measure the expectation for the Z Pauli operator for two of the Ansatz to represent the “keys” and “query”.

Remarks

- ☐ Qiskit has implemented parallel simulation using cuQuantum (Nvidia's GPU quantum circuit simulator):
  - cuTensorNet based on the tensor network approach
  - cuStateVec based on the state vector approach
- ☐ However, one should not rely only on GPUS to seek speed-up.
- ☐ Currently, the hybrid-quantum transformer runs for about two days for 50k events with the CEPC data.
- ☐ Optimising the hyperparameters of the transformer is tricky.
- ☐ Most of the applications of quantum machine learning do not show superiority to conventional learning.
- ☐ However, in some quantum generative models, it appears to perform better in data generation.





Summary

- ☐ Discussed the role of transformers in HEP event classification and pattern recognition.
- ☐ Highlighted the challenges in HEP and the potential benefits of integrating quantum computing.
- ☐ 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.
- ☐ Outlined the steps towards developing a pure-quantum transformer.