



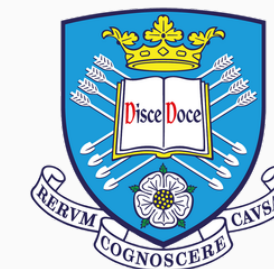
# QUANTUM ENHANCED GENOMIC SEQUENCE PREDICTION

**Personalized Medicine**

From Classical to Quantum Approaches

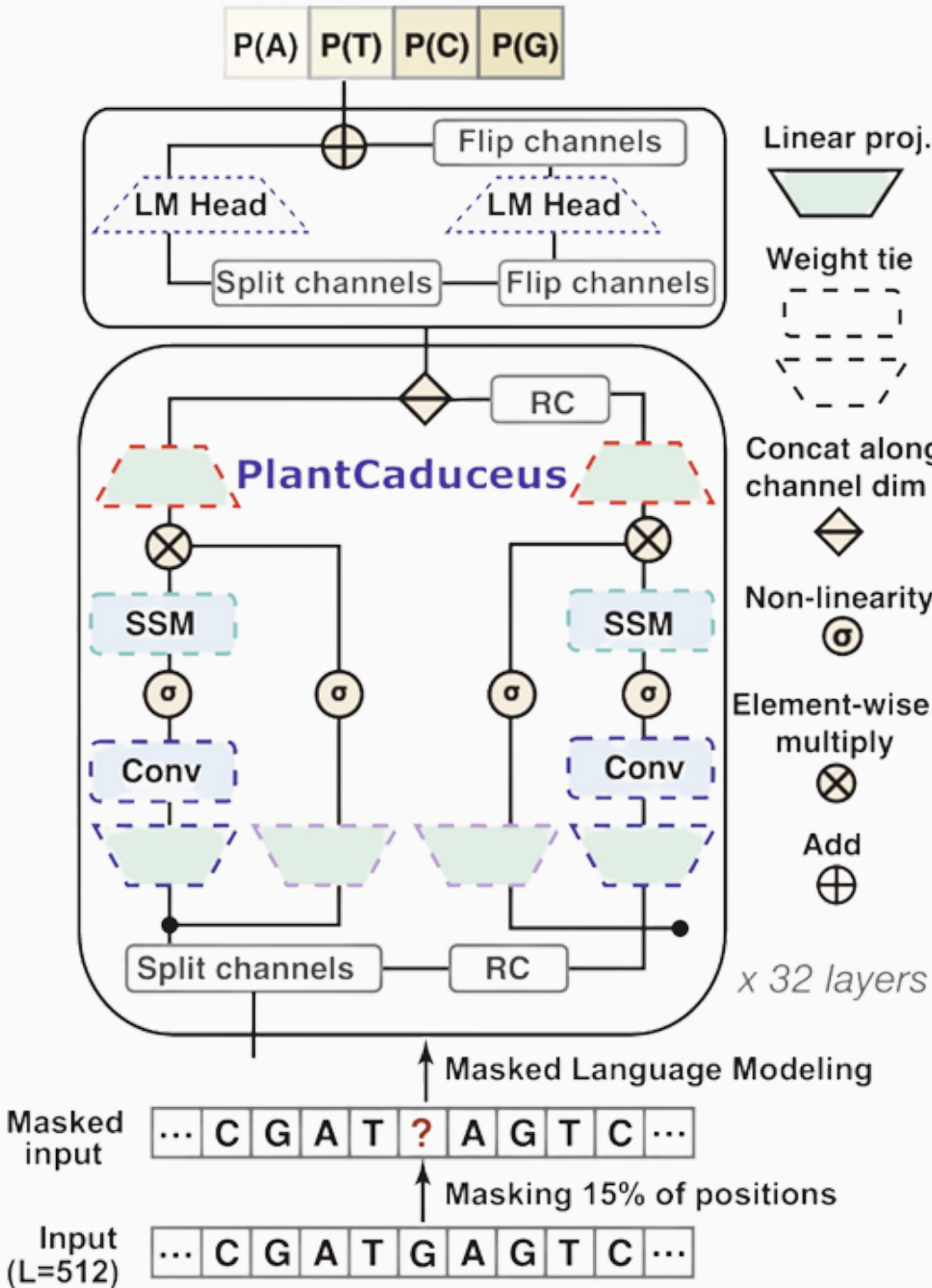
## RYOUSHI: QUANTUM BUDDIES

**Dat Chi (Ryan) Le, The University of Sheffield**  
**Abubakar (Sid) Iliyasu, Imperial College London**  
**Gyanateet Dutta, University of Leeds**



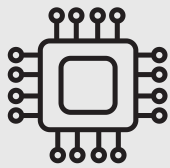


# Genomic Pretrained Network



## Importance

- 1. **Medical Diagnosis:** Identifying causative mutations for rare genetic diseases.
- 2. **Cancer Treatment:** Identifying therapeutic targets in tumors for personalised regimens.
- 3. **Pharmacogenomics:** Predicting patient response to medications based on genetic variants.
- 4. **Disease Prevention:** Assessing risks for conditions like diabetes and coronary artery disease.



## Quantum Advantages

- 1. **Computational Speed:** Quadratic speedup and parallel processing of massive datasets.
- 2. **Memory Efficiency:** Logarithmic scaling (constant space) vs. linear growth in classical methods.
- 3. **Enhanced Accuracy:** Fine-grained comparison at single nucleotide resolution.



# Research Questions

1. Can both **QMCMC** and **Quixer** learn genomic distributions more faithfully than classical baselines (Markov based and Transformer Based)?
2. Does the Quixer quantum transformer offer parameter efficiency over classical BERT/GPT models?
3. Does the compositional structure of QNLP offer superior interpretability for identifying biological motifs compared to classical 'black box' transformers?
4. Can Quantum Hamiltonian Monte Carlo (QD-HMC) demonstrate faster mixing times and better convergence for high-dimensional genomic distributions?

## Objectives & Success Criteria

- **Objective:** Develop a hybrid quantum-classical pipeline for genomic sequence prediction.
- **Success:** Achieve predictive accuracy comparable to classical baselines with fewer parameters.
- **Hypothesis:** Compositional QNLP captures biological structure better than flat embeddings.

# Impact: UN Sustainable Development Goals (SDGs)



*Driving earlier diagnosis and personalised medicine for better health outcomes.*

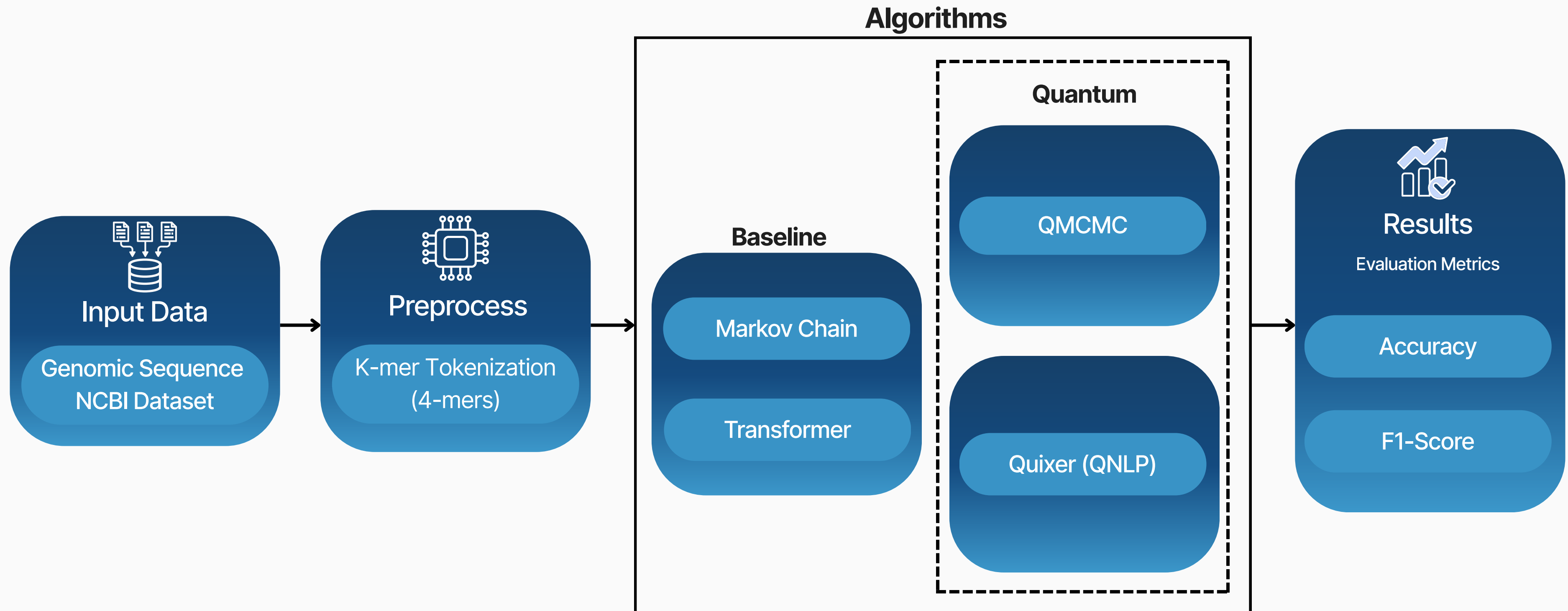


*Quantum-classical innovation for next-generation genomic modelling*



*Building Interdisciplinary Collaboration for Healthcare*

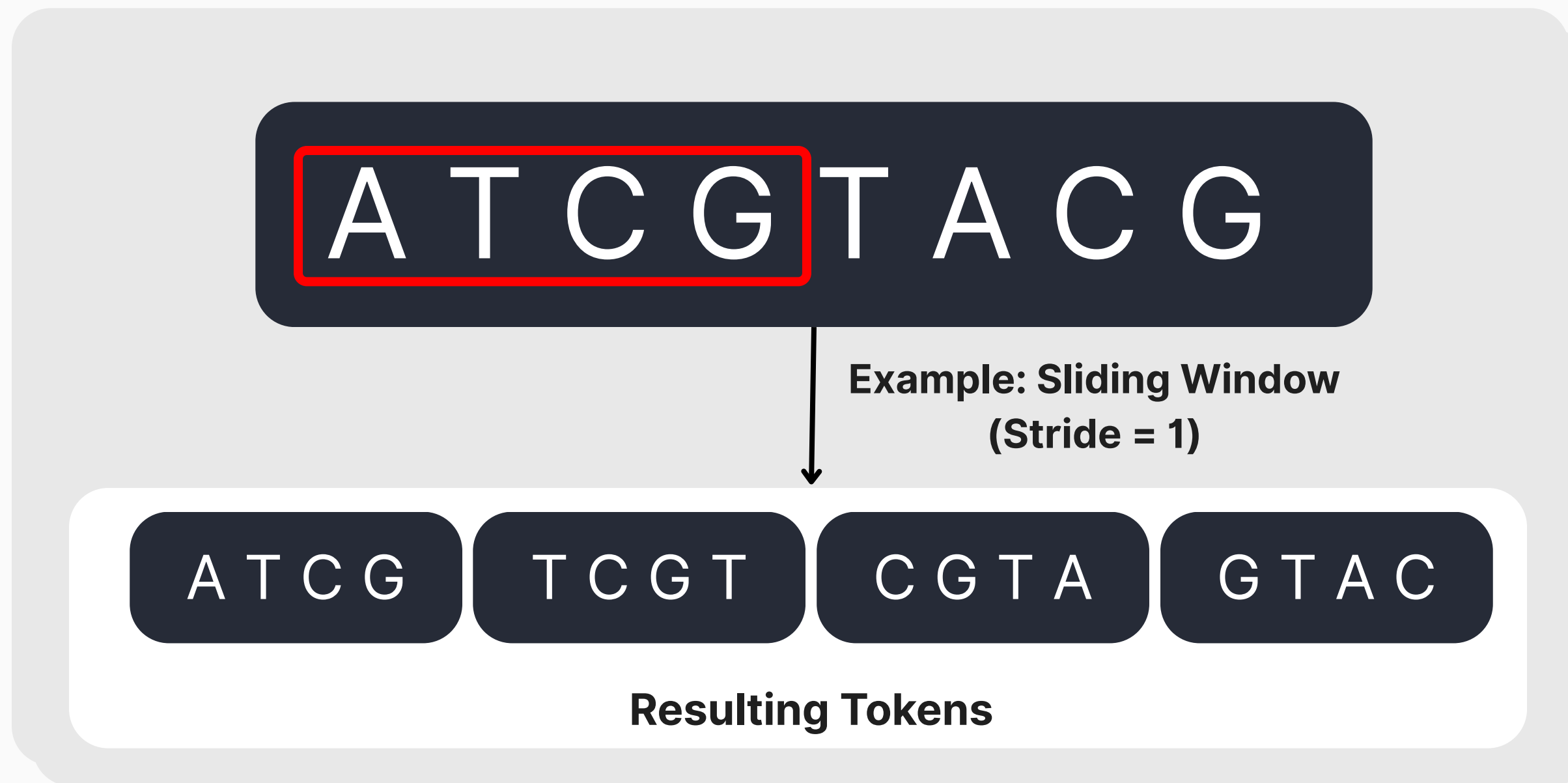
# Complete Pipeline



# Preprocessing

We break down long DNA chains into **4-mer** (4-letter tokens).

- **Vocabulary Size:** 4 bases (A, T, C, G) ^ 4 positions = **256 Unique Tokens**.
- **Context:** Allows the model to learn local biological motifs (like codons or regulatory sites).
- **Efficiency:** 256 tokens fits perfectly into 8 qubits ( $\log_2(256)=8$ ), making it NISQ-friendly, but computationally expensive for simulation.

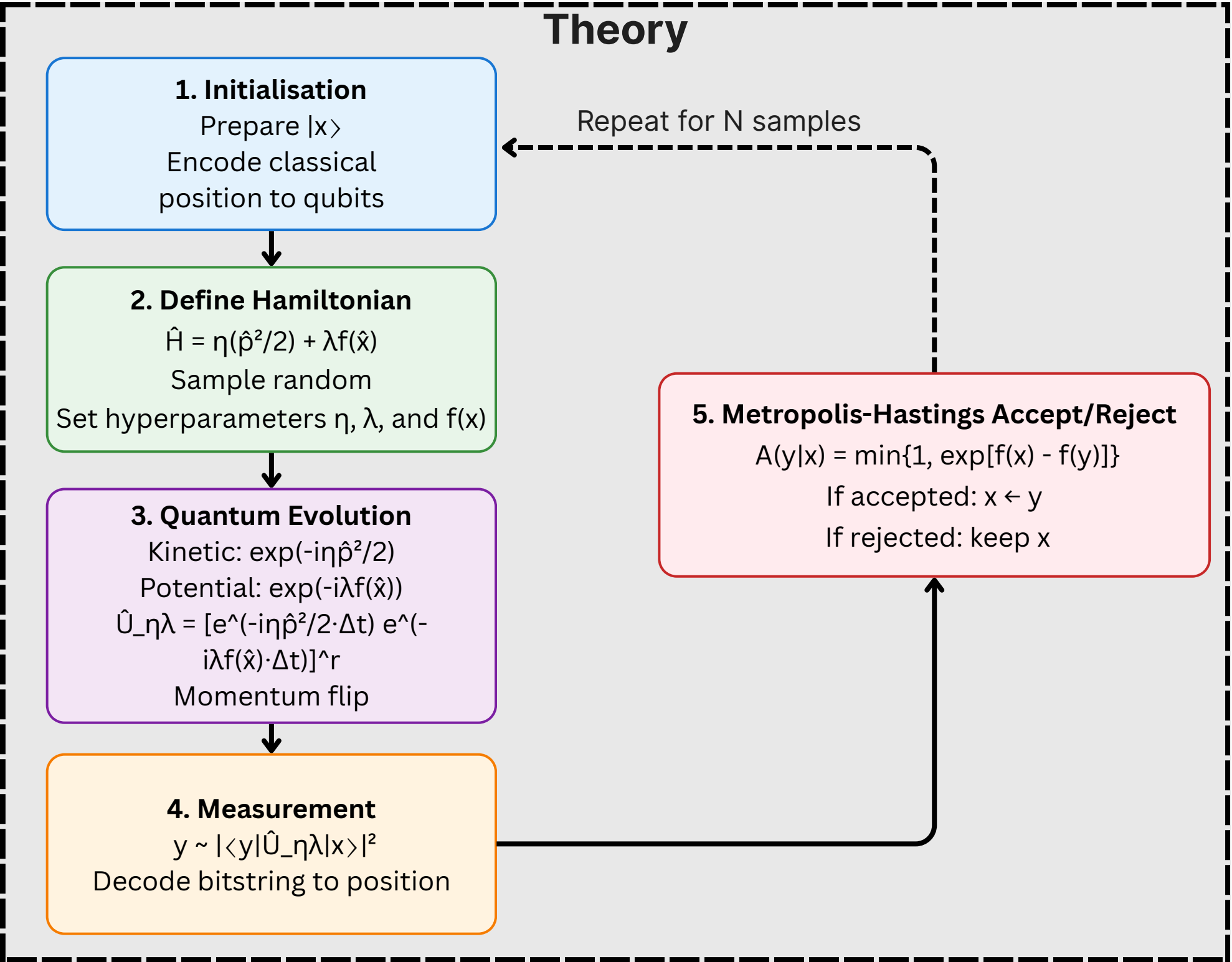
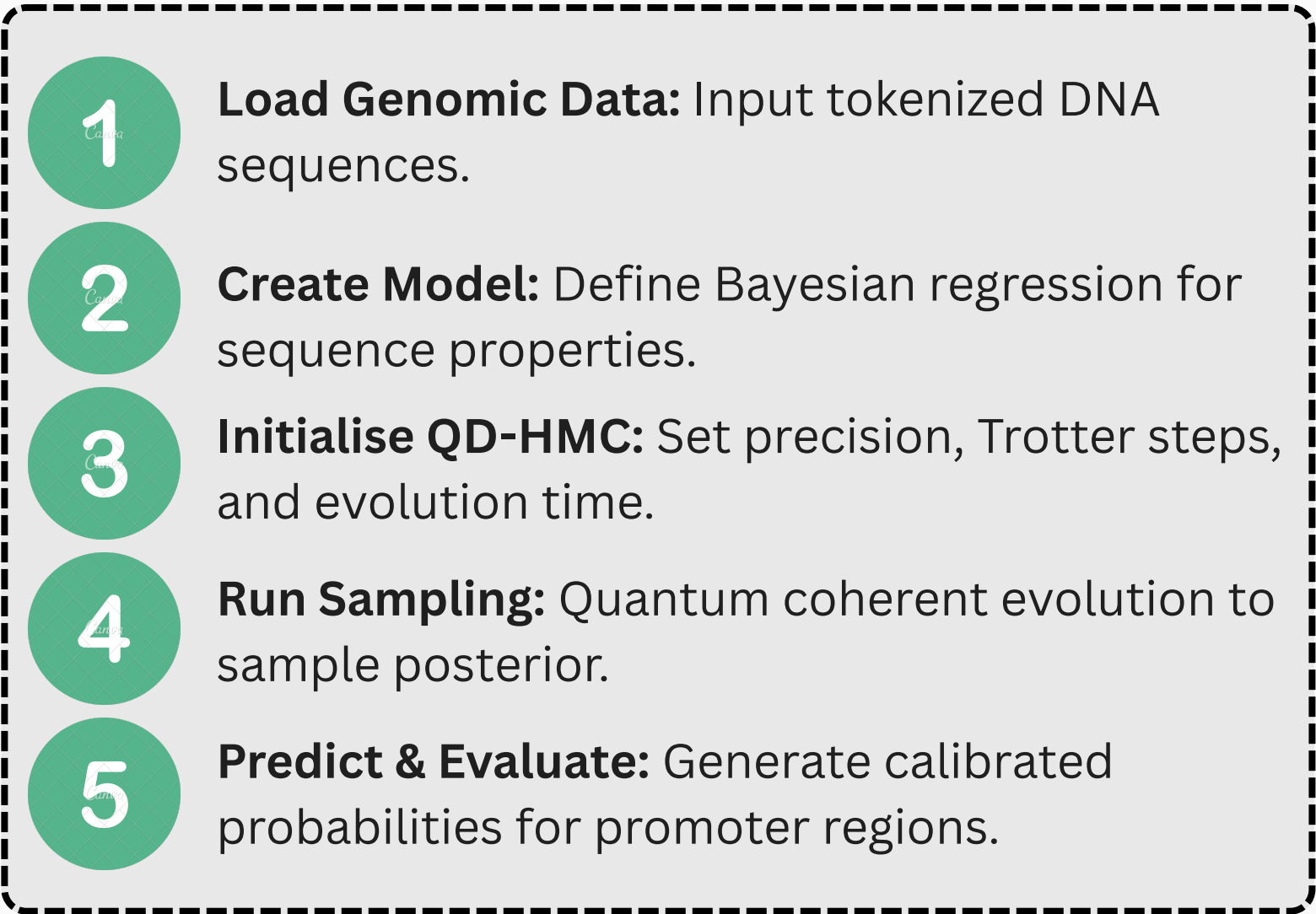


# Quantum Markov Chain Monte Carlo (QMCMC)

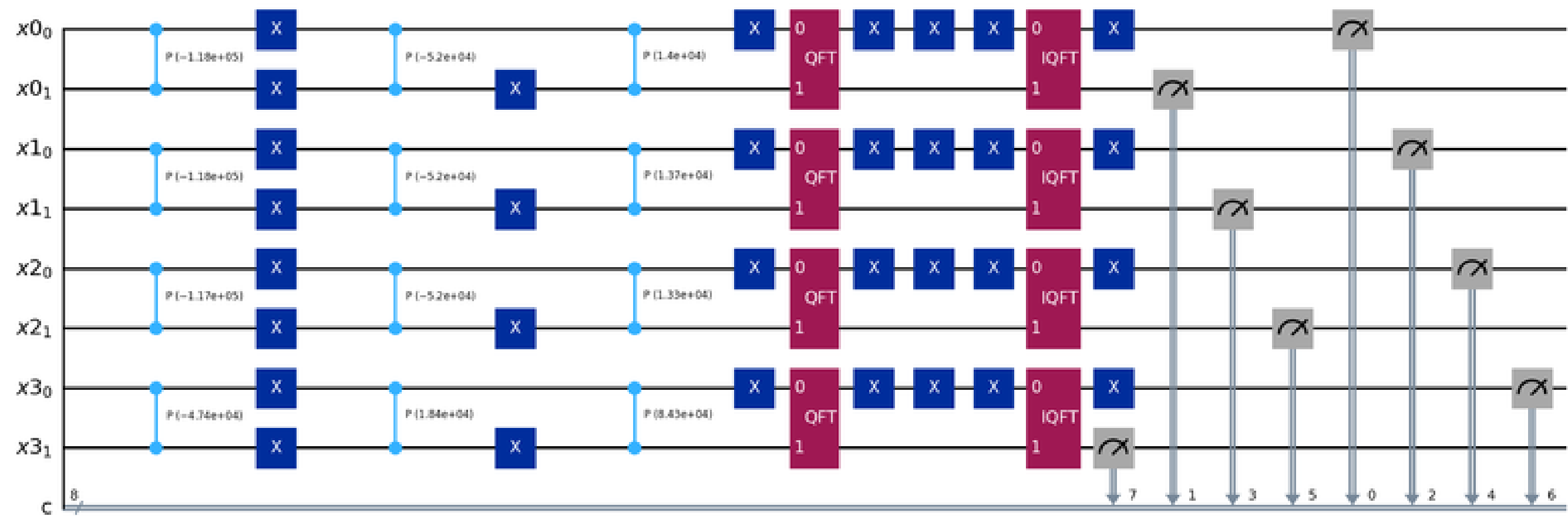
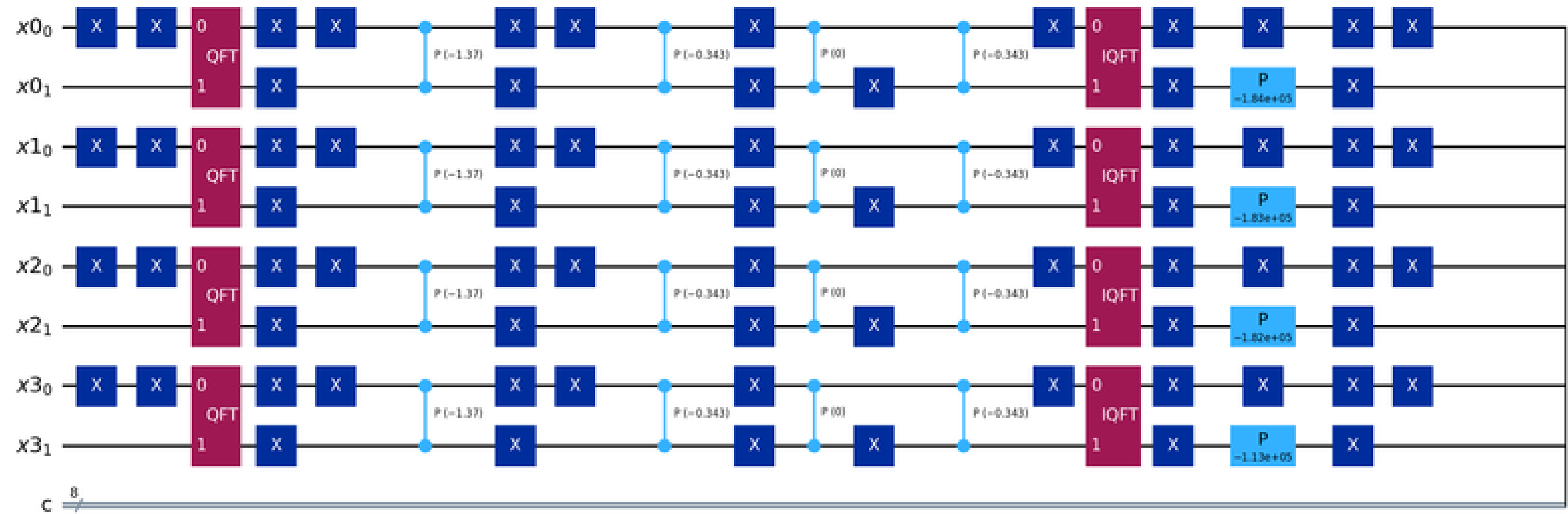
Quantum Dynamical Hamiltonian Monte Carlo (QD-HMC) is a hybrid quantum-classical Markov Chain Monte Carlo (MCMC) sampling algorithm that uses quantum Hamiltonian dynamics to perform Bayesian inference.

**Key Innovation:** Replace classical symplectic integration in Hamiltonian Monte Carlo with quantum coherent evolution on quantum hardware

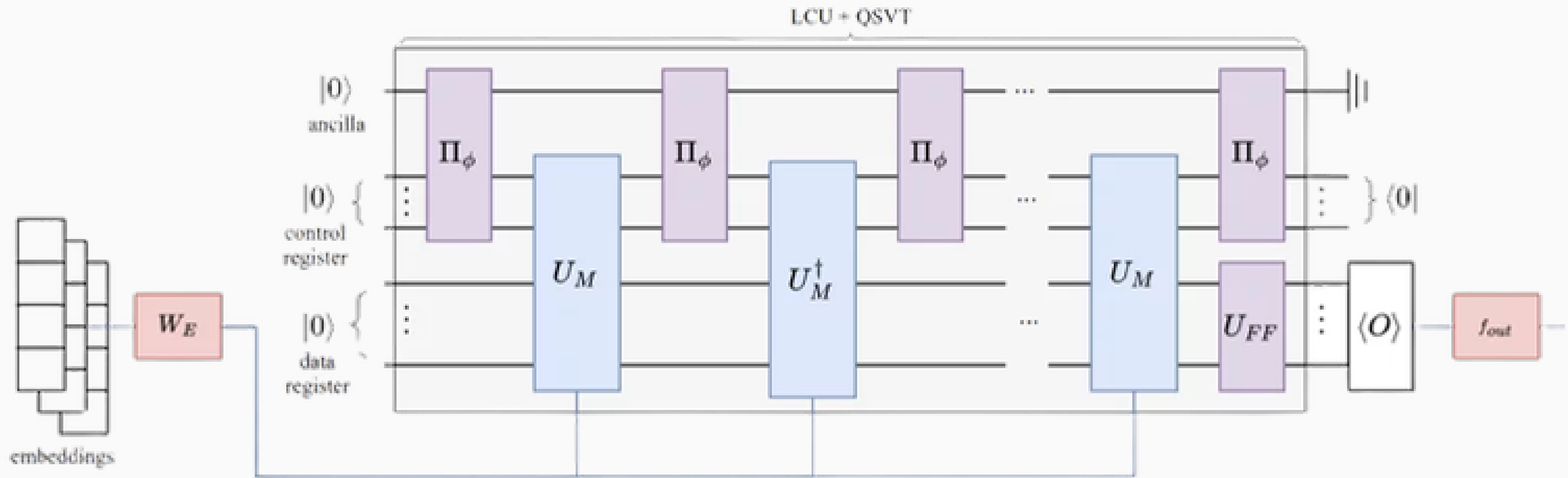
## Workflow:



# QMC MC Circuits

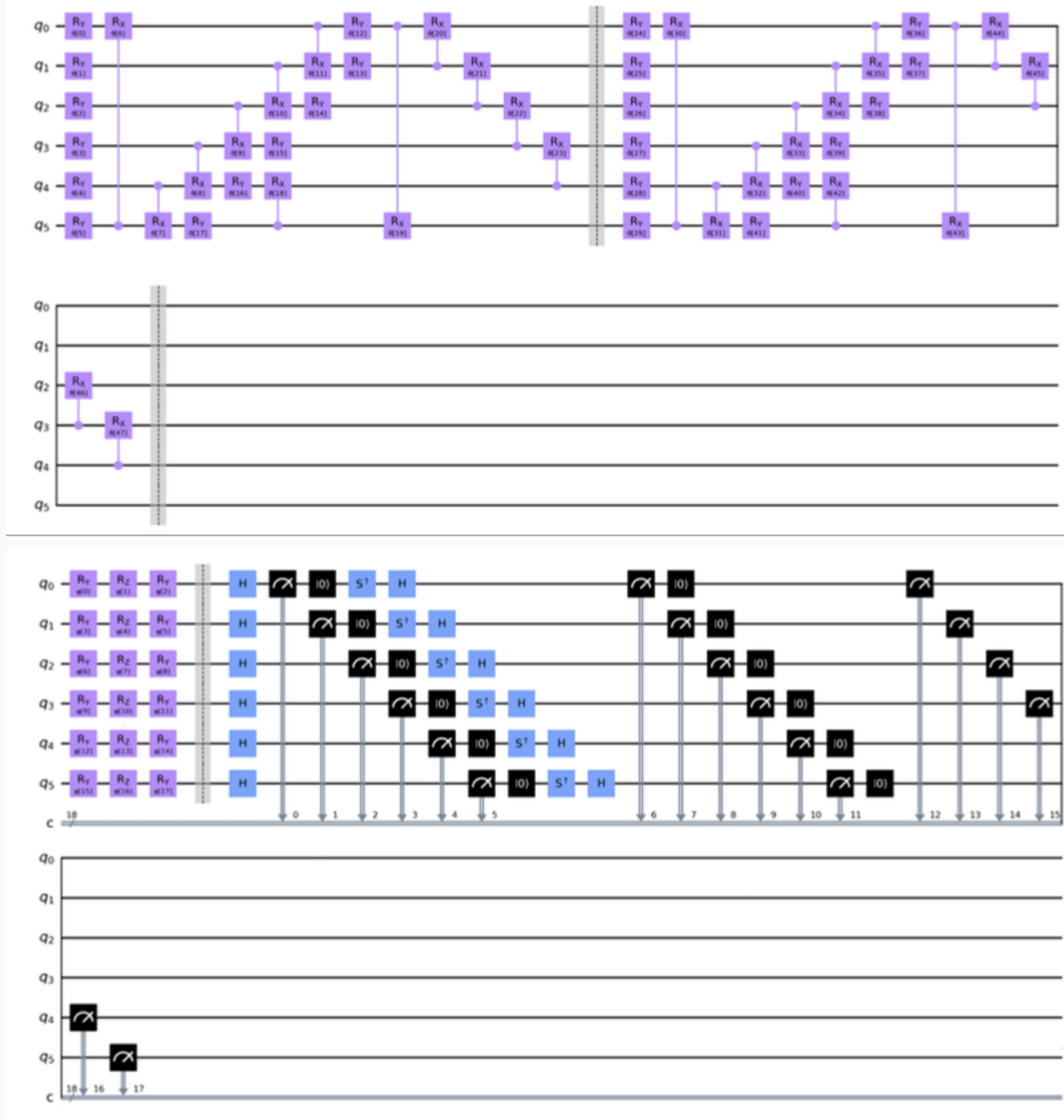


# Quixer (NLP)

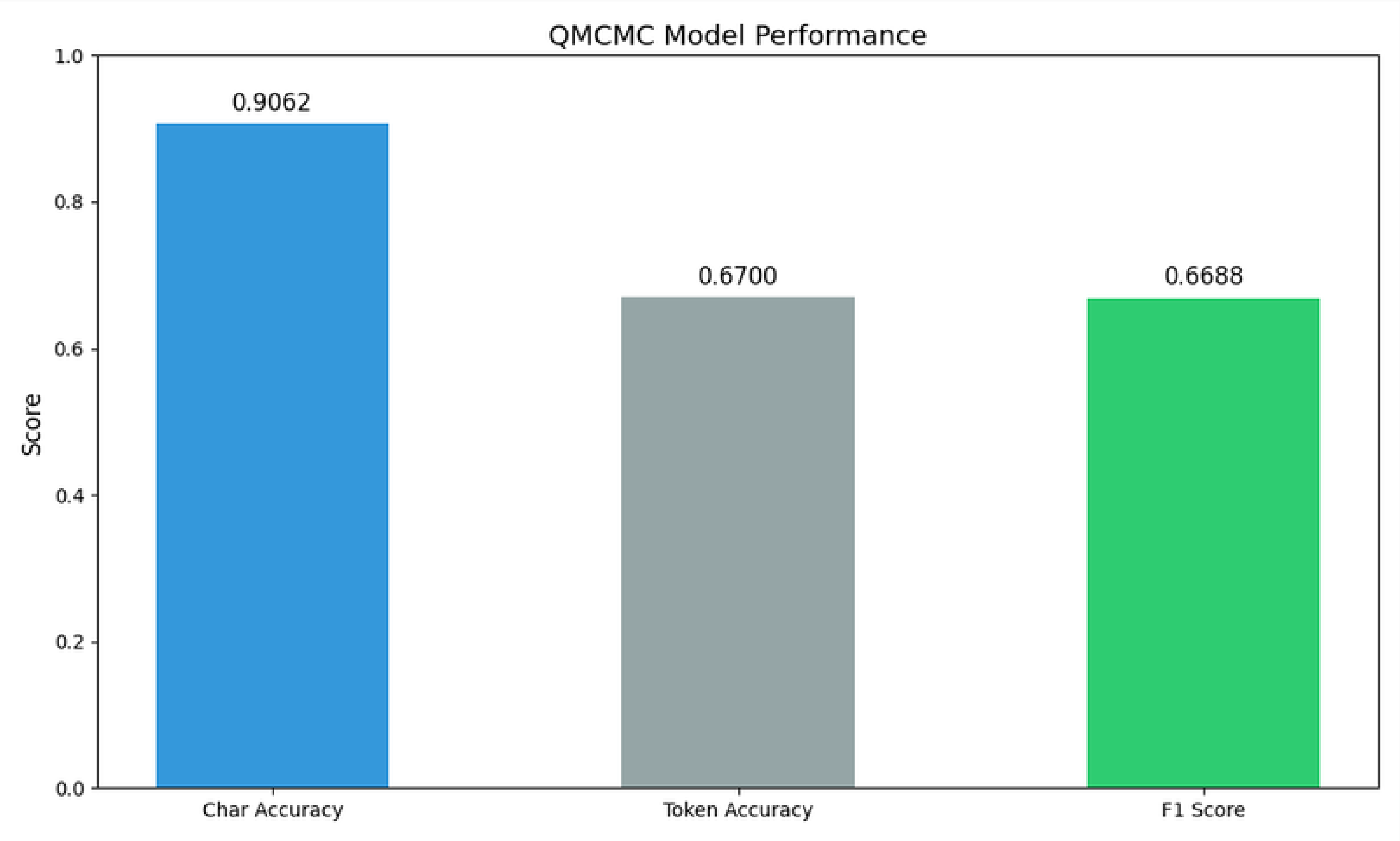


# Quixer

# Circuit

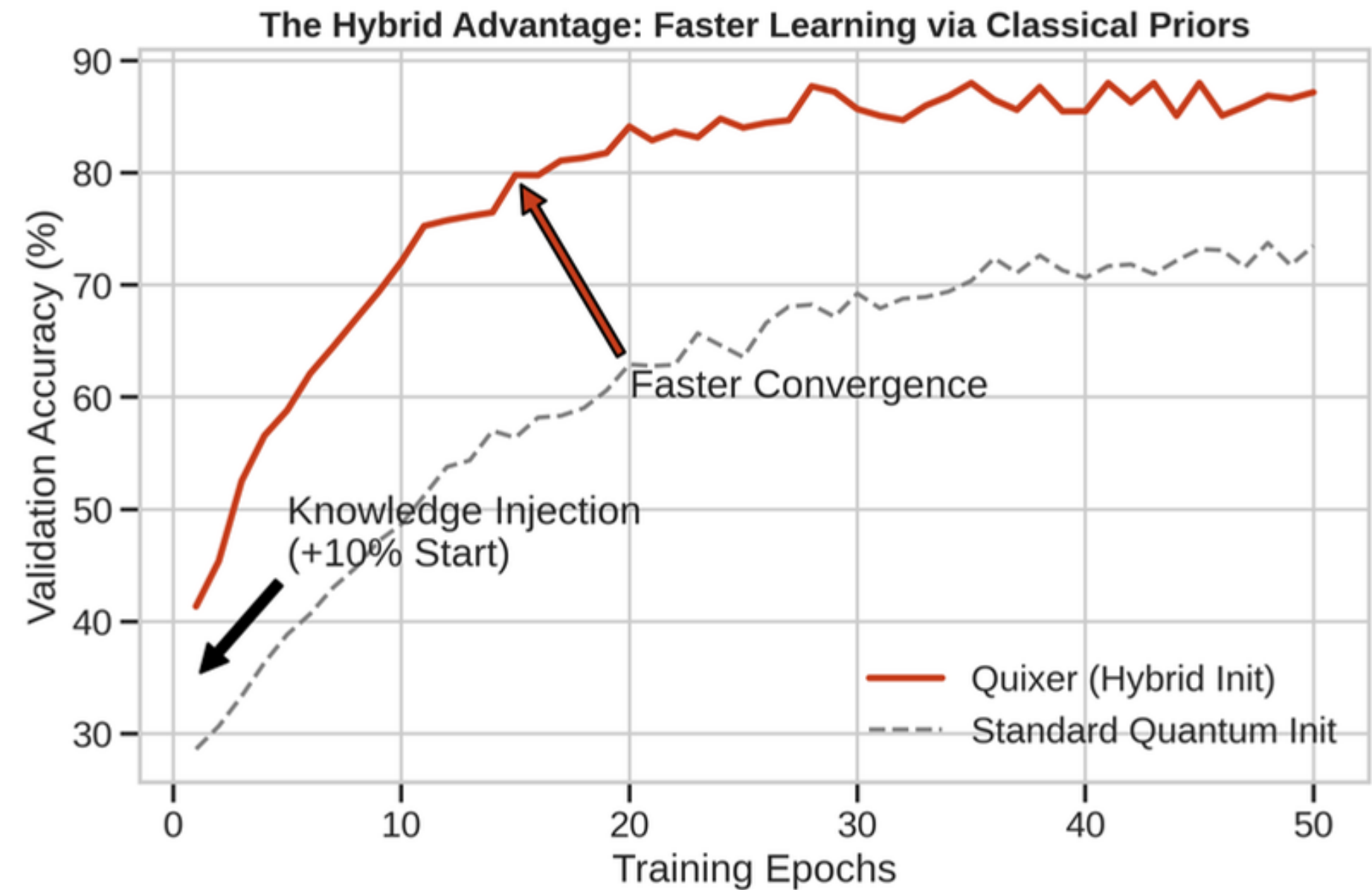
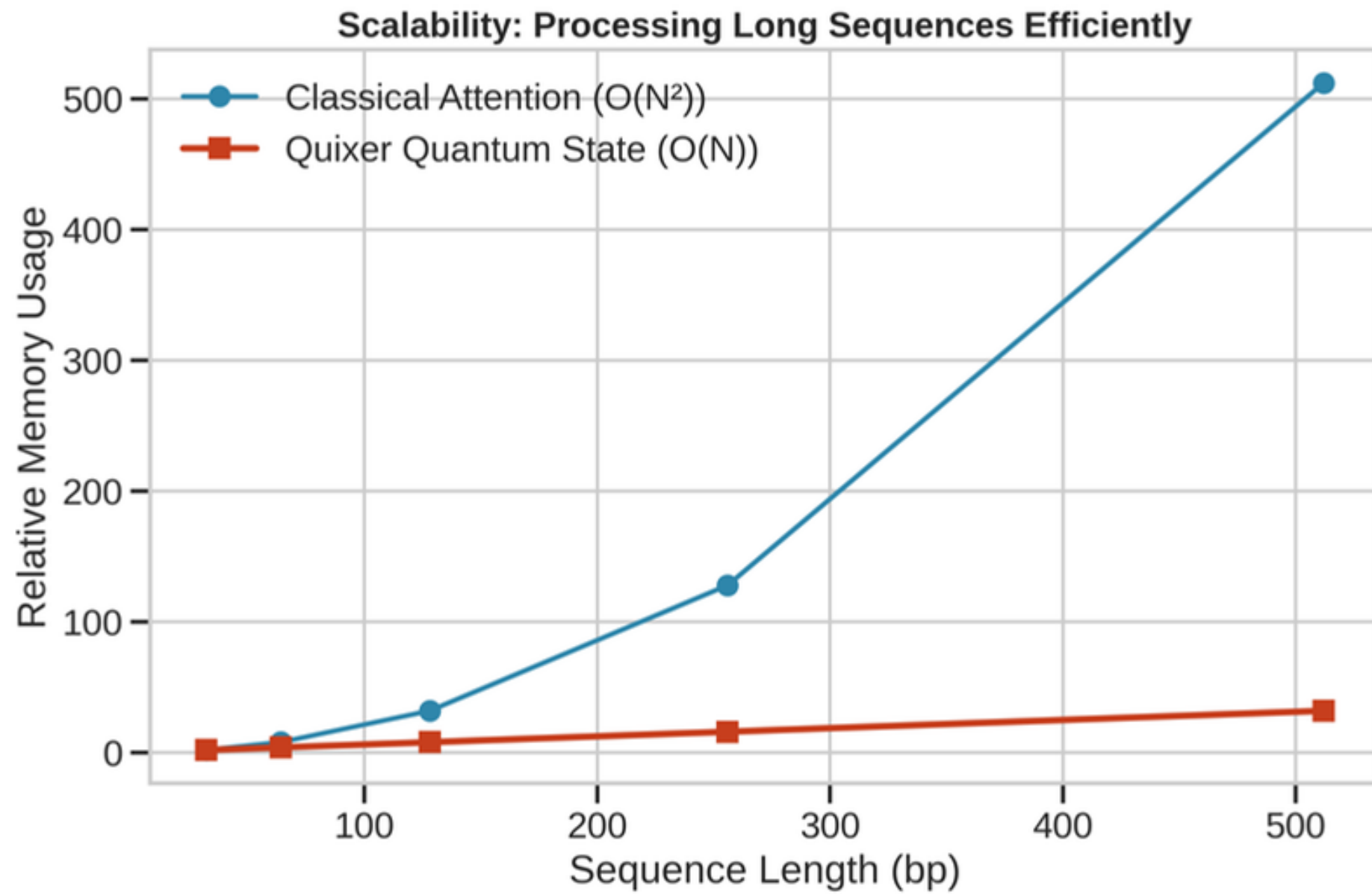


# Results (QMCMC)



CGTT,CGTT  
CTTC,CTTC  
TTAG,TTAG  
TATT,TATT  
AATC,AATT  
GACA,GACA  
GGCC,GGCA  
CTCA,CTCA  
ACAG,ACAG  
GCTG,GCTG  
TGGT,TGGT  
ATGA,ATGA  
ATCA,ATCA  
GGGG,GGGG  
TGGT,TGGT  
TGGT,TGGT  
TTAT,TTAT  
CCGC,CCGC  
TCAC,TCAA  
TTTT,TTTT  
AAGA,AAGT  
AGTT,AGTG  
TCTT,TCTT  
CATC,CATC  
AGGT,AGGG  
TTAA,TTAC  
TTCC,TTCA  
CCTG,CCTT  
GTTT,GTTT  
GGAA,GGAA  
GACC,GAAG

# More Results (Quixer)



# Conclusions

What we learned from the study

## **Objective: Did we develop a hybrid quantum-classical pipeline for genomic sequence prediction?**

- Yes, both QMCMC and Quixer were able to do genomic sequencing.

## **Success: How successful were our quantum algorithms?**

- Our models achieved approx. 90% accuracy for predicting genomic sequences.

## **Hypothesis: How did our algorithms compare to the classical methods?**

- Both QMCMC and Quixer achieved approx. 90% accuracy, exceeds Classical MCMC (near 80% accuracy).
- This demonstrates feasibility of quantum approaches for genomic sequence modeling

## **We demonstrated**

- **Speedup:** QD-HMC showed faster convergence compared to classical MCMC
- **Efficiency:** Quixer's compositional structure captured biological motifs effectively

# Future Work

## **Running our algorithms on quantum hardware**

- Run algorithms on actual quantum computers (H-Series from Quantinuum)
- Compare performance & efficiency on near-term NISQ devices

## **Scaling to Larger Genomes**

- Extend to longer k-mer sequences (6-mers, 8-mers) for richer biological context
- Apply to full chromosome-level sequence prediction
- Process population-level gene sequencing

## **Biological Applications**

- Drug-target interaction prediction and therapeutic design
- Integrate with protein structure prediction pipelines
- Cancer genomics for mutation prediction and therapeutic targeting
- Metagenomics for microbiome analysis

# Thank You!

