QUANTUM REINFORCEMENT LEARNING FOR PORTFOLIO OPTIMIZATION

Vincent Gurgul[✉,1], Ying Chen[2], Stefan Lessmann[1]
[1]Humboldt-Universität zu Berlin, [2]National University of Singapore
[✉]vincent.gurgul@hu-berlin.de

Abstract: This paper presents a fully quantum approach to portfolio optimization using Reinforcement Learning based on Variational Quantum Circuits (VQCs). We reimplement both Deep Deterministic Policy Gradient (DDPG) and Deep Q-Learning as a quantum circuit. Our quantum RL models outperform classical baselines at equivalent parameter counts, demonstrating both higher expressivity and more favorable scaling characteristics. However, due to limitations of current quantum hardware, inference still remains significantly slower than classical algorithms. Our findings suggest that while a quantum advantage is not yet realized for reinforcement learning in practice, it is likely to emerge as quantum hardware matures.

Keywords: quantum finance, quantum machine learning, reinforcement learning, portfolio optimization

INTRODUCTION

Portfolio optimization requires efficient modeling of high-dimensional asset dynamics under uncertainty. Classical Reinforcement Learning (RL) algorithms, while powerful, face challenges in terms of scalability and training efficiency due to large state-action spaces. Variational Quantum Circuits (VQCs), thanks to their entanglement-induced expressivity and potential for parallelism, offer a novel paradigm for approximating complex functions in policy and value-based learning.

Prior work on quantum portfolio optimization has focused largely on quantum annealing, often restricted to low-dimensional cases. Meanwhile, quantum RL remains mostly theoretical. This study implements VQC-based RL algorithms on actual quantum hardware, applies them to a real-world optimization problem with 15 financial assets, and systematically benchmarks the results against classical baselines under both static and dynamic portfolio optimization scenarios.

METHODOLOGY

Each quantum policy and critic network is constructed as a parameterized circuit $U(\theta)$, whose outputs are expectation values of observables $\widehat{B}$:

$$f_{\text{VQC}}(x;\theta) = \langle 0|U^{\dagger}(\theta)\widehat{B}U(\theta)|0\rangle$$

Classical inputs are encoded via amplitude encoding after transformation through a nonlinear feature map:

$$\phi(x) = \left[\tilde{x}, \tilde{x}^{\odot 2}, \sin(\tilde{x}), \cos(\tilde{x})\right], \quad \tilde{x} = \frac{x-\mu}{\sigma}$$

We train on a dataset of 5049 daily observations across 15 assets (equities, fixed income, and real assets). Actions are portfolio weights (allowing for negative weights, i.e. short selling). States include a 30-day window of asset prices and a 7-day Auto-ARIMA forecast. The reward function combines expected return and volatility-adjusted risk with a tunable risk preference parameter η:

$$r = R_p + \eta \cdot \sigma_p$$

where:

$$R_p = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} p_{t,i} \cdot w_{t,i} \quad , \quad \sigma_p = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} p_{t,i} \cdot w_{t,i} - R_p\right)^2}$$

Quantum networks are trained on simulators using parameter shift gradients and deployed on IBM Eagle r3 quantum computers for inference.

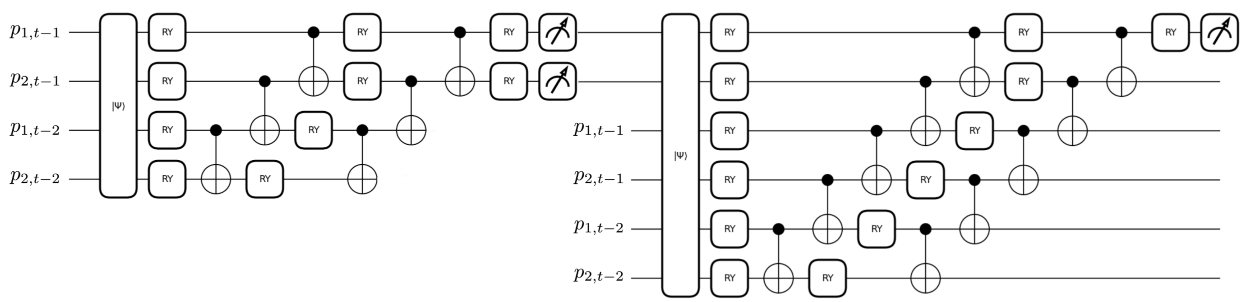

Figure 1: Example architecture of a quantum RL agent for two assets and two time steps.

FINDINGS

Quantum agents consistently outperform classical counterparts when matched by parameter count. For instance, the quantum DDPG agent with just 30 trainable parameters achieves a Sharpe ratio of 0.46, compared to 0.32 for its classical analog. This suggests that even small-scale quantum policies benefit from enhanced representational capacity.

Table 1: Selected Results for Static and Dynamic Portfolio Optimization

	SPO Sharpe	SPO Profit	DPO Sharpe	DPO Profit	Execution Time
Equal Weights Portfolio	0.4375	9.25%	0.4375	9.25%	–
Mean-Variance Optimization	0.5919	25.42%	0.4950	23.24%	0.3 s
Classical DDPG (30 params)	0.3254	18.79%	0.3600	17.63%	3 s[1]
Classical Q-Learning (30 params)	0.4131	10.87%	0.4209	10.68%	4 s[1]
Quantum DDPG (30 params)	0.4586	29.95%	0.4179	22.37%	23 min[2]
Quantum Q-Learning (30 params)	0.4598	29.69%	0.4776	25.43%	43 min[2]
Classical DDPG (160k params)	0.7880	21.57%	0.7926	21.76%	13 s[1]
Classical Q-Learning (160k params)	0.7969	26.85%	0.8237	27.78%	21 s[1]

[1]Hardware: Nvidia RTX 8000 GPU.
[2]Hardware: IBM Eagle r3 QPU.

Quantum models exhibit favorable scaling characteristics, since inference latency is dominated not by circuit width or depth, but by qubit initialization and measurement. However, on current quantum hardware, no speed advantage can be achieved. Classical inference remains several orders of magnitude faster, especially on GPU.

The main bottleneck lies in hardware latency. On systems like the IBM Eagle r3, each inference call incurs approximately 12 seconds of delay due primarily to system preparation and state initialization, rather than gate execution.

Nonetheless, the higher expressivity of quantum models at small sizes coupled with their more favorable scaling suggests a qualitative advantage that could become practically relevant as hardware improves.

CONCLUSIONS

Our findings suggest that quantum RL agents can offer higher performance than classical agents of similar size, due to their superior functional expressiveness. However, this advantage does not currently translate to execution speed, owing to severe latency constraints in present-day QPUs.

While VQCs show promise for reinforcement learning, a quantum advantage in portfolio optimzation will only emerge with significant improvements in quantum hardware.

REFERENCES

Benedetti, M., Lloyd, E., Sack, S., & Fiorentini, M. (2019). Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4), 043001.

Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S. C., Endo, S., Fujii, K., ... & Coles, P. J. (2021). Variational quantum algorithms. *Nature Reviews Physics*, 3, 625-644.

Conti, C. (2024). Quantum Machine Learning. *Quantum Science and Technology*, Springer International.

Kasirajan, V. (2021). *Fundamentals of Quantum Computing*. Springer Nature.

Liu, J., Liu, M., Liu, J. P. et al. (2024). Towards provably efficient quantum algorithms for large-scale machine-learning models. *Nature Communications*, 15, 434.

Mitarai, K., Negoro, M., Kitagawa, M., & Fujii, K. (2018). Quantum circuit learning. *Physical Review A*, 98(3), 032309.