Quantum Amplitude Estimation for Value at Risk

iQuHACK 2026 — State Street × Classiq

Team: Jasper Sands, Leila Erhili, Azain Khalid, Shay Manor, Mallin Kumar
Repository: `https://github.com/ShayManor/WhatTheDuck`

1 Problem Overview

Value at Risk (VaR) is a core financial risk metric defined by a tail-probability constraint. Given a confidence level $\alpha \in (0, 1)$, the Value at Risk is the threshold v such that

$$\mathrm{VaR}_\alpha = \inf \{v \; : \; P(\mathrm{loss} \leq v) \geq \alpha\} .$$

In practice, computing VaR requires repeatedly estimating tail probabilities during a bisection search over candidate thresholds. As a result, *probability estimation* dominates the computational cost of VaR evaluation.

The goal of this project is to implement *Iterative Quantum Amplitude Estimation (IQAE)* for VaR estimation and to rigorously benchmark its scaling behavior against a classical Monte Carlo baseline under identical problem formulations and stopping criteria.

2 Classical and Quantum Approaches

Both the classical and quantum methods solve the same VaR problem using an identical bisection procedure:

- the same confidence level α,
- the same stopping rule,
- the same discretized loss model.

The **only difference** lies in how the tail probability $P(\mathrm{loss} > v)$ is estimated at each bisection step.

2.1 Classical Monte Carlo

Classical Monte Carlo estimates probabilities by random sampling. For N samples, the estimation error scales as

$$\varepsilon = O\left(\frac{1}{\sqrt{N}}\right),$$

which implies a total cost scaling of

$$O\left(\frac{1}{\varepsilon^2}\right).$$

2.2 Iterative Quantum Amplitude Estimation

IQAE estimates probabilities using amplitude amplification and phase estimation techniques. Its estimation error scales as

$$\varepsilon = O\left(\frac{1}{N}\right),$$

leading to a total query complexity of

$$O\left(\frac{1}{\varepsilon}\right).$$

Our objective was not to demonstrate IQAE in isolation, but to empirically validate this scaling advantage within a realistic VaR workflow.

3 Experimental Challenges

Two key challenges made this problem non-trivial in a hackathon setting.

3.1 High-Throughput Evaluation

A single VaR computation requires many probability estimation calls:

- each VaR estimate performs a bisection search,
- each bisection step requires a probability estimate,
- experiments must be repeated across multiple distributions, random seeds, and precision targets ε.

To make any scaling claim credible, we required sufficient trials and consistent methodology across both classical and quantum baselines, all within strict time constraints.

3.2 Circuit Reuse and Compilation Overhead

Naïvely recompiling a quantum circuit for every bisection threshold is prohibitively slow. To address this, we:

- compiled the state preparation circuit once,
- dynamically constructed the threshold comparator and oracle during execution,
- reused the same compiled circuit across thousands of probability evaluations.

This architectural separation was critical for enabling large-scale benchmarking.

4 Error Decomposition

A central component of our analysis is the explicit separation of two independent sources of error.

4.1 Discretization (Modeling) Error

Discretization error arises from approximating continuous loss distributions with a finite grid:

- applies to both classical and quantum methods,
- defines a lower bound on achievable accuracy,
- controlled by the grid resolution (number of bins or qubits).

4.2 Probability Estimation Error

Probability estimation error arises from finite sampling or oracle calls. This is the regime in which IQAE provides a theoretical advantage:

$$\text{Monte Carlo: } O\left(\frac{1}{\sqrt{N}}\right), \qquad \text{IQAE: } O\left(\frac{1}{N}\right).$$

By controlling discretization resolution independently, we ensure that observed performance differences reflect estimation efficiency rather than modeling artifacts.

5 Sensitivity Analyses

We performed systematic sweeps over:

- estimation precision ε,
- confidence level α,
- discretization resolution (grid size or number of qubits).

These analyses show that:

- higher confidence levels significantly increase resource requirements,
- increasing discretization resolution reduces modeling error but does not alter scaling,
- IQAE maintains favorable scaling across realistic parameter ranges.

6 Results

Empirical log–log plots of estimation error versus computational cost confirm that:

- classical Monte Carlo follows the expected $O(1/\sqrt{N})$ trend,
- IQAE exhibits approximately linear-in-ε scaling,

- the scaling advantage persists when bisection, discretization, and multiple distributions are included.

Importantly, all results are benchmarked against optimized classical baselines rather than CPU-bound reference implementations.

7 GPU-Accelerated Optimization

While the asymptotic advantage of IQAE is well understood theoretically, observing this advantage empirically within hackathon constraints required aggressive optimization of both quantum and classical workflows. GPU acceleration was critical to enabling large-scale parameter sweeps and statistically meaningful comparisons.

7.1 Hyperparameter Tuning at Scale

We performed systematic sweeps over multiple hyperparameters, including:

- estimation precision ε,
- confidence level α,
- discretization resolution (grid size or number of qubits),
- loss distribution families and random seeds.

Each configuration required repeated probability estimation calls inside a VaR bisection loop. GPU acceleration enabled parallel evaluation of these configurations, allowing us to collect median and interquartile statistics rather than relying on single-run measurements. This throughput was essential for producing stable log–log scaling plots and avoiding conclusions driven by noise.

7.2 Improved Classical Monte Carlo Baseline

To ensure a fair comparison, we implemented an optimized classical Monte Carlo baseline rather than relying on a naïve CPU implementation. Our classical pipeline includes:

- GPU-native sampling,
- batched evaluation of loss functions,
- importance sampling variants for tail probability estimation.

This strengthened baseline ensures that any observed advantage of IQAE persists even when classical Monte Carlo is aggressively optimized, rather than serving as a strawman comparison.

8 GPU-Accelerated IQAE Execution

For the quantum workflow, GPU acceleration was used to reduce execution overhead in repeated IQAE calls. Probability estimation via IQAE requires many circuit executions with varying Grover iteration counts. Efficient batching and statevector simulation on GPUs made it feasible to evaluate IQAE across a wide range of precision targets and confidence levels.

By separating reusable state preparation from threshold-dependent oracle logic, we avoided recompilation overhead during bisection. This allowed a small set of compiled circuits to be reused tens of thousands of times across experiments, enabling high-throughput IQAE execution within the hackathon time budget.

9 Tail-Prioritized Discretization

Discretization plays a dual role in quantum VaR estimation: it introduces modeling error, but also determines how effectively tail probabilities are resolved. Uniform discretization can waste resolution in low-impact regions of the distribution while undersampling the tail.

To address this, we explored tail-prioritized grid discretization strategies, allocating more resolution to regions near the VaR threshold. This improves effective accuracy for fixed circuit width by concentrating representational power where estimation precision matters most.

While this approach does not change the asymptotic scaling of IQAE, it reduces constant factors and improves empirical performance under fixed qubit and depth constraints.

10 End-to-End Performance Summary

Combining GPU-accelerated execution, optimized classical baselines, circuit reuse, and tail-aware discretization enabled an end-to-end evaluation of quantum versus classical VaR estimation at scale. These system-level optimizations were essential for translating theoretical complexity advantages into observable empirical behavior.

11 Final Conclusion

This work presents a complete, system-aware benchmark of quantum amplitude estimation for Value at Risk. By holding the problem formulation fixed, explicitly separating modeling and estimation error, optimizing both classical and quantum pipelines, and leveraging GPU acceleration, we demonstrate that IQAE's theoretical scaling advantage can be meaningfully observed in a realistic VaR workflow.

Rather than a single proof-of-concept experiment, our results emphasize fairness, repeatability, and practical execution considerations—key requirements for evaluating quantum algorithms in financial risk analysis.