How these studies are made

Six method studies on quantum computing for finance, published in August and September 2026 — all out of the same machinery. We call it the **RQP Evidence Engine** — and here is the machinery itself, **so you can judge it before you judge the findings.**

6

method studies

all public, with raw results and scripts

1

hardware rung

ibm_kingston, 100,000 shots, raw

5

verdicts, closed

including “No advantage demonstrated”

THE CENTRAL RULE

No quantum method grades its own homework.

In the optimization studies, every readout is scored against a **classically certified optimum** — proven by exhaustive enumeration or branch-and-bound. “The circuit found the answer” is a checkable fact, not a claim.

In the machine-learning studies, every arm faces **ten tuned classical variants** — including an order-matched twin built to imitate the quantum model classically. Advantage is claimed only when the **whole 95% confidence interval** clears parity.

The referee is classical — and it is not ours to influence.

BEFORE THE FIRST RUN

Hypotheses and design, locked before execution.

Written down first

Every study's hypotheses, arms and analysis plan are **locked in writing before the first run** — and every deviation is disclosed in the study's appendix.

Publicly tagged — before the first run

From the newest study onward the locked design is **pushed and tagged in a public repository before execution**. The chronology is verifiable by anyone — not asserted by us.

Reproduced before printed

Every published number is **re-derived from the committed raw results by a committed script** before it appears in a document. If script and prose disagree, the build fails. Twice it did — both corrections are disclosed as dated revisions.

WHY YOU CAN TRUST IT

What the machinery publishes against itself.

Our warm start: 0 of 96 escapes

Our own products' classical warm start, given a slightly wrong seed, **never escaped** on the real books — printed as a headline finding, not buried.

Our best circuit on real hardware: 0 hits in 20,000

The adaptive circuit's entire simulated advantage is **erased on today's device** — and a prediction built from the device's own calibration data was optimistic by about **×3**.

Our own checklist: 83 of 84 wrong vetoes

The screening checklist we shipped wrongly vetoed 83 of 84 datasets on which the quantum kernel then **won**. We measured it, printed it, and replaced it.

A machine built for marketing does not produce these sentences.

THE RECEIPTS · AUGUST–SEPTEMBER 2026

Six studies, all public.

1 · Which quantum method — and when?

12,397 certified runs across 24 portfolio instances. The answer is a decision tree, and it now ships as the products' defaults.

2 · Does trained depth survive a real device?

A noise ladder from clean simulation to **ibm_kingston**. Today: essentially no — and measuring the device is cheaper than predicting it (44 seconds of processor time).

3 · Where can a quantum kernel win at all?

996 refereed arms map the territory between two walls — and the map now ships in the QML product as a printed diagnostic.

4 · Does the trainable model own territory?

65 fair-recipe trainings, design publicly tagged before the first run. Yes — size-gated, with a measured crossover.

5 · Which kernel settings can be measured at all?

7,200 measurements, 150 refereed, design deposited before the first run. Most settings above 8 qubits cannot be measured; one encoded nothing; four apparent wins were classical in disguise.

6 · Which trainable-model settings can be trained?

7,488 pre-training screens, 80 refereed, design tagged before the first run. The product's default region is trainable to 16 qubits and wins on its native labels; one measurement is dead at scale; a one-minute screen now ships.

WHY IT MATTERS

The studies are how the tools improve.

Every finding shipped back into the products: the decision tree became the **default guidance**, the noise ladder became a **printed hardware warning**, the island map became a **diagnostic panel**, the fair-run recipe became the **training defaults**.

And the machinery is not reserved for our own questions. **“Should we use method X on problem Y?”** — your data, your books, the same referees, the same publication standard. That path is described in the methodology paper.

Findings become defaults — which is why the verdicts stay honest.

THE PAPER

Judge the machinery, then the findings.

The 8-page methodology paper describes the engine in full — the referee, the registration chronology, the reproduction scripts, the receipts — and the four studies are public with **raw results and revision notes**.

The methodology paper — 8 pages **qubit-lab.ch/evidence**

The RQP Evidence Engine paper, the four study PDFs and the machine-readable run records are on qubit-lab.ch/evidence. Evidence, not promises.