QUBIT-LAB.CH · STUDY 5 OF 5

Which quantum-kernel settings can be measured at all

— and why measurable is not the same as winning. The product offers five encodings, repeats, scaling and shot budgets. We measured **every combination**, verified 150 of them against a full classical referee, and turned the answer into **guardrails in the product**. Design deposited before the first run.

1,200

configurations

× 6 datasets = 7,200 kernel matrices

150

refereed runs

ten classical variants, whole-CI win rule

0

genuine wins

4 at first sight — all classical in disguise

THE SETUP

A similarity score, a shot budget, and a referee.

A quantum kernel classifier compares data points through a quantum **similarity score**; a classical algorithm draws the boundary. Scores are read out by repeated measurement — **shots** — and the theory warns that many settings make the scores collapse toward one value, so that no realistic shot budget can tell rows apart: **concentration**.

The study measured the product's own concentration diagnostic on every setting it offers (in seconds, no training), then ran 150 settings through the full refereed pipeline: **ten tuned classical models** the quantum classifier must beat, with the whole 95% confidence interval above parity.

Cheap to screen — expensive to verify. Both layers, both printed.

FINDINGS 1–2 · THE MAP

Above 8 qubits, most settings cannot be measured.

Measurable at 512 shots: 100 / 86 / 33 / 25%

at 4 / 8 / 12 / 16 qubits. Even 32,768 shots reach only **36%** of the 16-qubit settings. More repeats and entangling encodings make it worse — exactly as the theory predicts (supported in 48 of 48 groups).

Input scaling is the lever — and our bet lost

We bet that the scaling knob (the *bandwidth*) would matter more than the choice of encoding at every size. It did not: one dead encoding swamps the comparison. Among live encodings, scaling leads from 12 qubits — printed as post-hoc, not as a result.

At 16 qubits, one corner survives

Only the **narrowest scaling with one or two repeats** stays measurable at 512 shots. The regime map ships in the product as configuration warnings.

FINDING 3 · DEAD ON ARRIVAL

One offered setting encoded nothing.

The angle encoding with a Z rotation gives **every row the same quantum state** — the similarity score is 1 for every pair, in all 1,440 cells. It was selectable in the product.

Worse: the diagnostic read it as “trivially measurable” — a blind spot for concentration **toward one**, the opposite direction from the theory's warning. Both are fixed: the setting is rejected with an explanation, and constant scores are flagged in every report.

A study that catches its own product’s mistake and prints it on the cover.

FINDINGS 4–5 · THE REFEREE

Measurable is not sufficient.

4 apparent wins → 0

All four used the simplest encoding, whose score has an **exact classical formula** (identical to 10^{-15}). Once that formula sits in the referee, the wins vanish. The product's referee now contains it: that encoding can at best reach **Parity**.

Most settings memorize

Training accuracy 0.99, new data near chance: a kernel that is nearly the identity matrix is perfectly measurable and perfectly useless. **Read both faces of the score matrix** — can it be measured, and does it still tell rows apart.

Two more registered bets, lost

The diagnostic did not predict which settings lose under shot noise (inconclusive), and winners did not sit near the shot-budget line (0 of 4). Printed with the same prominence as the findings.

WHAT IT CHANGES

Guardrails over freedom.

Do not run

settings that encode nothing or cannot be measured economically — the product now rejects the first and warns on the second.

Do not claim quantum value

where an exact classical twin exists — the referee now contains it.

Screen before training

17 seconds per setting against a minute-plus refereed verdict — screening rejects the dead; it does not prove the living.

A configuration screen should expose evidence-based choices, not every mathematically legal circuit.

THE STUDY

The full study, unedited.

Ten pages: the regime maps, five pre-registered hypotheses with their outcomes, the classical-twin identity, the product changes, and the registration — design tagged and deposited (DOI 10.5281/zenodo.22498143) **before the first run**.

The study — 10 pages **qubit-lab.ch/evidence**

Study PDF, published benchmark records and the QML Rapid Quantum Prototype are on qubit-lab.ch/evidence. Evidence, not promises.