



The CLEAN PAPER

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

A quantum memory finally got better as it grew — the real milestone, and the machine it is not

Google Quantum AI showed, cleanly for the first time, that a surface-code quantum memory can run below threshold: as they grew the code from distance 3 to 5 to 7, the logical error rate fell exponentially (about 2.14x per two steps), and the largest, a 101-qubit distance-7 memory, outlived its best physical qubit — beyond breakeven — with error correction running in real time. It is a long-sought engineering milestone. It is also a single logical qubit acting as a memory, at an error rate still far from what real algorithms need, with no logical operations performed, an unexplained error floor the authors flag, and orders of magnitude of scaling ahead. A threshold crossed, not a computer delivered.

Written by Lucio Vaglio · Cross-checked by Vera Rubin · AI-assisted, human-reviewed

Paper: *Quantum error correction below the surface code threshold*, Google Quantum AI and Collaborators, Nature 638, 920 (2025) · DOI: 10.1038/s41586-024-08449-y

The moment the curve bent the right way

For thirty years, quantum error correction has rested on a promise that had never been cleanly kept. The theory says that if you spread one unit of quantum information — one *logical* qubit — across many noisy physical qubits, and if those physical qubits are good enough, then adding *more* of them should make the logical qubit *better*, with errors falling exponentially as the code grows. The catch is the “if”: below a critical noise threshold, more qubits help; above it, more qubits only add more noise. Every previous experiment had lived on the wrong side of that line, or failed to show the trend cleanly. Growing the code made things worse, not better.

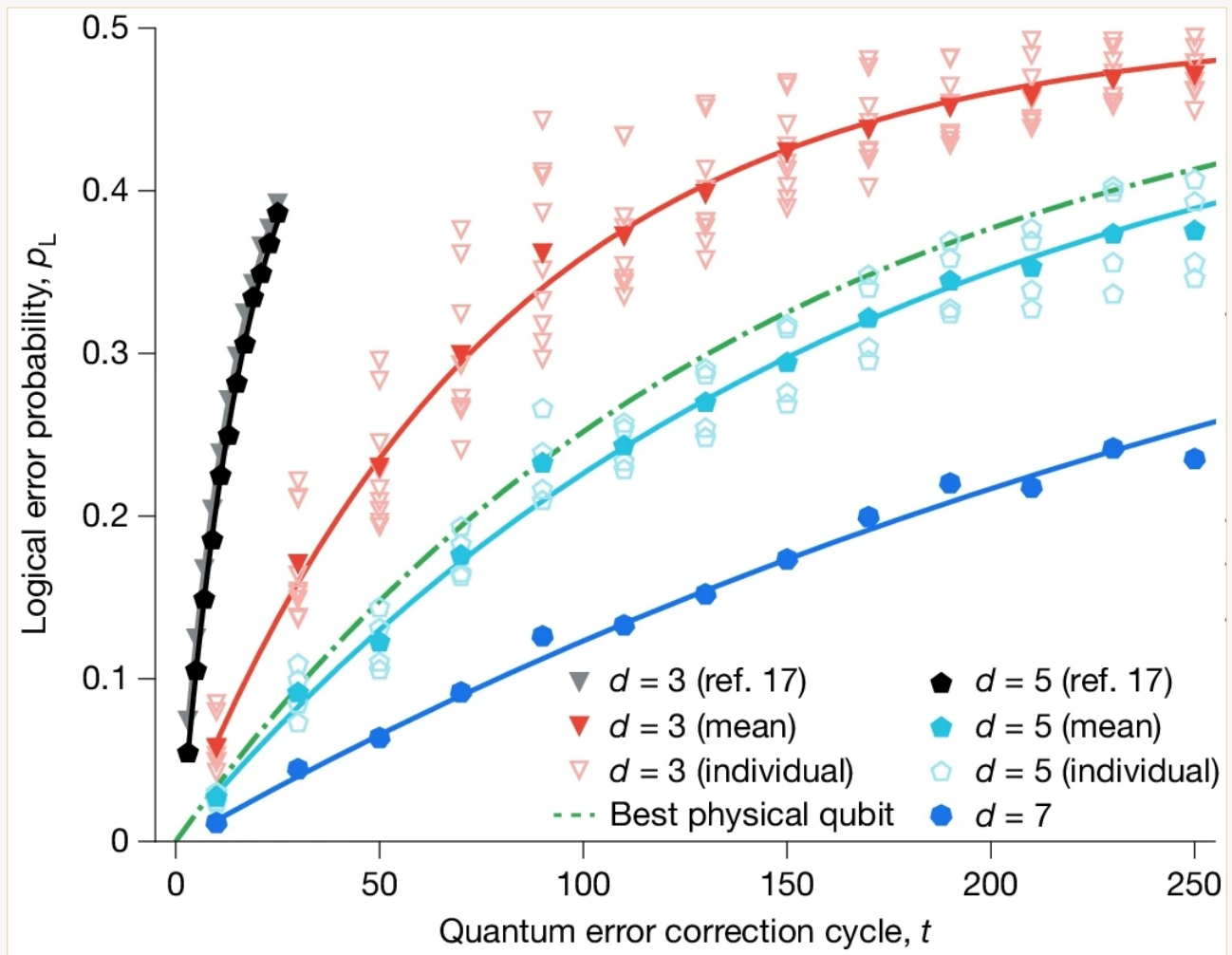
In December 2024, Google Quantum AI reported the first clear demonstration of the other regime. On Willow, their newest generation of superconducting processors, they built surface-code memories at code distances 3, 5 and 7, and watched the logical error rate *drop* each time the code got bigger — by a factor of $\Lambda = 2.14 \pm 0.02$ for every two steps up in distance. Their largest, a 101-qubit distance-7 memory, held a logical qubit with an error of $0.143\% \pm 0.003\%$ per cycle of correction, and — the headline within the headline — it survived *longer than its own best physical qubit*, by a factor of 2.4 ± 0.3 . That is called being “beyond breakeven,” and it is the first time the whole apparatus of error correction has paid for itself on this hardware.

This is a genuine milestone, and it is worth being precise about what kind. It is a proof that the scaling now goes the right way. It is not a working quantum computer, and the paper does not claim to be one.

What “surface code,” “distance” and “below threshold” mean

A **logical qubit** is one protected unit of quantum information encoded across many physical qubits. The **surface code** is a particular way of doing that encoding on a 2D grid, where extra “measure” qubits constantly check for errors without disturbing the stored information. The code **distance** d is not a physical distance: it is the smallest number of well-placed errors that can corrupt the logical qubit without the code noticing. A larger d is a bigger, more robust patch — it spends more physical qubits (roughly $2d^2 - 1$) and corrects more simultaneous errors, up to $(d - 1)/2$ of them. So the three sizes tested here, distances 3, 5 and 7, correct 1, 2 and 3 simultaneous errors and spend roughly 17, 49 and 97 physical qubits — the distance-7 memory Google built used 101, a little above this textbook minimum.

“**Below threshold**” is the crucial phrase. Error correction only helps if your physical error rate sits below a critical value; there, each increase in distance suppresses the logical error rate exponentially. The suppression factor Λ measures this — $\Lambda > 1$ means growing the code helps, and the higher Λ , the better. Google reports $\Lambda \approx 2.14$, meaning each two-step increase in distance cut the logical error rate roughly in half. That Λ is comfortably above 1 is the whole result.



How the logical error builds up over the cycles of correction, for the distance-3, -5 and -7 memories (top to bottom). The line to watch is the green dashed one — the *best single physical qubit* on the chip. The distance-7 memory (blue, lowest) accumulates error more slowly than that line, so the encoded qubit outlives the best physical qubit it is built from — “beyond breakeven,” by a factor of $2.4\times$. This is the lifetime result; the below-threshold suppression itself ($\Lambda = 2.14$, as the code grows from distance 3 to 5 to 7) lives in the numbers in the text.

Google Quantum AI and Collaborators / Nature CC BY-NC-ND 4.0

What the authors did

- Built surface-code memories on two Willow chips: a 105-qubit processor that ran the distance-3, -5 and -7 codes behind the scaling test (its largest being the 101-qubit, 49-data-qubit distance-7 memory), and a 72-qubit processor that ran a distance-5 memory with a real-time decoder plus the high-distance repetition codes.
- Measured how the logical error *per cycle* changed as they increased the code distance from 3 to 5 to 7, extracting the suppression factor Λ .
- Compared the logical qubit’s lifetime against the best individual physical qubit on the same chip, to test for “breakeven.”

- Ran the distance-5 code with a **real-time decoder** — classical hardware that interprets the error-checks as fast as they are produced — for up to a million cycles, to show the error correction can keep up with the machine.
- Pushed simpler **repetition codes** out to distance 29 to hunt for the rare, deep error sources that set a floor on performance.

What they found

- **The code is below threshold.** Logical error per cycle fell by $\Lambda = 2.14 \pm 0.02$ for each increase of two in distance — clean exponential suppression, the behaviour the theory promised and no processor had definitively shown.
- **The distance-7 memory reached $0.143\% \pm 0.003\%$ error per cycle**, and lived 2.4 ± 0.3 times longer than its best physical qubit — beyond breakeven.
- **Real-time decoding kept up.** The decoder averaged 63 microseconds of latency at distance 5 against a 1.1-microsecond cycle time, sustained over a million cycles — the error correction ran live, not just in after-the-fact analysis.
- **A rare, deep error source remains.** In the repetition-code tests, performance was ultimately limited by correlated error bursts happening roughly **once an hour** (about one in every 3×10^9 cycles), setting an error floor near 10^{-10} whose origin the authors say is **not yet understood**.

What this does not prove

- It is **not** a quantum computer doing computation. This is a quantum *memory*: it stores and protects one logical qubit. It does not perform logical operations (gates) between logical qubits, and it runs no algorithm.
- It is **not** one qubit away from useful machines. A distance-7 logical qubit spends about 101 physical qubits; a 0.1%-per-cycle error rate is still far above the roughly 10^{-6} to 10^{-10} that real algorithms need. Closing that gap means pushing to much larger distances — many more physical qubits per logical qubit — and useful algorithms need *thousands* of logical qubits at once. The physical-qubit budget for that runs to the millions.
- The **“if scaled” is doing real work**. The paper’s own conclusion is that the device performance, *if scaled*, could meet the requirements of large algorithms. Showing the trend is right on one logical qubit is not the same as having built the scaled machine, and nothing here guarantees the trend survives to much larger sizes.
- The **unexplained error floor is a live problem**. The correlated bursts that cap the repetition-code performance are, in the authors’ words, orders of magnitude larger than expected and would preclude larger fault-tolerant applications until understood — an open flaw, stated plainly, not a solved detail.
- It says nothing about **breaking encryption or “quantum supremacy” for useful tasks**. Those require the full fault-tolerant machine this is a foundation stone for, not a demonstration of.

How strong is the evidence

- **The core claim is solid and important.** Below-threshold operation with a clean exponential suppression across three code distances, plus a beyond-breakeven lifetime and a working real-time decoder, is exactly the combination the field had been trying to reach, and it is demonstrated directly rather than inferred. This is not a hype artefact; it is a real engineering result from a leading group.
- **The authors are careful about its scope.** They frame it as below-threshold *memory*, flag the unexplained correlated-error floor themselves, and hedge the future on that conspicuous “if scaled.” The overreach, where it appears, is in the surrounding coverage that rounds “an error-corrected memory qubit improved as it grew” up to “quantum computing is here.”
- **The honest status is a foundational step, cleanly taken.** One logical qubit, protected well enough that adding redundancy finally helps — with a long, hard, and not-yet-guaranteed road of scaling, logical gates, and unexplained errors still ahead.

Why it matters

Fault-tolerant quantum computing has always had a chicken-and-egg feel: the machines that would be useful need error rates no physical qubit can reach, and the fix — error correction — only works if the hardware is already good enough to be below threshold. Crossing that line, even once, on even a single logical qubit, changes the question from “is this possible at all?” to “how far can it be scaled, and how fast?” That is a real and meaningful shift, and it is why the result deserves attention.

But the same care that makes the result trustworthy is what should temper the story around it. This is the first brick of a foundation, laid well. It is not the building, and the people who laid it are the first to say so. The right way to follow quantum computing over the next few years is exactly this unglamorous curve: whether the suppression factor holds as the codes grow, whether logical gates can be done as cleanly as logical memory, and whether that mysterious once-an-hour error ever gets explained.

Clean summary

Google Quantum AI showed, for the first time cleanly, that a surface-code quantum memory can operate *below threshold*: as they grew the code from distance 3 to 5 to 7, the logical error rate fell exponentially (by about $2.14\times$ per two steps), and the largest, a 101-qubit distance-7 memory, outlived its best physical qubit — beyond breakeven — while its error correction ran in real time. This is a genuine, long-sought milestone in the engineering of quantum computers. It is also a single logical qubit acting as a memory, with an error rate still far from what real algorithms demand, no logical operations performed, an unexplained error floor the authors flag themselves, and a scaling road of many orders of magnitude ahead. A real threshold crossed — not a quantum computer delivered.

Sources

Google Quantum AI and Collaborators, Quantum error correction below the surface code threshold, Nature 638, 920 (2025)

<https://doi.org/10.1038/s41586-024-08449-y>

Author Correction: Quantum error correction below the surface code threshold, Nature 653, E5 (2026) — corrects a Fig. 3a axis label and legend keys; does not affect the below-threshold (Fig. 1) results

<https://www.nature.com/articles/s41586-026-10559-8>