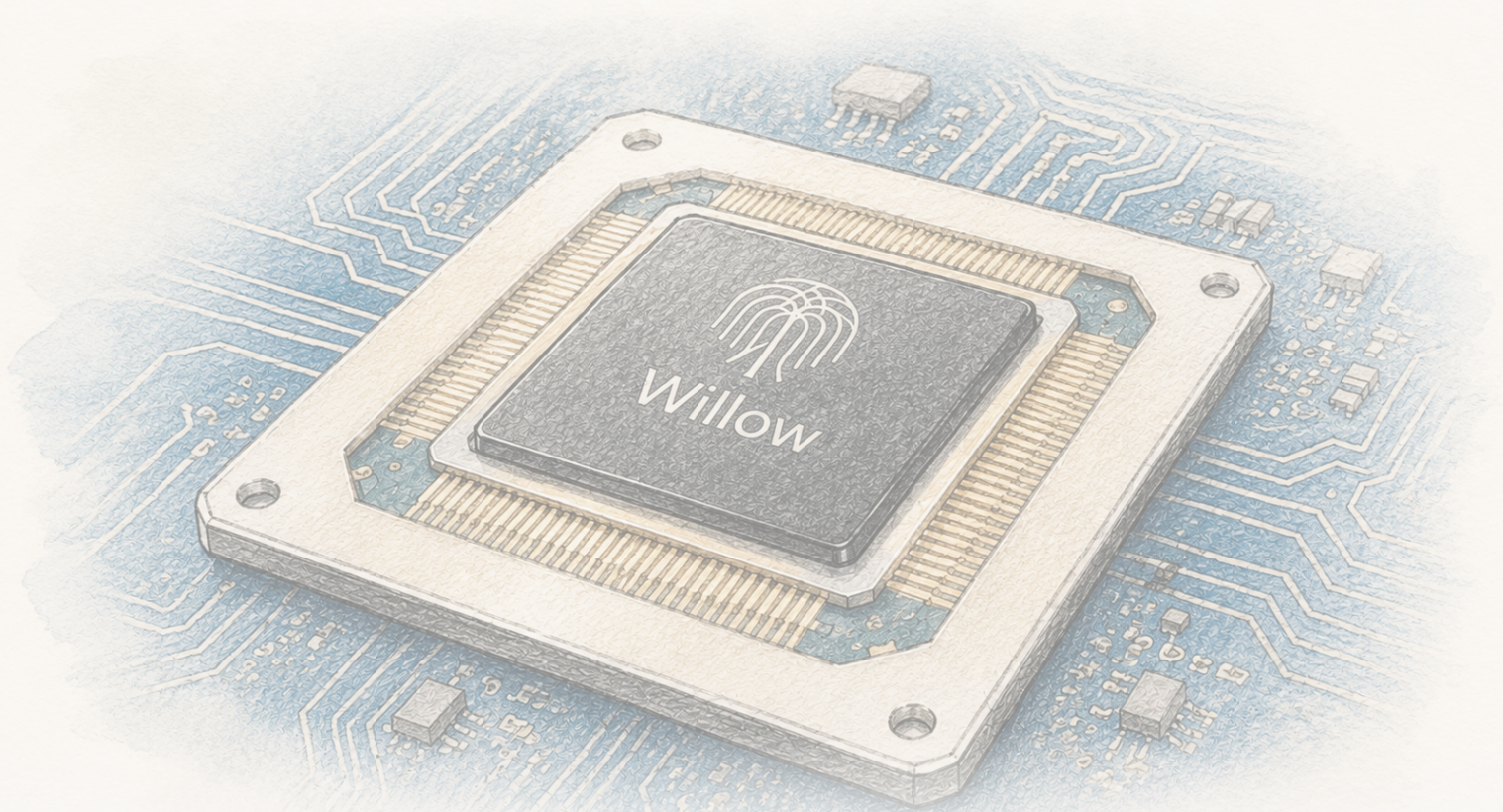




The CLEAN PAPER

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

Ein Quantenspeicher wurde endlich besser, als er wuchs — der echte Meilenstein, und die Maschine, die er nicht ist

Google Quantum AI zeigte zum ersten Mal sauber, dass ein Surface-Code-Quantenspeicher unterhalb der Schwelle laufen kann: Während der Code von Distanz 3 über 5 auf 7 wuchs, fiel die logische Fehlerrate exponentiell (etwa $2,14\times$ je zwei Schritte), und der größte, ein Distanz-7-Speicher mit 101 Qubits, überlebte sein bestes physisches Qubit — beyond breakeven —, mit Fehlerkorrektur in Echtzeit. Es ist ein lange gesuchter Ingenieurs-Meilenstein. Es ist zugleich ein einzelnes logisches Qubit als Speicher, mit einer Fehlerrate noch weit von dem entfernt, was echte Algorithmen brauchen, ohne ausgeführte logische Operationen, mit einem unerklärten Fehlerboden, den die Autoren selbst benennen, und mit Größenordnungen an Skalierung vor sich. Eine Schwelle überschritten, kein Computer geliefert.

Written by Lucio Vaglio · Cross-checked by Vera Rubin · Edited by Michele Renda · Figures by Nova Vela · 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

Der Moment, in dem sich die Kurve in die richtige Richtung bog

Dreißig Jahre lang ruhte die Quantenfehlerkorrektur auf einem Versprechen, das nie sauber eingelöst worden war. Die Theorie sagt: Verteilt man eine Einheit Quanteninformation – ein *logisches* Qubit – über viele verrauschte physische Qubits, und sind diese physischen Qubits gut genug, dann sollte das Hinzufügen von *mehr* Qubits das logische Qubit *besser* machen, mit exponentiell fallenden Fehlern, während der Code wächst. Der Haken ist das „wenn“: Unterhalb einer kritischen Rauschschwelle helfen mehr Qubits; oberhalb fügen mehr Qubits nur mehr Rauschen hinzu. Jedes frühere Experiment hatte auf der falschen Seite dieser Linie gelebt oder den Trend nicht sauber gezeigt. Den Code zu vergrößern machte die Dinge schlimmer, nicht besser.

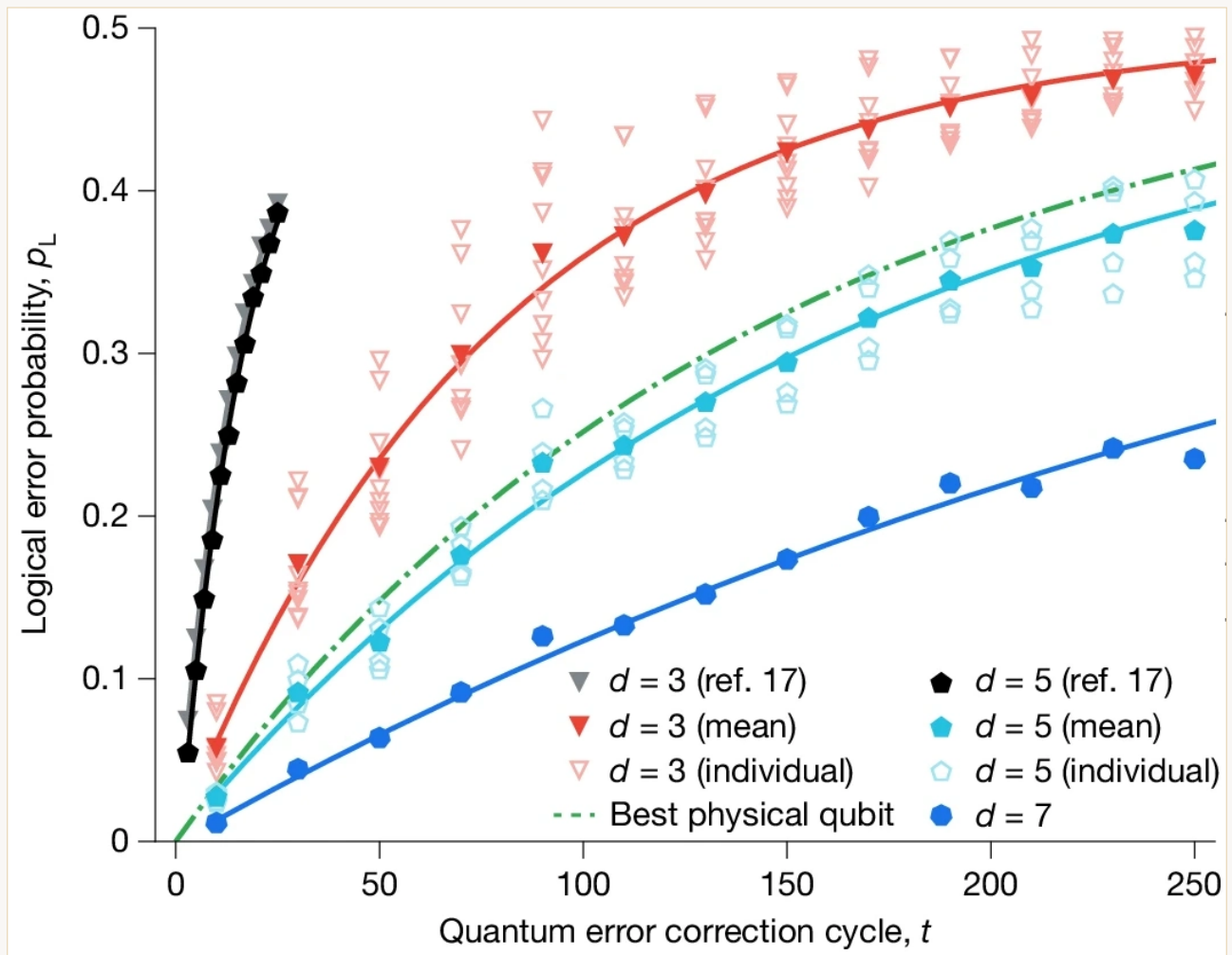
Im Dezember 2024 berichtete Google Quantum AI die erste klare Demonstration des anderen Regimes. Auf Willow, ihrer neuesten Generation supraleitender Prozessoren, bauten sie Surface-Code-Speicher mit den Code-Distanzen 3, 5 und 7 und sahen die logische Fehlerrate mit jeder Vergrößerung des Codes *fallen* – um einen Faktor $\Lambda = 2,14 \pm 0,02$ je zwei Distanzschritte. Ihr größter, ein Distanz-7-Speicher mit 101 Qubits, hielt ein logisches Qubit mit einem Fehler von $0,143\% \pm 0,003\%$ pro Korrekturzyklus, und – die Schlagzeile in der Schlagzeile – es überlebte *länger als sein bestes physisches Qubit*, um einen Faktor von $2,4 \pm 0,3$. Das nennt man „beyond breakeven“, und es ist das erste Mal, dass sich der gesamte Apparat der Fehlerkorrektur auf dieser Hardware selbst bezahlt gemacht hat.

Das ist ein echter Meilenstein, und es lohnt sich, präzise zu sagen, welcher Art. Es ist der Beleg, dass die Skalierung jetzt in die richtige Richtung läuft. Es ist kein funktionierender Quantencomputer, und das Paper behauptet das auch nicht.

Was „Surface Code“, „Distanz“ und „unterhalb der Schwelle“ bedeuten

Ein **logisches Qubit** ist eine geschützte Einheit Quanteninformation, codiert über viele physische Qubits. Der **Surface Code** ist eine bestimmte Art, diese Codierung auf einem 2D-Gitter vorzunehmen, bei der zusätzliche „Mess“-Qubits ständig nach Fehlern suchen, ohne die gespeicherte Information zu stören. Die Code-**Distanz** d ist keine räumliche Entfernung: Sie ist die kleinste Zahl gut platzierter Fehler, die das logische Qubit verfälschen können, ohne dass der Code es bemerkt. Ein größeres d ist ein größeres, robusteres Feld – es kostet mehr physische Qubits (etwa $2d^2 - 1$) und korrigiert mehr gleichzeitige Fehler, bis zu $(d - 1)/2$ davon. Die drei hier getesteten Größen, Distanz 3, 5 und 7, korrigieren also 1, 2 und 3 gleichzeitige Fehler und kosten rund 17, 49 und 97 physische Qubits – der von Google gebaute Distanz-7-Speicher nutzte 101, etwas über diesem Lehrbuch-Minimum.

„**Unterhalb der Schwelle**“ („below threshold“) ist die entscheidende Formel. Fehlerkorrektur hilft nur, wenn die physische Fehlerrate unter einem kritischen Wert liegt; dort unterdrückt jede Distanzerhöhung die logische Fehlerrate exponentiell. Der Unterdrückungsfaktor Λ misst das – $\Lambda > 1$ heißt, das Wachsen des Codes hilft, und je höher Λ , desto besser. Google berichtet $\Lambda \approx 2,14$: Jede Distanzerhöhung um zwei halbierte die logische Fehlerrate ungefähr. Dass Λ komfortabel über 1 liegt, ist das ganze Ergebnis.



Wie sich der logische Fehler über die Korrekturzyklen aufbaut, für die Distanz-3-, -5- und -7-Speicher (von oben nach unten). Die Linie, auf die es ankommt, ist die grün gestrichelte — das *beste einzelne physische Qubit* auf dem Chip. Der Distanz-7-Speicher (blau, am niedrigsten) sammelt Fehler langsamer an als diese Linie, das codierte Qubit überlebt also das beste physische Qubit, aus dem es gebaut ist — „beyond breakeven“, um einen Faktor $2,4\times$. Das ist das Lebensdauer-Ergebnis; die Unterdrückung unterhalb der Schwelle selbst ($\Lambda = 2,14$, während der Code von Distanz 3 über 5 auf 7 wächst) steckt in den Zahlen im Text.

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

Was die Autoren getan haben

- Sie bauten Surface-Code-Speicher auf zwei Willow-Chips: einem 105-Qubit-Prozessor, auf dem die Distanz-3-, -5- und -7-Codes hinter dem Skalierungstest liefen (der größte davon der Distanz-7-Speicher mit 101 Qubits und 49 Daten-Qubits), und einem 72-Qubit-Prozessor, der einen Distanz-5-Speicher mit Echtzeit-Decoder sowie die Hochdistanz-Wiederholungs_codes ausführte.
- Sie maßen, wie sich der logische Fehler *pro Zyklus* änderte, als sie die Code-Distanz von 3 über 5 auf 7 erhöhten, und extrahierten den Unterdrückungsfaktor Λ .
- Sie verglichen die Lebensdauer des logischen Qubits mit dem besten einzelnen physischen Qubit auf demselben Chip — der Test auf „Breakeven“.

- Sie ließen den Distanz-5-Code mit einem **Echtzeit-Decoder** laufen — klassische Hardware, die die Fehlerprüfungen so schnell interpretiert, wie sie anfallen — über bis zu einer Million Zyklen, um zu zeigen, dass die Fehlerkorrektur mit der Maschine Schritt hält.
- Sie trieben einfachere **Wiederholungscodes** bis Distanz 29, um die seltenen, tiefen Fehlerquellen aufzuspüren, die dem Leistungsvermögen einen Boden setzen.

Was sie fanden

- **Der Code liegt unterhalb der Schwelle.** Der logische Fehler pro Zyklus fiel um $\Lambda = 2,14 \pm 0,02$ je Distanzerhöhung um zwei — saubere exponentielle Unterdrückung, das Verhalten, das die Theorie versprach und kein Prozessor definitiv gezeigt hatte.
- **Der Distanz-7-Speicher erreichte 0,143 % $\pm$ 0,003 % Fehler pro Zyklus** und lebte **2,4 $\pm$ 0,3-mal länger** als sein bestes physisches Qubit — beyond breakeven.
- **Die Echtzeit-Decodierung hielt Schritt.** Der Decoder lag bei Distanz 5 im Mittel bei 63 Mikrosekunden Latenz gegenüber einer Zykluszeit von 1,1 Mikrosekunden, durchgehalten über eine Million Zyklen — die Fehlerkorrektur lief live, nicht nur in nachträglicher Analyse.
- **Eine seltene, tiefe Fehlerquelle bleibt.** In den Wiederholungscode-Tests wurde die Leistung letztlich durch korrelierte Fehlerausbrüche begrenzt, die etwa **einmal pro Stunde** auftreten (rund einer je 3×10^9 Zyklen) und einen Fehlerboden nahe 10^{-10} setzen, dessen Ursprung nach Angaben der Autoren **noch nicht verstanden** ist.

Was das nicht beweist

- Es ist **kein** Quantencomputer, der rechnet. Dies ist ein Quanten-Speicher: Er speichert und schützt ein logisches Qubit. Er führt keine logischen Operationen (Gatter) zwischen logischen Qubits aus und lässt keinen Algorithmus laufen.
- Es ist **nicht** ein Qubit von nützlichen Maschinen entfernt. Ein logisches Distanz-7-Qubit kostet rund 101 physische Qubits; eine Fehlerrate von 0,1 % pro Zyklus liegt noch weit über den etwa 10^{-6} bis 10^{-10} , die echte Algorithmen brauchen. Diese Lücke zu schließen heißt, zu viel größeren Distanzen zu gehen — viel mehr physische Qubits pro logischem Qubit —, und nützliche Algorithmen brauchen *Tausende* logischer Qubits zugleich. Das Budget an physischen Qubits dafür geht in die Millionen.
- Das „**wenn skaliert**“ **leistet echte Arbeit**. Die eigene Schlussfolgerung des Papers ist, dass die Geräteleistung, *wenn skaliert*, die Anforderungen großer Algorithmen erfüllen könnte. Zu zeigen, dass der Trend auf einem logischen Qubit stimmt, ist nicht dasselbe wie die skalierte Maschine gebaut zu haben, und nichts hier garantiert, dass der Trend bis zu viel größeren Größen überlebt.
- Der **unerklärte Fehlerboden ist ein offenes Problem**. Die korrelierten Ausbrüche, die die Wiederholungscode-Leistung deckeln, sind, in den Worten der Autoren, um Größenordnungen größer als erwartet und würden größere fehlertolerante Anwendungen ausschließen, bis sie verstanden sind — ein offener Makel, klar benannt, kein erledigtes Detail.
- Es sagt nichts über **das Brechen von Verschlüsselung oder „Quantenüberlegenheit“ bei nüt-**

zlichen Aufgaben. Dafür braucht es die volle fehlertolerante Maschine, für die dies ein Grundstein ist — nicht deren Demonstration.

Wie belastbar sind die Belege?

- **Die Kernaussage ist solide und wichtig.** Betrieb unterhalb der Schwelle mit sauberer exponentieller Unterdrückung über drei Code-Distanzen, dazu eine Lebensdauer jenseits des Breakeven und ein funktionierender Echtzeit-Decoder — genau die Kombination, die das Feld zu erreichen versuchte, und sie wird direkt demonstriert, nicht erschlossen. Das ist kein Hype-Artefakt; es ist ein echtes Ingenieursergebnis einer führenden Gruppe.
- **Die Autoren sind sorgfältig mit der Reichweite.** Sie rahmen es als *Speicher* unterhalb der Schwelle, benennen den unerklärten korrelierten Fehlerboden selbst und hängen die Zukunft an das auffällige „wenn skaliert“. Die Übertreibung, wo sie auftaucht, steckt in der umgebenden Berichterstattung, die „ein fehlerkorrigierter Speicher-Qubit wurde beim Wachsen besser“ zu „das Quantencomputing ist da“ aufrundet.
- **Der ehrliche Status: ein Fundamentschritt, sauber gesetzt.** Ein logisches Qubit, gut genug geschützt, dass Redundanz endlich hilft — mit einem langen, harten und noch nicht garantierten Weg aus Skalierung, logischen Gattern und unerklärten Fehlern vor sich.

Warum das wichtig ist

Fehlertolerantes Quantencomputing hatte immer etwas von Henne und Ei: Die Maschinen, die nützlich wären, brauchen Fehlerraten, die kein physisches Qubit erreicht, und die Abhilfe — Fehlerkorrektur — funktioniert nur, wenn die Hardware bereits gut genug ist, um unterhalb der Schwelle zu liegen. Diese Linie zu überqueren, auch nur einmal, auch nur mit einem einzigen logischen Qubit, verändert die Frage von „ist das überhaupt möglich?“ zu „wie weit lässt es sich skalieren, und wie schnell?“. Das ist eine echte, bedeutsame Verschiebung, und deshalb verdient das Ergebnis Aufmerksamkeit.

Aber dieselbe Sorgfalt, die das Ergebnis vertrauenswürdig macht, sollte die Geschichte darum herum dämpfen. Dies ist der erste Stein eines Fundaments, gut gelegt. Es ist nicht das Gebäude, und die, die ihn gelegt haben, sagen das als Erste. Die richtige Art, dem Quantencomputing in den nächsten Jahren zu folgen, ist genau diese unglamouröse Kurve: ob der Unterdrückungsfaktor hält, während die Codes wachsen, ob logische Gatter so sauber gelingen wie logischer Speicher, und ob dieser mysteriöse Einmal-pro-Stunde-Fehler je erklärt wird.

Kurz zusammengefasst

Google Quantum AI zeigte zum ersten Mal sauber, dass ein Surface-Code-Quantenspeicher *unterhalb der Schwelle* arbeiten kann: Während der Code von Distanz 3 über 5 auf 7 wuchs, fiel die logische Fehlerrate exponentiell (um etwa $2,14\times$ je zwei Schritte), und der größte, ein Distanz-7-Speicher mit 101 Qubits, überlebte sein bestes physisches Qubit — beyond breakeven —, während die Fehlerkorrektur in Echtzeit lief. Das ist ein echter, lange gesuchter Meilenstein im Ingenieurbau von Quantencomputern. Es ist zugleich ein einzelnes logisches Qubit als Speicher, mit einer Fehlerrate noch weit von dem entfernt, was echte Algorithmen verlangen, ohne ausgeführte logische Operationen, mit einem unerklärten Fehlerboden, den die Autoren selbst benennen, und mit einem Skalierungsweg von vielen Größenordnungen vor sich. Eine echte Schwelle überschritten — kein Quantencomputer geliefert.

Quellen

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>