



The CLEAN PAPER

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

量子存储器终于随着变大而变好——真正的里程碑，以及它还不是什么

Google Quantum AI 首次清楚展示，表面码量子存储器可以在错误阈值以下运作：当码距从 3 增加到 5、再到 7，逻辑错误率呈指数下降（每增加两级约改善 2.14 倍）；最大的一个码距 7 存储器用了 101 个量子比特，寿命超过同芯片上最佳的物理量子比特——跨过盈亏平衡点（*breakeven*）——而纠错还能实时运作。这是长期追求的工程里程碑。但它也只是单一逻辑量子比特作为存储器，错误率仍远高于实际算法需要的程度，没有执行逻辑门，作者自己还标出一个原因未明的错误底限，而且前面仍有好几个数量级的扩充。跨过的是阈值，不是交付了一台量子计算机。

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

曲线终于往对的方向弯了

三十年来，量子纠错一直建立在一个从未被明确兑现的承诺上。理论说，如果把一个量子信息单位——一个逻辑量子比特——分散编码到许多有噪声的物理量子比特上，而且这些物理量子比特已经够好，那么加入更多量子比特后，逻辑量子比特反而应该变得更好；随着纠错码变大，错误应呈指数下降。

问题就在那个“如果”。低于某个临界噪声阈值，增加量子比特会有帮助；高于它，更多量子比特只会带来更多噪声。过去所有实验不是还待在这条线的错误一侧，就是无法清楚展示所需趋势。把编码做大，结果通常反而更糟。

2024 年 12 月，Google Quantum AI 报告了首次清晰展示系统进入阈值以下区域。在最新一代超导处理器 Willow 上，他们建立了码距 3、5、7 的表面码存储器，并观察到每次增大码距，逻辑错误率都会下降——码距每增加两级，改善因子为 $\Lambda = 2.14 \pm 0.02$ 。最大的一个码距 7 存储器使用 101 个量子比特，每个纠错循环的逻辑错误率为 $0.143\% \pm 0.003\%$ ；而更关键的是，它比同一芯片上最佳的物理量子比特活得更久，寿命高出 2.4 ± 0.3 倍。这意味着跨过了盈亏平衡点（breakeven）：在这套硬件上，整套纠错的成本第一次真的换回了净收益。

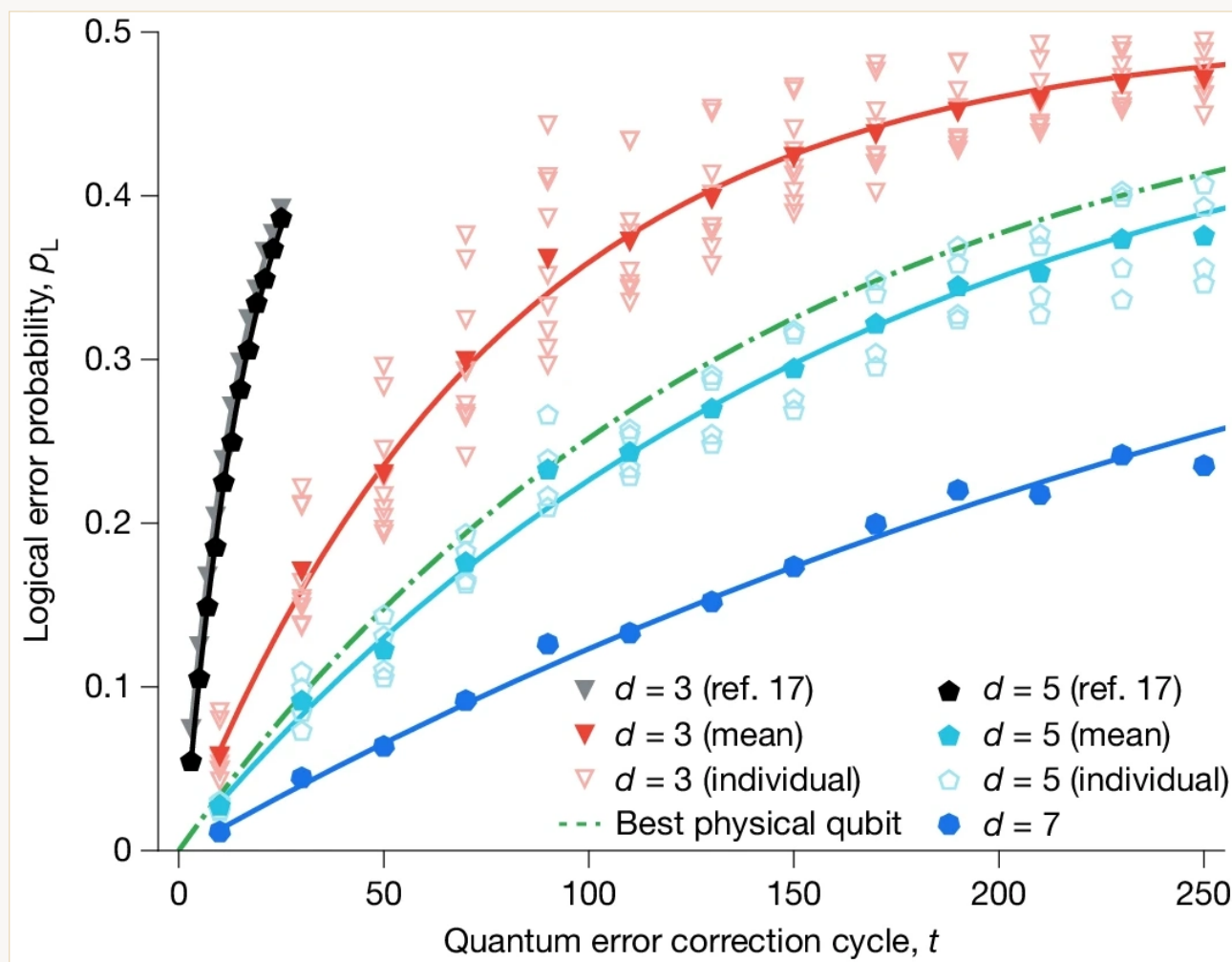
这是一个真正的里程碑，而重要的是把它属于哪一类里程碑说清楚。它证明扩充趋势现在终于朝正确方向前进。它不是一台可工作的量子计算机，论文也没有声称它是。

“表面码”、“码距”与“低于阈值”是什么意思

一个**逻辑量子比特**，是把一份需要保护的量子信息编码到许多物理量子比特上。**表面码 (surface code)** 是在二维网格上实现这种编码的一种方式；其中额外的“测量”量子比特会持续检查错误，同时尽量不扰动被储存的信息。

纠错码的**码距** d 不是物理距离，而是：至少需要多少个位置恰当的错误，才可能在编码没察觉的情况下破坏逻辑量子比特。较大的 d 代表更大、更耐错的编码区块，但也会花掉更多物理量子比特（约为 $2d^2 - 1$ ），并能修正更多同时发生的错误，最多为 $(d - 1) / 2$ 个。因此，这里测试的码距 3、5、7，分别可修正 1、2、3 个同时错误，理论上大约需要 17、49、97 个物理量子比特；Google 实际建立的码距 7 存储器用了 101 个，略高于这个教科书最低值。

“**低于阈值 (below threshold)**”才是整件事的核心。纠错只有在物理错误率低于某个临界值时才有用；在那个区域里，每增加码距，都会以指数方式压低逻辑错误率。抑制因子 Λ 就是在量化这件事： $\Lambda > 1$ 表示增大码距有帮助，而且 Λ 越大越好。Google 报告 $\Lambda \approx 2.14$ ，也就是码距每增加两级，逻辑错误率大约减半。真正的成果就是： Λ 明确而稳定地大于 1。



这张图显示码距 3、5、7 存储器（由上到下）在一轮轮纠错中如何累积逻辑错误。最值得看的，是绿色虚线——芯片上最佳的单一物理量子比特。码距 7 存储器（蓝色、最低）累积错误比这条线更慢，因此编码后的量子比特比组成它的最佳物理量子比特活得更久——也就是“跨过盈亏平衡点（breakeven）”，寿命约长 2.4 倍。这张图显示的是寿命结果；真正“低于阈值”的抑制，也就是码距从 3 增加到 5 再到 7 时 $\Lambda = 2.14$ ，则反映在正文数字里。

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

作者做了什么

- 在两颗 Willow 芯片上建立表面码存储器：一颗 105 量子比特处理器负责码距 3、5、7 的扩充测试（其中最大的是使用 101 个量子比特、49 个数据量子比特的码距 7 存储器）；另一颗 72 量子比特处理器则执行码距 5 存储器、实时解码器，以及高码距的重复码测试。
- 测量当码距从 3 增加到 5、再到 7 时，每一循环的逻辑错误率如何变化，并由此求出抑制因子 Λ 。
- 把逻辑量子比特的寿命与同芯片上最佳的单一物理量子比特相比，测试是否跨过盈亏平衡点（breakeven）。
- 让码距 5 的纠错码配合**实时解码器**运作——也就是能跟上错误检查产生速度的经典硬件——最多连续执行一百万个循环，以证明纠错不是只能事后离线分析，而是真能跟机器同步运作。
- 把较简单的**重复码**一路做到距离 29，用来寻找那些罕见、但会在极低错误率下限制表现的深层错

误来源。

他们发现了什么

- **代码确实低于阈值。**码距每增加两级，每循环逻辑错误率就以 $\Lambda = 2.14 \pm 0.02$ 的因子下降——呈现清晰的指数抑制，也就是理论长期预言、但先前处理器没有明确展示过的行为。
- **码距 7 存储器达到每循环 $0.143\% \pm 0.003\%$ 的错误率**，而且寿命比同芯片上最佳物理量子比特长 2.4 ± 0.3 倍——跨过盈亏平衡点（breakeven）。
- **实时解码跟得上。**在码距 5 上，解码器平均延迟为 63 微秒，而量子纠错循环本身是 1.1 微秒；系统持续运作达一百万循环。换言之，纠错是在线实时执行，不只是事后分析。
- **仍有罕见而深层的错误来源。**在重复码测试里，最终性能受到相关错误爆发限制，约每小时一次（大约每 3×10^9 个循环一次），造成约 10^{-10} 的错误底限；作者明确表示，这个来源目前仍不清楚。

这并没有证明什么

- **这不是一台正在进行计算的量子计算机。**这是一个量子存储器：它储存并保护一个逻辑量子比特。它没有在多个逻辑量子比特间执行逻辑运算（逻辑门），也没有跑任何算法。
- **它不是只差一个量子比特就能变成实用机器。**码距 7 的单一逻辑量子比特就花掉约 101 个物理量子比特；每循环约 0.1% 的错误率，仍远高于实际算法大约需要的 10^{-6} 到 10^{-10} 。要填平这个差距，就得把码距推得更大——也就是每个逻辑量子比特花更多物理量子比特——而有用的算法还需要成千上万个逻辑量子比特同时存在。物理量子比特的总预算很快就会进入数百万量级。
- **“如果能扩充”这个如果非常重要。**论文自己的结论是：目前的装置性能，如果能顺利扩充，可能达到大型算法需求。在单一逻辑量子比特上证明趋势正确，不等于已经打造出扩充后的机器，也没有任何事情保证这个趋势在大很多的规模仍然成立。
- **原因未明的错误底限是真正的现存问题。**这些限制重复码性能的相关爆发错误，按照作者自己的说法，比预期大好几个数量级；在找到原因前，它们会阻止更大型的容错应用。这是一个明白写在论文里的未解缺陷，不是已经处理完的小细节。
- **它也没有证明任何关于破解加密或在实用任务上达成“量子优越性”的事情。**那些需要的是完整的容错量子机器；这项工作只是为那台机器铺下的一块基石。

证据有多强

- **核心主张扎实而重要。**跨三种码距看到清晰的指数抑制、加上跨过盈亏平衡点（breakeven）的寿命、再加上能运作的实时解码器——这正是整个领域多年一直想达成的组合，而且是直接示范，不是间接推论。这不是炒作产物，而是顶尖团队做出的真实工程结果。
- **作者对范围相当谨慎。**他们把结果明确称为低于阈值的存储器，主动指出原因未明的相关错误底限，对未来也保留那个非常醒目的“如果能扩充”。若有夸大，多半出现在周边报道把“纠错后的存储器量子比特，随扩大而改善”直接四舍五入成“量子计算已经到来”。

- **最诚实的状态，是一个干净完成的基础步骤。**一个逻辑量子比特，终于被保护到增加冗余真的有帮助；但前面仍有漫长、艰难，而且没有保证能走完的路，包括更大规模、逻辑门，以及那个尚未解释的错误来源。

为什么重要

容错量子计算一直有一种先有鸡还是先有蛋的困境：真正有用的机器需要任何单一物理量子比特都达不到的低错误率；而解法——纠错——又只有在硬件本身已经好到低于阈值时才有效。哪怕只在一个逻辑量子比特上、只跨过一次这条线，也会把问题从“这到底做不做得得到？”改成“能扩到多远？能扩得多快？”这是真正有意义的改变，也是这个结果值得被注意的原因。

但也正因为研究做得仔细，周边叙事更应该克制。这是地基的第一块砖，而且铺得很好。它不是整栋建筑；而且最先这样说的人，就是铺砖的研究者自己。接下来几年追踪量子计算，最值得看的其实就是这条不华丽的曲线：码距继续增大时，抑制因子还能不能维持；逻辑门能不能像逻辑存储器一样干净；以及那个大约每小时出现一次的神秘错误，究竟能不能被解释。

简明总结

Google Quantum AI 首次清楚展示，表面码量子存储器可以在低于阈值的区域运作：当码距从 3 增加到 5、再到 7，逻辑错误率呈指数下降（每增加两级约改善 2.14 倍）；最大的一个码距 7 存储器使用 101 个量子比特，寿命超过同芯片上最佳的物理量子比特——跨过盈亏平衡点（breakeven）——而且纠错可以实时执行。这是量子计算工程上一个真正而且期待已久的里程碑。但它也只是单一逻辑量子比特作为存储器，错误率仍远高于实际算法需求，没有执行任何逻辑运算，作者还明确指出一个原因未明的错误底限，而且前方仍有好几个数量级的扩充路程。真正跨过的是阈值，不是交付了一台量子计算机。

来源

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>