The Clean Paper

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

양자 메모리가 커지면서 마침내 더 좋아졌다—진짜 이정표, 그리고 이것이 아닌 기계

Google Quantum AI 는 표면 부호 양자 메모리가 임계값 아래에서 작동할 수 있음을 처음으로 깔끔하게 보여 주었다. 부호를 거리 *3* 에서 *5*, *7* 로 키우자 논리 오류율은 지수적으로 떨어졌고 (두 단계마다 약 *2.14* 배), 가장 큰 *101* 큐비트 거리 *7* 메모리는 오류 정정이 실시간으로 돌아가는 가운데 자신의 가장 좋은 물리 큐비트보다 더 오래 버텼다—손익분기점을 넘어선 것이다. 이것은 오랫동안 추구해 온 공학적 이정표다. 동시에 이것은 메모리로 작동하는 논리 큐비트 하나이며, 오류율은 실제 알고리즘이 필요로 하는 수준과는 여전히 거리가 멀고, 논리 연산은 전혀 수행되지 않았으며, 저자들이 짚어 둔 설명되지 않은 오류 바닥이 있고, 여러 자릿수에 걸친 규모 확장이 앞에 놓여 있다. 임계값을 하나 넘은 것이지, 컴퓨터를 손에 쥔 것은 아니다.

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

곡선이 올바른 방향으로 꺾인 순간

지난 30 년 동안 양자 오류 정정은 한 번도 깔끔하게 지켜진 적이 없는 약속에 기대어 왔다. 이론에 따르면, 양자 정보 한 단위 — 즉 하나의 논리 큐비트 — 를 잡음이 많은 여러 물리 큐비트에 나누어 담고, 그 물리 큐비트들이 충분히 좋다면, 물리 큐비트를 더 늘릴수록 논리 큐비트는 더 좋아져야 하며, 부호가 커질수록 오류는 지수적으로 줄어들어야 한다. 함정은 그'~ 라면' 에 있다. 잡음이 임계값 아래에 있으면 큐비트를 늘리는 것이 도움이 되지만, 그 위에서는 큐비트를 늘려 봐야 잡음만 더해질 뿐이다. 이전의 모든 실험은 그 경계선의 잘못된 쪽에 머물러 있었거나, 그 경향을 깔끔하게 보여 주지 못했다. 부호를 키우면 상황은 나아지기는커녕 더 나빠졌다.

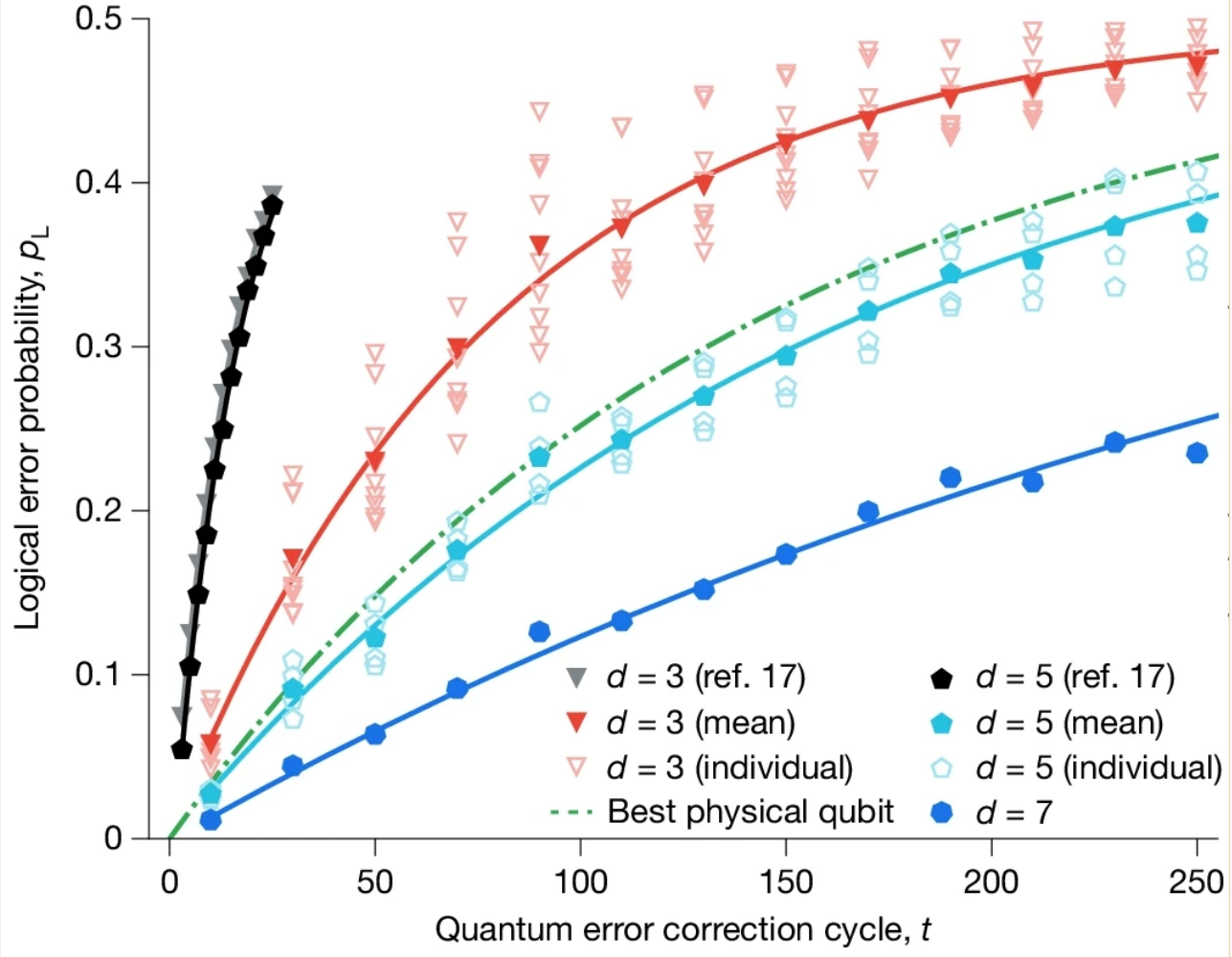
2024 년 12 월, Google Quantum AI 는 반대편 영역을 처음으로 명확하게 시연했다고 보고했다. 그들의 최신 세대 초전도 프로세서인 Willow 에서, 연구진은 부호 거리 3, 5, 7 에 해당하는 표면 부호 (surface code) 메모리를 만들었고, 부호가 커질 때마다 논리 오류율이 떨어지는 것을 관찰했다 — 거리가 두 단계 올라갈 때마다 $\Lambda = 2.14 \pm 0.02$ 배씩이다. 가장 큰 것은 101 큐비트 규모의 거리 7 메모리로, 정정 주기당 0.143% ± 0.003% 의 오류를 갖는 논리 큐비트를 유지했다. 그리고 — 헤드라인 속의 진짜 헤드라인은 — 이 메모리가 자신을 구성하는 가장 좋은 물리 큐비트보다 더 오래 살아남았다는 점이다, 그것도 2.4 ± 0.3 배만큼. 이것을' 손익분기점을 넘어섰다 (beyond breakeven)' 고 부르며, 오류 정정이라는 장치 전체가 이 하드웨어에서 제 값을 해낸 것은 이번이 처음이다.

이것은 진정한 이정표이며, 어떤 종류의 이정표인지 정확히 짚어 볼 가치가 있다. 이것은 이제 규모 확장 (scaling) 이 올바른 방향으로 간다는 증명이다. 이것은 작동하는 양자 컴퓨터가 아니며, 논문도 그렇다고 주장하지 않는다.

'표면 부호', '거리', ' 임계값 아래' 란 무엇인가

논리 큐비트는 여러 물리 큐비트에 걸쳐 부호화된, 보호받는 양자 정보 한 단위다. **표면 부호**는 그 부호화를 2 차원 격자 위에서 수행하는 특정한 방식으로, 여기서는 여분의 '측정' 큐비트가 저장된 정보를 교란하지 않으면서 끊임없이 오류를 검사한다. 부호 **거리** d 는 물리적 거리가 아니다. 그것은 부호가 알아채지 못한 채 논리 큐비트를 망가뜨릴 수 있는, 잘 배치된 오류의 최소 개수다. d 가 클수록 더 크고 더 튼튼한 패치가 된다 — 더 많은 물리 큐비트 (대략 $2d^2 - 1$ 개) 를 쓰고, 동시에 발생하는 오류를 더 많이, 최대 $(d - 1)/2$ 개까지 정정한다. 따라서 여기서 시험한 세 가지 크기, 즉 거리 3, 5, 7 은 각각 동시에 발생하는 오류를 1, 2, 3 개 정정하며 대략 17, 49, 97 개의 물리 큐비트를 쓴다 — Google 이 만든 거리 7 메모리는 101 개를 사용했는데, 이 교과서적 최솟값을 조금 웃도는 수치다.

' **임계값 아래**' 가 핵심 표현이다. 오류 정정은 물리 오류율이 임계값 아래에 있을 때만 도움이 된다. 그런 조건에서는 거리를 한 단계 높일 때마다 논리 오류율이 지수적으로 억제된다. 억제 인자 Λ 가 이것을 측정한다 — $\Lambda > 1$ 은 부호를 키우는 것이 도움이 된다는 뜻이고, Λ 가 클수록 좋다. Google 은 $\Lambda \approx 2.14$ 를 보고했는데, 이는 거리가 두 단계 올라갈 때마다 논리 오류율이 대략 절반으로 줄었다는 뜻이다. 이 Λ 가 1 을 넉넉히 웃돈다는 것이 결과의 전부다.



거리 3, 5, 7 메모리 (위에서 아래로) 에서 정정 주기가 거듭되는 동안 논리 오류가 어떻게 쌓이는지를 보여 준다. 주목할 선은 초록색 점선 — 칩에서 가장 좋은 단일 물리 큐비트 — 이다. 거리 7 메모리 (파란색, 가장 아래) 는 그 선보다 오류를 더 느리게 쌓아 가며, 그래서 부호화된 큐비트가 자신을 이루는 가장 좋은 물리 큐비트보다 더 오래 살아남는다 — '손익분기점을 넘어선', 2.4 배만큼의 결과다. 이것은 수명에 관한 결과다. 임계값 아래 억제 자체 (부호가 거리 3 에서 5, 7 로 커짐에 따른 $\Lambda = 2.14$) 는 본문의 수치 속에 담겨 있다.

연구진이 한 일

- 두 개의 Willow 칩 위에 표면 부호 메모리를 만들었다. 하나는 규모 확장 시험의 바탕이 된 거리 3, 5, 7 부호를 돌린 105 큐비트 프로세서 (그중 가장 큰 것이 101 큐비트, 데이터 큐비트 49 개의 거리 7 메모리) 이고, 다른 하나는 실시간 디코더로 거리 5 메모리와 고거리 반복 부호를 돌린 72 큐비트 프로세서다.
- 부호 거리를 3 에서 5, 7 로 늘림에 따라 주기당 논리 오류가 어떻게 변하는지 측정하여 억제 인자 Λ 를 뽑아냈다.
- ' 손익분기점' 을 확인하기 위해, 논리 큐비트의 수명을 같은 칩의 가장 좋은 개별 물리 큐비트와 비교했다.
- 오류 정정이 기계의 속도를 따라갈 수 있음을 보이기 위해, **실시간 디코더** — 오류 검사 결과를 생성되

는 속도만큼 빠르게 해석하는 고전 하드웨어 — 로 거리 5 부호를 최대 100 만 주기까지 돌렸다.
- 성능의 바닥을 결정짓는 드물고 깊은 오류원을 찾아내기 위해, 더 단순한 **반복 부호 (repetition code)** 를 거리 29 까지 밀어붙였다.

연구진이 발견한 것

- **부호가 임계값 아래에 있다.** 주기당 논리 오류는 거리가 2 씩 늘어날 때마다 $\Lambda = 2.14 \pm 0.02$ 배로 줄었다 — 깔끔한 지수적 억제로, 이론이 약속했지만 어떤 프로세서도 확실하게 보여 준 적 없던 거동이다.
- **거리 7 메모리는 주기당 0.143% ± 0.003% 의 오류에 도달했고**, 자신의 가장 좋은 물리 큐비트보다 **2.4 ± 0.3 배 더 오래** 살아남았다 — 손익분기점을 넘어선 것이다.
- **실시간 디코딩이 속도를 따라갔다.** 디코더는 거리 5 에서 1.1 마이크로초의 주기 시간에 대해 평균 63 마이크로초의 지연 시간을 보였고, 이를 100 만 주기 넘게 유지했다 — 오류 정정이 사후 분석에서만이 아니라 실시간으로 작동한 것이다.
- **드물고 깊은 오류원이 남아 있다.** 반복 부호 시험에서 성능은 결국 대략 **한 시간에 한 번**꼴로 (대략 3×10^9 주기당 한 번) 일어나는 상관된 오류 폭발에 의해 제한되었고, 이는 10^{-10} 근처의 오류 바닥을 형성했는데, 그 기원은 저자들의 말로는 **아직 밝혀지지 않았다**.

이것이 증명하지 못하는 것

- 이것은 연산을 수행하는 양자 컴퓨터가 **아니다**. 이것은 양자 메모리다. 하나의 논리 큐비트를 저장하고 보호한다. 논리 큐비트 사이의 논리 연산 (게이트) 을 수행하지 않으며, 어떤 알고리즘도 돌리지 않는다.
- 이것은 쓸모 있는 기계에 큐비트 하나 차이로 다가선 것이 **아니다**. 거리 7 논리 큐비트는 약 101 개의 물리 큐비트를 쓴다. 주기당 0.1% 의 오류율은 실제 알고리즘이 필요로 하는 대략 10^{-6} 에서 10^{-10} 수준보다 여전히 한참 높다. 그 격차를 좁히려면 훨씬 더 큰 거리로 — 논리 큐비트 하나당 훨씬 더 많은 물리 큐비트로 — 밀고 나아가야 하며, 쓸모 있는 알고리즘은 한 번에 논리 큐비트 수천 개를 필요로 한다. 그에 필요한 물리 큐비트 예산은 수백만 개에 이른다.
- **'규모를 키운다면' 이라는 단서가 실질적인 무게를 지닌다.** 논문 자체의 결론은, 이 소자의 성능이 규모를 키운다면 큰 알고리즘의 요구 조건을 충족할 수 있으리라는 것이다. 논리 큐비트 하나에서 경향이 옳다는 것을 보이는 일은 규모를 키운 기계를 실제로 만든 것과 같지 않으며, 여기서 그 무엇도 이 경향이 훨씬 더 큰 규모에서도 유지된다고 보장하지 않는다.
- **설명되지 않은 오류 바닥은 현재 진행 중인 문제다.** 반복 부호의 성능을 제한하는 이 상관된 폭발은, 저자들의 표현으로는 예상보다 여러 자릿수만큼 크며, 이해되기 전까지는 더 큰 결함 허용 응용을 가로막을 것이다 — 해결된 세부 사항이 아니라, 분명하게 밝혀 둔 미해결의 결함이다.
- 이것은 **암호 해독이나 쓸모 있는 작업에서의 '양자 우월성' 에 대해서는 아무것도 말해 주지 않는다.** 그런 것들에는 완전한 결함 허용 기계가 필요한데, 이 결과는 그 기계의 초석일 뿐 그것을 시연한 것은 아니다.

증거는 얼마나 강한가

- **핵심 주장은 탄탄하고 중요하다.** 세 가지 부호 거리에 걸친 깔끔한 지수적 억제를 동반한 임계값 아래 작동, 여기에 손익분기점을 넘어선 수명과 실제로 작동하는 실시간 디코더까지 — 이는 이 분야가 도달하려 애써 온 바로 그 조합이며, 추론된 것이 아니라 직접 시연되었다. 이것은 과장이 만들어 낸 산물이 아니라, 선도적인 연구 그룹이 내놓은 진짜 공학적 결과다.
- **저자들은 그 범위에 대해 신중하다.** 그들은 이 결과를 임계값 아래의 메모리로 규정하고, 설명되지 않은 상관 오류 바닥을 스스로 짚어 두며, 미래에 대한 전망은 그 눈에 띄는' 규모를 키운다면' 이라는 단서에 걸어 둔다. 과장이 나타나는 곳은, ' 오류 정정된 메모리 큐비트가 커지면서 나아졌다' 를' 양자 컴퓨팅의 시대가 왔다' 로 반올림해 버리는 주변 보도들이다.
- **정직하게 말하면 이것의 위상은, 깔끔하게 내디딘 기초적 한 걸음이다.** 여분을 더하는 것이 마침내 도움이 될 만큼 충분히 잘 보호된 논리 큐비트 하나 — 그리고 규모 확장, 논리 게이트, 설명되지 않은 오류라는 길고 험하며 아직 보장되지 않은 길이 여전히 앞에 놓여 있다.

왜 중요한가

결함 허용 양자 컴퓨팅에는 늘 닭이 먼저냐 달걀이 먼저냐 하는 느낌이 있었다. 쓸모 있을 만한 기계는 어떤 물리 큐비트도 도달할 수 없는 오류율을 필요로 하는데, 그 해법인 오류 정정은 하드웨어가 이미 임계값 아래에 있을 만큼 충분히 좋을 때만 작동한다. 그 경계선을, 단 한 번이라도, 그것도 논리 큐비트 단 하나에서 넘어서면, 질문은' 이것이 애초에 가능하기는 한가?' 에서' 얼마나 멀리, 얼마나 빠르게 규모를 키울 수 있는가?' 로 바뀐다. 이것은 실질적이고 의미 있는 전환이며, 그렇기에 이 결과는 주목받을 만하다.

하지만 이 결과를 믿을 만하게 만드는 바로 그 신중함이, 그 주변에 붙는 이야기의 온도를 낮춰 주어야 한다. 이것은 잘 놓인, 기초의 첫 벽돌이다. 이것은 건물 자체가 아니며, 그 벽돌을 놓은 사람들이 누구보다 먼저 그렇게 말한다. 앞으로 몇 년 동안 양자 컴퓨팅을 지켜보는 올바른 방법은 바로 이 화려하지 않은 곡선이다. 부호가 커져도 억제 인자가 유지되는지, 논리 게이트를 논리 메모리만큼 깔끔하게 해낼 수 있는지, 그리고 그 수수께끼 같은 한 시간에 한 번의 오류가 언젠가 설명되는지 말이다.

깔끔한 요약

Google Quantum AI 는 표면 부호 양자 메모리가 임계값 아래에서 작동할 수 있음을 처음으로 깔끔하게 보여 주었다. 부호를 거리 3 에서 5, 7 로 키우자 논리 오류율은 지수적으로 떨어졌고 (두 단계마다 약 2.14 배씩), 가장 큰 101 큐비트 거리 7 메모리는 오류 정정이 실시간으로 돌아가는 가운데 자신의 가장 좋은 물리 큐비트보다 더 오래 버텼다 — 손익분기점을 넘어선 것이다. 이것은 양자 컴퓨터 공학에서 오랫동안 추구해 온 진정한 이정표다. 동시에 이것은 메모리로 작동하는 논리 큐비트 하나이며, 오류율은 실제 알고리즘이 요구하는 수준과는 여전히 거리가 멀고, 논리 연산은 전혀 수행되지 않았으며, 저자들 스스로 짚어 둔 설명되지 않은 오류 바닥이 있고, 여러 자릿수에 걸친 규모 확장의 길이 앞에 놓여 있다. 진짜 임계값을 하나 넘은 것이지 — 양자 컴퓨터를 손에 쥔 것은 아니다.

출처

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