

The Scaling Paradox of Surface-Code Quantum Error Correction: Exponential Suppression, Logarithmic Overhead, and a Landauer Floor That Grows with Reliability

The Scaling Paradox of Surface-Code Quantum Error Correction: Exponential Suppression, Logarithmic Overhead, and a Landauer Floor That Grows with Reliability

Abstract

Surface-code quantum error correction (QEC) is widely regarded as the leading route to fault-tolerant quantum computation, yet its resource story contains an under-examined tension we call the *scaling paradox*: the code suppresses logical errors exponentially in code distance d , so the physical-qubit overhead per logical qubit grows only as the square of a quantity that is itself logarithmic in the target reliability $1/p_L$ — while at the same time the thermodynamic cost of correction, priced by Landauer's principle, grows linearly in d and therefore also logarithmically in $1/p_L$. Reliability is thus simultaneously cheap (polynomially, in qubits) and never free (the erasure floor per logical qubit per round, $\frac{n-k}{k}k_B T \ln 2$, increases with every increment of d). We make this tension quantitative. Using the standard threshold ansatz $p_L \approx 0.1 (p/p_{\text{th}})^{(d+1)/2}$ with $p_{\text{th}} = 10^{-2}$, we show that reaching $p_L = 10^{-12}$ at physical error rate $p = 10^{-3}$ requires $d = 21$ and $n = 881$ physical qubits per logical qubit, whereas $p = 10^{-4}$ requires only $d = 10$ and $n = 199$ — a $4.44\times$ overhead reduction. We compute the Landauer floor at $T = 20$ mK as 1.68×10^{-16} W per logical qubit at a 1 MHz cycle, roughly fifteen orders of magnitude below realistic control dissipation, and identify syndrome-decoding throughput (4.4×10^{14} bits/s at 10^6 logical qubits) as the operative bottleneck. We discuss falsifiability, failure modes, and the resource-commensurable alternative posed by bosonic encodings, which require $5\text{--}40\times$ fewer photons and $\sim 100\times$ fewer modes at matched $p_L = 10^{-6}$.

1. Introduction

The promise of surface-code quantum error correction is usually stated in one sentence: below a threshold physical error rate p_{th} , logical error rates fall exponentially with code distance d , so arbitrarily reliable computation is available at polynomial resource cost. This sentence is true, and it conceals a paradox of framing rather than of physics — a situation familiar across the sciences, where a formalism is internally consistent yet produces conclusions that clash with the intuition it is meant to serve. Paradoxes of this kind have been productive elsewhere: the Saint Petersburg paradox exposed the gap between expected value and real decision economics [5], apparent paradoxes in the spectral representation of stationary random processes exposed hidden assumptions in a mathematical formalism [6], and the identification of the firewall paradox with Wigner's friend showed that a seemingly gravitational puzzle reduces to a rule about combining observers' perspectives [8]. In each case, the resolution came from making explicit a quantity that the standard narrative left implicit.

Here the implicit quantity is twofold. First, the *marginal* cost of reliability: how many additional physical qubits must be deployed per additional decade of logical-error suppression? Second, the *thermodynamic* cost of correction itself: building on the observation that every cycle of measurement-based QEC is an erasure process priced by Landauer's principle at no less than $k_B T \ln 2$ per erased bit, so that an $[[n, k]]$ code carries a correction floor of $\frac{n-k}{k}k_B T \ln 2$ per logical qubit per round [11], within the broader program of mapping thermodynamic and informational bottlenecks of scalable fault tolerance [9]. The paradox is that these two costs pull in opposite

rhetorical directions: the qubit-overhead story says reliability gets cheaper as hardware improves, while the Landauer story says the erasure price of a logical qubit *rises* with every increment of distance — reliability is purchased with an ever-growing, strictly positive thermodynamic tax.

Our contributions are: (i) an explicit derivation of surface-code overhead as a function of target p_L and physical error rate p , with all arithmetic shown; (ii) a numerical evaluation of the Landauer floor and its comparison to realistic engineering dissipation; (iii) an identification of syndrome-decoding throughput, not thermodynamics, as the binding constraint at scale, drawing a rate-calibration analogy to large-scale detector readout systems [2]; and (iv) a discussion of bosonic encodings as a resource-commensurable alternative [10].

2. Background and Related Work

The surface code and its threshold behavior are the background against which our paradox is defined; we take as given the standard result that the threshold $p_{\text{th}} \approx 10^{-2}$ separates exponential suppression from exponential amplification of logical errors, and the heuristic ansatz for the logical error rate used in Section 3. Our quantitative framing of QEC overhead as a thermodynamic, not merely combinatorial, quantity follows the Landauer-machine analysis of error correction [11] and the broader bottleneck analysis of scalable fault-tolerant computation [9]. The claim that bosonic codes — cat, GKP, and binomial codes — may constitute the "native" encoding because they require $5\text{--}40\times$ fewer photons and roughly $100\times$ fewer modes than surface codes at matched $p_L = 10^{-6}$ [10] provides the principal alternative against which surface-code scaling should be judged; our analysis supplies the surface-code side of that comparison at additional target reliabilities.

The methodological lesson of this paper — that a paradox signals an implicit assumption, not a contradiction — is imported from several domains. The Saint Petersburg paradox showed that expected-value theory fails to capture real decision economics and motivated utility-based resolutions [5]; our Section 6 performs an analogous move, arguing that raw qubit counts are an incomplete utility metric for QEC architecture decisions. The paradox identified in the spectral representation of stationary random processes demonstrates that even mature mathematical formalisms can harbor inconsistent implicit premises [6]; we claim the surface-code narrative harbors an analogous implicit premise, namely that "overhead" means qubit count rather than the joint vector of qubits, decoding throughput, and erasure cost. The firewall/Wigner's-friend analysis shows that a celebrated paradox dissolves once one makes explicit the rule by which two observers' descriptions are combined [8]; similarly, the scaling paradox dissolves once one makes explicit the rule by which qubit overhead and thermodynamic overhead are combined into a single resource account. The analysis of QBism and the "objective paradox" clarifies when interpretational disputes reflect genuinely objective experimental facts rather than wording [3] — relevant because part of the scaling paradox, we argue, is interpretational rather than physical.

On the experimental side, large-scale detector systems offer the closest engineering analogy to QEC decoding. The ATLAS pixel project designed segmented silicon trackers with radiation-hard readout electronics operating at a 40 MHz collision rate, providing two-dimensional space-point information in the region nearest the interaction point [2]. We use this as a calibration point for what "large-scale, high-rate, low-latency readout" has historically meant in experimental physics, against the far higher syndrome rates that surface-code decoding demands (Section 4). The problem of reconciling inconsistent boundary data with a model's self-consistency requirements, as addressed for coronal magnetic-field extrapolation from vector magnetograms [4], has a direct QEC analogue: measured syndromes are noisy, inconsistent boundary data for the decoder's reconstruction of the error, and the weighted-self-consistency philosophy of [4] anticipates modern soft-information decoders. The antihydrogen program at CERN illustrates the discipline of stating fundamental-physics reach in terms of concrete, resource-bounded experimental

milestones [7], a discipline we apply to fault-tolerance roadmaps. Finally, the Araucaria Project's two-decade program of improving the cosmic distance scale exemplifies how a field's headline number (a distance, a qubit count) is only as credible as the ladder of calibrated intermediate steps beneath it [1].

3. Methods

Model. We adopt the standard phenomenological ansatz for the logical error rate per round of the rotated surface code under circuit-level depolarizing noise:

$$p_L(d, p) \approx A \left(\frac{p}{p_{\text{th}}} \right)^{(d+1)/2},$$

where d is the code distance, p the physical (circuit-level) error rate, and $p_{\text{th}} = 10^{-2}$ the threshold. We take the prefactor $A = 0.1$, the conventional heuristic fit; we treat it as a modeling assumption and propagate its uncertainty qualitatively in Section 6. The rotated surface code encodes $k = 1$ logical qubit using

$$n(d) = 2d^2 - 1$$

physical qubits (d^2 data qubits and $d^2 - 1$ measurement/ancilla qubits). Full fault-tolerant layouts with magic-state factories add multiplicative factors, which we handle as a labeled projection in Section 5.

Overhead metric. We define the qubit overhead ratio as

$$R_q = \frac{n - k}{k} = n - 1 = 2d^2 - 2,$$

which is also the per-logical-qubit multiplier appearing in the Landauer floor of [11]:

$$\dot{Q}_L = \frac{n - k}{k} k_B T \ln 2 \times f_{\text{cyc}},$$

with f_{cyc} the QEC cycle frequency and T the physical temperature of the erasing degrees of freedom.

Inputs. Every input number and its provenance:

- $p_{\text{th}} = 10^{-2}$: standard surface-code threshold value (literature consensus).
- Prefactor $A = 0.1$: conventional heuristic fit constant.
- Target reliabilities $p_L \in \{10^{-6}, 10^{-12}, 10^{-15}\}$: 10^{-6} matches the comparison point of [10]; 10^{-12} and 10^{-15} are representative algorithmic and cryptographic targets, chosen as round design targets, not measured values.
- Physical error rates $p \in \{10^{-3}, 10^{-4}\}$: representative of current and near-term improved hardware, used as scenario inputs.
- $k_B = 1.380649 \times 10^{-23}$ J/K (SI defined value); $\ln 2 = 0.693147$.
- $T = 20$ mK: representative dilution-refrigerator mixing-chamber temperature for superconducting QEC hardware (scenario input).
- $f_{\text{cyc}} = 1$ MHz: representative surface-code cycle time of $1 \mu\text{s}$ (scenario input).
- Control-dissipation comparison value 1 mW per physical-qubit channel (order-of-magnitude engineering estimate, labeled as such).

Procedure. For each (p, p_L) pair we solve the ansatz for the minimal integer d , compute n , R_q , the Landauer floor per logical qubit per round and per second, the marginal qubit cost per decade of reliability, and the aggregate syndrome-production rate for a 10^6 -logical-qubit machine. All arithmetic is shown in Section 4.

4. Analysis

Step 1: Distance from target reliability. Invert the ansatz:

$$\frac{d+1}{2} = \frac{\log_{10}(p_L/A)}{\log_{10}(p/p_{\text{th}})} \Rightarrow d = 2 \frac{\log_{10}(p_L/0.1)}{\log_{10}(p/p_{\text{th}})} - 1.$$

Case A: $p = 10^{-3}$, $p_L = 10^{-12}$. Here $p/p_{\text{th}} = 10^{-3}/10^{-2} = 0.1$, so $\log_{10}(p/p_{\text{th}}) = -1$, and $\log_{10}(10^{-12}/0.1) = \log_{10}(10^{-11}) = -11$. Then

$$d = 2 \times \frac{-11}{-1} - 1 = 21.$$

Check: $p_L \approx 0.1 \times (0.1)^{11} = 0.1 \times 10^{-11} = 10^{-12}$. ✓

Case B: $p = 10^{-4}$, $p_L = 10^{-12}$. $p/p_{\text{th}} = 10^{-2}$, $\log_{10} = -2$. Then $d = 2 \times (-11)/(-2) - 1 = 10$.

Check: $0.1 \times (10^{-2})^{11/2} = 0.1 \times 10^{-11} = 10^{-12}$. ✓

Case C: $p = 10^{-3}$, $p_L = 10^{-6}$. $\log_{10}(10^{-6}/0.1) = -5$; $d = 2 \times 5 - 1 = 9$. Check: $0.1 \times (0.1)^5 = 10^{-6}$. ✓

Case D: $p = 10^{-3}$, $p_L = 10^{-15}$. $\log_{10}(10^{-14}) = -14$; $d = 28 - 1 = 27$. Check: $0.1 \times (0.1)^{14} = 10^{-15}$. ✓

Step 2: Qubit counts and overhead. With $n = 2d^2 - 1$ and $R_q = 2d^2 - 2$:

- Case A ($d = 21$): $n = 2 \times 441 - 1 = 881$; $R_q = 880$.
- Case B ($d = 10$): $n = 2 \times 100 - 1 = 199$; $R_q = 198$.
- Case C ($d = 9$): $n = 2 \times 81 - 1 = 161$; $R_q = 160$.
- Case D ($d = 27$): $n = 2 \times 729 - 1 = 1457$; $R_q = 1456$.

Overhead ratio between Cases A and B: $R_q^{(A)}/R_q^{(B)} = 880/198 = 4.44$. A tenfold improvement in physical error rate buys a $4.44\times$ reduction in qubit overhead at fixed $p_L = 10^{-12}$.

Step 3: Marginal cost per decade of reliability. From Case A ($d = 21$) to Case D ($d = 27$): three decades cost $\Delta d = 6$, i.e., $\Delta d = 2$ per decade, and

$$\Delta n = (2 \times 27^2 - 1) - (2 \times 21^2 - 1) = 1457 - 881 = 576,$$

so $576/3 = 192$ physical qubits per decade of logical reliability at $p = 10^{-3}$ (locally, moving $d : 21 \rightarrow 23$ costs $\Delta n = 8d + 8 = 176$). This is the "cheap" side of the paradox: overhead grows linearly in d , hence logarithmically in $1/p_L$.

Step 4: Landauer floor. Single-bit erasure cost at $T = 20$ mK:

$$k_B T \ln 2 = (1.380649 \times 10^{-23}) \times (2.0 \times 10^{-2}) \times 0.693147.$$

First: $1.380649 \times 10^{-23} \times 2.0 \times 10^{-2} = 2.761298 \times 10^{-25}$ J. Then: $2.761298 \times 10^{-25} \times 0.693147 = 1.9139 \times 10^{-25}$ J.

Per logical qubit per round (Case A, $R_q = 880$):

$$Q_{\text{round}} = 880 \times 1.9139 \times 10^{-25} = 1.6842 \times 10^{-22} \text{ J.}$$

At $f_{\text{cyc}} = 1 \text{ MHz}$:

$$\dot{Q}_L = 1.6842 \times 10^{-22} \times 10^6 = 1.6842 \times 10^{-16} \text{ W}$$

per logical qubit. For Case B ($R_q = 198$): $Q_{\text{round}} = 198 \times 1.9139 \times 10^{-25} = 3.7895 \times 10^{-23} \text{ J}$, $\dot{Q}_L = 3.79 \times 10^{-17} \text{ W}$. The Landauer floor thus *falls* by the factor $880/198 = 4.44$ when hardware improves — but *rises* by $1456/880 = 1.65$ from Case A to Case D as reliability is pushed from 10^{-12} to 10^{-15} at fixed p .

Step 5: Gap to engineering reality. At an order-of-magnitude $1 \text{ mW} = 10^{-3} \text{ W}$ dissipation per physical-qubit control channel (labeled engineering estimate), Case A costs $\approx 881 \times 10^{-3} = 0.881 \text{ W}$ per logical qubit. Ratio to the Landauer floor:

$$\frac{0.881}{1.6842 \times 10^{-16}} = 5.23 \times 10^{15}.$$

The thermodynamic floor sits roughly fifteen to sixteen orders of magnitude below actual control dissipation.

Step 6: Syndrome throughput. Each round measures $d^2 - 1$ ancilla qubits, producing $d^2 - 1$ syndrome bits per logical qubit. Case A: $d^2 - 1 = 440$ bits. Per logical qubit at 1 MHz: 4.40×10^8 bits/s. For $N_{\text{log}} = 10^6$ logical qubits:

$$\dot{S} = 440 \times 10^6 \times 10^6 = 4.4 \times 10^{14} \text{ bits/s.}$$

Relative to the ATLAS pixel readout rate of $40 \text{ MHz} = 4.0 \times 10^7 \text{ events/s}$ [2]:

$$\frac{4.4 \times 10^{14}}{4.0 \times 10^7} = 1.1 \times 10^7.$$

Surface-code decoding at scale demands a raw syndrome decision rate about 10^7 times a flagship detector's event rate.

Step 7: Bosonic comparison at $p_L = 10^{-6}$ (projection). Following [10], bosonic codes (cat, GKP, binomial) require $5\text{--}40\times$ fewer photons and $\sim 100\times$ fewer modes than the surface code at $p_L = 10^{-6}$. Taking the surface-code patch at Case C ($n = 161$ physical qubits), the projection gives bosonic photon counts of $161/40$ to $161/5$, i.e., $4.0\text{--}32.2$ photons per logical qubit, and mode counts of order $161/100 \approx 1.6$, i.e., of order one to two modes per logical qubit (projection; uncertainty spans the published range and the mode-counting convention). Because the Landauer floor scales with the number of erased syndrome bits, which scales with the number of monitored degrees of freedom, a $\sim 100\times$ reduction in modes projects to a $\sim 100\times$ reduction in the per-round erasure burden — the thermodynamic face of the paradox is paradigm-dependent.

5. Results

All numbers below are computed in Section 4 from the stated inputs; none are measured or simulated.

Qubit overhead (computed). At $p = 10^{-3}$: $d = 9$, $n = 161$ for $p_L = 10^{-6}$; $d = 21$, $n = 881$ for $p_L = 10^{-12}$; $d = 27$, $n = 1457$ for $p_L = 10^{-15}$. At $p = 10^{-4}$: $d = 10$, $n = 199$ for $p_L = 10^{-12}$. Improving p from 10^{-3} to 10^{-4} reduces the $p_L = 10^{-12}$ overhead ratio from 880 to 198, a factor of 4.44.

Marginal cost (computed). At $p = 10^{-3}$, each additional decade of logical reliability costs $\Delta d = 2$ in distance and approximately 176–192 physical qubits per logical qubit (locally 176 at $d = 21 \rightarrow 23$; on average 192 across $d = 21 \rightarrow 27$).

Landauer floor (computed, given scenario inputs $T = 20$ mK, $f_{\text{cyc}} = 1$ MHz). Single-bit erasure: 1.91×10^{-25} J. Per logical qubit per round: 1.68×10^{-22} J at $d = 21$; 3.79×10^{-23} J at $d = 10$. Per logical qubit per second: 1.68×10^{-16} W and 3.79×10^{-17} W respectively. The floor grows by a factor of 1.65 when p_L is pushed from 10^{-12} to 10^{-15} at fixed $p = 10^{-3}$.

Thermodynamic–engineering gap (computed, given the labeled 1 mW/channel estimate). The Landauer floor is a factor of 5.2×10^{15} below realistic control dissipation per logical qubit.

Decoding throughput (computed). 4.4×10^8 syndrome bits/s per logical qubit at $d = 21$; 4.4×10^{14} bits/s for 10^6 logical qubits, i.e., 1.1×10^7 times the ATLAS 40 MHz event rate [2].

Machine-scale projection (labeled, with assumptions). For $N_{\text{log}} = 10^6$ logical qubits at $d = 21$, the aggregate Landauer floor power is $10^6 \times 1.6842 \times 10^{-16} = 1.68 \times 10^{-10}$ W, and the code-patch footprint is $10^6 \times 881 = 8.81 \times 10^8$ qubits; adding magic-state distillation factories (standard resource estimates multiply the footprint by a factor of a few to ~ 10) projects a total footprint of order 10^9 – 10^{10} physical qubits (projection, uncertainty factor ~ 10 dominated by the factory assumption). If decoding cost per logical qubit scales linearly in syndrome volume $d^2 - 1$, the Case B configuration ($d = 10$) requires $440/99 \approx 4.4\times$ less decoding throughput than Case A at equal $p_L = 10^{-12}$; uncertainty is dominated by the ansatz prefactor (factor-of-few) and by the linear-scaling assumption.

Bosonic projection (labeled). At $p_L = 10^{-6}$: 4.0–32.2 photons per logical qubit and of order one to two modes per logical qubit, with a projected $\sim 100\times$ reduction in per-round erasure burden relative to the surface code [10].

6. Discussion

The paradox, resolved. The "scaling paradox" is that two honest resource narratives about the same code point in opposite intuitive directions. The qubit narrative says reliability is astonishingly cheap: exponential suppression means each extra decade of p_L costs only ~ 2 in distance and ~ 176 – 192 qubits, and a tenfold hardware improvement cuts total overhead by $4.44\times$. The thermodynamic narrative says reliability is never free: the Landauer floor $\frac{n-k}{k} k_B T \ln 2$ [11] increases with d , so the erasure price of a logical qubit rises by $1.65\times$ as p_L goes from 10^{-12} to 10^{-15} . Both are correct; the apparent contradiction lives in the word "overhead," which silently denotes different quantities in the two narratives — precisely the structure of implicit-premise paradoxes documented in random-process spectral theory [6] and in observer-combination rules for the firewall problem [8]. Once the resource account is made explicit as the triple (qubits, decoding throughput, erasure cost), the paradox dissolves into a trade-off curve.

Where the real bottleneck sits. Our numbers show the Landauer floor is 5.2×10^{15} times smaller than plausible control dissipation. Thermodynamics, in the Landauer sense, is not the binding constraint on surface-code QEC at any foreseeable scale; the binding constraint is the 4.4×10^{14} bits/s syndrome stream for 10^6 logical qubits — a rate 1.1×10^7 times the event rate of the most demanding large-scale detector readout system built to date [2]. This reframes the bottleneck analysis of [9]: informational bottlenecks (decoder bandwidth, latency, and backlog) dominate thermodynamic ones by many orders of magnitude. It also sharpens the case for bosonic encodings [10]: if cat/GKP/binomial codes achieve $p_L = 10^{-6}$ with 5 – $40\times$ fewer photons and $\sim 100\times$ fewer modes, the implied reduction in syndrome volume — and hence decoding throughput — may be the decisive advantage, more than

photon count per se. As in the Araucaria Project's distance ladder [1], the headline number (qubits per logical qubit) is only credible when every rung beneath it — syndrome rate, decoder latency, erasure accounting — is independently calibrated.

Limitations and failure modes. (1) The ansatz $p_L \approx 0.1(p/p_{\text{th}})^{(d+1)/2}$ is a heuristic; the prefactor is uncertain by factors of a few and the effective exponent can differ under realistic noise models (leakage, correlated errors, measurement crosstalk). All derived distances could shift by ± 2 – 4 , changing n by tens of percent. (2) The 1 mW/channel and $T = 20$ mK inputs are scenario assumptions; the 5.2×10^{15} gap is robust to many orders of magnitude of error in them, but the conclusion "thermodynamics is irrelevant" would fail if cryogenic erasure were priced at far higher effective temperatures or if reversible-computing decoders approached the floor. (3) We treat $k = 1$; code deformation, concatenation, or block codes change R_q and the floor multiplicatively. (4) The linear decoding-cost projection is an assumption, not a result. (5) The analogy literature [1,2,4,6,7,8] is used heuristically, not as quantitative precedent.

What would falsify the claims. A demonstrated surface-code implementation achieving $p_L \leq 10^{-12}$ with $n \ll 400$ at $p \approx 10^{-3}$ would falsify the ansatz as used here. A decoder architecture with sub-linear scaling in d^2 would overturn the throughput bottleneck. Evidence that erasure in QEC cycles is irreducible at rates approaching $k_B T \ln 2 \times$ syndrome volume at effective temperatures well above 20 mK would revive the thermodynamic floor as binding, contradicting our gap estimate. Conversely, if bosonic encodings fail to deliver the published photon/mode advantages at scale [10], the comparative claim collapses.

Against ourselves. One may argue the "paradox" is merely a definitional artifact — that no practitioner ever conflated qubit overhead with thermodynamic overhead. We reply that resource roadmaps routinely report only qubit counts, and that the Saint Petersburg paradox teaches exactly this lesson: a formally correct expected-value (or qubit-count) accounting fails as a decision criterion until utilities — here, decoding capital and cooling capital — are added [5]. One may also argue the ATLAS analogy is superficial; we agree it is a rate calibration, not an architecture claim. Whether bosonic codes are the "native" encoding, as argued in [10], cannot be settled by our analysis; our results only show what the surface code must be compared against. The interpretational question of whether such resource trade-offs are objective facts or framing choices remains open in the sense discussed in [3], and the milestone-driven discipline advocated for antihydrogen physics [7] is the appropriate standard for settling it.

7. Conclusion

We have given an explicit, fully arithmetic account of the scaling paradox of surface-code QEC. Reliability is logarithmically cheap in qubits ($\Delta d = 2$ per decade; 176–192 qubits per decade at $p = 10^{-3}$) and a tenfold hardware improvement returns a $4.44 \times$ overhead reduction at fixed $p_L = 10^{-12}$ ($881 \rightarrow 199$ physical qubits per logical qubit). Yet the Landauer erasure floor per logical qubit per round grows linearly in d — rising by $1.65 \times$ as p_L improves from 10^{-12} to 10^{-15} — so the thermodynamic price of protection is proportional to the material price and can never be engineered to zero by better codes. At present scenario parameters the floor (1.68×10^{-16} W per logical qubit) sits some fifteen orders of magnitude below realistic control dissipation, and the operative bottleneck is instead the syndrome stream, 4.4×10^{14} bits/s for 10^6 logical qubits. The paradox therefore dissolves, once the resource account is made explicit, into a trade-off curve — and the resource-commensurable comparison with bosonic encodings [10], which need 5 – $40 \times$ fewer photons and $\sim 100 \times$ fewer modes at matched reliability, is the correct frame for evaluating it. Future work should quantify bosonic syndrome volumes under realistic imperfections, establish thermodynamic bounds directly in terms of p_L , and adopt milestone-driven resource accounting of the kind that has disciplined precision experiments in adjacent fields.

References

- [1] The Araucaria Project: Improving the cosmic distance scale. arXiv:2305.17247v1. <https://arxiv.org/abs/2305.17247v1> [2] The ATLAS Pixel Project. arXiv:hep-ex/9903035v1. <https://arxiv.org/abs/hep-ex/9903035v1> [3] QBism, Quantum Nonlocality, and the Objective Paradox. arXiv:1606.06286v2. <https://arxiv.org/abs/1606.06286v2> [4] Self-consistent Nonlinear Force-free Field Reconstruction from Weighted Boundary Conditions. arXiv:2004.12510v1. <https://arxiv.org/abs/2004.12510v1> [5] Relative Net Utility and the Saint Petersburg Paradox. arXiv:1910.09544v3. <https://arxiv.org/abs/1910.09544v3> [6] A paradox on the spectral representation of stationary random processes. arXiv:1212.6339v2. <https://arxiv.org/abs/1212.6339v2> [7] Particle Physics Aspects of Antihydrogen Studies with ALPHA at CERN. arXiv:0805.4082v1. <https://arxiv.org/abs/0805.4082v1> [8] The firewall paradox is Wigner's friend paradox. arXiv:2504.03835v1. <https://arxiv.org/abs/2504.03835v1> [9] <https://papers.qnfo.org/papers/thermodynamic-and-informational-bottlenecks-of-scalable-fault-tolerant-quantum-computation/>. QNFO: Thermodynamic and Informational Bottlenecks of Scalable Fault-Tolerant Quantum Computation. [10] DOI pending. QNFO: Bosonic Codes as the Native Encoding: Resource-Commensurable Comparison of Cat, GKP, Binomial, and Surface Codes. [11] <https://papers.qnfo.org/papers/jpcub-qec-landauer/>. QNFO: Error Correction Is a Landauer Machine: The Thermodynamic Floor of Quantum Error-Correction Overhead.

Appendix A. Divergence report

D1. Logical-error ansatz prefactor. Draft A used $p_L = (p/p_{\text{th}})^{(d+1)/2}$ with no prefactor, yielding $d = 11$ for $p_L = 10^{-6}$ at $p = 10^{-3}$. Drafts B and C used $p_L = 0.1(p/p_{\text{th}})^{(d+1)/2}$, yielding $d = 9$. *Convention behind the disagreement:* whether the standard heuristic fit includes the order-0.1 prefactor. **Resolution:** the main text adopts the prefactor convention of B and C (two of three drafts, and the more conservative literature practice); A's no-prefactor convention is documented here. Under A's convention all distances shift up by 2 and all n values upward accordingly.

D2. Physical-qubit count per patch. Draft A used $n \approx d^2$ (data qubits only, treated as total); drafts B and C used $n = 2d^2 - 1$ (data plus ancilla). **Resolution:** main text adopts $n = 2d^2 - 1$ (B/C, convergent), giving e.g. $n = 161$ at $d = 9$ rather than A's 121.

D3. Operating temperature. Draft A used $T = 0.1$ K; draft B used $T = 0.3$ K; draft C used $T = 20$ mK. All are defensible dilution-refrigerator stages. **Resolution:** main text adopts $T = 20$ mK (C), the mixing-chamber temperature typical of superconducting QEC hardware; the resulting Landauer floor (1.91×10^{-25} J/bit) is the smallest of the three and the gap-to-engineering conclusion is robust across all three choices (B's 0.3 K gives 2.87×10^{-24} J/bit; A's 0.1 K gives 9.57×10^{-25} J/bit).

D4. Emphasis of the thermodynamic conclusion. Draft A framed the Landauer floor as a fundamental barrier ("cannot be bypassed"); drafts B and C both computed the floor to be many orders of magnitude below engineering dissipation and framed it as strictly positive but not currently binding. **Resolution:** main text adopts the B/C framing (convergent), noting A's stronger claim as the minority position; the divergence is one of interpretation, not arithmetic, and A's qualitative point (the floor grows with d and cannot be coded away) is retained.

D5. Scope of target reliabilities. Draft A analyzed only $p_L = 10^{-6}$; drafts B and C both extended to $p_L = 10^{-12}$ (and C to 10^{-15}).