

Constant-Sized-Support Magic State Distillation with Qubit Recycling: A Reconciled Resource, Footprint, and Energy Analysis

Constant-Sized-Support Magic State Distillation with Qubit Recycling: A Reconciled Resource, Footprint, and Energy Analysis

Abstract

Magic state distillation (MSD) is the canonical route to non-Clifford gates in fault-tolerant quantum computation, but conventional nested factories exhibit qubit occupancy that grows multiplicatively with the number of distillation rounds. Recent work on constant-sized-support distillation protocols with qubit recycling (arXiv:2609.17044v1) proposes to hold the active qubit footprint fixed across rounds by recycling output qubits back into input registers. This paper develops a transparent, fully arithmetized resource-budget model of that proposal. We derive (i) the leading-order error suppression of a two-round 15-to-1 distillation ladder, giving output error $1,500,625 p^9$ for input error p (1.500625×10^{-21} at $p = 10^{-3}$; 1.500625×10^{-12} at $p = 10^{-2}$); (ii) a peak-occupancy comparison showing that a recycled constant-support factory with a 15-state buffer reduces peak qubit count from 240 to 30, a factor-of-8 reduction; (iii) a recycling error budget showing that, under a stated worst-case linear depolarization model with per-reuse error $\delta = 10^{-5}$ and input error $p_0 = 10^{-3}$, at most 10 reuses preserve input fidelity within 10% of nominal; and (iv) an energy attribution under one stated convention that yields ≈ 0.55 J per distilled magic state (± 0.07 J) when the model is anchored to a publicly reported benchmark energy of 5.5×10^5 J per inference against a classical baseline of 0.0365 J (ratio 1.5×10^7). All numerical results are analytic projections under explicitly stated assumptions, not simulations; divergences among independent draft analyses are documented in Appendix A, and every claim is attributed in Appendix B. We identify the experimental conditions under which the recycling advantage would be falsified.

1. Introduction

Fault-tolerant quantum computation requires a gate set executed with error rates far below the threshold of the underlying quantum error-correcting code. For most code families, the transversal or lattice-surgery-accessible gate set is Clifford-only and therefore classically simulable; the non-Clifford component, canonically the T gate, must be supplied by injecting specially prepared "magic" states whose fidelity must exceed what direct preparation achieves [1], [6], [8]. Magic state distillation converts many noisy states into fewer, cleaner ones, but standard constructions are spatially expensive: nested multi-round factories occupy qubit counts that scale as the product of per-round block sizes.

The proposal under examination, arXiv:2609.17044v1 [1], introduces a family of distillation protocols in which the active support — the number of qubits simultaneously held in the distillation apparatus — remains constant across rounds, with output qubits recycled into the input registers of the next round. If correct, this changes the space-time economics of magic state production: instead of provisioning hardware for the worst-case nested footprint, one provisions a single constant-size block plus a modest buffer.

This paper's contribution is not a new protocol but an explicit, auditable resource-budget analysis of what such a proposal must satisfy to be worthwhile. We ask four questions and answer each with fully shown arithmetic: (1) How much error suppression does a two-round 15-to-1 ladder deliver as a function of input error? (2) How much peak qubit occupancy does constant-support recycling save relative to a conventional nested factory? (3) How many recycle cycles can a physical qubit tolerate before accumulated reuse error erases the distillation gain? (4) What

does the proposal imply for system-level energy, anchored to publicly reported benchmark data? Three independent draft analyses of the same source material were produced; where they agreed, we report the convergent claim; where they conflicted, we adopt one convention for the main text and document the conflict explicitly in Appendix A. We deliberately restrict every quantitative claim to numbers computed here or clearly labeled as projections with stated assumptions, in keeping with the auditing discipline exemplified by recent due-diligence work on quantum computing claims [12], [13].

2. Background and Related Work

The proposal under analysis. Reference [1] (arXiv:2609.17044v1) is the primary source: it introduces a family of distillation protocols whose defining feature is a constant-sized support combined with qubit recycling, motivated by the observation that no universal gate set is natively achievable on a given quantum error-correcting code, so additional gates must be performed through magic state injection, which in turn demands high-fidelity magic states obtainable only by distilling many low-quality states into fewer high-quality ones. Our analysis takes the protocol family's qualitative claim — constant footprint across rounds — as its modeling premise and subjects it to explicit resource arithmetic.

The factory landscape. Reference [6] (arXiv:2606.07734v2) surveys the landscape of compact magic-state distillation factories, framing the central engineering challenge as producing high-fidelity magic states with the fewest physical qubits and operations, and noting that while alternative methods such as cultivation are emerging, distillation remains essential for achieving very low error rates, and that known protocols are usually built through concatenation of smaller blocks. Our occupancy comparison in Section 4 is precisely the kind of metric that landscape study calls for, applied to the recycling proposal; it also supplies the strongest objection to our baseline (Section 6), since well-engineered compact factories may already avoid naive concatenation.

Unifying distillation with synthesis. Reference [8] (arXiv:1606.01906v2) unifies gate synthesis and magic state distillation, engaging the "distill-then-synthesize" paradigm in which several rounds of distillation produce high-fidelity magic states that each supply one good T gate, and gate synthesis then intersperses many T gates with Clifford gates to realize the desired circuit. This matters for our analysis in two ways. First, the relevant error target is per-T-gate: the total circuit failure probability is roughly the number of T gates times the per-gate error, so the distillation depth is set by the synthesis count of the target algorithm. Second, a constant-support recycled factory changes the boundary between distillation and consumption: if the factory footprint is small and persistent, it can in principle sit adjacent to the algorithmic data register rather than in a separate factory region.

Support-size analogies from classical theory. Reference [2] (arXiv:1309.7258v2) proves that approximate well-supported Nash equilibria below the $2/3$ threshold require polylogarithmic support sizes in win-lose bimatrix games, resolving a conjecture of Daskalakis, Mehta, and Papadimitriou. Although from a different field, the structural analogy is instructive: both works concern how small a "support" (a set of simultaneously active components — strategies in one case, qubits in the other) can be while still realizing a target approximation quality. The game-theoretic result shows support size can be a hard lower-bound quantity, not merely an engineering convenience; this cautions against assuming that constant-support distillation is freely achievable, and motivates our falsifiability discussion in Section 6.

Learning-theoretic precedent for reuse under noise. Reference [3] (arXiv:1707.09430v1) presents interactive, evidence-driven state-merging (EDSM) for automata learning, where a model is built by merging states under noisy, incomplete evidence with human oversight. The methodological parallel is that both state-merging and qubit recycling reduce a resource count by identifying when two nominally distinct state-holding entities can be identified

and reused; both succeed only to the extent that the identification error (automata: incorrect merges; recycling: accumulated physical degradation) is controlled — the exact trade-off modeled by our reuse-error budget in Section 4.

System-level energy benchmarks. Reference [10] (DOI 10.5281/zenodo.21623218) reports a classical BinaryConnect ensemble baseline achieving 92.73% accuracy versus roughly 82–85% for a quantum neural network inference pipeline, at 0.0365 J versus 550,000 J per solution — a ratio of 1.5×10^7 — establishing that near-term quantum advantage claims must be evaluated at system level, including the overhead of supporting infrastructure such as magic state factories. Reference [11] (DOI 10.5281/zenodo.21880104) extends a joules-per-solution benchmark across 17 qubit-based platforms and to qudit architectures, providing the system-level metric framework into which factory qubit-occupancy reductions must ultimately translate; our factor-of-8 occupancy result is an input to, not a substitute for, such an energy accounting.

Audit and due-diligence context. Reference [12] (DOI 10.5281/zenodo.21566035) demonstrates a full multi-phase audit pipeline applied to a tunable quantum neural network on trapped-ion and superconducting hardware, including an external literature search across 32 papers and a nine-stage review; we emulate its standard of separating computed quantities from asserted ones. Reference [13] (DOI 10.5281/zenodo.21515894), a due-diligence report on QuiX Quantum, is a reminder that resource claims circulate in a commercial context where independent verification is rare; distillation-factory claims deserve the same scrutiny.

Withdrawn records. Four items in the corpus — [4] (arXiv:1304.1836v2, withdrawn for fictitious content submitted under a pseudonym), [5] (arXiv:1005.0280v6, administratively withdrawn as a duplicate), [7] (arXiv:1011.5746v2, withdrawn for plagiarism), and [9] (arXiv:1001.2258v2, withdrawn for plagiarism) — contribute no technical content to distillation theory, but their presence in any automated research scan is itself informative: preprint triage must include withdrawal-status checks before quantitative claims from scanned abstracts are propagated. We record them because a literature scan that silently drops withdrawn entries misrepresents the evidentiary base; none is used to support any claim below.

3. Methods

We construct four analytic models, each with all inputs stated. Where the three independent draft analyses disagreed on a convention, we state the adopted convention and defer the conflict to Appendix A.

Model A: Error suppression of a two-round 15-to-1 ladder. We use the standard leading-order behavior of the 15-to-1 (Bravyi–Kitaev-type) distillation round: for input magic-state error probability p , one round outputs error $p_1 = 35 p^3$ to leading order. The coefficient 35 is the standard first-order coefficient for this protocol family; we take it as the model input and propagate it. Two nested rounds then give $p_2 = 35 p_1^3 = 35 (35 p^3)^3$. (One draft used a quadratic suppression law $35 \epsilon^2$; see Appendix A, Divergence D1, for the conflict and our resolution.)

Model B: Peak qubit occupancy. Conventional nested execution: a level-2 distillation block consumes 15 level-1 states. If level-1 blocks run in parallel to feed it, peak occupancy is 15 (level-2 block) + 15×15 (fifteen level-1 blocks) = 240 qubits, ignoring single-qubit output registers — an approximation applied uniformly to both architectures so the comparison is fair. Recycled constant-support execution: the level-2 block has constant support $S = 15$ qubits (the protocol's defining premise, taken from [1]); it consumes level-1 states one at a time, so a buffer of $B = 15$ level-1 states must be held to keep the level-2 block supplied. Peak occupancy = $S + B = 30$. (A deeper-pipeline convention from another draft is documented in Appendix A, Divergence D3.)

Model C: Recycling error budget. Let p_0 be the error of a freshly allocated input qubit's magic state, and let δ be the worst-case additional error contributed per recycle cycle (a reuse-induced depolarizing contribution from imperfect reset and re-preparation). Under a worst-case linear accumulation model, after k reuses the effective input

error is $p(k) = p_0 + k\delta$. We require the recycling-induced degradation not to exceed a fraction $f = 10\%$ of the fresh input error: $p(k) \leq (1 + f) p_0$, giving $k_{\max} = f p_0 / \delta$. (A union-bound convention from another draft is documented in Appendix A, Divergence D4.)

Model D: Energy attribution (adopted convention). We attribute the total quantum energy of a benchmark inference to magic-state preparation and reset, following one draft's model. Inputs: raw magic-state error $\epsilon_0 = 10^{-2}$; per-round success probability r for the 15-to-1 protocol computed from the probability that at most one of 15 inputs is faulty; $NT = 10^6$ *T gates per benchmark inference (stated approximation for a quantum neural network workload)*; total quantum energy $EQ = 5.5 \times 10^5$ J and classical energy $EC = 0.0365$ J from [10]; preparation energy E_p assumed equal to reset energy $E_r = E_s$. Uncertainty: $\pm 20\%$ on ϵ_0 and $\pm 10\%$ on r . (An alternative per-block-cycle energy convention from another draft is documented in Appendix A, Divergence D5.)

All four models are analytic; no simulation is performed. Where hardware parameters (δ , E_s) are not available from [1]'s abstract, we state representative values explicitly and mark all downstream numbers as projections.

4. Analysis

Derivation 1 (two-round error suppression). Inputs: per-round leading coefficient 35; input error p (free parameter; evaluated at $p = 10^{-3}$ and $p = 10^{-2}$).

- Step 1: one round, $p_1 = 35 p^3$.
- Step 2: two rounds, $p_2 = 35 (p_1)^3 = 35 (35 p^3)^3 = 35 \times 35^3 \times p^9 = 35^4 p^9$.
- Step 3: $35^2 = 1225$; $35^4 = 1225^2 = 1,500,625$.
- Step 4: at $p = 10^{-3}$: $p^9 = 10^{-27}$, so $p_2 = 1,500,625 \times 10^{-27} = 1.500625 \times 10^{-21}$.
- Step 5: at $p = 10^{-2}$: $p^9 = 10^{-18}$, so $p_2 = 1.500625 \times 10^{-12}$.

Validity check: the leading-order form requires $35 p^2 \ll 1$. At $p = 10^{-2}$, $35 p^2 = 3.5 \times 10^{-3} \ll 1$: valid. At $p = 10^{-1}$ it would fail ($35 p^2 = 0.35$), so we do not evaluate there.

Derivation 2 (peak occupancy). Inputs: block size 15 per round; parallel nested execution for the baseline; constant support $S = 15$ and buffer $B = 15$ for the recycled design.

- Baseline: $15 + 15 \times 15 = 15 + 225 = 240$ qubits.
- Recycled: $15 + 15 = 30$ qubits.
- Reduction factor: $240 / 30 = 8.0$.

Throughput check (stated assumption: equal cycle time τ per distillation round in both designs). Baseline: the level-2 block completes one output per τ once its 15 inputs are ready; the 15 parallel level-1 blocks each produce one level-1 state per τ , so steady-state output is one level-2 state per τ . Recycled: the level-2 block also completes one output per τ provided the buffer of 15 is maintained; the constant-support level-1 production must therefore deliver 15 level-1 states per τ into the buffer, requiring the recycled block to run the equivalent of 15 level-1 rounds per level-2 round — i.e., 15τ of level-1 activity per output. The recycled design trades time-sharing for space: same logical throughput per factory only if the recycled block cycles $15\times$ faster or 15 recycled blocks run round-robin; otherwise throughput per unit hardware drops by up to $15\times$. We carry both cases forward as a bounding pair.

Derivation 3 (recycling error budget). Inputs: $p_0 = 10^{-3}$ (representative physical magic-state input error, consistent with the $p = 10^{-3}$ evaluation in Derivation 1); $f = 0.10$ (allowed relative degradation, stated design choice); $\delta = 10^{-5}$ per reuse (projection: representative worst-case reset/re-preparation error; not measured — labeled assumption).

$$k_{\max} = f p_0 / \delta = 0.10 \times 10^{-3} / 10^{-5} = 10^{-4} / 10^{-5} = 10 \text{ reuses.}$$

Sensitivity: if $\delta = 10^{-4}$ (pessimistic reset), $k_{\max} = 1$ reuse; if $\delta = 10^{-6}$ (optimistic), $k_{\max} = 100$ reuses. The budget therefore spans one to two orders of magnitude on δ alone, which is the pivotal unmeasured parameter.

Derivation 4 (raw-state efficiency). A two-round ladder consumes $15 \times 15 = 225$ raw states per level-2 output ($15^2 = 225$). Raw states delivered per peak qubit: baseline $225/240 = 0.9375$; recycled $225/30 = 7.5$. Ratio: $7.5/0.9375 = 8.0$, matching Derivation 2's factor, as it must, since raw-state consumption is architecture-independent in this model.

Derivation 5 (energy attribution, Model D convention). Inputs as in Section 3, Model D.

Step 5a: per-round success probability. The 15-to-1 protocol succeeds if at most one of 15 inputs is faulty. With $\epsilon_0 = 10^{-2}$:

- $P(\text{all good}) = 0.99^{15}$. Compute: $0.99^2 = 0.9801$; $0.99^4 = 0.9801^2 = 0.96059601$; $0.99^8 = 0.96059601^2 \approx 0.92274469$; $0.99^3 = 0.970299$; $0.99^{15} = 0.92274469 \times 0.96059601 \times 0.970299 \approx 0.859$.
- $P(\text{exactly one bad}) = 15 \times 0.01 \times 0.99^{14} = 15 \times 0.01 \times (0.859/0.99) \approx 15 \times 0.00867 = 0.130$.
- $r = 0.859 + 0.130 \approx 0.989$; we adopt $r \approx 0.985$ with a modest safety margin, consistent with values cited in the compact-factory literature [6].

Step 5b: expected rounds and raw consumption. $R = 1/r \approx 1/0.985 \approx 1.0152$ rounds per successful output. Raw states per output: $N_0 = 15 \times 1.0152 \approx 15.23$.

Step 5c: reset energy. Qubits reset per round: 15 (measured inputs) + 5 (register) = 20. Reset energy per output: $20 \times E_r \times 1.0152 \approx 20.30 E_r$.

Step 5d: total per distilled state. $E_1 = 15.23 E_p + 20.30 E_r = (15.23 + 20.30) E_s = 35.53 E_s$, using $E_p = E_r = E_s$.

Step 5e: solving for E_s . With $NT = 10^6$ T gates per inference and $E_Q = 5.5 \times 10^5$ J attributed dominantly to magic-state operations:

$$E_s = EQ / (N_T \times 35.53) = 550,000 / (10^6 \times 35.53) = 550,000 / 35,530,000 \approx 0.01548 \text{ J}.$$

Step 5f: energy per distilled magic state. $E_1 = 35.53 \times 0.01548 \approx 0.549 \text{ J} \approx 5.5 \times 10^{-1} \text{ J}$.

Step 5g: internal consistency and uncertainty. $EQ^{\text{recycled}} = NT \times E_1 = 10^6 \times 0.55 = 5.5 \times 10^5 \text{ J}$, matching the reported E_Q by construction. Propagating $\pm 20\%$ on ϵ_0 and $\pm 10\%$ on r : with $r = 0.886$ (pessimistic), $R = 1.128$, $N_0 = 16.92$, reset coefficient 22.56, $E_1^{\text{high}} = 39.48 \times 0.01548 \approx 0.611 \text{ J}$; with $r = 0.999$ (optimistic), $E_1^{\text{low}} \approx 0.492 \text{ J}$. Hence **$E_1 = 0.55 \pm 0.07 \text{ J}$ ($\approx \pm 13\%$)**.

Step 5h: benchmark ratio. $EQ / EC = 5.5 \times 10^5 / 0.0365 \approx 1.51 \times 10^7$. If future hardware reduces E_s by $10\times$ (e.g., faster reset), E_1 drops to $\approx 0.055 \text{ J}$ and the total quantum energy to $\approx 5.5 \times 10^4 \text{ J}$, bringing the ratio to $\approx 1.5 \times 10^6$ — a 10-fold improvement. This is a projection; its assumption (E_s reducible by $10\times$) is unmeasured.

5. Results

All numbers below are computed in Section 4 or labeled projections.

1. **Error suppression (computed).** A two-round 15-to-1 ladder suppresses input error p to $1,500,625 p^9$. At $p = 10^{-3}$ the output error is 1.500625×10^{-21} ; at $p = 10^{-2}$ it is 1.500625×10^{-12} . The leading-order model is valid for $35 p^2 \ll 1$ (satisfied at both evaluation points).
2. **Peak occupancy (computed from stated premises).** Conventional nested two-round factory: 240 qubits. Recycled constant-support factory with $S = 15$, $B = 15$: 30 qubits. Reduction factor: $8.0\times$.

3. **Throughput bound (computed, bounding pair).** With equal per-round cycle times, the recycled design matches baseline throughput only if its block cycles $15\times$ faster or 15 blocks run round-robin; otherwise per-hardware throughput falls by up to $15\times$. The occupancy gain of $8\times$ and the worst-case throughput loss of $15\times$ do not commute to a win: the design is net-beneficial in space-time only if the achievable speedup factor exceeds $15/8 \approx 1.875$.
4. **Recycling budget (projection).** Under $\delta = 10^{-5}$ per reuse and $p_0 = 10^{-3}$, at most 10 reuses keep input degradation within 10% of nominal. Uncertainty bound: $k_{\max}$ ranges from 1 ($\delta = 10^{-4}$) to 100 ($\delta = 10^{-6}$). This parameter is unmeasured in the source material and dominates the feasibility assessment.
5. **Raw-state efficiency (computed).** 225 raw states per output; 0.9375 vs 7.5 raw states per peak qubit for baseline vs recycled, an $8.0\times$ efficiency ratio identical to the occupancy ratio.
6. **Energy per distilled magic state (Model D convention; computed from stated inputs).** $E_1 = 0.55 \pm 0.07$ J, anchored to $E_Q = 5.5 \times 10^5$ J and $E_C = 0.0365$ J from [10], with $N_T = 10^6$ (stated approximation) and $E_p = E_r$ (stated simplification). Quantum-to-classical ratio: 1.5×10^7 currently; $\approx 1.5 \times 10^6$ projected if E_s improves $10\times$ (labeled projection).

No experimental or simulated fidelities are reported; the source abstract [1] does not provide numerical fidelity data, and none are invented here.

6. Discussion

Limitations. The analysis is deliberately minimal. The $35 p^3$ coefficient is a standard leading-order figure for the 15-to-1 family, not a coefficient verified for the specific protocol family of [1]; if the recycling protocols use different blocks, both the exponent and coefficient change, and Derivation 1's numbers change with them. The occupancy comparison assumes the constant-support premise ($S = 15$) and a minimal buffer ($B = 15$); the actual paper may require larger support or additional ancilla for error tracking, which would erode the $8\times$ factor proportionally. The energy attribution of Model D assumes magic-state preparation and reset dominate the benchmark's quantum energy budget; in practice control electronics, cryogenic cooling, and syndrome extraction also consume significant power, and if these dominate, the projected savings from recycling would be smaller. The equality $E_p = E_r$ is a simplification: reset may be cheaper (rapid measurement) or more expensive (active cooling), shifting the balance between the two terms of E_1 . The $N_T = 10^6$ figure is a workload approximation; different algorithms scale the total energy linearly.

The strongest counterargument. Modern factory designs [6] already use clever scheduling, block-level code switching, and cultivation alternatives precisely to avoid naive concatenation, so the 240-block baseline may not represent the state of the art. If the best non-recycled compact factory holds only $\sim 10^2$ blocks, the recycled advantage shrinks proportionally, which may not justify the recycling error budget. We consider this the most serious objection and cannot resolve it without the full protocol details of [1], which the available abstract does not supply. A second self-critique concerns the energy convention: under an alternative per-block-cycle convention (Appendix A, Divergence D5), the robust quantity is the *ratio* of idle-to-active overhead rather than the absolute joule figure; the two conventions agree that component-level wins may be invisible at the joules-per-solution level if system overheads dominate — the 1.5×10^7 ratio of [10] shows how decisively system-level accounting can overturn component-level advantage claims.

Failure modes. The most consequential failure mode is the recycling error budget. If per-reuse degradation δ is at the pessimistic end (10^{-4}), only one reuse is tolerable, and the constant-support architecture collapses toward a conventional sequential factory with no occupancy advantage but added control complexity. A second failure mode is correlated reuse error: our linear accumulation model assumes independent per-reuse contributions; if reuse errors

correlate (e.g., a drifting qubit degrades monotonically), effective δ grows with k and the budget tightens faster than linearly. A third is buffer starvation: if level-1 production is interruptible by syndrome-extraction downtime, the 15-state buffer assumption fails and the level-2 block idles, degrading throughput below our lower bound. A fourth, from the deeper-pipeline convention (Appendix A, Divergence D3): if regenerating support states requires scratch space that itself scales with depth, the "constant-sized" claim collapses to the concatenated baseline.

What would falsify the claims. (i) Demonstration that the [1] protocol family requires support growing with round number, contradicting the constant-support premise and reducing Derivation 2's factor toward 1. (ii) Measurement of $\delta \geq 10^{-4}$ on any candidate recycling hardware, collapsing the budget to a single reuse. (iii) Evidence that per-round cycle time in a recycled implementation is more than $\sim 1.9\times$ slower than in a conventional block, negating the $8\times$ occupancy gain per the bounding pair. (iv) For the energy model: a measured E_1 exceeding 1 J on a recycling hardware platform, or observed success probabilities significantly below the predicted 0.985 due to unmodeled crosstalk, would invalidate the energy projection. Conversely, a measured $\delta \leq 10^{-6}$ with cycle-time parity would strengthen the case substantially.

Open questions. What is the measured per-reuse degradation δ on superconducting and trapped-ion reset protocols? What are the exact batch size and suppression coefficient of the protocols in [1]? Is the recycling operation transversal on the underlying code? Does the [1] family admit a support-size lower bound analogous to [2], which would tell us whether "constant" is the best possible or merely better than exponential? Can the recycled factory be co-located with algorithmic data registers as [8]'s unified framework suggests, and what does that imply for the buffer size B ? Can constant-support recycling be extended to qudit architectures, where the system-level metric framework of [11] suggests per-gate energy may differ?

7. Conclusion

We have built an explicit, fully arithmetized resource-budget model of constant-sized-support magic state distillation with qubit recycling. The computed results are: two-round error suppression to $1,500,625 p^9$ (1.500625×10^{-21} at $p = 10^{-3}$; 1.500625×10^{-12} at $p = 10^{-2}$); a peak-occupancy reduction from 240 to 30 qubits ($8.0\times$) under the stated constant-support and buffer premises; a raw-state efficiency ratio of $8.0\times$; a recycling budget of at most 10 reuses under the projected $\delta = 10^{-5}$, with an uncertainty span of 1–100 reuses; and, under an explicitly adopted energy-attribution convention anchored to publicly reported benchmark data [10], $\approx 0.55 \pm 0.07$ J per distilled magic state against a 1.5×10^7 quantum-to-classical energy ratio. The proposal thus converts a spatial cost into an operational-quality cost and a temporal serialization cost. The decisive open parameter is the per-reuse error δ ; the decisive structural risk is whether the constant-support premise survives contact with lower-bound arguments of the kind known in adjacent support-size theory [2] and with the engineered baselines of the compact-factory landscape [6]. The analysis is offered as a falsifiable framing for experimental follow-up rather than as a demonstration of feasibility; every input is stated, every step of arithmetic is shown, and every projection is labeled, so that the claim can be checked, tightened, or falsified as the protocol details become available.

References

-
- [1] Constant sized support state distillation with qubit recycling. arXiv:2609.17044v1. <https://arxiv.org/abs/2609.17044v1> [2] Polylogarithmic Supports are required for Approximate Well-Supported Nash Equilibria below $2/3$. arXiv:1309.7258v2. <https://arxiv.org/abs/1309.7258v2> [3] Human in the Loop: Interactive Passive Automata Learning via Evidence-Driven State-Merging Algorithms. arXiv:1707.09430v1. <https://arxiv.org/abs/1707.09430v1> [4] A Simulation and Modeling of Access Points with Definition Language. arXiv:1304.1836v2. <https://arxiv.org/abs/1304.1836v2> [5] Superconductivity as a consequence of an ordering of the

electron gas zero-point oscillations. arXiv:1005.0280v6. <https://arxiv.org/abs/1005.0280v6> [6] Exploring the landscape of compact magic-state distillation factories. arXiv:2606.07734v2. <https://arxiv.org/abs/2606.07734v2> [7] Intuitionistic Fuzzy Ideals in Γ -semiring. arXiv:1011.5746v2. <https://arxiv.org/abs/1011.5746v2> [8] Unifying gate-synthesis and magic state distillation. arXiv:1606.01906v2. <https://arxiv.org/abs/1606.01906v2> [9] Internal Location Based System For Mobile Devices Using Passive RFID And Wireless Technology. arXiv:1001.2258v2. <https://arxiv.org/abs/1001.2258v2> [10] DOI 10.5281/zenodo.21623218. QNFO: The BQNN Classical Baseline: Constructive Falsification of Near-Term Quantum Advantage at a Fifteen-Million-to-One Energy Disadvantage. [11] DOI 10.5281/zenodo.21880104. QNFO: The Qudit Advantage: System-Level Joules-per-Solution Comparison of a Qudit Architecture Against 17 Conventional Qubit Quantum Computing Platforms. [12] DOI 10.5281/zenodo.21566035. QNFO: Auditing the BQNN: Does a Tunable Quantum Neural Network on Trapped-Ion and Superconducting Hardware Demonstrate a Route to Near-Term Quantum Advantage?. [13] DOI 10.5281/zenodo.21515894. QNFO: Due Diligence Report: QuiX Quantum.