Exact Shallow Quantum Computation with Copies of the Input

Shiva Kintali*

September 5, 2026

Abstract

We study constant-depth quantum circuits supplied with multiple computational-basis copies of a classical input. In the model with arbitrary single-qubit gates and generalized Toffoli gates, we give exact polynomial-size computation of the weight-equality predicate EX_j using $O(1 + \min(j, n - j))$ copies. A structure-sensitive sum gives exact computation of arbitrary symmetric functions with $O(n^2)$ copies. The construction combines coherent bounded-load compression with input-preserving, known-probability amplitude filters. We give an exact polynomial-time formula for the additional gate parameters, and first present an elementary quadratic-copy equality construction that already implies exact copied-input simulation of TC^0 and a total-language separation $\mathsf{EQAC}^0 \not\subseteq \mathsf{AC}^0[p]$ for every prime p. On the lower-bound side, two entangling layers require supplied-copy counts exponential in input length for parity, and exponential in transition radius for general symmetric functions, even with bounded error and unrestricted product-state ancillary width. We also determine exact copy minima in a further restricted fresh-output architecture.

1 Introduction and relation to prior work

Overview and motivation. A classical circuit may route one input bit to many gates without charging depth for copying. A quantum circuit with disjoint physical gate supports cannot make that assumption. Supplying $|x\rangle^{\otimes m}$ isolates the cost of input access from the remaining computational restrictions. The input is classical, but the circuit is unitary and may create coherent intermediate states. We ask how the supplied-copy count interacts with exact computation and with very small depth.

This question separates two obstacles: how many disjoint physical occurrences of each input bit are available, and what constant-depth interference and unbounded-arity controlled operations can accomplish with those occurrences. Symmetric Boolean functions, whose values depend only on Hamming weight, provide a natural test case. Input copies supply additional access to the same classical data, not additional unknown information about the input.

Copies versus depth: the central contrast. The paper's organizing claim is that *polynomially many classical input copies enable exact computation of every symmetric function at constant depth, whereas two entangling layers still require exponentially many copies for parity and symmetric functions with a central transition.* The constructive side combines coherent compression with input-preserving amplitude filters to obtain zero error and weight-sensitive copy bounds. It also

*Email: shiva.kintali@gmail.com Homepage: shivakintali.github.io

yields exact simulation of each fixed-depth TC^0 family with polynomially many copies, with quantum depth depending on the classical depth. The lower-bound side allows bounded error and unrestricted product-state ancillary width. For general symmetric functions, the exponential lower bound is in the transition radius defined below, not necessarily in the input length. Thus the main comparison is between polynomial-copy exact upper bounds at constant depth and exponential-copy lower bounds at depth two; it is not a characterization of optimal copy complexity at every constant depth.

Conventions and scope. We use arbitrary one-qubit unitaries and generalized Toffoli gates, initially zero ancillas, and disjoint physical gate supports within each entangling layer. There is one final measured output bit; garbage is allowed, but intermediate measurements, postselection, and free unbounded fanout are not. Upper bounds are nonuniform and have polynomial size and total width in the original input length n. Exactness means correctness with probability one on every basis input $|x\rangle^{\otimes m}$; no action on inconsistent copies is prescribed. Unless stated otherwise, lower bounds allow arbitrary finite pure product-state ancillary width and worst-case error at most $1/3$. Pure product-state ancillas can be absorbed into the initial local state preparations; pre-entangled ancillas are not free in the two-layer results. Section 2 gives the formal definitions.

Foundational context and neighboring resources. Moore introduced the QAC framework and constant-depth connections among parity, cat-state preparation, and fanout [1]; Green, Homer, Moore, and Pollett developed counting and modular-gate equivalences [2]. Here coherent fanout maps $|b, t_1, \ldots, t_s\rangle$ to $|b, t_1 \oplus b, \ldots, t_s \oplus b\rangle$, extended linearly to quantum inputs. This is stronger than receiving copies of a classical input. With fanout as a primitive, Høyer and Špalek gave constant-depth approximate counting and symmetric-function constructions [3]; Takahashi and Tani later proved the exact collapse $\mathsf{QNC}^0_{\mathrm{f}} = \mathsf{QAC}^0_{\mathrm{f}} = \mathsf{QTC}^0_{\mathrm{f}}$, where the subscript f denotes access to fanout [4]. Grier and Morris showed that quantum threshold gates can approximate fanout [6]. Joshi and Vasconcelos study constant-depth unitary preparation of Dicke states, including exact arbitrary-weight preparation with fanout [7]. These are neighboring resources, not implementations of free unbounded fanout in our copied-input model.

Previous copied-input upper bounds. The closest predecessor is Grier, Morris, and Wu [11]. Their Theorems 21 and 24 establish bounded-error copied-input TC^0 simulation and a total BQAC^0 language outside $\mathsf{AC}^0[p]$ for every prime p. Their Corollary 30 gives bounded-error copy bounds $n/\operatorname{polylog} n$ for equality and $n^2/\operatorname{polylog} n$ for arbitrary symmetric functions. Their Theorem 31 gives a randomized-circuit ensemble for parity using $n^{3/2}/\operatorname{polylog} n$ copies, with a per-input success guarantee averaged over the ensemble and measurement; it is not an exact evaluator. Their Corollary 16 already gives one-copy exact computation of every symmetric AC^0 family, including equality within any fixed-polylogarithmic distance of an endpoint. The underlying classical characterization is due independently to Brustmann and Wegener and to Moran [8, 9]; related small-wire constructions are due to Håstad, Wegener, Wurm, and Yi [10]. We import the exact endpoint and limited-fanout primitives, rather than claim them as new.

What is novel: upper bounds and exact consequences. The constructive half of this comparison is an exact, structure-sensitive compiler (Theorem 2.2): supplied copies support parallel evaluation, while compression and interference make the cost depend on which weights must be recognized. Write f_w for the value of f at weight w, and set $E_\beta = \{w \in \{0, \ldots, n\} : f_w \neq \beta\}$ for $\beta \in \{0, 1\}$. Define

$$D(f) = \min_{\beta \in \{0,1\}} \sum_{w \in E_\beta} (1 + \min\{w, n - w\}).$$

The compiler uses $O(D(f))$ copies; constants need none. For every fixed $a > 0$, the imported limited-fanout primitive reduces the supplied count for a nonconstant function to

$$O\left(\max\left\{1, \frac{D(f)}{\log^a(n+2)}\right\}\right),$$

with depth and polynomial-size exponent allowed to depend on a. In particular, $\mathsf{EX}_j(x) = \mathbf{1}[|x| = j]$ costs $O(1+\min\{j, n-j\})$ before this saving. For fixed $0 < \alpha < 1$ and $j = \lfloor n^\alpha \rfloor$, this is $O(n^\alpha)$ rather than the generic near-linear bounded-error equality bound above. This example is outside the already known fixed-polylogarithmic endpoint regime. For arbitrary symmetric functions $D(f) = O(n^2)$, so the worst-case copy exponent remains two: the improvement is exactness and weight-sensitive accounting, not a subquadratic worst-case bound or a matching optimality theorem.

The exact compiler also yields $\mathsf{TC}^0 \subseteq \mathsf{EQAC}^0 \circ \mathsf{NC}^0$ and a single total language $L \in \mathsf{EQAC}^0$ with $L \notin \mathsf{AC}^0[p]$ for every prime p (Corollaries 2.3 and 8.1). Here EQAC^0 denotes exact decision on ordinary physical inputs, and the NC^0 preprocessing explicitly duplicates classical bits into polynomially many output positions. These consequences strengthen the corresponding error guarantees in [11] to zero error; they do not give one-copy exact TC^0 simulation. The total language has exponentially small acceptance density in its physical input length, but is not sparse in the standard sense of having only polynomially many accepted strings per length. The separation is worst-case, not average-case, and does not separate EQAC^0 from all of ACC^0.

Overview of the upper-bound proof. We first give an elementary exact equality construction using $O(n^2)$ copies in Section 4.2; this already suffices for the qualitative class consequences. For the sharper bound, coherent hashing sends active input positions into buckets while retaining both input wires and color records. Threshold guards reject overloaded branches coherently, without postselection. Clean occupation-parity oracles let many unary rows interrogate one compressed string. Retained orthogonal color records and column-permutation symmetry make the actual guarded acceptance probability depend exactly on input weight, not merely approximately so.

Exactification uses phase gates on initialized private work and a source that preserves every physical input-basis sector, including inconsistent copies (Section 4). A matched phase filter kills a prescribed acceptance probability in $[0, 3/4]$ while fixing probabilities zero and one; a companion filter makes a target probability in $[1/4, 1]$ exactly one. The repeated sources satisfy these respective hypotheses. For target $k = \min\{j, n-j\}$ outside the imported small-weight regime, the compressed capacity satisfies $L \le 5k$. The wrong-weight modules use

$$R_w = \left\lceil \frac{21L}{5(w-k)^2} \right\rceil, \qquad 0 \le w \le L, \quad w \ne k.$$

Thus $\sum_{w \ne k} R_w \le L + (42L/5)\sum_{d \ge 1} d^{-2} = O(L) = O(k)$, including the cost of rounding; weights above L are rejected by the guard. Modules act in parallel on private input roles, and each role is reused sequentially inside its constant-depth filter. The exact probabilities are evaluated by a fixed three-variable generating function with polynomially bounded degrees and polynomial-bit integer coefficient tables (Section 7). This supplies polynomial-time algebraic descriptions of the additional phase-gate entries relative to the imported primitives, not an exact finite-gate-alphabet or stronger uniformity theorem. Pooling equality modules over E_β, clean compute–copy–uncompute, private formula expansion, and a sequential consistency check then give the stated symmetric-function and total-language consequences.

Known-probability amplitude amplification and exact phase matching are established techniques [13, 12], already used for QAC constructions in [11]. Our contribution is their input-preserving implementation and composition with bounded-load compression, distance-adapted

repetition, and exact parameter extraction. Neither amplitude amplification itself nor an isolated scalar phase-matching identity is claimed as a new principle.

Previous lower bounds. The ordinary-input literature constrains different resources and sometimes uses a stronger quantum-output interface. Fang, Fenner, Green, Homer, and Zhang prove logarithmic-depth exact parity and fanout lower bounds with constantly many ancillas, and depth–ancilla tradeoffs [5]. Rosenthal develops generalized-reflection normal forms, constant-depth approximate parity upper bounds using superpolynomial size, and depth-two lower bounds for approximation of the parity operation on quantum inputs [14]. Fenner, Grier, Padé, and Thierauf exclude exact depth-two parity of more than three non-target input bits, even with garbage and arbitrarily many ancillas [15]. For one-output classical decision, Joshi, Tal, Vasconcelos, and Wright exclude exact depth-three parity for $n > 100$, independently of size, and prove depth-two total influence $O(\log n)$ and parity-correlation bounds [18, Theorems 1.2–1.3 and Corollary 1.4].

For unrestricted constant depth but restricted ancillary width, Nadimpalli, Parham, Vasconcelos, and Yuen obtain average-case lower bounds in the $n^{O(1/d)}$-ancilla regime by Pauli analysis [16]. Operator approximations give the degree exponent $1 - 3^{-d}$ in Anshu, Dong, Ou, and Yao [17], improved to $1 - 2^{-d}$ by Dong, Ou, and Yao [21], with superlinear width consequences. Gretta, Gupta, and Joshi establish a quantitative equivalence between parity hardness and a quantum Fourier-concentration conjecture, and give one-copy circuits approximating majority on uniformly random inputs [19, Theorems 1.1 and 1.4]. The latter average-case guarantee is distinct from exact, worst-case symmetric-function evaluation here.

What is novel: copied-input lower bounds. The complementary half asks how much depth can be replaced by additional copies. Ordinary-input theorems do not automatically apply to a circuit constrained only on the copy code $\{|x\rangle^{\otimes m} : x \in \{0,1\}^n\}$. Identifying physical coordinates changes influence and Fourier mass, and off-code behavior is unconstrained; Section 11 illustrates this explicitly. Our second principal contribution is a direct bound for this interface: two entangling layers require $m = \Omega(2^{n/2}/\sqrt{n+1})$ for worst-case bounded-error parity (Theorem 2.4). For a nonconstant symmetric function define

$$\rho(f) = \max_{\substack{0 \le i < n \\ f_i \ne f_{i+1}}} \min\{i+1, n-i\},$$

and put $\rho(f) = 0$ for constants. Theorem 10.1 gives, when $\rho = \rho(f) \ge 5$,

$$m \ge \frac{2^{(2\rho-2)/9}}{3000(2\rho+1)^{10/9}}.$$

Both bounds allow any output wire and unrestricted product-state ancillary width. A central transition ($\rho = \Theta(n)$) therefore forces exponentially many copies in n; no such conclusion is asserted for a bounded endpoint radius. We use the radius in the form of Xu and Li's ordinary-input fanout characterization [20, Theorem 1]; it is also an affine reparameterization of Paturi's earlier jump parameter [22]. The contribution here is a quantitative two-layer copy bound, not the parameter itself or that fanout characterization. Appendix A additionally determines exact copy minima for the specified equality, threshold, and interval families in a restricted architecture: the output starts in a fixed state and is untouched by the first entangling layer. Those optima do not apply to an unrestricted output source or to bounded error.

Overview of the lower-bound proofs. For two layers, local basis changes put each generalized Toffoli into product-reflection form. A source is a first-layer block, including idle singleton wires. Fix every logical label represented in the output's source, hence all copies of each such label. If $k \geq 1$ labels remain, the output probability has the form $p(x) = a + b\prod_g c_g(x)$, with $0 \leq c_g \leq 1$. A source missing any remaining label has exactly zero even–odd mean difference; a source containing all k labels has difference at most $6/2^k$ (Lemma 9.1). Product-factor deficit inequalities then force many full-label sources, each charged a separate occurrence of every remaining bit. Deleting a large first-layer output gate, and explicitly paying for the resulting exceptional inputs, completes the unrestricted-output parity bound.

For a symmetric transition, nonnegative deficits $1-c_g$ are transported between adjacent Hamming slices. A source contraction is a scalar product function plus the diagonal of an operator of trace norm at most two. Only the scalar product part is represented as a positive mixture of subcube indicators; no such representation is asserted for the signed quantum correction. Small sources obey a direct slice-marginal comparison. For large sources, the subcube argument transports the product part, while the correction is suppressed by the exponentially small maximum marginal atom. After controlling the output source and restricting to a deepest transition, this proves the radius-dependent bound.

The fresh-output proofs use exact certainty to constrain the pre-reflection positive defect, compressed to the span of local certainty words, to rank at most one. Rank alone does not imply rectangular support: the additional product structure and the certainty words supplied by the specified interior slices force the actual source deficit to be supported on a Boolean rectangle (a subcube obtained by fixing some bits). Boundary inputs can then be charged to distinct physical copies. Endpoint predicates admit additional star-shaped defects, whose optimum reduces to a tournament-load calculation. The all-depth appendix follows a different route: Chebyshev polynomials approximate product projectors, Pauli degree is propagated backwards through disjoint entangling layers, and parity degree or symmetric approximate degree gives a physical-width bound.

What is derived. The exact total-language separation invokes Smolensky's modular lower bound [23]; the transition-degree comparison invokes Paturi's theorem [22]. The all-depth physical-width estimate is a copied-input consequence of the operator-approximation method and the $1-2^{-d}$ exponent of Dong, Ou, and Yao [21]; no novelty is claimed for that exponent. Its copy consequence additionally assumes that total workspace is $O(nm)$. Accordingly, this paper proves neither an unrestricted all-depth supplied-copy lower bound nor an optimal worst-case copy exponent.

In summary, the claimed new results relative to the cited literature are the structure-sensitive exact compiler, the direct two-layer copy lower bounds, and the restricted fresh-output optima. Their technical contributions are the copy-preserving composition of exact filters with compression and the copy-sensitive source and defect-transport analysis. Exact copied-input TC^0 simulation and the total-language separation are consequences of the compiler and established circuit and modular-lower-bound techniques, not independent new lower-bound methods. The fresh-output optima sharpen the copies-versus-depth picture in a restricted architecture; the all-depth width consequence, derived from prior operator approximation, provides supporting context under an additional workspace premise. Neither replaces the central contrast or establishes optimal copy complexity at arbitrary constant depth. These distinctions concern mathematical content relative to the cited results, not a certification of exhaustive publication priority.

Organization of the paper. Section 2 fixes the model and states the quantitative conclusions. Section 3 records imported exact primitives. Section 4 constructs the work-only filters and the

elementary compiler. Sections 5, 6, and 7 develop compression, weight-conditioned batches, and distance-adapted exact compilation, respectively. Section 8 proves the class consequences. Sections 9 and 10 prove the parity and symmetric-transition lower bounds. Section 11 distinguishes physical-input influence from behavior on the copy code, and Section 12 records limitations. Appendix A gives fresh-output optima; Appendix B derives the workspace-conditioned all-depth bound. Independent subset-state samples supplied to an unrestricted receiver form a different statistical model and are not treated as free shallow-computation resources here.

2 Model and quantitative conclusions

All logarithms are natural unless a base is specified. The input length is an integer $n \geq 1$, and $|x| = \sum_{i=1}^{n} x_i$ denotes Hamming weight.

Circuit and resource conventions. The input is $|x\rangle^{\otimes m}$, where $x \in \{0,1\}^n$ and $m \geq 0$ is an integer. The circuit also receives $A \geq 0$ input-independent ancillas in $|0^A\rangle$, with $nm + A \geq 1$. A generalized Toffoli gate with controls $c_1, \ldots, c_t$ and a distinct target b acts by

$$|c_1, \ldots, c_t, b\rangle \longmapsto |c_1, \ldots, c_t, b \oplus (c_1 \cdots c_t)\rangle, \qquad t \geq 1.$$

Arbitrary one-qubit unitaries are allowed. Entangling depth counts layers of generalized Toffoli gates with pairwise disjoint physical supports, allowing arbitrary one-qubit layers before, between, and after them. A Hadamard on the target before and after a generalized Toffoli turns it into a controlled-Z product reflection, without changing this depth. One designated physical qubit, possibly an input qubit, is measured in the computational basis at the end. Garbage is allowed. There is no intermediate measurement, postselection, free classical control depending on x, or free unbounded fanout.

Upper bounds concern nonuniform families with constant entangling depth and polynomial total gate count and width $nm + A$, measured in the original length n. Constants in these bounds may depend on explicitly fixed parameters, but not on n. Arbitrary exact one-qubit gates are part of the model; no finite gate alphabet is assumed. Exact decision has zero error, independently of any positive error tolerances used to analyze intermediate constructions. Supplied copies are charged at the input, whereas any internally generated roles and their preparation circuits are charged to width, size, and depth.

When an intermediate bounded-error construction has a freely chosen positive accuracy parameter, its constant-depth polynomial-resource guarantee assumes that parameter is at least $(n+2)^{-C}$ for some fixed $C > 0$; the resource constants may depend on C. This convention does not restrict the final zero error of an exact construction or replace its explicitly specified internal tolerances.

The two-layer lower bound below allows arbitrary finite size and ancillary width. It also allows any input-independent pure product ancillary state, since its preparation is a one-qubit layer. This does not allow pre-entangled ancillas. The different initialization and workspace assumptions in Appendix B are stated there separately.

Definition 2.1 (Copy code and decision)**.** For $m \geq 1$, the copy code is the set of basis vectors

$$\mathcal{C}_{n,m} = \{|x\rangle^{\otimes m} : x \in \{0,1\}^n\}.$$

An exact evaluator for $f : \{0,1\}^n \to \{0,1\}$ is a unitary circuit U, with designated output o, such that

$$U|x\rangle^{\otimes m}|0^A\rangle = |f(x)\rangle_o \otimes |g_x\rangle \quad \text{for every } x \in \{0,1\}^n.$$

The right side places the output factor first as a notational convention, not as a physical routing operation. The vector $|g_x\rangle$ is normalized and is subject to unitarity, but no input preservation or cleanup is required. In particular, no initialized action on inconsistent copies is prescribed. Correctness on superpositions in the span of the code follows by linearity; it does not assert cloning of arbitrary quantum states.

For $0 \le \varepsilon < 1/2$, a worst-case ε-error evaluator instead satisfies

$$\Pr[\text{output} = f(x) \mid |x\rangle^{\otimes m}] \ge 1 - \varepsilon \quad \text{for every } x.$$

At $m = 0$ the supplied input is the same empty tensor for every x, so only constant functions can be evaluated with error below $1/2$; they need no input copy. The all-sector input preservation required of filtered sources is a separate, stronger hypothesis in Section 4.

Language classes and preprocessing. All language classes here are nonuniform. The class EQAC^0 consists of total Boolean decision families exactly computed by constant-depth polynomial-size circuits of the stated gate basis, with polynomial total width and one ordinary physical input string at each length. The class $\mathsf{AC}^0[p]$, for a fixed prime p, permits unbounded-fanin AND, OR, and modulo-p gates, NOT gates, and classical wire reuse. The class TC^0 has polynomial size and constant depth with unweighted threshold gates $[|y| \ge t]$, constants, NOT gates, and classical wire reuse.

A classical NC^0 preprocessing map has polynomial output length and polynomial size, with constant bounded-fanin depth and unrestricted classical wire reuse. In $\mathsf{EQAC}^0 \circ \mathsf{NC}^0$, a total quantum decision family is applied to the output of such a map. Our containment uses the explicit map $x \mapsto (x, \dots, x)$ with polynomially many output blocks; it grants no free copying inside the quantum circuit. The consistency check in Section 8 converts the copy promise into a total language by rejecting inconsistent blocks.

For comparison, Grier, Morris, and Wu give a *bounded-error* supplied-copy upper bound $n^2/\operatorname{polylog} n$ for arbitrary symmetric functions [11, Corollary 30]. Their exact limited-fanout and small-threshold primitives are imported here [11, Corollary 10 and Lemma 12], not claimed as copy lower bounds or optimality results.

Symmetric-function parameters. A function is symmetric if its value depends only on $|x|$; write f_w for its value at weight $w \in \{0, \dots, n\}$. Define $\mathsf{EX}_j(x) = \mathbf{1}[|x| = j]$ for integers $0 \le j \le n$. For $\beta \in \{0, 1\}$ set

$$E_\beta = \{w \in \{0, \dots, n\} : f_w \ne \beta\},$$
$$D_\beta(f) = \sum_{w \in E_\beta} (1 + \min\{w, n - w\}), \qquad D(f) = \min_{\beta \in \{0,1\}} D_\beta(f).$$

For a nonconstant function, its transition radius is

$$\rho(f) = \max_{\substack{0 \le i < n \\ f_i \ne f_{i+1}}} \min\{i + 1, n - i\};$$

put $\rho(f) = 0$ for constants. Thus $1 \le \rho(f) \le \lceil n/2 \rceil$ for nonconstants. We use the formulation in Xu and Li [20, Section 2.2]. For a nonconstant function, Paturi's jump parameter [22] is

$$\Gamma(f) = \min_{\substack{0 \le i < n \\ f_i \ne f_{i+1}}} |2i - n + 1|; \qquad n - \Gamma(f) = 2\rho(f) - 1.$$

The identity follows from $|2i - n + 1| = n + 1 - 2\min\{i + 1, n - i\}$ at each transition. Fanout gates in the cited fanout reductions are unitary resources, not supplied classical copies.

Theorem 2.2 (Exact structure-sensitive copy upper bound)**.** *Every family of symmetric Boolean functions has exact constant-depth polynomial-size, polynomial-width evaluators in the model above using $m = O(D(f))$ supplied copies. Constants need none. In particular:*

1. *Each endpoint predicate $\mathsf{EX}_0, \mathsf{EX}_n$ uses one copy. Every interior predicate EX_j, $1 \le j \le n-1$, uses at most $90\min\{j, n-j\}$ copies, hence $O(1+\min\{j, n-j\})$ copies for all weights.*
2. *A function with radius $r = \rho(f)$ uses $O(r + r^2)$ copies.*
3. *For every fixed $b > 0$, equality at weights satisfying $\min\{j, n-j\} = O(\log^b(n+2))$ has an exact one-copy implementation, with depth and size exponent allowed to depend on b and the fixed constant in this bound.*

For every fixed $a > 0$, exact limited fanout further gives

$$m = O\left(\max\left\{1, \frac{D(f)}{\log^a(n+2)}\right\}\right)$$

for nonconstant functions. The depth and polynomial-size exponent may depend on a. These are nonuniform arbitrary-angle statements. Section 7 gives polynomial-time algebraic descriptions of the additional filter-gate entries relative to the imported primitive families, not an exact finite-alphabet or stronger uniformity claim.

Derivation from the equality construction. Sections 5–7 construct the exact equality modules: the large-weight branch uses fewer than $90k$ copies for $k = \min\{j, n-j\}$, and the small-weight branch uses one. Endpoints are direct AND or NOR gates. Run private modules for $j \in E_\beta$, OR their flags, and complement the result when $\beta = 1$. Their total cost is at most $90D_\beta(f)$; minimizing over β gives the first bound. The number of modules is at most $n+1$, so parallel composition preserves constant depth and polynomial resources.

For $r \ge 1$, the weights in $[r, n-r]$, if that interval is nonempty, share a common value β: a transition between two weights in that interval would have radius at least $r+1$. If the interval is empty, choose either value for β. All exceptional weights then lie in $\{0, \dots, r-1\} \cup \{n-r+1, \dots, n\}$. Consequently

$$D(f) \le D_\beta(f) \le 2\sum_{w=0}^{r-1}(1+w) = r(r+1).$$

This also covers the maximal odd-length radius, where the middle interval is empty; constants are handled separately. Item 3 follows from the imported exact threshold and limited-fanout primitives in Section 3. Finally, for a pool of $S > 0$ requested copies, generating $s = \lceil \log^a(n+2) \rceil$ roles per supplied bit reduces the supplied count to $\lceil S/s \rceil$. The pooling circuit is outside all filtered sources, as proved in Section 7. □

For example, fix $0 < \alpha < 1$ and put $j = \lfloor n^\alpha \rfloor$. The equality bound gives $O(n^\alpha)$ copies before logarithmic savings, or $O(\max\{1, n^\alpha/\log^a(n+2)\})$ for any fixed $a > 0$. This improves on the generic near-linear bounded-error equality bound [11, Corollary 30] and lies outside the fixed-polylogarithmic endpoint regime. A symmetric family with $\rho(f) = O(n^\alpha)$ similarly uses $O(n^{2\alpha})$ copies before pooling. For arbitrary symmetric functions,

$$D(f) \le \frac{1}{2}\sum_{w=0}^{n}(1+\min\{w, n-w\}) = \frac{n+1+\lfloor n^2/4 \rfloor}{2}.$$

Thus the worst-case exponent in this upper bound remains two. None of these upper bounds is an optimality claim.

Corollary 2.3 (Exact copied-input power)**.** *For every nonuniform* TC^0 *Boolean decision family there is an exact constant-depth polynomial-size, polynomial-width evaluator with polynomially many supplied copies. Its depth and size exponent may depend on the fixed classical circuit family. With explicit classical duplication and total quantum decision families this gives*

$$\mathsf{TC}^0 \subseteq \mathsf{EQAC}^0 \circ \mathsf{NC}^0.$$

There is a single total language $L \in \mathsf{EQAC}^0$ *such that* $L \notin \mathsf{AC}^0[p]$ *for every prime* p*.*

Proof reference and scope. Section 8 cleans exact symmetric-gate evaluators by compute–copy–uncompute and expands each fixed-depth classical circuit into private gate occurrences. The resulting copy count and all other resources are polynomial. It then checks consistency before evaluation, rejecting inconsistent blocks and nonencoded lengths, to obtain total quantum decision families. The single-language separation is Corollary 8.1, using Smolensky's modular lower bound [23]. Both qualitative conclusions already follow from the elementary polynomial-copy construction of Proposition 4.1, with a larger polynomial padding. Neither conclusion removes the classical duplication step or asserts a separation from all of ACC^0. □

Theorem 2.4 (Two-layer copy lower bound)**.** *Let* $n \geq 2$ *and let a finite circuit of entangling depth at most two compute* $\mathrm{PARITY}_n(x) = x_1 \oplus \cdots \oplus x_n$ *from* m *supplied copies with worst-case error at most* $1/3$*. Its designated output may be any physical qubit, and its first-layer output gate, if present, is unrestricted. The number of input-independent pure product ancillas is unrestricted. If* G_1 *counts generalized Toffoli gates in the first entangling layer (padding missing layers with empty layers), then*

$$m(G_1 + 2) \geq c_0\, 2^n, \qquad m(nm + 1) \geq c_0\, 2^n,$$

and hence

$$m \geq c\, \frac{2^{n/2}}{\sqrt{n+1}}.$$

Here the absolute positive constants are

$$a_0 = \frac{2/3 + \log(2/3)}{6}, \qquad c_0 = \frac{1 + \log(2/5)}{12 + 9216/a_0}, \qquad c = \sqrt{c_0}.$$

There is no polynomial-size assumption. Thus polynomially many supplied copies cannot suffice at this depth, even with arbitrarily large finite product-ancillary workspace.

Proof reference and final constant calculation. Section 9 proves the copy-sensitive source estimate, controls a large output block by deletion, and obtains

$$c_* 2^n \leq (12 + 9216/a_0) m(M+1), \qquad c_* = 1 + \log(2/5),$$

where the number M of nonconstant sources after deletion and merging idle wires obeys $M \leq G_1 + 1$ and $M \leq nm$. These inequalities give the two product bounds. Since parity is nonconstant, $m \geq 1$; thus $nm + 1 \leq (n+1)m$, giving the square-root bound. The constants are positive: $\log(3/2) < 1/2$ gives $a_0 > 1/36$, and $e > 1 + 1 + 1/2 + 1/6 = 8/3 > 5/2$ gives $c_* > 0$. □

Ordinary-input depth-two results and product-reflection normal forms have important precedents [14, 15]. Their initialization and approximation interfaces must not be identified with the repetition-code promise; Section 11 explains this distinction. The lower bound just stated is proved directly on the supplied-copy interface. The radius-dependent two-layer extension is stated and proved in Section 10; it is separate from the parity bound and its constants.

3 Exact primitives and circuit interfaces

This section specifies the exact interfaces used later. The existence results are imported from Grier, Morris, and Wu [11, Corollary 10, Lemma 12, Fact 14, and Theorem 17]; the constructions below explain their composition and resource accounting. Fix an ambient integer $n \geq 2$ and put $\lambda_n = \log(n+2)$. All polynomial bounds in this section refer to n, not necessarily to a module's own input length. The number of data wires may be polynomial in n, while a threshold value or fanout length is a fixed power of λ_n.

Clean functional oracles. For a total Boolean function $f : \{0,1\}^q \to \{0,1\}$, a clean functional oracle is a unitary circuit satisfying

$$O_f|y\rangle_Y|b\rangle_B|0^\ell\rangle_W = |y\rangle_Y|b \oplus f(y)\rangle_B|0^\ell\rangle_W \quad (y \in \{0,1\}^q,\ b \in \{0,1\}).$$

The data, target, and private-work registers are disjoint. This identity extends by linearity to arbitrary superpositions and reference-entangled states of YB, provided W starts in its indicated zero state. It does not specify the circuit's action on nonzero private work.

If an exact decision circuit U of entangling depth d has initialized action $U|y\rangle|0^\ell\rangle = |f(y)\rangle_o|g_y\rangle$, use a separate target B not acted on by U. The chronological sequence U, $\mathrm{CNOT}_{o\to B}$, $U^\dagger$ gives

$$|y\rangle|0^\ell\rangle|b\rangle_B \longmapsto |f(y)\rangle_o|g_y\rangle|b \oplus f(y)\rangle_B \longmapsto |y\rangle|0^\ell\rangle|b \oplus f(y)\rangle_B.$$

All garbage and input-dependent phases reverse. The depth is at most $2d+1$, and no additional supplied input copy is needed. This remains valid if o is one of the original data wires. The guarantee is initialized clean action, not block diagonality on arbitrary private-work states.

Proposition 3.1 (Imported exact primitives)**.** *For each pair of fixed constants $a, C > 0$, nonuniform constant-depth circuits with polynomial gate count and total width in n provide the following operations in the arbitrary-one-qubit-gate model:*

1. *For every integer $0 \leq s \leq \lceil \lambda_n^a \rceil$, clean fanout on one control and s targets:*

$$|b, t_1, \ldots, t_s\rangle|0_{\mathrm{work}}\rangle \longmapsto |b, t_1 \oplus b, \ldots, t_s \oplus b\rangle|0_{\mathrm{work}}\rangle.$$

   This is a unitary interface for arbitrary target bits, not only zero targets.

2. *For integers $0 \leq q \leq n^C$ and $0 \leq h \leq \lambda_n^a$, a clean oracle for the total threshold $[|y| \geq h]$ on q data bits. For $0 \leq h \leq q$ in this range, clean oracles also exist for EX_h and EX_{q-h} on those bits and for their output complements.*

3. *For each integer $0 \leq q \leq \lambda_n^a$, a clean oracle for every symmetric Boolean function on q bits.*

Depth and polynomial-resource constants may depend on a, C, but not on n or the particular legal arities. Every data vector is supplied once; extra roles used internally are generated by the circuit.

Imported results and their interfaces. The fanout unitary is defined in [11, Section 2], and its fixed-polylogarithmic implementation is Corollary 10 there. Lemma 12 supplies exact small-threshold decision, and Fact 14 supplies exact small-arity symmetric decision; its Appendix A proof gives the construction. The clean-oracle conversion above applies to the decision primitives.

The parameters can be reconciled with the source's ambient convention as follows. Choose a fixed polynomial $P(n) = \lceil (n+2)^B \rceil$, with constant B large enough to dominate the data lengths under consideration, before allocating primitive workspace. Pad a threshold's data vector with

zeros to length $P(n)$, so its Hamming weight is unchanged. Since $\log P(n) = \Theta(\log(n+2))$, each fixed-power cutoff here is covered by a suitable fixed-power cutoff in the source. Enlarging that exponent absorbs fixed multiplicative constants and integer rounding. The polynomial resource bounds in $P(n)$ remain polynomial in n. The ambient length is not defined recursively in terms of the workspace produced by a primitive. Fact 14 already permits small data arity relative to a larger ambient size.

If the available fanout has $S \geq s$ targets, prepare the $S - s$ unused targets in $|+\rangle = H|0\rangle$. Since $X|+\rangle = |+\rangle$, fanout leaves them unchanged and unentangled with the data; applying H afterwards returns them to zero. Thus reducing the arity preserves the clean interface even on superpositions. Padding those targets with zeros and leaving them untouched afterwards would not suffice: they would retain the control bit.

For equality, compute the clean flags $[|y| \geq h]$ and $[|y| \geq h+1]$ sequentially on the same data, combine them as

$$\mathsf{EX}_h(y) = [|y| \geq h] \wedge \neg[|y| \geq h+1],$$

and reverse the flag computations. Using the flags as controls toggles an arbitrary output target and leaves both flags intact for uncomputation. The extra value $h+1 \leq 2\lambda_n^a$ is covered by the fixed-constant enlargement just described. Conjugating the data by $X^{\otimes q}$ gives $\mathsf{EX}_{q-h}(y) = \mathsf{EX}_h(\neg y)$ and restores the data afterwards. Complementing the output function requires one additional X on the oracle target.

Fanout with $s = 0$ is the identity. A threshold with $h = 0$ is constant one; if $h > q$ it is constant zero. The case $q = 0$ and all other constant predicates are implemented by either the identity or X on the target. Finitely many small ambient lengths can be handled by finite circuits without changing the asymptotic bounds. Item 2 is a total-input statement, not a promise that $|y|$ is small. □

A total truncated-parity oracle. Let $u \in \{0,1\}^q$ be an occupation column of polynomial length, and let $J \geq 0$ be an integer with $J = O(\lambda_n)$. Define

$$h_J(c) = \min\{c, J\} \bmod 2, \qquad 0 \leq c \leq q.$$

For every such c,

$$h_J(c) = \bigoplus_{v=1}^{J} \mathbf{1}[c \geq v],$$

because exactly $\min\{c, J\}$ of these flags are one. The empty XOR at $J = 0$ is zero.

For $J \geq 1$, let F generate J fresh roles for each occupation bit using item 1. Roles belonging to different original bits have disjoint supports. Compute, in parallel, the clean threshold flags $\theta_v = [|u| \geq v]$ on their respective private role vectors; denote this parallel circuit by T. Item 2 applies even when $v > q$, in which case the flag is zero. Use item 3 to compute the parity of the J flags into a separate target by a clean oracle P. The chronological sequence

$$F, \quad T, \quad P, \quad T^\dagger, \quad F^\dagger$$

therefore implements a clean oracle O_J for $h_J(|u|)$, for either initial value of its target. The threshold inputs and flags are restored by the clean parity oracle, so the two inverse stages are legal. Each stage has constant depth; their private work and role count are polynomial. For basis input u, the generated roles all encode u; the identity for quantum inputs follows by linearity, not by assuming that the roles are independent quantum copies.

To use this oracle as a phase, its flag b must start at zero. For a separate compressed-input qubit x, apply O_J, then $\mathrm{CZ}_{b,x}$, then $O_J^\dagger$. On initialized scratch this gives

$$|u\rangle|x\rangle|0\rangle_b|0_{\mathrm{work}}\rangle \longmapsto (-1)^{x h_J(|u|)}|u\rangle|x\rangle|0\rangle_b|0_{\mathrm{work}}\rangle.$$

This identity holds for all occupation strings, including $|u| > J$, and extends to reference-entangled superpositions. It also holds at $J = 0$ with the trivial oracle. Different mode columns execute in parallel only when their occupation wires, compressed-input bits, and work are disjoint, as in the batch construction. Each compressed-input bit participates in one CZ in this stage. No fresh supplied input copy or serial chain of J gates on a common parity target is implicit.

Unary-state preparation and the full return test. For an integer $q \geq 1$, write

$$e_i = 0^{i-1}10^{q-i}, \qquad |W_q\rangle = q^{-1/2}\sum_{i=1}^{q}|e_i\rangle.$$

The basis vectors e_i are orthogonal, so this state is normalized. Grier, Morris, and Wu [11, Theorem 17] supply a constant-depth clean preparer with polynomial resources in q:

$$U_q|0^{q+\ell_q}\rangle = |W_q\rangle|0^{\ell_q}\rangle, \qquad \ell_q = \mathrm{poly}(q).$$

For polynomial q these are polynomial resources in the ambient n. The case $q = 1$ is a single X gate, and no state W_0 is used. Section 5 also gives a warm-start construction.

The preparation identity implies the operator identity

$$U_q|0^{q+\ell_q}\rangle\langle 0^{q+\ell_q}|U_q^\dagger = |W_q, 0^{\ell_q}\rangle\langle W_q, 0^{\ell_q}|.$$

Thus applying $U_q^\dagger$ and testing the *entire* preparation register for zero realizes precisely this rank-one return effect. Testing just the data wires would instead conjugate $|0^q\rangle\langle 0^q| \otimes I_{\mathrm{work}}$, a generally higher-rank projector, and is not justified by the preparation identity. The zero test may be recorded coherently in a flag; it is not an intermediate measurement or postselection. Preparing a single known state does not supply an arbitrary change of basis on the unary subspace.

Product reflections and work-only phases. For normalized one-qubit vectors $v_1, \ldots, v_r$ with $r \geq 1$, choose one-qubit unitaries B_i satisfying $B_i|1\rangle = |v_i\rangle$ and put $B = \bigotimes_i B_i$. Then

$$I - 2\bigotimes_{i=1}^{r}|v_i\rangle\langle v_i| = B(I - 2|1^r\rangle\langle 1^r|)B^\dagger.$$

For $r \geq 2$ the middle gate is a generalized controlled-Z, so this costs one entangling layer; for $r = 1$ it is a one-qubit operation and costs no entangling layer.

For a unit complex number z, the desired phase on a private register W is

$$S_{0,W}(z) = I + (z-1)|0_W\rangle\langle 0_W|.$$

Using a fresh scratch bit a initially zero, compute $\mathbf{1}[W = 0]$ into a by a negative-control generalized Toffoli, apply $\mathrm{diag}(1, z)$ to a, and reverse the indicator computation. The result is $S_{0,W}(z)$ with a returned to zero, for arbitrary states of W, including states entangled with a reference. Negative controls are implemented by local X conjugations. The cost is at most two entangling layers. An empty W gives a scalar phase and can be handled with a locally prepared scratch bit.

In every filtered source, W includes all source workspace, including preparation ancillas, flags, and scratch from earlier nested filters, but excludes all unknown input roles. The fresh scratch a used for the present zero phase is not among its own controls. This work-only phase and initialized clean oracles do not by themselves imply the all-sector source preservation required in Section 4; that condition is verified for the actual source construction.

4 Exact filters and an elementary polynomial-copy construction

The scalar identity in this section specializes established arbitrary-phase amplitude amplification [12, 13]. We prove the specialization directly, including its input-access conditions, and then combine it with the unary-state test of [11, Lemma 18] to obtain an elementary exact equality circuit. Throughout, phases are exact gates in the arbitrary-one-qubit-gate model, and circuit products act from right to left. No intermediate measurement or postselection is used.

4.1 A work-only exact zero filter

Let X include every physical supplied-input wire and let W include all private work, including a designated output flag o. Suppose the full source unitary A has the form

$$A = \sum_y |y\rangle\langle y|_X \otimes A_y.$$

Here each A_y is a unitary on W, and the index y ranges over *all* physical input strings, including inconsistent copies. This is an operator identity on arbitrary work states, not merely a statement about initialized action. It is a sufficient hypothesis for the construction; no necessity claim is made. Work starts at zero.

Let Π be the projector onto flag value $o = 1$, tensored with the identity on the other work qubits. In a fixed input sector put

$$|\psi\rangle = A_y|0_W\rangle, \qquad |\psi_{\mathrm{g}}\rangle = \Pi|\psi\rangle, \qquad |\psi_{\mathrm{b}}\rangle = (I-\Pi)|\psi\rangle, \qquad p = \||\psi_{\mathrm{g}}\rangle\|^2.$$

These unnormalized components are orthogonal and have squared norms p and $1-p$. This notation includes $p = 0$ and $p = 1$ without normalizing a zero vector. For a known parameter $a \in [0, 3/4]$, chosen independently of the actual input, set

$$z = \frac{1 - 2a + i\sqrt{3-4a}}{2(1-a)}.$$

The square root is real and nonnegative, and $(1-2a)^2 + (3-4a) = 4(1-a)^2$, so $|z| = 1$. In particular, $\Re z = (1-2a)/(2(1-a))$. Define

$$S_{\mathrm{good}}(z) = I + (z-1)\Pi, \qquad S_{0,W}(\bar{z}) = I + (\bar{z}-1)|0_W\rangle\langle 0_W|,$$

with both operators acting as the identity on X. The first is a one-qubit phase on o. Consider the initialized circuit

$$F_a(A) = AS_{0,W}(\bar{z})A^\dagger S_{\mathrm{good}}(z)A.$$

In the fixed input sector, conjugating the work-zero phase gives

$$A_y S_{0,W}(\bar{z})A_y^\dagger = I + (\bar{z}-1)|\psi\rangle\langle\psi|.$$

Moreover, $\langle\psi|S_{\text{good}}(z)|\psi\rangle = 1 - p + pz$. Hence the accepting component after $F_a(A)$ is

$$\begin{aligned}\Pi F_a(A) * y|0_W\rangle &= [z + (\bar{z} - 1)(1 - p + pz)]|\psi * \mathrm{g}\rangle \\ &= [2\Re z - 1 + 2p(1 - \Re z)]|\psi_{\mathrm{g}}\rangle \\ &= \frac{p - a}{1 - a}|\psi_{\mathrm{g}}\rangle.\end{aligned}$$

Its acceptance probability is therefore

$$T_a(p) = \frac{p(p - a)^2}{(1 - a)^2}. \tag{1}$$

Thus $T_a(a) = 0$, $T_a(0) = 0$, and $T_a(1) = 1$, including the endpoints of the allowed parameter interval. Unitarity also gives $0 \le T_a(p) \le 1$ for every $p \in [0, 1]$. The formula concerns each input sector's acceptance probability; it does not assert a nonlinear action on superpositions of different inputs.

The zero phase does not test an unknown input. As in Section 3, compute the negative-control AND of all W wires into a fresh zero scratch bit, phase the scratch, and uncompute. The scratch is not among its own controls and returns to zero on arbitrary work states, using at most two entangling layers. The displayed formula for $F_a(A)$ suppresses this clean implementation scratch. If A has entangling depth d, the filter has depth at most $3d + 2$. The three sequential calls use the *same* supplied copies; they do not require three private input bundles. Polynomial source width and gate count give polynomial filtered width and gate count.

Equivalently, in each preserved input sector the conjugated work-zero projector is

$$A_y|0_W\rangle\langle 0_W|A_y^\dagger,$$

the rank-one projector onto that sector's initialized source state. This is the starting-state phase used in the two-dimensional calculation above. If a filtered circuit is filtered again, its entire private workspace, including its earlier zero-phase scratch, belongs to the new source workspace; only the new phase scratch is freshly appended.

Every factor in the implemented filter preserves the input sectors on arbitrary work and phase-scratch states, so the filtered source satisfies the same block-diagonal hypothesis for subsequent use.

Perfectification and its bound. Let $p_* \in [1/4, 1]$ be a known target probability. Apply the filter to the rejecting subspace $I - \Pi$, with $a = 1 - p_* \in [0, 3/4]$. Phasing this subspace is also a one-qubit operation on o. The new rejection probability is $T_{1-p_*}(1 - p)$, so the resulting acceptance is

$$P_{p_*}(p) = 1 - \frac{(1 - p)(p_* - p)^2}{p_*^2}. \tag{2}$$

It sends p_* to certainty and fixes both zero and one. To bound its effect on other probabilities, let z' be the phase chosen above for $a = 1 - p_*$. Its real part is $1 - 1/(2p_*)$. The accepting component of this rejecting-subspace filter equals

$$[1 + (\bar{z}' - 1)(p + (1 - p)z')]|\psi_{\mathrm{g}}\rangle = (2 - p/p_* - z')|\psi_{\mathrm{g}}\rangle.$$

For $p \in [0, 1]$, we have $0 \le p/p_* \le 4$ and therefore

$$|2 - p/p_* - z'| \le |2 - p/p_*| + 1 \le 3.$$

Taking squared norms proves $0 \le P_{p_*}(p) \le \min\{1, 9p\}$.

Why initialized preservation is insufficient. On two input bits r, s and one work bit u, let C be the reversible map

$$(r, s, u) \longmapsto (r, r \oplus u, r \oplus s)$$

and put $A = (I_{rs} \otimes H_u)C$. Applying C twice gives the identity, so it is indeed a basis permutation and a unitary. On a valid repeated input with initialized work,

$$A|r, r, 0\rangle = |r, r\rangle|+\rangle,$$

so A preserves the supplied bits there and has acceptance probability $p = 1/2$ if $u = 1$ is accepting. However, for

$$Q_0 = I_{rs} \otimes |0\rangle\langle 0|_u, \qquad \Pi_{\text{code}} = |00\rangle\langle 00| * rs + |11\rangle\langle 11| * rs,$$

the image of $\operatorname{ran} Q_0$ under C is the whole space $\operatorname{ran} \Pi_{\text{code}} \otimes \mathbb{C}^2_u$. Consequently

$$AQ_0A^\dagger = \Pi_{\text{code}} \otimes I_u.$$

On each valid input sector this is the identity on work, not the rank-one projector $|+\rangle\langle +|$. For example, use filter parameter $a = 1/2$, so $z = i$. The physical circuit $F_{1/2}(A)$ sends

$$|0, 0, 0\rangle \longmapsto |0, 0\rangle \frac{-i|0\rangle + |1\rangle}{\sqrt{2}}.$$

Its acceptance remains $1/2$, whereas the inapplicable scalar formula would give $T_{1/2}(1/2) = 0$. This example disproves the inference from initialized preservation on the copy code to the required conjugated rank-one projector. The sources used below satisfy the stronger hypothesis directly: in the elementary construction, unknown input wires occur only as phase controls; in the later bounded-load construction, original input wires occur only as controls in the color preparation.

4.2 An elementary exact equality construction

The following proof gives exact polynomial-copy computation without the bounded-load compression or probability-extraction argument used later. Its one-copy source is a padded version of the unary return test in [11, Lemma 18]; the additional step is distance-dependent repetition followed by the exact zero filter.

Proposition 4.1. *For integers $n \geq 1$ and $0 \leq j \leq n$, EX_j has an exact nonuniform constant-depth circuit with polynomial gate count and total width, using fewer than $n + 8n^2$ supplied copies. The depth and polynomial-resource bounds are independent of j. Consequently every symmetric Boolean function on n bits has an exact $O(n^3)$-copy circuit in the same model.*

Proof. Let A_{2n} be a clean constant-depth preparer on a work register V of $2n + \ell$ qubits, with $\ell = \text{poly}(n)$, such that

$$A_{2n}|0_V\rangle = |W_{2n}\rangle|0^\ell\rangle, \qquad |W_{2n}\rangle = (2n)^{-1/2} \sum_{i=1}^{2n} |e_i\rangle.$$

Its existence follows from [11, Theorem 17]; an explicit warm-start proof is also given in Section 5. Fix a known string $\kappa \in \{0, 1\}^n$ of weight $n - j$. After preparation, apply CZ from each unknown bit x_i to unary qubit i, and the fixed local gate Z^{κ_i} to unary qubit $n + i$. The n CZ gates have disjoint supports and form one entangling layer; the known padding needs no supplied input copies. Let

$D_{x,\kappa}$ denote this diagonal phase operation on V in the sector x. Apply $A_{2n}^\dagger$. The amplitude of the *entire* preparation register being zero is

$$\langle 0_V|A_{2n}^\dagger D_{x,\kappa} A_{2n}|0_V\rangle = \frac{\sum_{i=1}^n (-1)^{x_i} + \sum_{i=1}^n (-1)^{\kappa_i}}{2n} = \frac{j-|x|}{n}.$$

Compute the OR of all V qubits into a separate zero flag; this is a coherent nonreturn test, including all preparer ancillas, not a measurement. A negative-control generalized Toffoli followed by a local X implements the OR in at most one entangling layer. The flag's one-copy acceptance probability is

$$p_j(w) = 1 - \frac{(w-j)^2}{n^2}, \qquad w = |x|.$$

Each unknown input wire is only a control of one phase gate; the preparer, its inverse, and the flag computation act only on private work. Thus the full source unitary is block diagonal on all input sectors even for arbitrary initial work, and independent trial bundles preserve all physical input strings, including inconsistent copies.

For every designated wrong weight $w \in \{0,\dots,n\}\setminus\{j\}$, choose, at circuit-design time, $d = |w-j|$ and

$$R_w = \left\lceil \frac{2n^2}{d^2} \right\rceil, \qquad a_w = \left(1 - \frac{d^2}{n^2}\right)^{R_w} \le e^{-2} < \frac{1}{4}.$$

The inequality follows from $(1-t)^R \le e^{-Rt}$ for $0 \le t \le 1$ and $R \ge 1$, including the case $t = 1$ separately. Also $e^2 > 1+2+2 = 5 > 4$, so $e^{-2} < 1/4$. Run R_w trials on distinct supplied copies with disjoint private work, and compute the AND of their nonreturn flags into a new zero output flag. For each fixed classical input of actual weight v, the trial states are a tensor product, so this source's acceptance probability is $p_j(v)^{R_w}$. It equals one when $v = j$ and a_w when $v = w$. No independence assertion about a superposition of classical inputs is needed.

This entire source, including the final AND, satisfies the block-diagonal premise. Apply the work-only filter (1) with the fixed parameter a_w. Its private register includes every trial's preparer ancillas and nonreturn flag, as well as the source's AND output. The true-weight probability stays one, and the designated wrong-weight probability becomes zero, including when $a_w = 0$.

Give different designated weights disjoint input bundles and work, run their filters in parallel, and AND the final flags into a fresh zero output. On a true input every flag is certainly one; on any wrong input its designated flag is certainly zero. A certain flag value means the state is supported entirely on that flag subspace, so arbitrary values and garbage in the other modules cannot change the final AND. The decision is exact, without a union bound or a cleanup requirement for the final garbage. There are n wrong weights, with each positive distance represented at most twice, so

$$\sum_{w\ne j} R_w \le n + 4n^2 \sum_{d=1}^{\infty} d^{-2} < n + 8n^2.$$

Here $\lceil t \rceil \le t+1$ gives the first inequality, and $\sum_{d=1}^{\infty} d^{-2} < 1 + \int_1^\infty t^{-2}\,dt = 2$ gives the strict second bound. If a single trial has depth d_0, each repeated source has depth at most d_0+1, its filtered version at most $3(d_0+1)+2$, and the final AND adds at most one layer. Thus depth is independent of n, j, w; all repeated operations inside a trial stage act on disjoint registers. The three source calls per module reuse the same bundle sequentially, so they do not multiply the displayed copy count. Since $R_w \le 2n^2$ and each trial has polynomial resources, total gate count and width are polynomial.

The additional phase parameters have exact finite descriptions. In fact

$$a_w = \frac{(n^2-d^2)^{R_w}}{n^{2R_w}}$$

has a numerator and denominator of $O(n^2 \log(n+2))$ bits, obtainable by exact integer arithmetic. The displayed formula for z then gives a rational real part and an imaginary part that is the square root of a nonnegative rational, with polynomial description length. There is no numerical limiting process; this assertion concerns the additional filter gates, relative to the imported exact preparer, and does not claim exact implementation over a fixed finite gate alphabet.

Finally, for a symmetric function with accepting weight set $S \subseteq \{0, \dots, n\}$, allocate private equality modules for $j \in S$ and OR their exact flags. There are at most $n+1$ modules, yielding $O(n^3)$ copies with constant depth and polynomial resources. The empty accepting set is the constant-zero function and requires no copies. □

For endpoints a direct wide AND or NOR uses one copy. Proposition 4.1 alone already suffices for the qualitative complexity-class conclusions of Section 8, using a larger polynomial padding. The following construction improves the quantitative copy budget.

5 Coherent bounded-load compression

We construct a constant-depth unitary that, from one supplied copy of x, produces a bounded-length thermometer representation on branches certified by a guard. The color records are retained, so this is not an irreversible map that discards information. The preparation uses the standard one-step amplitude-amplification identity [12, 11] and the exact small-threshold and limited-fanout primitives of Proposition 3.1. We give the initialized and full-unitary interfaces separately, since the latter are needed by the filters of Section 4.

5.1 A controlled vacuum-to-unary isometry

Let $B \geq 1$ be an integer, polynomially bounded in the ambient input length. On a B-qubit data register D, let e_b be the string with its only one in position b, and set

$$|W_B\rangle = B^{-1/2} \sum_{b=1}^{B} |e_b\rangle, \qquad |v\rangle = |0^B\rangle.$$

These are orthonormal vectors. Therefore

$$|\nu_B\rangle = \frac{|v\rangle - |W_B\rangle}{\sqrt{2}}$$

is normalized, and its reflection $R_{\nu_B} = I - 2|\nu_B\rangle\langle\nu_B|$ satisfies

$$R_{\nu_B}|v\rangle = |W_B\rangle, \qquad R_{\nu_B}|W_B\rangle = |v\rangle.$$

For example, the first identity follows from $\langle\nu_B||v\rangle = 1/\sqrt{2}$; the second uses $\langle\nu_B||W_B\rangle = -1/\sqrt{2}$. We first prepare the normal vector cleanly and then implement the controlled reflection by conjugation.

Exact preparation of the normal vector. The normalized product state

$$|\zeta_B\rangle = (1 + 1/B)^{-B/2} \bigotimes_{b=1}^{B} (|0\rangle - B^{-1/2}|1\rangle)$$

has weight-at-most-one component

$$\Pi_{\leq 1}|\zeta_B\rangle = (1 + 1/B)^{-B/2}(|v\rangle - |W_B\rangle) = \sqrt{q_B}|\nu_B\rangle,$$

where $\Pi_{\le 1}$ projects onto computational strings of weight at most one and

$$q_B = 2(1 + 1/B)^{-B} \ge 2/e.$$

The estimate follows from $B \log(1 + 1/B) \le 1$. Also $(1 + 1/B)^B \ge 2$ by the binomial theorem, so $q_B \le 1$. In particular $q_B > 1/4$, and the following one-qubit coin state is well defined:

$$|\chi_B\rangle = \sqrt{1 - \frac{1}{4q_B}}|0\rangle + \frac{1}{2\sqrt{q_B}}|1\rangle.$$

Let F include the threshold flag and all its private oracle work, initialized to zero. The full starting state

$$|s\rangle = |\zeta_B\rangle_D |\chi_B\rangle_c |0_F\rangle$$

is a known product state. For the good predicate "data weight at most one and coin one," put $\Pi_g = \Pi_{\le 1} \otimes |1\rangle\langle 1|_c \otimes I_F$. Its component in $|s\rangle$ is exactly

$$\Pi_g |s\rangle = \tfrac{1}{2}|g\rangle, \qquad |g\rangle = |\nu_B\rangle * D|1\rangle_c |0_F\rangle.$$

Thus, for a normalized vector $|b_\perp\rangle$ orthogonal to the good subspace,

$$|s\rangle = \tfrac{1}{2}|g\rangle + \tfrac{\sqrt{3}}{2}|b* \perp\rangle.$$

One amplification step gives

$$(2|s\rangle\langle s| - I)(I - 2\Pi_g)|s\rangle = |g\rangle.$$

Indeed, the good reflection changes the sign of the $|g\rangle$ component, and its resulting inner product with $|s\rangle$ is $-1/4 + 3/4 = 1/2$. The second reflection therefore subtracts this sign-flipped vector from $|s\rangle$, leaving $|g\rangle$ exactly.

To implement the good reflection, compute the clean flag $t = [|D| \ge 2]$, phase by -1 when $t = 0$ and $c = 1$, and reverse the flag computation. Proposition 3.1 supplies the threshold oracle, including when $B = 1$ and the predicate is constant zero. On the subspace with $F = 0$, this circuit agrees with $I - 2\Pi_g$ and restores F; no assertion about its action on dirty oracle work is needed for this preparation.

For the second reflection, choose a product unitary P with $P|0_{DcF}\rangle = |s\rangle$. Then

$$2|s\rangle\langle s| - I = -P(I - 2|0_{DcF}\rangle\langle 0_{DcF}|)P^\dagger.$$

The middle zero reflection includes *all* preparation work and costs at most one entangling layer; the minus sign can be supplied exactly by a local $-I$ gate. Finally, applying X to the now-known coin state $|1\rangle$ yields a unitary preparer A_B satisfying

$$A_B|0_D\rangle|0_{\text{work}}\rangle = |\nu_B\rangle * D|0 * \text{work}\rangle, \qquad \text{work} = cF.$$

There are constantly many primitive calls and polynomially many work qubits. This proves constant depth and polynomial size and width without an approximation or postselection. The case $B = 1$ also has the direct preparation $ZH|0\rangle = |\nu_1\rangle$.

Controlled reflection and the color record. The clean preparation identity implies the full operator identity

$$A_B(I - 2|0_{\text{all}}\rangle\langle 0_{\text{all}}|)A_B^\dagger = I - 2|\nu_B, 0_{\text{work}}\rangle\langle \nu_B, 0_{\text{work}}|.$$

Here "all" means D together with every qubit of preparation work. The preparer need not preserve that work on arbitrary initial states for this conjugation identity to hold. For a separate supplied-input qubit x_i, apply $A_B^\dagger$, then phase the joint condition $x_i = 1$ and $D\,mwork = 0$, then apply A_B. The resulting full unitary is

$$I - 2|1\rangle\langle 1|_{x_i} \otimes |\nu_B, 0_{\text{work}}\rangle\langle \nu_B, 0_{\text{work}}|,$$

which preserves x_i on *all* data and work states. Its initialized action is

$$|x_i\rangle|0^B\rangle|0_{\text{work}}\rangle \longmapsto \begin{cases} |0\rangle|0^B\rangle|0_{\text{work}}\rangle, & x_i = 0, \\ |1\rangle|W_B\rangle|0_{\text{work}}\rangle, & x_i = 1. \end{cases} \tag{3}$$

The central conditional phase is one generalized controlled-Z with local basis changes for its negative controls; choose a work or data qubit as its target, leaving x_i as a control. Only this central entangling gate touches x_i. All n instances have disjoint input bits, private data rails, and private work, and run in parallel. They require one supplied copy of x, not one copy per bucket.

More explicitly, for a fixed classical x with support $S = \{i : x_i = 1\}$ and weight w, the resulting state is

$$|x\rangle\, B^{-w/2} \sum_{f:S\to\{1,\dots,B\}} |C_f\rangle|0_{\text{prep}}\rangle,$$

where row i of the $n \times B$ binary color record C_f is $e_{f(i)}$ for $i \in S$ and is 0^B otherwise. Distinct maps f give orthogonal records. Hence the squared amplitudes have the distribution of w independent uniform bucket choices. This is a description of a coherent fixed circuit on each basis input, not external randomization or a claim of independent quantum copies. For $w = 0$ the sum has one term, indexed by the empty map.

An auxiliary clean unary preparer. For an integer $N \geq 1$, use the *positive-sign* product state

$$(1 + 1/N)^{-N/2} \bigotimes_{i=1}^{N} (|0\rangle + N^{-1/2}|1\rangle).$$

Its weight-exactly-one component is $(1 + 1/N)^{-N/2}|W_N\rangle$, with probability $r_N = (1 + 1/N)^{-N} \geq 1/e > 1/4$. Use a coin of amplitude $1/(2\sqrt{r_N})$ on $|1\rangle$ and the exact equality predicate $[|D| = 1]$ from Proposition 3.1. The same calculation and workspace cleanup give $|W_N\rangle$ with positive overall phase, in constant depth and polynomial resources. For $N = 1$ a single X suffices. This is a direct warm-start variant of the clean unary preparation in [11, Theorem 17], not an independent new amplitude-amplification principle.

5.2 Exact weight preservation on guarded branches

Fix integers $1 \leq k \leq n$ and $K \geq 1$ with $K = O(\log(n + 2))$, the implicit constant being fixed, and set

$$B = \lceil 4k/K \rceil, \qquad L = BK.$$

The case $k = 0$ is an endpoint predicate treated separately, not a choice $B = 0$. The compression circuit below is defined on *all* inputs; the restriction $|x| \le k$ is used only for its probability estimate. On a color record C, define the load $\ell_b = \sum_{i=1}^{n} C_{i,b}$ and compute the thermometer bits and guard

$$z_{b,u} = \mathbf{1}[\ell_b \ge u] \quad (1 \le u \le K), \qquad V = \mathbf{1}[\max_b \ell_b \le K].$$

For each color bit $C_{i,b}$, use exact limited fanout to generate $K + 1$ fresh zero-initialized role targets. For each pair (b, u), $1 \le u \le K+1$, the corresponding role column is a private n-bit input to a clean threshold oracle. All these modules run in parallel on disjoint inputs and work. The outputs for $u \le K$ are the bits $z_{b,u}$; the output for $u = K + 1$ is the overflow flag $o_b = [\ell_b \ge K + 1]$. Thresholds larger than n are simply constant zero.

The oracles restore their role inputs and private work even when entangled with the original records. Reverse the role fanouts after computing the flags; this resets their targets and work while leaving C, z, and the overflow flags intact, by the basis-state identities and linearity. A negative-control generalized Toffoli computes $V = \bigwedge_b \neg o_b$ into a fresh zero bit in at most one entangling layer. The records C are retained and are not mixed by the subsequent forward batch computation; inverse calls to the whole source may of course uncompute them. Keeping these records makes the many-to-one load encoding compatible with unitarity.

The threshold cutoff $K + 1$ and fanout length $K + 1$ are logarithmic, so Proposition 3.1 applies. There are $nB(K + 1)$ role targets and $B(K + 1)$ threshold modules. Since $B \le 4n$ and $K = O(\log(n + 2))$, all resources are polynomial in n, and the constant number of parallel stages has constant depth. None of these stages touches the original supplied-input wires. Together with the controlled color construction, this verifies the full input-sector preservation required for later filters, not merely preservation on initialized valid copies.

For every color record in the support of the initialized circuit, each active input row contains exactly one one and each inactive row is zero, so $\sum_b \ell_b = |x|$. On every such branch the thermometer string has $\min(\ell_b, K)$ ones in bucket b. Thus

$$V = 1 \quad \Longrightarrow \quad |z| = \sum_b \min(\ell_b, K) = \sum_b \ell_b = |x|. \tag{4}$$

Here $|z|$ counts ones in the L thermometer bits, not the length of a binary encoding of the loads. The implication is exact for inputs of every weight. Invalid branches may have truncated weights, but their guard is zero and must be included in the final decision. If $|x| > L$, then $\sum_b \ell_b > BK$, so at least one bucket overflows and every branch has $V = 0$. These assertions concern the supported color records, not arbitrary inconsistent assignments to internal registers.

For a fixed x of weight $w \le k$, the squared-amplitude distribution just derived gives $\ell_b \sim \mathrm{Bin}(w, 1/B)$, with mean $w/B \le k/B \le K/4$. A direct factorial-moment estimate suffices for its tail: if $w \ge K$,

$$\Pr[\ell_b > K] \le \Pr[\ell_b \ge K] \le \mathbb{E}\binom{\ell_b}{K} = \binom{w}{K} B^{-K} \le \left(\frac{ew}{BK}\right)^K \le (e/4)^K.$$

The expectation identity counts the K-subsets of active rows assigned to bucket b; the binomial estimate uses $K! \ge (K/e)^K$. If $w < K$, the overflow probability is zero, so the same bound holds. A union bound over buckets yields

$$\Pr[V = 0] \le B(e/4)^K. \tag{5}$$

These are Born probabilities of the coherent guard flag. No independence between different bucket loads is assumed. Finally, $B \le 4k/K + 1$ gives

$$4k \le L = BK \le 4k + K,$$

and hence $L \le 5k$ when $K \le k$. When $k < K$, the later equality compiler uses the exact small-threshold fallback instead of claiming this last bound. The compression is therefore a guarded weight-preserving construction; it does not claim that every branch is valid or that $L < n$ for every parameter choice.

6 A weight-conditioned batch and copy accounting

This section combines guarded compression with truncated parallel repetition, whose unary-state and occupation-counting precedent is [11, Appendix C.1, Theorem 27 and Claims 28–29]. We distinguish an ideal parallel phase operation from its explicit constant-depth approximation using one compressed string. The ideal acceptance profile, together with a uniform initialized-state error bound, yields the bounded-error constructions here. Exact statements about the actual truncated circuit are proved separately in Section 7.

For the resource claims, $n \ge 1$ is the ambient input length. Positive accuracy parameters η and ε below lie in $(0, 1/4]$ and, when they vary with n, are at least $(n+2)^{-C}$ for a fixed $C > 0$, as stipulated in Section 2. Depth and polynomial-resource constants may depend on C. An accuracy parameter is fixed at circuit-design time, not supplied as an additional input.

6.1 One batch needs one compressed string

Let $L \ge 1$ be an integer polynomially bounded in n, and let $j \in \{0, \dots, L\}$. For a compressed string $z \in \{0,1\}^L$, choose a fixed padding string $\kappa \in \{0,1\}^L$ of weight $L - j$. These are $L - j$ negative phase columns and j positive phase columns; their phases are local gates and cost no supplied input copies. A row has $N = 2L$ unary-mode qubits, whose one-excitation subspace has dimension N. Its phased uniform unary state is

$$|\phi_{z,\kappa}\rangle = N^{-1/2} \left(\sum_{i=1}^{L} (-1)^{z_i} |e_i\rangle + \sum_{i=1}^{L} (-1)^{\kappa_i} |e_{L+i}\rangle \right).$$

It has return amplitude

$$\langle W_N || \phi_{z,\kappa} \rangle = \frac{N - 2(|z| + L - j)}{N} = \frac{j - |z|}{L}.$$

Prepare $T = N$ unary rows independently on private registers using clean preparers. For fixed z, the ideal phase operation leaves these rows in the product state $|\phi_{z,\kappa}\rangle^{\otimes T}$. Thus the ideal probability that *none* returns is

$$p_j(|z|) = \left(1 - \frac{(|z| - j)^2}{L^2} \right)^{2L}. \tag{6}$$

It equals one at $|z| = j$. More generally, for every integer $v \in [0, L]$,

$$0 \le p_j(v) \le \exp(-2(v-j)^2/L),$$

using $(1-t)^{2L} \le e^{-2Lt}$ for $t \in [0,1]$. In particular it is at most $e^{-2/L}$ when $v \ne j$. This product formula is an *ideal* profile, not an assertion that the physical truncated-phase rows are independent.

Legal phase operation and its error. Let $u_{r,i}$ denote the occupation bit of row r in mode i and let $c_i = \sum_{r=1}^{T} u_{r,i}$. Choose an integer cutoff $J \geq 0$ with $J = O(\log(n+2))$, and use the total function $h_J(c) = \min(c, J) \bmod 2$ from Section 3. In each unknown mode column, compute $h_J(c_i)$ into a fresh zero flag, apply CZ once to z_i, and uncompute. Columns have disjoint supports and work, so they run in parallel. Each known negative column uses local Z gates on all of its occupation bits. The resulting induced diagonal unitary D_J on z and the occupation registers has phase

$$(-1)^{\sum_{i=1}^{L} z_i h_J(c_i) + \sum_{i=1}^{L} \kappa_i c_{L+i}}.$$

The ideal diagonal unitary D_{id} replaces each $h_J(c_i)$ by c_i in this exponent. Directly applying its T phase gates on each shared z_i would not give the required disjoint-support depth. The truncated implementation uses only one copy of z, with all column-oracle scratch returned to zero.

On the initialized product of unary rows, every c_i has distribution $\mathrm{Bin}(T, 1/N)$ and mean one. Define Π_{bad} on the occupation registers by the condition that some $c_i > J$, including the known columns for convenience. For integer $J \geq e^2$,

$$\delta = \Pr[\text{some column load} > J] \leq N e^{-J}.$$

Indeed, if $J \leq T$, then

$$\Pr[c_i > J] \leq \mathbb{E}\binom{c_i}{J} = \binom{T}{J} N^{-J} \leq (e/J)^J \leq e^{-J}.$$

If $J > T$, the tail is zero. The last inequality uses $\log J \geq 2$, and the union bound uses no independence between columns. Both induced phase operators are diagonal, agree off the bad subspace, and have norm one. Consequently

$$D_J - D_{\mathrm{id}} = (D_J - D_{\mathrm{id}})\Pi_{\mathrm{bad}}, \qquad \|D_J - D_{\mathrm{id}}\| \leq 2.$$

The row preparation is independent of the compressed register and any retained reference. For any normalized joint state $|\xi\rangle$ of those latter registers, writing $|\Omega\rangle$ for the initialized product of unary rows therefore gives

$$\|(D_J - D_{\mathrm{id}})(|\xi\rangle \otimes |\Omega\rangle)\| \leq 2\|\Pi_{\mathrm{bad}}|\Omega\rangle\| = 2\sqrt{\delta}.$$

Preparation ancillas and column-oracle scratch are zero at this interface and have been suppressed. The estimate applies uniformly to reference-entangled compressed inputs, but is not a small global operator-norm bound for arbitrary row or scratch states.

Invert every clean unary preparation and compute a nonreturn flag by OR of its *entire* preparation register. Before inversion the return effect is $|W_N, 0\rangle\langle W_N, 0|$, including preparer work. The phase stage acts only on the data modes and leaves preparation ancillas zero, while returning its own scratch to zero. Thus on this initialized subspace the return effect reduces to $|W_N\rangle\langle W_N|$ on the data. Distinct row tests have disjoint supports. AND their nonreturn flags into a fresh zero output. These are coherent computations, not measurements or postselection. The ideal acceptance is (6); applying the same subsequent unitaries in the ideal and physical experiments preserves their norm difference.

6.2 Independent random compressions are sufficient

Here "random" describes the Born distribution of retained coherent color records, not an externally randomized choice of circuit. Fix integers $1 \leq k \leq n$, $0 \leq j \leq k$, and $R \geq 1$, with compression parameters from Section 5.

Take R supplied copies of x and perform the complete color, compression, and batch circuit independently on each. The final output is the AND of all guards and all no-return flags. The compressed strings need not be identical: by (4), every valid one has the same weight w.

For a fixed classical input x of weight w, retained color records are orthogonal and the forward batch leaves them unchanged. The final accepting projector does not mix them, so cross terms between different records vanish. All valid records have weight w after compression and hence the same ideal scalar (6). The record distribution depends only on w, not on the positions of the ones; write $v_w = \Pr[V = 1 \mid |x| = w]$. At fixed x, the complete private batch states are tensor products across supplied copies. For $w \le L$, ideal final acceptance is therefore

$$\big(v_w p_j(w)\big)^R.$$

For $w > L$ it is zero, since every guard is zero; the polynomial expression $p_j(w)$ is not used outside its domain $[0, L]$. The physical truncated circuit also rejects such inputs exactly, because its phase and row-test stages do not alter the guard. This equal-weight property is essential: nonidentical compressed strings cannot in general replace repeated identical strings in an arbitrary multi-copy algorithm.

Parameters and the equality error bound. Target $j = 0$ is a direct endpoint gate; in particular $k = 0$ needs no compression. For $1 \le j \le k \le n$ and accuracy $0 < \eta \le 1/4$ in the range specified above, define the following parameters in the stated order:

$$R_* = \left\lceil \frac{5k}{2} \log \frac{4}{\eta} \right\rceil, \qquad K = \left\lceil 8 \log \frac{20 n R_*}{\eta} \right\rceil.$$

We have $R_* = O(n \log(n+2))$ and therefore $K = O(\log(n+2))$. If $k < K$, then $j \le k < K$ is a logarithmic threshold, and exact equality uses one supplied copy by Proposition 3.1. Otherwise use $B = \lceil 4k/K \rceil$, $L = BK \le 5k$, and

$$R = \left\lceil \frac{L}{2} \log \frac{4}{\eta} \right\rceil \le R_*.$$

These parameters are noncircular: R_* is chosen before K, B, L, R. The specified K is at least five. In the branch $k \ge K$, it follows that $B \le 4k/K + 1 \le k \le n$. Put $S = 20nR_*/\eta > 1$. Since $8(\log 4 - 1) > 1$, equation (5) gives

$$1 - v_j \le B(e/4)^K \le n S^{-8(\log 4 - 1)} \le n/S = \eta/(20R_*) \le \eta/(4R_*).$$

At $w = j$, the ideal row acceptance is one, so the completeness failure is $1 - v_j^R \le R(1 - v_j) \le \eta/4$. On any wrong weight $w \le L$, ideal acceptance is at most $e^{-2R/L} \le \eta/4$; for $w > L$ it is zero. No concentration estimate is needed on wrong high-weight inputs.

Set

$$\delta_0 = (\eta/(8R))^2, \qquad J = \left\lceil \max\{e^2, \log(N/\delta_0)\} \right\rceil.$$

Then $Ne^{-J} \le \delta_0$, and $J = O(\log(n+2))$. Replace the physical phase stage by the ideal one in one private batch at a time. In each hybrid comparison that batch's rows are freshly prepared and its other registers may be treated as references, so the uniform bound above applies. The triangle inequality bounds final vector error by $2R\sqrt{\delta_0} = \eta/4$.

For completeness, if $|\psi\rangle, |\phi\rangle$ are normalized and $0 \le M \le I$ is any measurement effect, then

$$\begin{aligned} |\langle\psi|M|\psi\rangle - \langle\phi|M|\phi\rangle| &= |\Re\langle\psi - \phi|(M - I/2)|\psi + \phi\rangle| \\ &\le \tfrac{1}{2}\|\psi - \phi\| \, \|\psi + \phi\| \le \|\psi - \phi\|. \end{aligned}$$

The equality uses normalization to subtract $I/2$, and the inequality uses $\|M - I/2\| \le 1/2$. References are included in the pure states, or by purification. Adding the ideal error to this probability discrepancy gives worst-case error at most $\eta/2$.

There are constantly many stages: controlled color preparation; role fanout; parallel thresholds and role cleanup; unary-row preparation; column parities; one phase per compressed bit; phase-flag uncomputation; row return tests; guarded AND. Within each stage the modules have disjoint supports. Since $T = N = 2L = O(n)$, there are polynomially many occupation wires, threshold modules, and ancillas across the R copies. The cutoff oracles have constant depth by Section 3. This proves constant depth, polynomial size and width, and a supplied-copy bound of $\max\{1, R_*\}$. The later exact compiler uses the same type of source with its separately specified repetition parameters.

6.3 Symmetric functions

Fix a default bit β and its exceptional set E_β from Section 2, and let $t = |E_\beta|$. If $t = 0$, the function is constant β and needs no supplied copies. Assume $1 \le t \le n+1$ and $0 < \varepsilon \le 1/4$ in the stipulated accuracy range. The copy count avoids a multiplicative $\log t$ by using the full distance dependence in (6); $\log t$ still enters internal precision cutoffs.

Put $q = \log(16/\varepsilon)$. For each $j \in E_\beta$, choose x when $j \le n/2$ and its bitwise complement otherwise, and set $k_j = \min(j, n-j)$. Writing $k = k_j$, the transformed actual weight v satisfies $|v - k| = |w - j|$. The complement is a fixed local operation on this module's private input bundle, outside any later filtered source. If $k = 0$, use the exact endpoint gate with one copy. Otherwise set

$$R_* = \lceil 5kq/2 \rceil, \quad K = \left\lceil 8 \log \frac{64 n R_* t}{\varepsilon} \right\rceil .$$

Here $K = O(\log(n+2))$. If $k < K$, use the exact one-copy small-threshold fallback. The remaining parameters apply only to the branch $k \ge K$. Use $B = \lceil 4k/K \rceil$, $L = BK \le 5k$, and $R = \lceil Lq/2 \rceil \le R_*$. As above, $K \ge 5$ gives $B \le n$. Setting $S = 64 n R_* t/\varepsilon$, and using $K \ge 8 \log S$ and $8 \log(4/e) > 1$, gives the per-batch overflow bound

$$B(e/4)^K \le \frac{\varepsilon}{64 R_* t}.$$

The true-weight ideal failure over R batches is therefore at most $\varepsilon/(64t)$. On every wrong transformed weight $v \le L$, all valid branches have the ideal profile, and complete-module ideal acceptance is at most

$$\exp(-2R(v-k)^2/L) \le e^{-q(w-j)^2}.$$

For $v > L$ the guarded acceptance is zero, so the same upper bound holds without evaluating the profile outside its domain. Exact endpoint and fallback modules also satisfy these true- and wrong-weight error bounds.

Independently choose the coherent phase precision

$$\delta_0 = (\varepsilon/(32Rt))^2, \qquad J = \left\lceil \max\{e^2, \log(2L/\delta_0)\} \right\rceil .$$

The R-batch hybrid has vector error at most $2R\sqrt{\delta_0} = \varepsilon/(16t)$, hence outcome-probability error at most that value and, conservatively, at most $\varepsilon/(8t)$. The finer precision changes the cutoff J and internal resources, not the supplied count R. Thus each module's true-weight failure is at most $\varepsilon/(64t) + \varepsilon/(8t) < \varepsilon/(4t)$, and its wrong-weight acceptance is at most $e^{-q(w-j)^2} + \varepsilon/(8t)$.

Allocate private supplied-copy bundles, OR all equality flags into a fresh output, and XOR β. If $w \in E_\beta$, rejection by the test for $j = w$ is necessary for a wrong final answer, so the error is less

than $\varepsilon/(4t)$. If $w \notin E_\beta$, a union bound and the fact that there are at most two integer targets at each positive distance give

$$\Pr[\text{error}] \le 2\sum_{d\ge 1} e^{-qd^2} + \varepsilon/8 \le \frac{2e^{-q}}{1-e^{-q}} + \varepsilon/8 < \varepsilon/2.$$

For the last inequality, $d^2 \ge d$ and $e^{-q} = \varepsilon/16$ give $2e^{-q}/(1-e^{-q}) + \varepsilon/8 = \varepsilon(2/(16-\varepsilon)+1/8) < \varepsilon/2$ throughout the specified range. No independence assumption is needed for this union bound. Thus the worst-case error is below $\varepsilon/2$. The explicit supplied-copy count is

$$m \le \sum_{j\in E_\beta} \max\{1, \lceil 5k_j q/2\rceil\} \le t + \tfrac{5}{2}\log(16/\varepsilon)\sum_{j\in E_\beta} k_j.$$

Indeed, an endpoint or fallback module costs one, and a large-weight module costs $R \le \lceil 5k_j q/2\rceil$. The inequality $\max\{1,\lceil u\rceil\} \le 1+u$ for $u \ge 0$ proves the second bound. Since $t \le n+1$ and the accuracy is inverse polynomial, all cutoffs remain $O(\log(n+2))$ and the aggregate size and width remain polynomial. The next section removes the error with a different repetition allocation, rather than asserting that truncation itself is exact.

Transition radius and pooling. For a nonconstant function of radius $r = \rho(f)$, choose β to be its value on the central interval $[r, n-r]$ when nonempty, and choose either bit otherwise. As proved in Section 2, its exceptional weights lie in the two endpoint intervals of length r. Thus

$$t \le 2r, \qquad \sum_{j\in E_\beta} k_j \le 2\sum_{u=0}^{r-1} u = r(r-1),$$

including the maximal odd-length radius, where the central interval is empty. The preceding bound gives $m \le 2r + (5/2)\log(16/\varepsilon)r(r-1)$. Constants, with $r = 0$, need no copies.

Finally fix $a > 0$ and put $\lambda_n = \log(n+2)$ and $s = \lceil \lambda_n^{a+1}\rceil$. If the preceding construction requests $M > 0$ full-input roles, apply exact limited fanout to every supplied bit, using disjoint work, to prepare s zero-initialized targets encoding that bit. Allocate the resulting full-input role vectors to the complete construction. Then $\lceil M/s\rceil$ supplied copies suffice. This pooling is performed before any source; generated unknown roles must not be included in its work-zero tests if the source is later filtered. The pool need not be uncomputed for a decision circuit with garbage.

Because $\log(16/\varepsilon) = O(\lambda_n)$, the coarse bound is $M = O(\lambda_n D_\beta(f))$. Pooling and optimizing β therefore give

$$m = O\left(\max\left\{1, \frac{D(f)}{\log^a(n+2)}\right\}\right)$$

for nonconstant functions with the present bounded-error guarantee. The extra logarithmic power in s absorbs the accuracy logarithm; it does not assume free fanout. Depth and polynomial-resource exponents may depend on the fixed a and accuracy exponent C.

7 Distance-adapted exact compilation

We now apply the two scalar transformations from Section 4 to the actual guarded batch. The essential points are exact symmetry despite truncation, finite support supplied by the guard, and exact evaluation of the probabilities used to choose the phases. All parameters are fixed at circuit-design time. We first construct EX_k for an integer $1 \le k \le \lfloor n/2\rfloor$, with $n \ge 2$; endpoints and input

complementation are treated separately. Each repeated source preserves every physical supplied-input basis sector on arbitrary private work, as established in Sections 5–6. Its only operations touching the supplied inputs use them as controls in the color preparation. Thus it satisfies the full block-diagonal hypothesis of Section 4, not merely initialized preservation on the copy code. Each filter's zero-state phase test includes the entire source workspace, including retained color records and earlier filter scratch, but excludes the supplied-input register.

7.1 Exact symmetry and finite guard support

Fix the *same total* truncated-parity function in each unknown mode column, for example

$$h_J(c) = \min\{c, J\} \pmod 2.$$

It is the parity of the exact threshold flags $[c \geq 1], \ldots, [c \geq J]$, so its functional oracle is clean for every occupation string, even above the cutoff.

Fix a bucket count B, capacity K, length $L = BK \geq k$, and cutoff J as specified below. The padding has exactly $L - k$ negative and k positive phase columns. For a retained valid coloring, let z be its length-L thermometer string. Simultaneously permuting the L unknown coordinates and the corresponding columns preserves the product of uniform row states, the actual truncated phase transformation, and all row return effects. The L known padding columns remain fixed. The row preparer itself need not be permutation equivariant on arbitrary work: its initialized state and its return projector are invariant, which suffices. Thus the *actual* row-AND probability depends only on $|z|$. This symmetry concerns the initialized row experiment; it does not require a permutation to preserve the thermometer layout of a bucket.

For every integer $0 \leq v \leq L$, let q_v denote the actual probability that *none of the rows returns*, on any length-L string of weight v. This is a joint probability, not a product of individual row probabilities. For $0 \leq v \leq n$, let ν_v be the valid-color probability when the original weight is v. Retained color labels are orthogonal and unchanged by the row tests, and the accepting effect does not mix them, so cross terms between distinct labels vanish. Their distribution consists of v independent uniform bucket labels, regardless of the locations of the active input bits. Thus the actual one-batch guarded probability is exactly $b_v = \nu_v q_v$ for $v \leq \min(n, L)$, even if $\nu_v = 0$; no conditional state of probability zero is needed. Since each valid coloring has $|z| = v$, and no coloring with $v > L = BK$ is valid,

$$b_v = 0 \quad (L < v \leq n). \tag{7}$$

For $R \geq 1$ private batches, the initialized state at each fixed basis input is a tensor product across batches. Their final AND therefore has actual probability exactly b_v^R. Rows within one truncated batch need not be independent. Approximate proximity to an ideal symmetric test would not by itself prove these exact assertions; identical total column action and the retained orthogonal color records do.

7.2 Distance-dependent repetitions

Take a fixed $\eta = 1/1024$ and target $k \leq n/2$. Target zero is an exact endpoint gate and is treated before the parameters below. To make parameter extraction purely integer arithmetic, put $H = 42/5 > \log 4096$ and use

$$R_* = 21k, \qquad K = 8\lceil \log_2(20nR_*/\eta) \rceil.$$

If $k < K$, the exact small-threshold implementation needs one supplied copy: since $k \leq n/2$,

$$K \leq 8\lceil \log_2(215040n^2) \rceil = O(\log(n+2)).$$

Compute equality as $[|x| \geq k] \wedge \neg[|x| \geq k+1]$, with two internally generated roles and clean threshold oracles. Otherwise set $B = \lceil 4k/K \rceil$ and $L = BK \leq 4k + K \leq 5k$. Choose the common phase cutoff using

$$\delta_0 = \left(\frac{\eta}{8R_*}\right)^2, \qquad J = \max\{8, \lceil \log_2(2L/\delta_0) \rceil\}.$$

Both logarithm arguments are positive integers, since $1/\eta = 1024$ and $1/\delta_0 = (8192R_*)^2$. For a positive integer A, $\lceil \log_2 A \rceil$ is the binary bit length of $A - 1$, with bit length zero for zero; this convention includes powers of two. In the large-weight branch $K \geq 8$ and $k \geq K$, so $B \leq 4k/K + 1 \leq k \leq n$. Writing $S = 20nR_*/\eta$, using $K \geq 8 \log S$, and using $8 \log(4/e) > 1$ gives

$$B(e/4)^K \leq nS^{-8\log(4/e)} \leq n/S = \eta/(20R_*) \leq \eta/(4R_*).$$

Also $J \geq e^2$ and $2Le^{-J} \leq \delta_0$, since the chosen binary-log cutoff is at least the required natural logarithm. For each wrong weight $w \in \{0, \ldots, \min(n, L)\} \setminus \{k\}$, build a private source A_w using

$$R_w = \left\lceil \frac{21L}{5(w-k)^2} \right\rceil \leq R_* \tag{8}$$

independent guarded batches. Its supplied-copy cost before external pooling is R_w. Indeed, $L \leq 5k$ and $|w - k| \geq 1$ imply that the unrounded expression in (8) is at most the integer $21k$, so taking its ceiling preserves $R_w \leq R_*$. The batches run in parallel on private bundles; increasing R_w increases size, not depth.

In the notation above the two actual probabilities used to choose phases are

$$p_*^{(w)} = (\nu_k q_k)^{R_w}, \qquad p_w^{(w)} = (\nu_w q_w)^{R_w}.$$

On the true weight, every valid ideal row test accepts with certainty, and color failure over the batches is at most $R_w\eta/(4R_*) \leq \eta/4$. The phase-truncation initialized-vector error is at most $2R_w\sqrt{\delta_0} \leq \eta/4$, by the reference-uniform estimate and hybrid argument of Section 6. The event-probability difference is at most this vector distance, as proved there. Hence the actual true acceptance $p_*^{(w)}$ is at least $1 - \eta/2$.

For the designated wrong weight w, each valid ideal batch accepts with probability

$$\left(1 - \frac{(w-k)^2}{L^2}\right)^{2L} \leq \exp\left(-\frac{2(w-k)^2}{L}\right).$$

The ideal repeated acceptance is at most $e^{-H} < \eta/4$, and the same truncation bound gives actual acceptance $p_w^{(w)} \leq \eta/2$. Guarding can only reduce wrong-weight acceptance, so this argument needs no overflow concentration estimate at w. No upper bound on this source's other wrong-weight probabilities is needed.

First apply (2) with $p_*^{(w)}$, then (1) with

$$a_w = P_{p_*^{(w)}}(p_w^{(w)}) \leq 9\eta/2 < 1/4.$$

The first parameter satisfies $p_*^{(w)} \geq 1 - \eta/2 > 1/4$, and the second satisfies $0 \leq a_w < 1/4 < 3/4$, so both phase formulas are defined, including $p_*^{(w)} = 1$ and $a_w = 0$. Their composition accepts weight k with certainty, rejects its designated weight w with certainty, and preserves probability zero at every weight in (7). This last conclusion follows from $P_{p_*}(0) = T_a(0) = 0$; it does not require the guard bit to stay zero during inverse source calls. AND all private filters' flags in a final generalized

Toffoli. At the true weight all flags are deterministically one. A wrong weight at most $\min(n, L)$ has its designated zero flag, and a weight above L has every flag zero. There is at least one filter since $1 \le k \le n/2$ and $L \ge k$. The decision is exact; no error union bound across these final filters is involved.

There are $Q = \min(n, L) \le L$ designated weights. Each positive distance occurs at most twice, so

$$\sum_w R_w \le L + \frac{LH}{2} \sum_{w \ne k} \frac{1}{(w-k)^2} < L + 2LH = \tfrac{89}{5} L \le 89k < 90k.$$

Here $\sum_{d \ge 1} d^{-2} < 2$, for example by bounding the terms with $d \ge 2$ strictly by $1/(d(d-1))$ and telescoping. The rounding cost is included in the L term. The first scalar transformation has depth at most $3d_A + 2$, the second at most $3(3d_A + 2) + 2$, and the final AND adds one, giving $9d_A + 9$ for a common constant source-depth bound d_A. Its nine sequential source or inverse calls reuse the same input bundles. All different private modules have disjoint physical input bundles and work, and their schedules can be padded by idle layers. There are polynomially many mode columns, row registers, threshold calls, and supplied bundles; $K, J = O(\log(n+2))$, and each primitive has polynomial size in the fixed ambient scale. Thus their total gate and qubit count is polynomial.

Complementing the input before the source, when needed, yields the bound with $k = \min(j, n-j)$. These complements are outside the filtered source, whose input roles remain preserved. Every interior weight obeys the numerical bound $90k$: the large-weight branch uses fewer copies, and the small-weight branch costs $1 \le 90k$. Endpoint equality needs one copy. OR private exact indicators for an exceptional set E_β and adjust the default bit. This proves the exact upper bound in Theorem 2.2.

7.3 Exact polynomial-time extraction of the filter phases

The actual probabilities needed above have polynomial-time exact evaluations in n, not merely numerical approximations. Fix the design parameters L, k, J and an integer $0 \le w \le L$, let $N = T = 2L$, and define the formal three-variable exponential generating function

$$G(X, Y, Z) = \sum_{c,u,u' \ge 0} \frac{(-1)^{h_J(c+u)+h_J(c+u')}}{c!u!u'!} X^c Y^u Z^{u'}.$$

Let $a_r(w)$ be the probability that a fixed set of r rows all return, ignoring the others, on a compressed string of weight w. We derive its coefficient formula from the physical row state.

Write a mode assignment as $\alpha \in [N]^T$, with $|\alpha\rangle = \bigotimes_{t=1}^{T} |e_{\alpha_t}\rangle$ on the physical unary row registers. The initialized product of rows is $N^{-T/2} \sum_\alpha |\alpha\rangle$. For each mode i, let $n_i(\alpha)$ be its occupation count and write its phase as $\sigma_i(n_i(\alpha))$, where $\sigma_i(c) = (-1)^{h_J(c)}$ on the w active unknown modes, $\sigma_i(c) = 1$ on inactive or positive-padding modes, and $\sigma_i(c) = (-1)^c$ on negative-padding modes. Hence the phased amplitude of α is

$$N^{-T/2} \prod_{i=1}^{N} \sigma_i(n_i(\alpha)).$$

For the effect that returns the first r rows and ignores the rest, the matrix element between assignments α, β is

$$\langle \beta | \big((|W_N\rangle\langle W_N|)^{\otimes r} \otimes I \big) |\alpha\rangle = N^{-r} \prod_{t=r+1}^{T} \mathbf{1}[\alpha_t = \beta_t].$$

Thus the ignored rows have common bra/ket mode labels. Let c_i count these $T-r$ labels in mode i, and let u_i, u_i' count the returned ket and bra labels. They obey

$$\sum_i c_i = T - r, \qquad \sum_i u_i = \sum_i u_i' = r.$$

There are $(T-r)!/\prod_i c_i!$ common ignored assignments, and $r!/\prod_i u_i!$ and $r!/\prod_i u_i'!$ returned assignments. It follows exactly that

$$a_r(w) = \frac{(T-r)!(r!)^2}{N^{T+r}} \sum_{\substack{\sum c_i = T-r \\ \sum u_i = \sum u_i' = r}} \prod_{i=1}^{N} \frac{\sigma_i(c_i+u_i)\sigma_i(c_i+u_i')}{c_i! u_i! u_i'!}.$$

An active unknown mode contributes G. Inactive unknown and positive-padding modes contribute e^{X+Y+Z}, whereas each negative-padding mode contributes e^{X-Y-Z}, since its common ignored count occurs twice. There are $L-w+k$ positive modes and $L-k$ negative modes. Taking their product therefore gives

$$\begin{aligned} a_r(w) &= \frac{(T-r)!(r!)^2}{N^{T+r}} [X^{T-r}Y^r Z^r]\, G(X,Y,Z)^w e^{(N-w)X+(2k-w)(Y+Z)}, \\ q_w &= \sum_{r=0}^{T} (-1)^r \binom{T}{r} a_r(w). \end{aligned}$$

Here q_w is the probability that none of the rows returns. The return effects on distinct rows commute, so expanding their complementary product gives inclusion–exclusion. Permuting rows preserves the phased state, making every fixed r-row term equal to $a_r(w)$; this gives the binomial coefficient in the second line. The derivation keeps the actual truncated phases throughout, not ideal independent-row phases.

As a normalization check, take $L = k = w = 1$ and $J = 1$ in this row experiment, independently of the compiler's large-weight schedule. Writing $|i,j\rangle = |e_i\rangle \otimes |e_j\rangle$, the phased state is

$$\tfrac{1}{2}(-|1,1\rangle - |1,2\rangle - |2,1\rangle + |2,2\rangle).$$

The formula gives $a_0 = 1$, $a_1 = 1/2$, and $a_2 = 1/4$, hence $q_1 = 1 - 2a_1 + a_2 = 1/4$. This differs from the ideal value one at the target weight and illustrates why the actual probability must be used. If $J \geq T$, no physical occupation is truncated, and the formula reduces instead to $q_w = (1-(w-k)^2/L^2)^T$.

Only exponents at most T in each variable are needed. Store the factorial-scaled coefficients $F[a,b,c] = a!b!c![X^a Y^b Z^c]F$, whose product rule is

$$(F \star G)[a,b,c] = \sum_{i \leq a,\, j \leq b,\, \ell \leq c} \binom{a}{i}\binom{b}{j}\binom{c}{\ell} F[i,j,\ell] G[a-i, b-j, c-\ell].$$

The base factorial-scaled coefficients of G are signs. For each w, the factorial-scaled coefficients of the remaining exponential are

$$(N-w)^a (2k-w)^{b+c};$$

their signs are retained when $2k - w < 0$. Cache the powers $G^0, \ldots, G^L$ and multiply each by its corresponding exponential table. There are $(T+1)^3$ entries per table, and each multiplication uses

$O(T^6)$ naive integer operations; all needed powers, products, and extractions therefore use $O(LT^6)$ operations. Truncation of these coefficient tables does not alter the physical total function h_J.

The factorial-scaled coefficients of G^w have absolute value at most w^{a+b+c}, by assigning labeled tokens of the three kinds to the w columns. Since $|2k-w| \le N-w$ for $0 \le k, w \le L$, multiplication by the remaining exponential bounds absolute coefficient sums by N^{a+b+c}. Thus intermediate integers, including the factorial and binomial factors in the final sums, have $O(T \log N)$ bits. Alternating sums are performed exactly, not by floating-point subtraction.

To make normalization and rational arithmetic explicit, let $C_w[a,b,c]$ denote the factorial-scaled coefficient table of $G^w e^{(N-w)X+(2k-w)(Y+Z)}$. The factorials in the formula for a_r cancel its coefficient scaling, giving

$$a_r(w) = \frac{C_w[T-r,r,r]}{N^{T+r}}, \qquad q_w = \frac{1}{N^{2T}} \sum_{r=0}^{T} (-1)^r \binom{T}{r} N^{T-r} C_w[T-r,r,r].$$

Thus each a_r has denominator dividing N^{T+r}, and q_w has denominator dividing N^{2T}. In particular $a_0(w) = 1$, including at cutoffs that correlate the rows. Negative intermediate coefficients are allowed; the completed values are probabilities by their derivation from orthogonal projections. The exact valid-color probability is

$$\nu_w = \frac{w!}{B^w}[z^w] \left(\sum_{c=0}^{K} \frac{z^c}{c!} \right)^B .$$

A one-variable version of the same integer calculation evaluates it: $w!$ labels the active positions, the coefficient restricts each of the B bucket loads to at most K, and B^{-w} is the probability of each ordered coloring. Its denominator divides B^w; it equals one for $w = 0$ and zero for $w > BK$. The expression can also be evaluated for formal weights larger than n, but such weights are not input cases of the compiler. Hence $b_w = \nu_w q_w$, its powers b_w^R, and $a_w = P_{p_*^{(w)}}(p_w^{(w)})$ are polynomial-bit rationals obtained in polynomial time. In particular, $R \le 21k = O(n)$ increases their bit lengths only polynomially. There are $O(n)$ weights and, for a symmetric function, at most $n+1$ targets, so computing all additional parameters remains polynomial-time. Each chosen phase parameter has the form $u + i\sqrt{v}$ with rational u, v, $v \ge 0$, of polynomial bit length. It is a root of $(z-u)^2 + v$, with the nonnegative-imaginary-part choice specifying the root; the conjugate phases use the other sign. Thus each added gate entry has algebraic degree at most two; no degree bound on the field generated by all gate entries together is asserted.

This establishes efficient extraction of the additional parameters, given the symmetric truth table and integer circuit parameters, relative to the imported exact primitive families. The arbitrary-one-qubit-gate model is the convention of [11, Section 2]; the exact coefficient algorithm above is the justification for the additional filter parameters, not an inference from that model convention. The main family statements are nonuniform: an arbitrary sequence of symmetric truth tables need not itself be uniformly computable. No stronger logspace-uniformity or exact implementation over a finite gate alphabet is established; numerical approximations to the exact phases give only approximate computation.

7.4 External pooling and structure-sensitive sums

For a default value β, run exact equality modules on separate bundles for each $j \in E_\beta$, OR their flags, and complement the result if $\beta = 1$. Constants have an empty exceptional set for the appropriate default and require no copies. Otherwise the coarse cost is at most $90 D_\beta(f)$, with endpoint modules charged one copy each. Minimize over β. Since $\sum_{j=0}^{n} \min(j, n-j) = \lfloor n^2/4 \rfloor$, this gives $O(n^2)$

copies for arbitrary symmetric functions. For radius $r \geq 1$, the interval of weights $[r, n-r]$, when nonempty, has one common value: an internal transition would have radius at least $r+1$. Choose that value as β, or either value when the interval is empty. Then

$$E_\beta \subseteq \{0, \ldots, r-1\} \cup \{n-r+1, \ldots, n\}, \qquad D_\beta(f) \leq 2\sum_{j=0}^{r-1}(1+j) = r(r+1).$$

This includes the empty central interval at odd maximum radius and proves $O(r+r^2)$; constants have radius zero and were handled separately.

For any fixed $a > 0$, put $s = \lceil (\log(n+2))^a \rceil$ and use exact limited fanout from Proposition 3.1, imported from [11, Corollary 10], to prepare s fresh zero-initialized roles from each supplied bit, independently across physical bits. Retain the original controls outside the private modules. Allocate the resulting full-input role vectors to the entire coarse copy pool before any filtered source starts. If the pool has $S > 0$ full-input roles, $\lceil S/s \rceil = O(\max\{1, S/(\log(n+2))^a\})$ supplied copies suffice; unused roles may be ignored. All generated unknown roles are part of the preserved input register, not zero workspace. Original pool controls, unused roles, and fanout auxiliaries also stay outside every later private work-zero test. On initialized valid inputs the pool is exactly a collection of classical basis copies; on superpositions the same identity extends linearly, without cloning arbitrary unknown states. The subsequent source circuits are block diagonal on their whole role spaces, so no assertion about the pooling circuit on arbitrary dirty work is required. The pool need not be uncomputed for a one-output decision with garbage. There are polynomially many fanout modules of polynomial width, and extra depth depends only on the fixed a.

8 Exact threshold-circuit simulation and a total language

Two qualitative consequences clarify the computational power supplied by polynomially many copies. The quantum evaluators have deterministic final outputs on every valid classical input. Formula expansion and consistency lifting follow the methods of [11, Theorem 21 and Lemma 22]; exact symmetric-gate modules supply zero error here. The constructions below make the private-register interfaces, total-language convention, and resource accounting explicit. These consequences do not require the sharper linear-copy equality bound: the elementary polynomial-copy construction in Proposition 4.1 also suffices.

Here EQAC^0 denotes the class of total languages exactly decided by the stated nonuniform, polynomial-size QAC families with arbitrary one-qubit gates and initially zero ancillas. The class TC^0 uses ordinary Boolean gates and unweighted majority or threshold gates with classical wire reuse; fan-in counts wire occurrences, including repeated occurrences of one value. All families in this section are nonuniform. In $\mathsf{EQAC}^0 \circ \mathsf{NC}^0$, the preprocessing is a classical polynomial-output, bounded-fanin, constant-depth map: a language A belongs to this composition if there are such maps g_n and one total language $L \in \mathsf{EQAC}^0$ with $x \in A$ if and only if $g_n(x) \in L$ for every $x \in \{0,1\}^n$. The preprocessing outputs may reuse an input bit freely. This notation does not grant such reuse to the subsequent quantum circuit. We describe positive input lengths; any value at the empty string is handled separately by a constant circuit.

8.1 Cleaning and private formula expansion

Let X_x denote the ordered supplied-copy encoding of x. If a unitary U exactly decides a function on these states with zero-initialized private work, its output qubit is supported on one basis value,

so we may write

$$U|X_x, 0\rangle = |f(x)\rangle_o |g_x\rangle.$$

The normalized garbage vector may contain changed input wires and input-dependent phases; the output o may itself be an original input wire. With a separate target $b \in \{0, 1\}$ untouched by U, applying U, then CNOT from o to b, then $U^\dagger$ gives

$$|X_x, 0, b\rangle \longmapsto |X_x, 0, b \oplus f(x)\rangle.$$

Indeed, on a valid basis input the CNOT changes only b, leaving the state $U|X_x, 0\rangle$ available for exact reversal. The entire garbage vector, including its phase, is reversed. The identity extends by linearity to arbitrary target states and reference-entangled superpositions of valid copied inputs. It needs the same copies, and changes entangling depth D to at most $2D+1$. No initialized correctness on inconsistent copies, or block diagonality on arbitrary work, is required for this cleaning step.

For every interior weight $1 \le j \le a-1$, the large-weight compiler or its one-copy small-threshold branch uses at most $90 \min(j, a-j)$ copies: in the latter branch $1 \le 90 \min(j, a-j)$. Each endpoint equality costs one. Therefore an arbitrary symmetric gate on $a \ge 1$ inputs has the coarse exact copy bound

$$M(a) = 90\lfloor a^2/4 \rfloor + 2,$$

because $\sum_{j=0}^{a} \min(j, a-j) = \lfloor a^2/4 \rfloor$; constants require no copies. One may simply compute all needed equality flags on private bundles and OR them. The identity follows by summing the two arithmetic progressions on either side of the middle weight, and includes $a = 1$, where $M(1) = 2$ is a deliberately loose bound.

Now take a polynomial-size classical decision circuit of symmetric gates, maximum fan-in $F \ge 1$, and fixed depth $d \ge 1$. Constant-only and depth-zero circuits are handled directly; constant child values are prepared as known basis qubits. Let D be a common constant upper bound on the depth of the unclean symmetric macros for all arities at most F in a fixed polynomial ambient scale. Theorem 2.2 and the ambient convention in Proposition 3.1 give such a bound, with polynomial resources in the original input length. Pad a macro's requested copy count to $M(a)$ if necessary by leaving surplus bundles untouched. For each gate occurrence and each of these $M(a)$ bundles, instantiate a *private* computation of each of its a child inputs. Repeated child occurrences, even of the same gate or input label, are instantiated separately. Repeat down to the original input leaves.

Since $M(a)$ is nondecreasing, the resulting tree has branching at most $B = FM(F) \ge 2$ and height at most d. It has at most B^d input leaves and at most $\sum_{h=0}^{d} B^h$ total occurrences. In particular, if $\ell(g)$ counts input leaves below a private occurrence, then $\ell(g) = M(a) \sum_{i=1}^{a} \ell(g_i)$, with $\ell = 1$ for an original input and $\ell = 0$ for a constant. This recurrence includes all physical copy requests.

Assign each input-leaf occurrence its own complete supplied copy of the original input and use only its desired coordinate; unused coordinates are not private zero workspace. Process tree heights bottom-up. Inductively, at a fixed original basis input, each clean child macro returns its exact Boolean output in a separate fresh target and restores its own inputs and work. Its output is retained, not uncomputed as part of a larger subtree. Each parent therefore receives $M(a)$ identical classical fan-in vectors on distinct physical wires, even when some children have the same value. Macros at one height have disjoint supports, and lower-height outputs may wait idle until their parent is run. This is not a cloning assumption for arbitrary unknown qubit states; coherent correctness follows from the basis identities by linearity. Consequently

$$m \le [FM(F)]^d, \qquad D_{\text{total}} \le d(2D+1).$$

Since $M(F) = O(F^2)$, the number of occurrences is polynomial for fixed d and polynomial F. Every occurrence has polynomial width and gate count in the fixed ambient scale, so the total resources, including the complete supplied copies, are polynomial. No iteration of an input-size-dependent number of cleaning steps is hidden in the depth bound: one parallel stage of cleaned gate macros is run at each tree height.

Choose a fixed integer $C \geq 1$ large enough that $m(n) = n^C$ dominates the evaluator's copy requirement for every $n \geq 2$, and ignore surplus supplied blocks. Such a C exists because the requirement is bounded by one fixed polynomial. At $n = 1$, any Boolean function has a direct one-copy exact implementation, so $m(1) = 1$ suffices as well. The lengths $N(n) = nm(n) = n^{C+1}$ are strictly increasing. Let L at length $N(n)$ be the total consistency lifting, defined below, of the original decision function f_n, and reject at all other positive lengths. Then

$$g_n(x) = x^{\otimes n^C} \quad \text{(classical concatenation)}, \qquad \mathbf{1}[g_n(x) \in L] = f_n(x).$$

Each output coordinate of g_n is one input coordinate, so g_n is a polynomial-output NC^0 map. The circuit for L at an encoded length has polynomial resources in n, hence in $N(n)$; nonencoded lengths use a constant-zero circuit. The length-dependent choice and parsing are fixed in the nonuniform circuit family, not performed by an uncharged quantum subroutine. Combining this evaluator with the total consistency check gives

$$\mathsf{TC}^0 \subseteq \mathsf{EQAC}^0 \circ \mathsf{NC}^0.$$

Thus the quantum class here is a total-language class, not an unstated copy promise. This proves the simulation assertion of Corollary 2.3; it does not assert $\mathsf{TC}^0 \subseteq \mathsf{EQAC}^0$ on one-copy inputs. Unweighted threshold gates give the stated TC^0 presentation; arbitrary binary-encoded gate weights are not treated as free multiplicities.

8.2 Four-layer consistency and exact total lifting

Fix integers $n, m \geq 1$. For arbitrary physical basis blocks $Y = (y^{(1)}, \ldots, y^{(m)})$, with each $y^{(r)} \in \{0,1\}^n$, compute for $1 \leq i \leq n$ in at most four sequential entangling layers

$$a_i = \bigwedge_r y_i^{(r)}, \quad o_i = \bigvee_r y_i^{(r)}, \quad d_i = o_i \wedge \neg a_i, \quad c = \mathrm{NOR}_i\, d_i.$$

In each of the first three layers, gates for different i have disjoint supports; the last layer is one gate. All computed flags have fresh zero targets. AND, OR and negative-control variants each require at most one generalized Toffoli and local X gates; any temporary changes to control wires are restored within that stage. The first two layers are sequential because they touch the same input wires. The bit d_i is one exactly when coordinate i contains both a zero and a one across the blocks. Thus $c = 1$ exactly when all blocks agree, including $m = 1$, when every $d_i = 0$. The check preserves Y and uses $3n + 1$ fresh flags and $3n + 1$ entangling gates at most.

Run this check *before* a depth-D_E copied-input evaluator. They share input wires but operate at separate times; the evaluator has its own zero-initialized work and does not touch any consistency flag. Finally AND c with the evaluator's flag into a fresh zero output, using one additional entangling layer; only c from the consistency work is used at this last stage, as a control. This exactly decides the total function

$$\mathrm{Lift}_m(f)(Y) = \begin{cases} f(y^{(1)}), & y^{(1)} = \cdots = y^{(m)}, \\ 0, & \text{otherwise.} \end{cases}$$

On a consistent basis input, the check has left the evaluator's data unchanged, $c = 1$, and the evaluator's flag equals $f(y^{(1)})$ with certainty. On an inconsistent basis input, $c = 0$ remains a

separate zero bit; regardless of any superposition or changed input wires produced by the evaluator, the final AND leaves its fresh output zero. The overall entangling depth is at most $D_E + 5$, with $3n + 2$ additional flags and at most $3n + 2$ additional entangling gates. Total lifting therefore requires only correct initialized action on the valid copy code, not full-sector input preservation by an arbitrary evaluator.

The order matters. For example, with $n = 1, m = 2$ and $f(x) = x$, a reversible evaluator may have initialized action

$$|r, s\rangle|0, 0, 0\rangle \longmapsto |0, 0\rangle|r, s, r\rangle.$$

This is implemented by copying r, s to private work, clearing the original wires using those records, and copying the saved r to the output. It correctly evaluates f whenever $r = s$. Checking only the changed input wires afterwards would wrongly accept the inconsistent input $(r, s) = (1, 0)$. The preceding construction avoids this by retaining the original consistency flag.

Corollary 8.1. *There is a single total language $L \in \mathsf{EQAC}^0$ such that $L \notin \mathsf{AC}^0[p]$ for every prime p.*

Proof. For each integer $r \geq 1$, at input length $N = 400r^2$, parse $200r$ consecutive blocks of length $2r$ and accept exactly when they are all equal and their common weight is r. At all other lengths, including zero, reject. Distinct positive r give distinct lengths, so this defines one total language, independently of any prime. The equality compiler needs at most $90r \leq 200r$ copies; use those blocks for evaluation and all blocks for the preceding consistency check. The other blocks can otherwise be ignored. The resulting family has constant entangling depth and polynomial gate count and width in the physical length N.

Fix a prime p and suppose this language had $\mathsf{AC}^0[p]$ circuits of depth at most d_0 and size at most AN^c, with constants independent of N. Substituting one $2r$-bit string into all blocks, using ordinary classical wire reuse, gives EX_r on $2r$ bits with no increase in depth or gate count. For every integer $0 \leq t \leq r$ and every $z \in \{0, 1\}^r$,

$$\mathsf{EX}_r(z, 1^{r-t}, 0^t) = \mathsf{EX}_t(z),$$

because the padded weight is $|z| + r - t$, which equals r precisely when $|z| = t$. Choose a fixed prime $q \neq p$, for example $q = 3$ if $p = 2$ and $q = 2$ otherwise. OR private copies of these classical restricted circuits over $t \in \{0, \ldots, r\}$ divisible by q. The result computes

$$\mathsf{MOD} * q(z) = \mathbf{1}[|z| \equiv 0 \pmod q] = \bigvee *^{\substack{0 \leq t \leq r \\ q \mid t}} \mathsf{EX}_t(z)$$

at every positive input length r, with depth at most $d_0 + 1$ and size at most $(r + 1)A(400r^2)^c + 1$, still polynomial in r. This contradicts Smolensky's modular lower bound: for prime p, a modulus not a power of p cannot be computed by polynomial-size constant-depth circuits over AND, OR, NOT and MOD_p [23, Theorem 2, printed page 81]. The distinct prime q satisfies that hypothesis. The source's zero-output-on-divisible-inputs convention is converted to ours by an output NOT; a different fixed accepted residue is obtained by padding with a constant number of ones. These changes preserve constant depth and polynomial size. Since p was arbitrary and the language was fixed before choosing it, the same L separates from every prime-modulus class. □

The corollary is a worst-case separation, not containment in the opposite direction or an average-case hardness statement, and it does not separate L from ACC^0. At length $N = 400r^2$, exactly $\binom{2r}{r}$ strings are accepted, one repeated string for each weight-r block. Their density is at most 2^{2r-N}, so the constant-zero predictor is highly accurate on uniformly random physical inputs.

For completeness, the elementary exact W-filter construction also yields a separating language with a different polynomial padding, without Section 7. At logical length $2r$, Proposition 4.1 uses

fewer than $2r + 32r^2 \le 34r^2$ copies. Taking $40r^2$ blocks gives physical length $80r^3$; the same consistency check and classical restrictions prove the analogous total-language separation. Its arbitrary symmetric-gate bound is polynomial as well, so private formula expansion also gives the qualitative TC^0 simulation without the sharper copy accounting.

9 A copy-sensitive two-layer lower bound

We prove Theorem 2.4. Fix $n \ge 2$ and a finite circuit with at most two entangling layers that decides parity with worst-case error at most $1/3$. Missing layers may be padded by empty layers. Ancillas start in an input-independent pure product state; no size or width bound is imposed. Since parity is nonconstant, $m \ge 1$. A logical label is an original input coordinate, not a physical wire: restricting one label always fixes all its m supplied occurrences. All expectations below are over explicitly specified classical assignments, not over independent physical copies.

9.1 Output normal form

A generalized Toffoli is the rank-one product reflection $I - 2|1, \dots, 1, -\rangle\langle 1, \dots, 1, -|$. Conjugation by a product of one-qubit unitaries preserves this form and its physical support; this is the product-reflection normal form of [14, Proposition 2.1]. Moving intervening one-qubit layers through adjacent reflections therefore leaves two disjoint-support reflection layers, a product initial preparation, and a terminal one-qubit effect. No gate support or first-layer gate count is increased.

Before the second entangling layer, the state at each fixed classical input factors across first-layer blocks, counting idle wires as singleton blocks. Let B_0 contain the output, and fix every logical label represented in B_0. Relabel the k free labels by $x \in \{0,1\}^k$. Write the state as $\rho_0 \otimes \sigma_x$, where ρ_0 is fixed and $\sigma_x = \bigotimes_{g \ne 0} \rho_g(x)$. For purposes of the output probability, remove every second-layer gate except the one touching the output. Indeed, these other gates commute with that gate by disjointness and with the terminal one-qubit effect, so they cancel in its expectation. They may act on non-output wires of B_0: removing them need not preserve the entire B_0 marginal. The partial trace calculated next is that of the circuit after this removal, which has the same output probability.

If an output-touching second-layer gate exists, write its reflection as $D = I - 2P_0 \otimes P_E$, where $P_E = \bigotimes_{g \ne 0} P_g$. For each block, P_g is the tensor product of its one-qubit triggering projections on the intersection with this gate and identities on the other wires. In particular, $P_g = I$ on an empty intersection; P_0 need not have rank one on the whole block. Put $R_0 = I - 2P_0$. The exact identity

$$D = I \otimes (I - P_E) + R_0 \otimes P_E$$

and $\mathrm{Tr}(P_E \sigma_x (I - P_E)) = 0$, by cyclicity and $(I - P_E)P_E = 0$, give

$$\mathrm{Tr}_{\ne 0}[D(\rho_0 \otimes \sigma_x)D] = (1 - F(x))\rho_0 + F(x) R_0 \rho_0 R_0, \qquad F(x) = \mathrm{Tr}(P_E \sigma_x).$$

Here the trace removes all blocks other than B_0. Let E be the terminal one-qubit effect tensored with the identity on the other wires of B_0, so $0 \le E \le I$. Put $a = \mathrm{Tr}(E\rho_0)$ and $a + b = \mathrm{Tr}(E R_0 \rho_0 R_0)$. This proves

$$p(x) = a + bF(x), \qquad F(x) = \prod_{g \ne 0} c_g(x), \quad 0 \le a, a + b \le 1, \quad 0 \le c_g \le 1. \tag{9}$$

Here $c_g(x) = \mathrm{Tr}(P_g \rho_g(x))$. Subsequent source products and sums range only over non-output blocks retained in the factorization. If the output has no second-layer gate, its probability is constant

after the restriction; take $b = 0$ and an empty source product, equal to one. Input-independent factors may be absorbed into b: multiplication of b by a scalar in $[0, 1]$ moves $a + b$ along the segment between the old endpoints and preserves the displayed bounds, even for a zero factor. This argument uses quantum factorization at each fixed x, not probabilistic independence between repeated occurrences of a logical variable.

Lemma 9.1 (One-source parity bias)**.** *Let $k \geq 1$, and let $\mathbb{E}_{\mathrm{even}}$ and $\mathbb{E}_{\mathrm{odd}}$ be the uniform averages on the two parity classes of the k free logical labels. Every non-output source factor satisfies*

$$|\mathbb{E}_{\mathrm{even}} c_g - \mathbb{E}_{\mathrm{odd}} c_g| \leq 6/2^k.$$

If a free label is absent from that block, the difference is zero. A block containing every free label will be called full-label*.*

Proof. In the full-label case, the initialized pre-reflection copy encoding is the isometry

$$V|z\rangle = \bigotimes_{i=1}^{k} \bigotimes_{w \in W_i} |\alpha_w(z_i)\rangle \otimes |\alpha_{\mathrm{anc}}\rangle,$$

where W_i is the nonempty set of physical occurrences of label i in this block, $\langle \alpha_w(0), \alpha_w(1) \rangle = 0$, and all displayed states are normalized. The last state is a fixed product state containing fixed logical inputs as well as ancillas. Thus $V^\dagger V = I$. Let $P = P_g$ be the product test inherited from the second-layer gate. Grouping its tensor factors by logical label gives

$$A = V^\dagger P V = \kappa \bigotimes_{i=1}^{k} A_i, \qquad 0 \leq \kappa \leq 1, \quad 0 \leq A_i \leq I_2.$$

Each A_i is the compression of the physical product test to the two-dimensional encoding of label i; κ is its expectation on the fixed product factors. With $Z = \mathrm{diag}(1, -1)$ and $Z_k = Z^{\otimes k}$, each $|\operatorname{Tr}(Z A_i)| = |(A_i)_{00} - (A_i)_{11}| \leq 1$, so $|\operatorname{Tr}(Z_k A)| \leq 1$. Let $R = I - 2|v\rangle\langle v|$, with $\|v\| = 1$, be the first-layer reflection on this block. Explicitly,

$$RPR - P = -2|v\rangle\langle v|P - 2P|v\rangle\langle v| + 4\langle v|P|v\rangle\, |v\rangle\langle v|.$$

This difference has range in $\mathrm{span}\{v, Pv\}$ and operator norm at most one, since $0 \leq RPR, P \leq I$ implies $-I \leq RPR - P \leq I$. Isometric compression preserves the rank bound and does not increase operator norm. Writing $\|\cdot\|_1$ for the trace norm, its bound by rank times operator norm and trace-norm duality with $\|Z_k\| = 1$ give

$$\|V^\dagger (RPR - P) V\|_1 \leq 2, \qquad |\operatorname{Tr}(Z_k V^\dagger RPRV)| \leq 3.$$

Since $c_g(z) = \langle z|V^\dagger RPRV|z\rangle$, its even-minus-odd average is exactly $2^{1-k} \operatorname{Tr}(Z_k V^\dagger RPRV)$: each class has 2^{k-1} strings. This proves the bound. An idle product source has $R = I$ and satisfies the stronger bound $2/2^k$ by the product-test calculation alone; this remains true if several idle wires are grouped into one product source. A missing label makes c_g independent of that bit, and pairing assignments that differ in it gives exact cancellation. □

9.2 From small source bias to many full-label blocks

Assume first that $k \geq 1$ labels remain and that the restricted circuit has worst-case error $0 \leq \epsilon < 1/2$ for parity or its complement. Then $b = 0$ is impossible, because a constant acceptance probability cannot satisfy the two disjoint required ranges. Complement the output event if necessary so $b > 0$; the physical endpoints remain in $[0, 1]$. Let H, L be respectively the desired-one and desired-zero parity classes, each with the uniform distribution, so $p|_H \geq 1 - \epsilon$ and $p|_L \leq \epsilon$. Put $u = \epsilon - a$ and $d = 1 - 2\epsilon$. Since $a \leq p \leq a + b$, the two nonempty classes and the physical endpoints imply

$$0 \leq u \leq \epsilon, \qquad u + d \leq b \leq 1 - \epsilon + u, \qquad F|_H \geq h = \frac{u+d}{b}, \quad F|_L \leq \ell = \frac{u}{b}.$$

In particular $0 \leq \ell < h \leq 1$. For $d_g = 1 - c_g$ and $\mathcal{D} = \sum_g d_g$, the inequalities

$$1 - F \leq \mathcal{D}, \qquad \mathcal{D} \leq -\log F \quad (F > 0)$$

follow respectively by induction from $1 - c_1 c_2 \leq (1 - c_1) + (1 - c_2)$ and by summing $1 - c_g \leq -\log c_g$. The source set is finite; $F > 0$ ensures that every factor is positive. No logarithm is taken at low-class points where $F = 0$. Since $h > 0$, these inequalities imply $\mathcal{D}|_H \leq -\log h$ and $\mathcal{D}|_L \geq 1 - \ell$, hence

$$\sum_g (\mathbb{E}_L d_g - \mathbb{E}_H d_g) \geq 1 - \ell + \log h.$$

The derivative of $1 - u/b + \log((u+d)/b)$ with respect to b is $(u - b)/b^2 < 0$. Its minimum therefore has $b = 1 - \epsilon + u$. At this endpoint its derivative in u has a positive denominator and numerator

$$(1 - \epsilon)(3\epsilon - 1) - (1 - 2\epsilon)u.$$

The denominator is $(1 - \epsilon + u)^2(u + 1 - 2\epsilon) > 0$. For $0 \leq \epsilon \leq 1/3$ the numerator is nonpositive, so the minimum is at $u = \epsilon$, hence

$$1 - \ell + \log h \geq 1 - \epsilon + \log(1 - \epsilon).$$

Since $d_g = 1 - c_g$, Lemma 9.1 bounds each source's absolute difference by $6/2^k$; only full-label source blocks contribute. Their number is therefore at least

$$\frac{1 - \epsilon + \log(1 - \epsilon)}{6}\, 2^k.$$

Choose any free label. Each full-label block consumes a distinct supplied occurrence of that label, and there are only m such occurrences. Thus the full-label count is at most m, independently of the ancillary width. If r distinct labels occur in the output's first block, then $k = n - r$. Restricting them to any fixed values leaves parity or its complement, and at error $1/3$ this proves

$$m \geq a_0 2^{n-r}. \tag{10}$$

For $r = n$, the same inequality follows trivially from $m \geq 1$ and $a_0 < 1$.

9.3 Deleting a large output block

Return to the original circuit; do not keep the r-label restriction from the preceding argument. Delete only the output's first-layer reflection, retaining every one-qubit preparation and the rest of the circuit. If there was no such gate, take the deleted circuit to be the original one and set

the deletion error to zero. If the output initially carries an input bit, fix that label to any one value in both circuits and put $r' = 1$; otherwise put $r' = 0$. The deleted circuit now has an idle, input-independent output first block and $k = n - r' \geq n - 1 \geq 1$ free labels. Here r still counts labels in the original output block, so $0 \leq r' \leq r$. Fixing an output label may interchange the two parity classes, which has no effect on the argument.

When a gate is deleted, let $|\alpha_x\rangle$ be its block's normalized product state immediately before that reflection and write its triggering probability as $q(x) = |\langle v, \alpha_x\rangle|^2$. Conditional on the possible fixed output label, distinct assignments to its $r - r'$ other labels give orthogonal block states. Writing p, p' for the original and deleted acceptance probabilities and taking the uniform expectation over all k free labels, we obtain

$$\mathbb{E}q(x) \leq 2^{-(r-r')}, \qquad |p(x) - p'(x)|^2 \leq 4q(x).$$

For the first inequality, the sum of squared overlaps with the $2^{r-r'}$ orthonormal initialized block vectors is at most one; labels outside the block merely repeat those vectors equally often in the average. The other first-layer blocks are in a tensor factor independent of this block at each fixed input. The second inequality follows from pure-state trace distance in the convention $\frac{1}{2}\|\rho - \sigma\|_1$: the overlap with the reflected vector is $1 - 2q$, giving distance $\sqrt{1 - (1-2q)^2} = 2\sqrt{q(1-q)} \leq 2\sqrt{q}$. Subsequent computation and the terminal measurement cannot increase this distance. If no gate is deleted, the displayed bounds hold with $q = 0$.

Call an input bad if $|p - p'| > 1/24$, and let $\beta_H, \beta_L \in [0, 1]$ be its fractions in the two parity classes. Markov's inequality bounds its uniform fraction by $4 \cdot 24^2\, 2^{-(r-r')}$. Since $k \geq 1$, the classes have equal cardinality, so $\beta_H + \beta_L$ is twice this fraction. Using $r' \leq 1$ gives

$$\beta_H + \beta_L \leq 9216\, 2^{-r}. \tag{11}$$

This upper bound may exceed two and is still valid. Every good input has deleted-circuit error at most $1/3 + 1/24 = 3/8$.

The deleted circuit again has form (9), now with its own affine coefficients and source factors. Merge all idle non-output source wires into at most one product factor, and absorb input-independent factors into the affine coefficient. The merged factor is still a product test on product states and is covered by the idle-source case of Lemma 9.1. Physical endpoint bounds persist. Let M count the retained nonconstant sources. Each contains at least one occurrence of a free input label, and their physical supports are disjoint. There are at most nm supplied input wires; the other source blocks are at most G_1 surviving first-layer gate blocks, with at most one merged idle factor. Therefore

$$M \leq nm, \qquad M \leq G_1 + 1.$$

The number M_{full} containing every free label is at most m, by charging any one of the $k \geq 1$ free labels. Ancillary-only sources do not enter either count. An empty source set has $M = M_{\text{full}} = 0$, $F = 1$ and $\mathcal{D} = 0$.

For general $0 \leq \epsilon < 1/2$, the derivative numerator in the previous endpoint minimization decreases with u. It can change sign only from positive to negative, giving an interior maximum, not an interior minimum. The minimum is consequently at $u = 0$ or $u = \epsilon$, and is

$$c(\epsilon) = \min\left\{1 - \epsilon + \log(1-\epsilon), 1 + \log\frac{1-2\epsilon}{1-\epsilon}\right\}.$$

At $\epsilon = 3/8$ the second endpoint is smaller: the first minus the second is $\log(25/16) - 3/8 > 0$, since $\log(25/16) = 2\log(5/4) > 2/5 > 3/8$. Here $\log(1+t) > t/(1+t)$ for $t > 0$, by integrating $1/u$ on $[1, 1+t]$. Thus $c(3/8) = c_* := 1 + \log(2/5)$, with $0 < c_* < 1$ because $e > 5/2$.

Suppose both parity classes have a good input. The deleted affine coefficient cannot be zero, since those two inputs have acceptance probabilities in disjoint intervals. If this coefficient is negative, complement both p and p' and exchange the desired-one and desired-zero classes; this leaves $|p-p'|$ and the sum of bad fractions unchanged. Use a, b for the resulting deleted coefficients, with $b > 0$, and H, L for the resulting high and low classes. Their good points imply the same endpoint restrictions and thresholds h, ℓ as above, with $\epsilon = 3/8$. These pointwise bounds apply only on good inputs. Nevertheless $0 \le \mathcal{D} \le M$ everywhere, so whole-class averages satisfy

$$\mathbb{E}_H \mathcal{D} \le -\log h + M\beta_H, \qquad \mathbb{E}_L \mathcal{D} \ge (1-\beta_L)(1-\ell) \ge 1-\ell-\beta_L.$$

Comparing the whole-class averages, rather than averages conditioned on goodness, permits Lemma 9.1 to cancel every missing-label source exactly. Therefore

$$\frac{6M_{\text{full}}}{2^k} + M\beta_H + \beta_L \ge c_*.$$

If a class has no good input, then $\beta_H + \beta_L \ge 1$ and $M + 1 \ge 1 > c_*$, without any need to normalize the affine coefficients. Combining this case with the preceding one gives the unconditional inequality

$$\frac{6M_{\text{full}}}{2^k} + (M+1)(\beta_H + \beta_L) \ge c_*.$$

This includes $b = 0$ and $M = 0$ cases. It explicitly pays for exceptional inputs; the multiplier must not be replaced by the full-label count. Cancellation of missing-label sources is valid in the whole parity classes, not necessarily after conditioning on goodness.

Using (11), $k \ge n-1$, and $M_{\text{full}} \le m$ yields

$$c_* \le 12m\, 2^{-n} + 9216(M+1)2^{-r}.$$

The independent restriction argument (10), applied to the original circuit, gives $2^{-r} \le m/(a_0 2^n)$. Consequently, using $M + 1 \ge 1$ to absorb the term $12m$,

$$c_* 2^n \le (12 + 9216/a_0)m(M+1).$$

Dividing by $12 + 9216/a_0$ and using $M + 1 \le G_1 + 2$ and $M + 1 \le nm + 1$ gives respectively $m(G_1 + 2) \ge c_0 2^n$ and $m(nm+1) \ge c_0 2^n$, with c_0 as in Theorem 2.4. Finally $m \ge 1$ implies $nm + 1 \le (n+1)m$, and hence $m \ge \sqrt{c_0}\, 2^{n/2}/\sqrt{n+1}$. The constants are positive: $\log(3/2) < 1/2$ gives $a_0 > 1/36$, and $e > 1 + 1 + 1/2 + 1/6 = 8/3 > 5/2$ gives $c_* > 0$. This proves the theorem without a size or ancillary-width restriction beyond finiteness and product initialization. At depth three, the corresponding sources need not be single-reflection blocks on a product initial state; the rank-two perturbation estimate used here has not been established for those sources.

10 A two-layer bound for every deep symmetric transition

The parity argument uses alternating signs on the full cube. A different invariant, transport of nonnegative defects between adjacent Hamming slices, gives a lower bound for every symmetric function with a deep transition. We use the circuit model and transition radius of Section 2. Circuits have at most two entangling layers, with arbitrary finite size and input-independent pure product ancillas; missing layers may be empty. Fixing a logical label always fixes all its supplied occurrences. The transition makes the function nonconstant, so $m \ge 1$.

Theorem 10.1. *If a symmetric function has transition radius $\rho \geq 5$, every two-entangling-layer circuit computing it with worst-case error at most $1/3$ satisfies*

$$m \geq \frac{2^{(2\rho-2)/9}}{3000(2\rho+1)^{10/9}}.$$

The output may be any physical wire, the first output gate has unrestricted arity, and ancillary width is unrestricted subject to pure product initialization.

We first prove a bound for an input-independent first output block at an even central transition. Restriction and deletion then remove the output-block condition. Finally, a restriction places a transition attaining ρ in this central form. No correctness assumption away from the two selected slices is needed in the central-transition argument.

10.1 Defect transport with an input-independent output source

First take $N = 2t \geq 8$ logical labels, and suppose the first-layer block containing the output has no free input labels. In this section N denotes the number of free logical labels, not total physical width. Assume the circuit has error at most $1/3$ on two opposite prescribed constant values on slices $t-1, t$. The normal form (9) holds without further fixing labels. Absorb every input-independent factor into its affine coefficient, preserving the physical endpoints. For a retained source on $k \geq 1$ distinct logical labels, its scalar factor has the form

$$c(z) = q(z) + \langle z|\Delta|z\rangle, \qquad q(z) = \kappa \prod_i a_i(z_i), \qquad \|\Delta\| * 1 \leq 2,$$

where $0 \leq \kappa, a_i(b) \leq 1$ and Δ is Hermitian. More explicitly, use the initialized copy-code isometry V and product test P from the proof of Lemma 9.1. If $S = I - 2|v\rangle\langle v|$ is the first-layer reflection, put

$$A = V^\dagger P V = \kappa \bigotimes * i = 1^k A_i, \qquad a_i(b) = \langle b|A_i|b\rangle, \qquad \Delta = V^\dagger(SPS - P)V.$$

Here $0 \leq A_i \leq I_2$, and $c(z) = \langle z|V^\dagger SPSV|z\rangle \in [0,1]$. The range of $SPS - P$ lies in $\mathrm{span}\{v, Pv\}$, and its operator norm is at most one because it is a difference of two projections. Compression preserves these rank and norm bounds, giving $\|\Delta\|_1 \leq 2$. This argument includes repeated occurrences of each label and fixed product ancillary factors. Idle product sources have $S = I$ and $\Delta = 0$.

Let μ_s be the uniform distribution on $\{x \in \{0,1\}^N : |x| = s\}$. View every local source function as a function of all N labels by ignoring the absent labels. For every scalar product contraction q, its defect $d_q = 1 - q$ satisfies

$$\mathbb{E}_{\mu_{s-1}} d_q \leq \frac{N-s+1}{N-s} \mathbb{E}_{\mu_s} d_q, \qquad \mathbb{E}_{\mu_s} d_q \leq \frac{s}{s-1} \mathbb{E}_{\mu_{s-1}} d_q, \tag{12}$$

for integers $2 \leq s \leq N-1$. To prove this, express each soft one-bit factor as a convex mixture of the full set, a singleton and the empty set, with respective probabilities $\min(a_i(0), a_i(1))$, $|a_i(0) - a_i(1)|$, and $1 - \max(a_i(0), a_i(1))$. The singleton is the maximizing bit value; a zero singleton weight makes its choice irrelevant. Choose these sets independently, and independently replace their product by the empty subcube with probability $1 - \kappa$. Absent labels are unrestricted. Thus $q(x) = \mathbb{E}_A \mathbf{1}[x \in A]$ for a random subcube A, including the empty subcube.

Fix A. Couple $X \sim \mu_{s-1}$ to $Y \sim \mu_s$ by adding one of its $N - s + 1$ zeros uniformly. From $X \notin A$, at most one added coordinate can repair all violated literals: two wrong positive literals cannot both be repaired, and a wrong negative literal cannot be repaired by adding a one. Hence

$$\Pr[Y \notin A] \geq \frac{N-s}{N-s+1} \Pr[X \notin A].$$

The reverse coupling removes a uniformly chosen one from Y and has marginal μ_{s-1}. At most one of the s one-coordinates can repair a point $Y \notin A$: only a unique violated negative literal could be repaired. This gives

$$\Pr[X \notin A] \geq \frac{s-1}{s} \Pr[Y \notin A].$$

The empty subcube has defect identically one and obeys both inequalities as well. Averaging over A proves the claim.

Put $K = \lfloor N/8 \rfloor$ and let $R_0 = 4/3$ denote a scalar transport coefficient throughout this section, not the output-block reflection in Section 9. The marginal probability of a specified k-bit pattern z of weight j is

$$p_s(z) = \frac{\binom{N-k}{s-j}}{\binom{N}{s}}.$$

For $k \leq K \leq t/4$ every such pattern has positive probability on both central slices, and

$$\frac{p_{t-1}(z)}{p_t(z)} = \frac{(t-j)(t+1)}{t(t-k+j+1)}.$$

This ratio is largest at $j = 0$; its reciprocal is largest at $j = k$. Their respective upper bounds are

$$\frac{t+1}{t-k+1} \leq R_0, \qquad \frac{t}{t-k} \leq R_0.$$

Here and below, H, L denote the two central slices in either chosen order, and their expectations are uniform on the indicated slice. Consequently every nonnegative function of those $k \leq K$ labels obeys $\mathbb{E}_L d \leq R_0 \mathbb{E}_H d$. For arbitrary $0 \leq k \leq N$, with out-of-range binomial coefficients interpreted as zero, use $\binom{N}{t} \geq 2^N/(N+1)$ and $\binom{N}{t-1} = t\binom{N}{t}/(t+1) \geq (4/5)2^N/(N+1)$. Every atom of either marginal is therefore at most

$$\frac{2^{N-k}}{\binom{N}{s}} \leq \frac{5}{4}(N+1)2^{-k}, \qquad s \in \{t-1, t\}.$$

For Hermitian Δ, the diagonal sign matrix and trace-norm duality give $\sum_z |\langle z|\Delta|z\rangle| \leq \|\Delta\|_1 \leq 2$. Thus, on either slice,

$$\mathbb{E}|c-q| \leq \max_z p_s(z) \sum_z |\langle z|\Delta|z\rangle| \leq \tfrac{5}{2}(N+1)2^{-k}.$$

The two factors in (12) at $s = t$ are at most R_0, since $t \geq 4$. Comparing c to q on both slices, and using $(1+R_0)(5/2) < 6$, a large source obeys

$$\mathbb{E}_L(1-c) \leq R_0 \mathbb{E}_H(1-c) + 6(N+1)2^{-k}. \tag{13}$$

Small sources obey the same inequality without its additive term.

The two opposite required values rule out $b = 0$. Complement the output event and the prescribed values together if necessary so the affine normal form has $b > 0$. Now let H be the desired-one slice and L the desired-zero slice. Set $h = (2/3 - a)/b$, $\ell = (1/3 - a)/b$. The physical endpoints and correctness on both nonempty slices imply $0 \leq \ell < h \leq 1$. High-slice inputs have $F \geq h$, low-slice inputs have $F \leq \ell$, so the product-defect inequalities from Section 9 give, for $D = \sum_g (1 - c_g)$,

$$\mathbb{E}_H D \leq -\log h, \qquad \mathbb{E}_L D \geq 1 - \ell.$$

The physical affine endpoints imply the positive gap

$$1-\ell+R_0\log h\geq c_s:=1-\tfrac{4}{3}\log 2>0.$$

Here is the minimization, also used below with a different error. For $0\leq\epsilon<1/2$ and a coefficient $R\geq 1$, feasibility gives

$$u=\epsilon-a\in[0,\epsilon],\qquad u+1-2\epsilon\leq b\leq 1-\epsilon+u.$$

In particular $b>u$, so the derivative of $1-u/b+R\log((u+1-2\epsilon)/b)$ with respect to b is $(u-Rb)/b^2<0$. Thus the minimum has $b=1-a$. With $v=1-a\in[1-\epsilon,1]$, the remaining function and its derivative are

$$g_{\epsilon,R}(v)=\frac{1-\epsilon}{v}+R\log\frac{v-\epsilon}{v},\qquad g'_{\epsilon,R}(v)=\frac{((1+R)\epsilon-1)v+\epsilon(1-\epsilon)}{v^2(v-\epsilon)}.$$

The denominator is positive throughout this interval. For $\epsilon=1/3$ and $R=4/3$, the numerator is $2(1-v)/9\geq 0$. The minimum is consequently at $v=2/3$, giving c_s. For later numerical estimates, note the elementary rational bound

$$e^{7/10}>1+\frac{7}{10}+\frac{(7/10)^2}{2}+\frac{(7/10)^3}{6}=\frac{12013}{6000}>2.$$

Hence $\log 2<7/10$ and $c_s>1/15>0$.

If M_{large} sources contain at least $K+1$ labels, summing the transport inequalities gives

$$c_s\leq 6(N+1)M_{\text{large}}2^{-(K+1)}.$$

Their physical supports are disjoint and each consumes at least $K+1$ supplied wires. Thus $(K+1)M_{\text{large}}\leq Nm$. Since $K+1>N/8$, we have $M_{\text{large}}<8m$ and $2^{-(K+1)}<2^{-N/8}$, yielding

$$m>\frac{c_s2^{N/8}}{48(N+1)}.\tag{14}$$

Ancillary-only factors were absorbed into the affine coefficient, so their number is not charged to input copies.

10.2 Removing the output-source restriction

Still take even $N\geq 8$ and the same correctness requirement on its central slices, but require only that the output wire is initially in an input-independent pure state. Write B_0 for its first-layer block and $0\leq r\leq N$ for the number of free labels represented there. Fix these labels and, if r is odd, one additional label, with exactly half the fixed labels set to one. The extra label exists because N is even. The remaining length $q=N-2\lceil r/2\rceil$ is even. If $q\geq 2$, the original selected weights become $q/2-1,q/2$, and the first output block is input-independent. If $q\geq 8$, apply (14); if $q<8$, use the original circuit's $m\geq 1$, without asserting any residual transition when $q=0$. Both cases imply

$$2^{-r}\leq A(N+1)^8m^8 2^{-N},\qquad A=2(48/c_s)^8.\tag{15}$$

Indeed, if $q\geq 8$, raising (14) to the eighth power yields $2^q<(48/c_s)^8(q+1)^8m^8$, and $N-r\leq q+1$ and $q\leq N$ give (15). If $q<8$, its evenness implies $q\leq 6$, so $2^{N-r}\leq 2^7<A\leq A(N+1)^8m^8$. Here $0<c_s<1$ makes $A>2^7$ immediate.

Return to the original N-label circuit, before this auxiliary restriction, and delete only the first reflection containing the output; retain all one-qubit preparations and all other gates. Write its new acceptance probability as p'. For the initialized copy-code isometry V on the deleted block's r labels and normalized product reflection normal v,

$$|p(x) - p'(x)| \le 2|\langle v, Vx\rangle|, \qquad \mathbb{E}_{\mu_s}|\langle v, Vx\rangle|^2 \le \tfrac{5}{4}(N+1)2^{-r}.$$

Here $\langle v, Vx\rangle$ abbreviates $\langle v|V|x_{B_0}\rangle$, and $s \in \{t-1, t\}$. The other first-layer blocks factor from this block at each fixed input. As in Section 9, the trace distance between its original and reflected pure states is $2\sqrt{w(1-w)} \le 2\sqrt{w}$, where $w = |\langle v, Vx\rangle|^2$; subsequent gates and the terminal measurement cannot increase distinguishability. This proves the first estimate. The second follows by bounding each marginal mass by $\frac{5}{4}(N+1)2^{-r}$ and using $\sum_z |\langle v|V|z\rangle|^2 = \|V^\dagger v\|^2 \le 1$. The argument includes $r = 0$, with a one-dimensional code domain. If there is no first output reflection, take $p' = p$ and zero deletion error instead. On each central slice Markov's inequality bounds the fraction of points with $|p - p'| > 1/240$ by

$$\beta = 288000(N+1)2^{-r}.$$

Indeed the mean squared change is at most $5(N+1)2^{-r}$ and $5 \cdot 240^2 = 288000$. The symbol β is a common upper bound on the two bad fractions, not their sum, and need not be less than one. On good points the deleted error is at most $\epsilon' = 1/3 + 1/240 = 27/80$.

The deleted output wire is now an input-independent singleton first-layer block. Merge all idle non-output wires, including its former block companions, into at most one product source, and absorb every input-independent factor. Idle sources still have correction $\Delta = 0$. Let M count retained input-dependent factors and M_{large} those on at least $K+1$ free labels. Their supports are disjoint and each contains at least one supplied input occurrence, so

$$M \le Nm, \qquad M_{\text{large}} \le Nm/(K+1) < 8m.$$

The deleted circuit has its own affine coefficients. If both slices have good points, its coefficient is nonzero, since the two required acceptance intervals are disjoint. If negative, complement both p, p' and the prescribed values, preserving $|p - p'|$ and the common bad-fraction bound. Reuse a, b for the resulting deleted coefficients, with $b > 0$, and H, L for the desired-one and desired-zero slices. Put $h = (1 - \epsilon' - a)/b$ and $\ell = (\epsilon' - a)/b$. Existence of a good point in each slice gives the same feasibility inequalities as above, in particular $0 \le \ell < h \le 1$. For $D = \sum_g (1 - c_g)$, the good-point bounds and $0 \le D \le M$ everywhere imply the whole-slice bounds

$$\mathbb{E}_H D \le -\log h + M\beta, \qquad \mathbb{E}_L D \ge 1 - \ell - \beta.$$

At error ϵ', the gap is bounded below by

$$1 - \ell + R_0 \log h \ge 1 + \tfrac{4}{3}\log(26/53) > 1/25.$$

Apply the displayed formula for $g'_{\epsilon,R}$: on $v \in [53/80, 1]$ its numerator is $-(17/80)v + 1431/6400 \ge 71/6400 > 0$, so the minimum is at $v = 53/80$. Finally $\log(53/26) = \log 2 + \log(53/52) < 7/10 + 1/52 < 18/25$ proves the strict numerical bound.

The source transport inequalities apply to whole-slice averages, not to averages conditioned on goodness. Summing them and combining with these defect bounds yields

$$\frac{1}{25} \le 6(N+1)M_{\text{large}}2^{-(K+1)} + (R_0 M + 1)\beta.$$

If a slice has no good point, then $\beta \geq 1$, and the same inequality is trivial; no affine normalization is asserted in that case. This also covers a zero deleted coefficient and an empty retained source set, for which $F = 1$ and $D = M = 0$. The exceptional-input penalty uses the total count M, not merely M_{large}. Since $Nm \geq 8$, we have $R_0 M + 1 \leq (3/2)Nm$. Using this, the source counts, and the independent original-circuit estimate (15) gives

$$\frac{1}{25} \leq 48(N+1)m2^{-N/8} + 432000AN(N+1)^9 m^9 2^{-N}.$$

At least one term is at least $1/50$. If it is the first, then $m \geq 2^{N/8}/(2400(N+1))$. If it is the second, then

$$m^9 \geq \frac{2^N}{21600000AN(N+1)^9}.$$

The bound $c_s > 1/15$ proved above gives $A < 2 \cdot 720^8$. Moreover,

$$\frac{3000^9}{43200000 \cdot 720^8} = \frac{(25/6)^8}{14400} > \frac{4^8}{14400} > 1.$$

Using $N \leq N+1$, the second alternative therefore implies the bound below. The first alternative is stronger as well: its ratio to the claimed lower bound is $(5/4)2^{N/72}(N+1)^{1/9} > 1$. Thus

$$m \geq \frac{2^{N/9}}{3000(N+1)^{10/9}}. \tag{16}$$

10.3 Restricting an arbitrary symmetric transition

Choose a transition $i, i+1$ attaining $\rho = \min\{i+1, n-i\}$. If necessary, complement all logical input bits, on all their physical occurrences, so that the selected transition is between weights $\rho - 1, \rho$. This is a local unitary change and does not change copy count or entangling depth. We have $2\rho - 1 \leq n$.

If $n > 2\rho$, retain 2ρ labels and fix all others to zero. When the output is an input wire, choose the retained set to exclude its label. If $n = 2\rho$ and the output is ancillary, no fixing is needed. If $n = 2\rho$ and the output is an input wire, fix its label and one other label to values zero and one; the remaining selected weights are $\rho - 2, \rho - 1$, central for length $2\rho - 2$.

The only remaining case is $n = 2\rho - 1$. Fix one label to one, choosing the output label when necessary. This again leaves length $2\rho - 2$ with a central transition and an initially known output. Every fixed logical label means fixing all its physical copies; these become product ancillas already allowed by the theorem.

In all cases the output is initially input-independent, and the remaining even length satisfies $2\rho - 2 \leq N \leq 2\rho$ and $N \geq 8$. Keep fixed physical wires as product ancillas, with their one-qubit preparations; no gate simplification or extra entangling layer is needed. Equation (16), together with

$$2^{N/9} \geq 2^{(2\rho-2)/9}, \qquad (N+1)^{10/9} \leq (2\rho+1)^{10/9},$$

proves Theorem 10.1. The restricted function need not globally be majority or equality: only its opposite values on the two chosen slices are used.

11 Independent physical inputs versus a repetition promise

The ordinary-input lower bounds of [18] are consistent with polynomial-copy computation at a larger constant depth: the initialization interfaces differ. In particular, their depth-two influence

bound concerns the uniform distribution on independent physical inputs. One cannot simply replace the physical input length by the number of logical labels. The following construction makes this distinction explicit; it does not preclude using those results together with an additional transfer argument.

A read-once DNF whose copy restriction is parity. Fix $n \geq 2$, let $\mathcal{O}_n = \{z \in \{0,1\}^n : |z| \text{ is odd}\}$, and put $m = |\mathcal{O}_n| = 2^{n-1}$ and $N = nm$. Here n is the logical input length and N is the number of physical input qubits, excluding ancillas. For every $z \in \mathcal{O}_n$, allocate a private block of n independent physical variables $y_{i,z}$ and put

$$F(y) = \bigvee_{z \in \mathcal{O}_n} C_z(y), \qquad C_z(y) = \bigwedge_{i=1}^{n} [y_{i,z} = z_i].$$

There is an exact two-entangling-layer implementation with $m+1$ ancillary qubits. The first layer computes each C_z into a separate zero-initialized flag with one generalized Toffoli, using local X gates for negative literals. All these gate supports are disjoint. For the second layer, initialize a fresh output to one and flip it precisely when every flag is zero, using a negative-control generalized Toffoli. Its resulting value is $1 - \prod_z (1 - C_z) = F$. Negative controls require only local conjugations, not additional entangling layers; no uncomputation is required. The total width is $N + m + 1$ and the construction has $m + 1$ entangling gates and $O(N)$ one-qubit gates. Thus it is polynomial-size in N, but exponential-size in n.

This is the standard read-once-DNF implementation discussed in [18, Section 2.4]; after negating selected physical variables, F is a tribes function. Its role here is the particular identification of physical occurrences. On the repetition code $y_{i,z} = x_i$, the clause C_z is true exactly when $x = z$. Consequently precisely one clause is true for odd x and none is true for even x, so F restricts to logical parity using m supplied copies.

Influence and its normalization. Define pivotal total influence by

$$I(F) = \sum_{z \in \mathcal{O}_n} \sum_{i=1}^{n} \Pr_y[F(y) \neq F(y \oplus e_{i,z})],$$

where y is uniform on $\{0,1\}^N$ and $e_{i,z}$ flips one physical coordinate. A coordinate is pivotal exactly when the other $n-1$ literals in its own clause are satisfied and all other clauses are false. The bit being flipped is unrestricted. Each other clause is true with probability 2^{-n}, and the disjoint blocks are independent. Therefore

$$I(F) = nm\, 2^{-(n-1)}(1 - 2^{-n})^{m-1} = n(1 - 2^{-n})^{m-1}.$$

Bernoulli's inequality gives $(1 - 2^{-n})^{m-1} \geq 1 - (m-1)2^{-n} = 1/2 + 2^{-n}$, so $n/2 < I(F) \leq n$. Since $\log N = (n-1)\log 2 + \log n = \Theta(n)$, this proves $I(F) = \Theta(\log N)$ as $n \to \infty$.

Index the physical coordinates by $[N]$ and define

$$\widehat{F}(S) = \mathbb{E}_y[F(y)(-1)^{\sum_{j \in S} y_j}].$$

The spectral convention of [18, Section 3.2] is

$$\begin{aligned} \mathrm{Inf} * \mathrm{sp}(F) &:= \sum * S \subseteq [N] |S| \widehat{F}(S)^2, \\ &= \frac{1}{4} \sum_{j=1}^{N} \mathbb{E}_y (F(y) - F(y \oplus e_j))^2 = \frac{1}{4} I(F). \end{aligned}$$

The middle equality follows from Parseval applied to each discrete derivative; the last uses $F \in \{0,1\}$. For the sign-valued function $1-2F$, spectral influence instead equals $I(F)$. Thus the example is compatible, in either normalization, with the $O(\log N)$ independent-input bound of [18, Theorem 1.3].

Identification does not in general preserve influence even up to a universal constant. On the same physical inputs, consider $G(y) = \bigwedge_{z\in\mathcal{O}_n} y_{1,z}$. Its pivotal influence is $m2^{-(m-1)}$, since each of its m relevant coordinates is pivotal exactly when the other $m-1$ are one. On the repetition code, however, G becomes the dictator x_1, whose logical influence is one. Since $m2^{-(m-1)} \to 0$, no dimension-independent bound of logical influence by a constant times full-cube influence holds for all such identifications.

Average-case statements and the size of the promise. Let $\mathcal{C} = \{y : \exists x\ \forall i, z,\ y_{i,z} = x_i\}$ be the classical repetition code in the physical input cube. It has 2^n points, so

$$\Pr_y[y \in \mathcal{C}] = \frac{2^n}{2^N} = 2^{-n(m-1)}.$$

For any nonnegative error function e on that cube,

$$\mathbb{E}[e \mid \mathcal{C}] = \frac{\mathbb{E}[e\mathbf{1}_{\mathcal{C}}]}{\Pr[\mathcal{C}]} \le 2^{n(m-1)}\mathbb{E}[e].$$

This loss can be attained: $e = \mathbf{1}_{\mathcal{C}}$ has full-cube mean $\Pr[\mathcal{C}]$ but conditional mean one. This elementary example concerns transfer of averages, not realizability of that particular error function by a shallow circuit. Hence a full-cube average guarantee alone cannot simply be conditioned onto the supplied-copy promise.

A valid small-copy consequence of an ordinary-input theorem. Two classical basis copies can be prepared from one ordinary input using n disjoint CNOT gates:

$$|x\rangle|0^n\rangle \longmapsto |x\rangle|x\rangle.$$

This adds one entangling layer. On superpositions it prepares the correlated encoding $\sum_x \alpha_x|x\rangle|x\rangle$, not two independent copies of an unknown quantum state. Prepending this layer to an exact two-layer parity evaluator with two supplied copies gives an ordinary-input exact parity circuit of depth at most three. With one supplied copy no preparation is needed; with zero copies nonconstant parity is impossible. The designated output may be any wire, and all additional ancillas start in an input-independent product state, as permitted in the cited model. Theorem 1.2 of [18] therefore excludes $m \le 2$ for exact two-layer parity when $n > 100$, irrespective of finite size or ancillary width. This is a derived small-copy obstruction, not a new depth-three theorem or a bounded-error extension of that theorem. The exponential lower bound in Theorem 2.4 is asymptotically stronger and also permits worst-case error $1/3$.

Finally, identifying occurrences as above is not a coordinate-fixing restriction on the independent physical cube. When a logical label is fixed in a supplied-copy lower-bound argument, every one of its physical occurrences must be fixed together.

12 Consequences and limitations

Upper bounds and exact gate specifications. Theorem 2.2 gives structure-sensitive sufficient copy counts, not a characterization of the minimum. Its families are nonuniform and use arbitrary

exact one-qubit gates. The depth and polynomial-resource exponent in the logarithmic pooling bound may depend on the fixed saving exponent; the theorem does not justify letting that exponent grow with the input length. Section 7 describes the additional filter parameters algebraically relative to the imported exact primitives, but establishes neither exact synthesis over a fixed finite gate alphabet nor stronger uniformity.

Exactness is not preserved in general by replacing those gates with approximations. A bounded-error implementation requires a quantitative precision budget. For example, for corresponding unitary gates with $\|U_j - \widetilde{U}_j\| \le \eta_j$ in operator norm, telescoping their products gives

$$\|U_s \cdots U_1 - \widetilde{U}_s \cdots \widetilde{U}_1\| \le \sum_{j=1}^{s} \eta_j.$$

For the same initialized state and terminal measurement effect, the acceptance probabilities then differ by at most $2\sum_j \eta_j$, by expanding the difference of the two quadratic forms. This elementary error estimate does not supply exact finite-alphabet synthesis or an implementation of the required approximants with any additional resource guarantee.

Scope of the lower bounds. Theorems 2.4 and 10.1 allow arbitrary finite ancillary width, with input-independent pure product initialization and any designated output wire. They depend on the single-reflection source structure at two entangling layers. At greater depth, intermediate correlations need not admit the same source decomposition and rank-two correction, so these proofs do not by themselves give the corresponding depth-three lower bounds. Appendix A treats exact copy minima for its listed functions under the additional requirement that the output be an ancillary qubit untouched by the first entangling layer; those minima do not automatically apply to bounded-error computation or unrestricted output wires.

The separate all-depth result in Appendix B is a physical-width bound. Its copy consequence requires the additional premise $N \le Cnm$, where $N = nm + a$ is total width, including a ancillary qubits, and C is fixed for the circuit family. Without that premise, a width lower bound does not give the stated copy lower bound: ancillary qubits can contribute to N without contributing supplied input copies. Polynomial ancillary width in n alone does not imply $N = O(nm)$. That appendix permits any fixed input-independent ancillary state, a different initialization assumption from the two-layer results. No extension of the two-layer supplied-copy lower bounds to arbitrary constant depth with unrestricted ancillary width is proved here.

The total-language separation and its density. For the particular language L in Corollary 8.1, the number of accepted strings at physical input length $P = 400r^2$, $r \ge 1$, is exactly

$$|L \cap \{0,1\}^P| = \binom{2r}{r}, \qquad 2^r \le \binom{2r}{r} \le 2^{2r}.$$

Each accepted input is determined by one weight-r block, repeated $200r$ times. The lower inequality follows from $\binom{2r}{r} = \prod_{j=1}^{r}(r+j)/j \ge 2^r$; the upper follows by counting all $2r$-bit strings. Thus the acceptance density is at most $2^{-P+2r} = 2^{-P+\sqrt{P}/10}$, exponentially small in P. Nevertheless L is not polynomially sparse: 2^r eventually exceeds every fixed polynomial in $P = 400r^2$. At nonencoded lengths the language is empty. In particular, the constant-zero predictor has exponentially small error under uniform physical inputs. The separation proved in Section 8 is a worst-case statement, not an average-case hardness result.

The same fixed L lies outside $\mathsf{AC}^0[p]$ for every prime p, using the distinct-prime case of Smolensky's modular lower bound [23]. This does not establish $L \notin \mathsf{ACC}^0$, which also allows fixed composite moduli. Likewise the simulation consequence is precisely $\mathsf{TC}^0 \subseteq \mathsf{EQAC}^0 \circ \mathsf{NC}^0$, with the explicit classical duplication map of Section 8. It is not an ordinary-input containment $\mathsf{TC}^0 \subseteq \mathsf{EQAC}^0$. Any limited fanout used inside the quantum constructions is implemented and charged; arbitrary unbounded fanout is not a free gate in the model.

Further directions within the circuit model. Natural targets are improved supplied-copy upper bounds for central-transition symmetric families, with $\rho(f) = \Theta(n)$, and a copy-sensitive lower-bound argument that survives a third entangling layer without the relative-workspace premise. An auxiliary-state sampling approach would also have to account for state preparation, supplied input copies, and the gates and depth needed to realize its measurement. An unrestricted collective measurement is not a free operation in the model of Section 2; a statistical sample bound alone does not provide a shallow circuit construction.

A Exact copy optima with a fresh depth-two output

Throughout this appendix, $n \geq 1$ and the circuit has at most two entangling layers in the model of Section 2. All ancillary qubits start in an input-independent pure product state. We impose a further restriction: the output is an ancillary qubit untouched by the first entangling layer. Arbitrary one-qubit gates remain allowed. Computation is exact, but input copies may be overwritten and work need not be cleaned. The copy minimum is over nonnegative integers m and finite circuits of arbitrary size; a superpolynomial copy minimum therefore excludes polynomial-size families at this depth. Every nonconstant function requires $m \geq 1$, since a circuit supplied with no input wires has input-independent output.

Theorem A.1 (Exact fresh-output minima)**.** *In this restricted two-layer model,*

$$m_{\min}(\mathsf{EX}_j) = \binom{n}{j} \quad (2 \leq j \leq n-2).$$

For $n \geq 2$,

$$m_{\min}(\mathsf{EX}_1) = m_{\min}(\mathsf{EX}_{n-1}) = \min\{n-1, \lceil n/2 \rceil + 2\}.$$

For a threshold $\mathsf{TH}_t(x) = \mathbf{1}[|x| \geq t]$,

$$m_{\min}(\mathsf{TH} * t) = \binom{n-1}{t-1} \quad (3 \leq t \leq n-2).$$

For $n \geq 2$, TH_2 and TH_{n-1} require $\min\{n-1, \lceil (n-1)/2 \rceil + 2\}$ copies. The nonconstant extreme thresholds and $\mathsf{EX}_0, \mathsf{EX}_n$ each need one copy. For $2 \leq a \leq b \leq n-2$, the interval predicate satisfies

$$m * \min(\mathbf{1}[a \leq |x| \leq b]) = \binom{n-1}{a-1} + \binom{n-1}{b}.$$

A.1 Certainty forces a rectangle, except at an endpoint

The fresh-output form is (9) without any fixed input labels. Exact nonconstant output takes both values $0, 1$, while every value $a + bF$ lies on the segment between the physical endpoints

$a, a+b \in [0,1]$. Consequently those endpoints must be $0, 1$ in one order. Thus F is the intended predicate or its complement, and $F \in \{0,1\}$. Since every factor lies in $[0,1]$, every factor is one on the accepting set of F, and at every rejecting input at least one factor is exactly zero. Here and below, a factor is *certain* at y when $c(y) = 1$, and *rejects* y when $c(y) = 0$. Input-independent factors must equal one and can be removed. An idle first-layer wire is also a source; such wires may either be kept separate or grouped into one product source.

For a nonconstant source containing $k \geq 1$ logical labels, let V be its pre-reflection product isometry as in Lemma 9.1, $R = I - 2|v\rangle\langle v|$ its reflection with normalized product normal v, and P its product test. An idle source instead has $R = I$. Write $A = V^\dagger P V = \bigotimes_i A_i$, with $0 \preceq A_i \preceq I_2$, absorbing any scalar contraction into one factor. Then

$$I - A \succeq (I - A_i) \otimes I.$$

Let C be any collection of local computational words on which the factor is certain, and let Π_C project onto their logical span. Since P is a projection, certainty means $\|(I-P)RV|y\rangle\|^2 = 0$ for each $y \in C$. By linearity, $(I-P)RV\Pi_C = 0$, giving

$$(I-P)V\Pi_C = 2(I-P)|v\rangle\langle v|V\Pi_C.$$

Its Gram matrix is $\Pi_C(I-A)\Pi_C$, so this compressed positive defect has rank at most one. For an idle source, certainty instead gives $(I-P)V\Pi_C = 0$, and the same rank bound holds with rank zero. The displayed domination follows from

$$I - \bigotimes_j A_j - (I - A_i) \otimes I = A_i \otimes \left(I - \bigotimes_{j \neq i} A_j \right) \succeq 0.$$

Suppose C contains r different words with bit i fixed to b. On their span the compression of $(I - A_i) \otimes I$ is $(I - A_i)_{bb} I_r$. If this scalar were positive, positive-semidefinite domination would force the compressed defect to have rank at least r. In particular, on the local weight-s slice, $r = \binom{k-1}{s-b}$. Whenever certainty supplies at least two such words for each i, b, all diagonal entries of each positive matrix $I - A_i$ vanish; its off-diagonal entries then vanish as well. Thus $A = I$. Because $V^\dagger(I-P)V = 0$, this implies $(I-P)V = 0$, or $PV = V$, globally. In that case

$$1 - c(y) = 4\|(I-P)v\|^2 |\langle v|V|y\rangle|^2. \tag{17}$$

Indeed, $(I-P)RV|y\rangle = -2(I-P)v\langle v|V|y\rangle$, whose squared norm is the failure probability. An idle source with $PV = V$ is identically certain and can be discarded. The overlap in (17) is a product of functions of individual logical bits, even when a bit has several physical occurrences in the source. The support of $1 - c$ is therefore a Boolean rectangle (a subcube), although its nonzero values need not be Boolean. Extend this rectangle to all n logical labels by leaving absent labels unrestricted. In particular, every fixed coordinate requires a physical occurrence in the source. Empty-support factors are irrelevant. A nonempty rectangle with u fixed ones and z fixed zeros contains exactly the consecutive weights $u, u+1, \ldots, n-z$; this observation will be used to charge boundary points.

For $F = \mathsf{EX}_j$ with $2 \leq j \leq n-2$, a factor rejecting a weight-$(j+1)$ input must contain all its one-labels: otherwise flipping an absent one to zero would preserve the local input but produce a global certainty word. Hence $k \geq j+1$. If $k \geq j+2$, the certainty slice j makes both defect multiplicities exceed one, since $1 \leq j-1 < k-1$ and $1 \leq j < k-1$. If $k = j+1$, at least one global label is absent, so local slices j and $j-1$ are both certainty slices, by setting an absent bit to zero or one. For bit value zero the second slice supplies $\binom{j}{j-1} = j \geq 2$ words; for bit value one the

first supplies the same number. Thus (17) holds. Complementing all input bits proves the same assertion for factors rejecting weight $j-1$.

A rectangle avoiding weight j can meet at most one adjacent side. On the upper side it fixes at least $j+1$ ones and contains only one weight-$(j+1)$ point. On the lower side it fixes at least $n-j+1$ zeros and contains only one weight-$(j-1)$ point. Choose an exactly-zero factor for every wrong adjacent input. For each logical variable i, the upper points with $x_i = 1$ charge $\binom{n-1}{j}$ distinct first-layer source occurrences; the lower points with $x_i = 0$ charge $\binom{n-1}{j-1}$ further occurrences. Disjoint physical sources therefore require $m \geq \binom{n}{j}$.

If instead $F = 1 - \mathsf{EX}_j$, any nontrivial factor must contain every label: an absent label would give each local assignment an extension outside weight j, forcing certainty everywhere. For an all-label source, a positive defect entry at bit b would have rank at least

$$2^{n-1} - \binom{n-1}{j-b} > 1$$

on its certainty subspace. Thus its defect is a rectangle contained entirely in weight j, necessarily a singleton. Covering the $\binom{n}{j}$ target points requires that many all-label sources. This proves the interior lower bound in both orientations. The same argument gives the opposite-orientation lower bound n at $j = 1$ for $n \geq 3$: the displayed multiplicities are then at least two. It does not give that endpoint bound at $n = 2$, which is treated separately below.

The matching circuit computes one minterm for every weight-j input in a disjoint first layer and ORs the flags into a fresh output in the second. Alternatively it computes violations for every $(j+1)$-subset of ones and every $(n-j+1)$-subset of zeros, then NORs them. Each construction uses exactly $\binom{n}{j}$ occurrences of every variable, assigned statically to distinct supplied wires. Negative controls and a final output complement use only one-qubit X gates. A single generalized Toffoli gate, with such complements when needed, also computes each nonconstant extreme threshold and $\mathsf{EX}_0, \mathsf{EX}_n$ with one copy. Together with $m \geq 1$, this proves their claimed minima.

A.2 Endpoint modules and their exact graph cost

Suppose $n \geq 3$ and $F = \mathsf{EX}_1$. Every source is certain on all global one-hot inputs. A source with at most one present label is trivial, because every assignment to its labels extends to a global one-hot input. For a source on $k \geq 2$ labels, temporarily retain the physical decomposition $A = \kappa \bigotimes_i A_i$. Its certainty span has dimension at least two, and $I - A \succeq (1-\kappa)I$. The rank-one bound therefore forces $\kappa = 1$. It also forces $(A_i)_{00} = 1$ for every label: when $k = n$, use the $k - 1 \geq 2$ one-hot words with $x_i = 0$; when $k < n$, use the two local certainty words $0, e_j$ with $j \neq i$.

Crucially, the test P is a product of rank-one physical qubit projectors on its tested wires, and the initialized vectors are products across wires. An expectation equal to one forces every tested physical projector to accept its local zero codeword with certainty. Such a projector is exactly that codeword's projector. Each compressed A_i is consequently I if no occurrence of label i is tested, or $|0\rangle\langle 0|$ otherwise, because the physical zero and one codewords are orthogonal. There can be at most one tested logical label: two of them would give two independent diagonal defects on the local one-hot certainty span, contradicting rank at most one. For an idle source the compressed defect has rank zero, so no label can be tested; the source is identically certain and can be discarded. For a reflection source, call the unique tested label, when it exists, the center. With no center, (17) holds. A rectangle avoiding all one-hot strings covers at most one weight-two point, or covers zero by fixing every bit to zero and then covers no weight-two point.

With a unique center i, put $u = V|e_i\rangle$, $a = \langle v, u\rangle$, and $Q = I - P$. Since $Pu = 0$ whereas certainty gives $Ru \in \operatorname{ran} P$, we have $\langle u, Ru\rangle = 0$. As $\langle u, Ru\rangle = 1 - 2|a|^2$, it follows that

$$|a|^2 = \tfrac{1}{2}, \qquad Qv = \frac{u}{2a}, \qquad \|Qv\|^2 = \|Pv\|^2 = \tfrac{1}{2}.$$

Choose any tested physical center wire w, and let Q_w be failure of its one-qubit test, tensor the identity elsewhere. Since $Q_w \preceq Q$ and $Q_w u = u$, applying Q_w to $Qv = u/(2a)$ gives

$$Q_w v = \frac{u}{2a}.$$

Both sides are nonzero product vectors. Thus every normal factor on a wire other than w equals, up to phase, the corresponding factor in u. If another center wire w' were tested, its normal factor would be its one-codeword, orthogonal to its zero-codeword test, forcing $Pv = 0$, a contradiction. There is exactly one tested center wire. Its normal has squared overlap $1/2$ with each of its two codewords. On every other center occurrence the normal is the one-codeword; on every other logical label it is the zero-codeword; on ancillary wires it is the initialized ancillary vector. This conclusion allows arbitrary complex local phases.

Let t_i be the number of physical center occurrences in this source, including untested ones. Directly evaluating the reflection now gives the full scalar table

$$c(y) = \begin{cases} 1 - y_i, & \text{some other present label is one,} \\ 1, & y_i = 1 \text{ and every other present label is zero,} \\ 0, & y = 0 \text{ and } t_i = 1, \\ 1, & y = 0 \text{ and } t_i \geq 2. \end{cases}$$

For the first case, the normal overlap vanishes on a noncenter one, so the reflection has no effect and the test accepts exactly when $y_i = 0$. The second is the certainty word u. In the last two cases, an untested center occurrence makes the normal overlap zero when $t_i \geq 2$; if $t_i = 1$, reflection about a balanced one-qubit normal interchanges its orthogonal zero and one codewords up to phase, so the zero input fails. A source of this last kind must contain every logical label, because otherwise a global one-hot word outside the source would induce its rejected local zero word.

Thus every nontrivial source is either a rectangle-defect source or one of the following centered sources, with the indicated behavior on zero and weight two:

- a zero-rejecting full star, containing every label and one center copy, rejecting zero and every collision edge incident to its center;
- a zero-accepting collision star, costing at least two center copies and one copy of each represented neighbor.

For a tournament on r vertices, define its load at i to be $\operatorname{indeg}(i) + \min\{2, \operatorname{outdeg}(i)\}$. Its minimum possible maximum is

$$L(0) = L(1) = 0, \qquad L(r) = r - 1 \ (2 \leq r \leq 4), \qquad L(r) = \lceil (r-1)/2 \rceil + 2 \ (r \geq 5).$$

For $r \leq 1$, take the load to be zero. For the lower bound when $r \geq 2$, take a vertex of outdegree at most $\lfloor (r-1)/2 \rfloor$, which exists because the average outdegree is $(r-1)/2$. Outdegree zero or one gives load $r - 1$, and larger outdegree gives at least the displayed balanced load. For $r \leq 4$ the chosen outdegree is at most one. For $r \geq 5$, $r - 1 \geq \lceil (r-1)/2 \rceil + 2$. Transitive tournaments attain

the small cases. For odd r, a cyclic tournament is regular; for even r, delete one vertex of a regular tournament on $r+1$ vertices. These give indegrees and outdegrees differing by at most one and attain the remaining cases.

Let a be the number of distinct labels that are centers of at least one zero-rejecting full star, and select one such source for each of these centers. These selected sources cost at least a copies of every variable. On the remaining $r = n - a$ labels, every collision edge must still be rejected by another source: a full star centered outside this set rejects none of its weight-two inputs. Assign each remaining edge to one rejecting source, orienting collision-star edges away from their center and rectangle edges arbitrarily. Incoming edges at a vertex require different sources: a star with a different center covers only its one edge to that vertex, and a rectangle covers only one weight-two point. They therefore require different physical copies. Outgoing edges require at least $\min(2, \text{outdeg})$ additional copies: a collision star costs at least two center copies, while distinct rectangle sources each cost a copy. No incoming source is also an outgoing source at that vertex. Hence $m \geq a + L(r)$. If $a = 0$, an exactly-zero factor at the zero input must instead be a rectangle fixing all n zeros. This source costs at least one full copy and rejects no weight-two input, giving $m \geq 1 + L(n)$.

When $a \geq 1$ and $r \leq 4$, $a + L(r) \geq n - 1$. When $a \geq 1$ and $r \geq 5$,

$$a + L(r) = \left\lceil \frac{n + a - 1}{2} \right\rceil + 2 \geq \lceil n/2 \rceil + 2.$$

For $a = 0$, $1 + L(n)$ is also at least the smaller of these two bounds. Therefore

$$m \geq \min\{n - 1, \lceil n/2 \rceil + 2\}.$$

Both bounds are attainable. For $n - 1$ copies, private first-layer gates use the physical x_i wire as target and every other variable as a negative control, for $i = 1, \ldots, n - 1$:

$$z_i = x_i \oplus \mathsf{NOR}(x_{-i}).$$

One second-layer NOR accepts exactly when all $z_i = 0$, which is precisely weight one. For the other bound, use a full star centered at variable 1 and a balanced tournament on the other labels. At each remaining vertex, two physical copies $a_i = b_i = x_i$ and negative controls on its outneighbors give, in one destructive Toffoli,

$$a_i' = x_i \oplus [x_i \wedge \mathsf{NOR}(x_{N_i^+})] = x_i \wedge \mathsf{OR}(x_{N_i^+}).$$

NOR the full-star flag and all collision flags. Every other variable uses $3 + \text{indeg}(i) \leq \lceil n/2 \rceil + 2$ copies; variable 1 uses just its occurrence in the full star. All occurrences are physically distinct. The case $n = 2$ is attained by the first construction and needs at least one input copy. Complementing every supplied input wire exchanges EX_1 and EX_{n-1} without changing the number of copies or entangling layers.

A.3 Thresholds and bands

For a threshold with $3 \leq t \leq n - 2$, first orient F to accept weights below t. Any factor rejecting a weight-t input contains its t one-labels: flipping an absent one would give a certainty input with the same source state. In particular its number of present labels is at least three. For every present i, choose another present j. The two local words $0, e_j$ and the two local words $e_i, e_i + e_j$ all extend with outside zeros to certainty inputs, since their weights are at most two. They force both diagonal defect entries for i to vanish. Thus the defect is a rectangle fixing at least t ones, and contains

only one weight-t point. The boundary points containing any fixed variable i charge $\binom{n-1}{t-1}$ different sources containing i. In the other affine orientation, complement all inputs: the certainty region becomes weights below $t' = n - t + 1$, where again $3 \le t' \le n-2$, and $\binom{n-1}{t'-1} = \binom{n-1}{t-1}$. A first layer of all t-subset AND flags followed by OR attains this count; each variable is in exactly that many subsets.

For $t = 2$ and $n \ge 3$, the orientation with certainty on zero and one-hot inputs permits collision stars but forbids zero-rejecting full stars. The same edge argument gives $L(n)$. For $n \ge 5$, a balanced tournament with two copies at each center attains this bound: its flags are $x_i \wedge \mathsf{OR}(x_{N_i^+})$, and their second-layer OR is TH_2. For $n \le 4$, use one AND flag for each pair and OR the flags, costing $n-1$ copies per variable.

For the opposite orientation and $n \ge 4$, consider a factor rejecting the one-hot word e_i. It must contain all $n-1$ zero-labels of that word, since changing an absent zero to one would produce a certainty word of weight two. Its number of present labels is therefore $k \in \{n-1, n\}$. If $k = n-1$, an outside one shows that every local word of weight at least one is certain. For each present bit, there are two such certainty words with that bit zero, and two with that bit one, because $k \ge 3$. If $k = n$, all local words of weight at least two are certain; with $n \ge 4$ the same two-word property holds for either value of any bit. The rank argument therefore forces a rectangle defect in either case. This rectangle is contained in weights zero and one, and any weight-one point it contains requires at least $n-1$ fixed zeros. For a fixed variable j, all $n-1$ points e_i with $i \ne j$ charge different sources containing its zero occurrence. Thus this orientation needs at least $n-1$ copies. At $n = 3$, input complementation exchanges the two orientations; at $n = 2$ the threshold is a single AND. These observations, and $L(n) = \min\{n-1, \lceil (n-1)/2 \rceil + 2\}$ for $n \ge 2$, prove the endpoint threshold formula. Complementing inputs and the output also handles TH_{n-1}.

For the interior band $[a, b]$, first let F accept the band. A factor rejecting a weight-$(b+1)$ input is certain on slice b, so the interior-equality argument with $j = b$ proves it is a rectangle. The complemented argument with $j = a$ applies to factors rejecting weight $a-1$. A rectangle avoiding the whole band cannot meet both sides. A lower-side rectangle meeting weight $a-1$ fixes all $n-a+1$ zero positions of that unique boundary word. Thus the $\binom{n-1}{a-1}$ lower boundary words with $x_i = 0$ charge that many copies of variable i. Upper boundary words of weight $b+1$ with $x_i = 1$ require $\binom{n-1}{b}$ further copies.

In the complementary orientation, certainty holds on global weights $0, 1, n-1, n$. A source on at most one label is trivial. In a source on at least two labels, for any present i choose another present j. The local words $0, e_j$ extend to the first two certainty weights, whereas the local all-ones word and the word with only j changed to zero extend to the last two. These two pairs force both diagonal defect entries for i to vanish. Every nontrivial factor consequently has rectangle defect. The rectangle is contained in the band and fixes at least a ones and $n-b$ zeros. Charge weight-a boundary points with $x_i = 1$ and weight-b boundary points with $x_i = 0$. There are respectively $\binom{n-1}{a-1}$ and $\binom{n-1}{b}$ such points. A rectangle can meet both boundaries, but cannot receive both charges for the *same* variable: the first charge requires its ith bit fixed to one, and the second requires it fixed to zero. Within either boundary it covers at most one point. This proves the same sum in both orientations. A first layer computing all $(b+1)$-ones violations and all $(n-a+1)$-zeros violations, followed by NOR, attains it. Every occurrence is assigned a private supplied wire, and each variable occurs exactly $\binom{n-1}{b} + \binom{n-1}{a-1}$ times.

These exact minima do not imply stable bounded-error bounds: certainty, not a small failure probability, drove the rank-one defect collapse. Nor may the fresh-output restriction be silently dropped.

B An all-depth lower bound with an explicit workspace condition

There is a useful all-depth consequence if ancillary width is charged. This is a physical-operator approximation argument, with an important predecessor in the operator-degree theorem of Dong, Ou, and Yao [21, Theorem 3.1]. The proof below gives its own approximation estimates and constants rather than invoking that theorem's parameter range; no novelty is claimed for the exponent $1-2^{-d}$. All norms in this appendix are operator norms, and all logarithms are natural. The *physical Pauli degree* of an operator on N qubits is the largest number of nonidentity tensor factors in a Pauli string with nonzero coefficient in its Pauli expansion; scalar multiples of the identity have degree zero. Logical degree, used after restricting to the copied inputs, is the degree of the resulting multilinear polynomial on $\{0,1\}^n$.

Theorem B.1 (Physical width and copied inputs)**.** *Let $n, m \geq 1$ and $a \geq 0$ be integers. Let a circuit in the gate model of Section 2, with $d \geq 1$ entangling layers and $M \geq 0$ entangling gates, act on $N = nm + a$ qubits. In this appendix only, its ancillary state τ may be any fixed input-independent density operator, including a mixed or internally entangled state. Let A denote a Hermitian contraction on the output qubit, tensored with the identity on all other qubits, and write*

$$h(x) = \mathrm{Tr}\Big[AU\big((|x\rangle\langle x|)^{\otimes m} \otimes \tau\big)U^\dagger\Big].$$

If, for $0 < \gamma \leq 1$,

$$\left|2^{-n} \sum_{x\in\{0,1\}^n} (-1)^{|x|} h(x)\right| \geq \gamma,$$

then

$$N \geq \frac{n^{2^d/(2^d-1)}}{36\log^2(256\max\{1,M\}/\gamma)}.$$

Under the additional premise $N \leq Cnm$ for $C \geq 1$, this implies

$$m \geq \frac{n^{1/(2^d-1)}}{36C\log^2(256\max\{1,M\}/\gamma)}.$$

The second inequality is asserted only under its relative-workspace premise.

Proof. We first construct an approximation for any desired accuracy $0 < \delta \leq 1$. Put $\overline{M} = \max\{1, M\}$, $\eta = \delta/(64\overline{M})$, and $L = \log(2/\eta)$. Thus $0 < \eta \leq 1/64$ and $L > 1$. Absorbing the local basis changes into the one-qubit layers, each entangling gate is a reflection about a normalized product vector. For its product projector P and a Hermitian contraction Q satisfying $\|Q - P\| \leq \eta$, put

$$\mathcal{T}_Q(A) = A - 2[Q,[Q,A]].$$

This map fixes the identity exactly and approximates conjugation by $R = I - 2P$:

$$\|\mathcal{T}_Q(A) - RAR\| \leq 16\eta\|A\|, \qquad \|\mathcal{T}_Q(A)\| \leq (1+16\eta)\|A\|.$$

Since $P^2 = P$, conjugation by R is exactly $\mathcal{T}_P$. Subtract the two double commutators as $[Q-P,[Q,A]] + [P,[Q-P,A]]$ and use $\|[X,Y]\| \leq 2\|X\|\|Y\|$. The proof also holds after tensoring with an identity on any environment. The approximation map need not be positive or physically implementable.

For an arity-b product projector $P = \bigotimes_{i=1}^b P_i$, each P_i is a rank-one qubit projector. Regard these projectors as operators on the whole gate support, and write $K = \sum_{i=1}^b (I - P_i)$. The summands

commute, K has spectrum $0, \ldots, b$, and P is its zero-eigenspace projector. For $b \geq 2$, let T_r be the degree-r Chebyshev polynomial of the first kind. The polynomial

$$q_r(K) = \frac{T_r((b+1-2K)/(b-1))}{T_r((b+1)/(b-1))}$$

equals one at zero and has magnitude at most $2e^{-2r/\sqrt{b}}$ at every other eigenvalue. To see this, use $|T_r(t)| \leq 1$ on $[-1,1]$ for the numerator, and $T_r(\cosh\theta) = \cosh(r\theta) \geq e^{r\theta}/2$ for the denominator, with $\theta = \operatorname{arcosh}((b+1)/(b-1)) = 2\operatorname{artanh}(1/\sqrt{b}) \geq 2/\sqrt{b}$. Taking $r = \lceil \sqrt{b}L/2 \rceil$, use $Q = q_r(K)$ if $r < b$, and $Q = P$ otherwise. Since K has physical Pauli degree at most one, $q_r(K)$ has degree at most r. Its eigenvalue at zero is one and all other eigenvalues have magnitude at most $\eta < 1$, so Q is a Hermitian contraction with

$$\|Q - P\| \leq \eta, \qquad \deg Q \leq 2\sqrt{b}\log(2/\eta).$$

Here $\deg Q \leq \min\{b, r\} \leq \sqrt{b}L/2 + 1 \leq 2\sqrt{b}L$, using $L > 1$. The arity-one case is exact and obeys the same bound.

Fix $1 \leq \ell \leq N$. A Pauli monomial of degree at most ℓ touches at most ℓ gates in a disjoint layer. Gates not touched fix it exactly, because $\mathcal{T}_Q(I) = I$; another disjoint gate cannot activate them. Products of operators have Pauli degree at most the sum of their degrees. The expansion

$$\mathcal{T}_Q(A) = A - 2Q^2A - 2AQ^2 + 4QAQ$$

therefore bounds the new degree by

$$\ell + 4L \sum_{g \text{ touched}} \sqrt{b_g} \leq \ell + 4L\sqrt{N\ell} \leq 6L\sqrt{N\ell}.$$

The first inequality uses disjoint gate supports, so $\sum_g b_g \leq N$, and Cauchy–Schwarz. The last uses $\ell \leq N$ and $L > 1$. Linearity extends the estimate to all degree-at-most-ℓ operators; scalar operators are fixed exactly. No bound on the sum of their Pauli coefficients is used in the norm estimate.

Exact unitary conjugations have induced operator norm one; each approximate map has norm at most $1+16\eta$ and differs from its exact conjugation by at most 16η in that induced norm. Telescoping all M approximate conjugations in the backward evolution of an output contraction therefore gives norm error at most

$$(1+16\eta)^M - 1 \leq e^{\delta/4} - 1 \leq \delta.$$

The first bound is zero when $M = 0$. The last inequality follows from convexity, $e^{\delta/4} - 1 \leq \delta(e^{1/4} - 1) < \delta$. Local changes of basis preserve Pauli degree, and all the maps preserve Hermiticity. Starting at degree at most one, iteration over d layers sums the geometric series $1+1/2+\cdots+2^{-(d-1)} = 2(1-2^{-d})$. It yields a Hermitian approximation B with

$$\|U^*AU - B\| \leq \delta, \qquad \deg B \leq (6L)^{2\alpha}N^{\alpha}, \quad \alpha = 1 - 2^{-d}, \quad L = \log(128\max\{1, M\}/\delta). \tag{18}$$

Intermediate degrees can be capped at N without weakening this bound.

Define $p_B(x) = \operatorname{Tr}[B((|x\rangle\langle x|)^{\otimes m} \otimes \tau)]$. An input X or Y Pauli factor has zero diagonal expectation. The surviving Z factors reduce modulo two separately for each repeated logical label: they give characters $\prod_{i \in S}(-1)^{x_i}$ with $|S| \leq \deg B$. Ancillary expectation only multiplies each character by a scalar, even for internally entangled ancillas. Thus p_B is a real logical polynomial of

degree at most $\deg B$, and the norm estimate implies $|h(x) - p_B(x)| \le \delta$ for every x. If its degree were below n, character orthogonality would give

$$2^{-n} \sum_x (-1)^{|x|} p_B(x) = 0.$$

Taking $\delta = \gamma/2$ would then bound the parity correlation of h by $\gamma/2$, a contradiction. Thus $(6L)^{2\alpha} N^\alpha \ge n$, with $L = \log(256 \max\{1, M\}/\gamma)$. Raising to the power $1/\alpha$ gives $36L^2 N \ge n^{1/\alpha}$, which is the first inequality. Dividing by Cn under the width premise gives the second. □

The same argument applies to approximate degree. Let $\deg_\varepsilon(f)$ denote the minimum degree of a real polynomial approximating f uniformly on the Boolean cube within ε. For a symmetric nonconstant function with transition radius ρ as in Section 2, Paturi's theorem gives $\deg_{3/8}(f) = \Theta(\sqrt{n\rho})$ [22, Theorem 1]. Indeed, his parameter is $\Gamma(f) = \min_{f_i \ne f_{i+1}} |2i - n + 1|$, and

$$n - |2i - n + 1| = 2\min\{i + 1, n - i\} - 1, \qquad n - \Gamma(f) = 2\rho(f) - 1.$$

For nonconstant f, $\rho \ge 1$, so $2\rho - 1$ and ρ differ by at most a factor of two. Paturi also explicitly permits replacing error $1/3$ by any fixed error below one half. For a circuit with worst-case error at most $1/3$, apply (18) to its acceptance-probability observable with $\delta = 1/24$. The induced logical polynomial approximates f within $3/8$. Consequently

$$N = \Omega\left(\frac{(n\rho)^{2^{d-1}/(2^d-1)}}{\log^2(M+2)} \right).$$

With $N \le Cnm$, divide this bound by Cn to obtain a supplied-copy lower bound. For central-transition functions, fixed depth, polynomial size, and fixed C, it is

$$m = \Omega\left(\frac{n^{1/(2^d-1)}}{\log^2 n} \right).$$

This does not exclude polynomial-copy constant-depth computation. Unrestricted polynomial ancillary workspace can invalidate the premise $N = O(nm)$, so these inequalities do not close the full-model gap.

Use of AI: We prepared a list of problems and proof strategies and techniques related to Copy-complexity (how many copies of a classical input make shallow quantum computation powerful). We fed this to multiple AI tools (Claude, ChatGPT and two open-weight models with open-source harnesses) and prompted these AI tools to improve the existing bounds. When an AI tool was hallucinating, stuck, or going in the wrong direction, we stopped that AI tool, corrected the intermediate `.md` files, added more specific new ideas and switched to a different AI tool with more detailed fresh prompts. Equations written by one AI tool were critiqued, verified and expanded by a completely different AI tool in an iterative manner, till we were satisfied with their correctness and details. The author is solely responsible for the correctness of the proofs.

References

[1] Cristopher Moore. *Quantum Circuits: Fanout, Parity, and Counting*. arXiv:quant-ph/9903046, 1999. https://arxiv.org/abs/quant-ph/9903046.

[2] Frederic Green, Steven Homer, Cristopher Moore, and Christopher Pollett. *Counting, Fanout, and the Complexity of Quantum ACC*. Quantum Information and Computation 2(1), 35–65, 2002. https://arxiv.org/abs/quant-ph/0106017.

[3] Peter Høyer and Robert Špalek. *Quantum Fan-Out Is Powerful*. Theory of Computing 1, 81–103, 2005. https://theoryofcomputing.org/articles/v001a005/v001a005.pdf.

[4] Yasuhiro Takahashi and Seiichiro Tani. *Collapse of the Hierarchy of Constant-Depth Exact Quantum Circuits*. Computational Complexity 25(4), 849–881, 2016. https://arxiv.org/abs/1112.6063.

[5] Maosen Fang, Stephen Fenner, Frederic Green, Steven Homer, and Yong Zhang. *Quantum Lower Bounds for Fanout*. Quantum Information and Computation 6(1), 46–57, 2006. https://arxiv.org/abs/quant-ph/0312208.

[6] Daniel Grier and Jackson Morris. *Quantum Threshold Is Powerful*. Proceedings of the 40th Computational Complexity Conference, 2025. https://arxiv.org/abs/2411.04953.

[7] Malvika Raj Joshi and Francisca Vasconcelos. *Constant-Depth Unitary Preparation of Dicke States*. arXiv:2601.10693v3, 2026. https://arxiv.org/abs/2601.10693.

[8] Bettina Brustmann and Ingo Wegener. *The Complexity of Symmetric Functions in Bounded-Depth Circuits*. Information Processing Letters 25(4), 217–219, 1987. https://doi.org/10.1016/0020-0190(87)90163-390163-3).

[9] Shlomo Moran. *Generalized Lower Bounds Derived from Håstad's Main Lemma*. Information Processing Letters 25(6), 383–388, 1987.

[10] Johan Håstad, Ingo Wegener, Norbert Wurm, and Sang-Zin Yi. *Optimal Depth, Very Small Size Circuits for Symmetrical Functions in AC^0*. Information and Computation 108(2), 200–211, 1994. https://doi.org/10.1006/inco.1994.1008.

[11] Daniel Grier, Jackson Morris, and Kewen Wu. *QAC^0 Contains TC^0 (with Many Copies of the Input)*. arXiv:2601.03243v1, 2026. https://arxiv.org/abs/2601.03243.

[12] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. *Quantum Amplitude Amplification and Estimation*. Contemporary Mathematics 305, 53–74, 2002. https://arxiv.org/abs/quant-ph/0005055.

[13] Peter Høyer. *On Arbitrary Phases in Quantum Amplitude Amplification*. Physical Review A 62, 052304, 2000. https://arxiv.org/abs/quant-ph/0006031.

[14] Gregory Rosenthal. *Bounds on the QAC^0 Complexity of Approximating Parity*. Innovations in Theoretical Computer Science, 2021. https://arxiv.org/abs/2008.07470.

[15] Stephen Fenner, Daniel Grier, Daniel Padé, and Thomas Thierauf. *Tight Bounds on Depth-2 QAC-Circuits Computing Parity*. arXiv:2504.06433, 2025. https://arxiv.org/abs/2504.06433.

[16] Shivam Nadimpalli, Natalie Parham, Francisca Vasconcelos, and Henry Yuen. *On the Pauli Spectrum of QAC^0*. Proceedings of the 56th Annual ACM Symposium on Theory of Computing, 1498–1506, 2024. https://arxiv.org/abs/2311.09631.

[17] Anurag Anshu, Yangjing Dong, Fengning Ou, and Penghui Yao. *On the Computational Power of QAC^0 with Barely Superlinear Ancillae*. arXiv:2410.06499v4, 2025. https://arxiv.org/abs/2410.06499.

[18] Malvika Raj Joshi, Avishay Tal, Francisca Vasconcelos, and John Wright. *Improved Lower Bounds for QAC^0*. Proceedings of the 58th ACM Symposium on Theory of Computing, 2026. https://doi.org/10.1145/3798129.3800922. Full author version: https://arxiv.org/abs/2512.14643v4, August 2026.

[19] Lucas Gretta, Meghal Gupta, and Malvika Raj Joshi. *Parity $\notin$ QAC^0 $\iff$ QAC^0 Is Fourier-Concentrated*. arXiv:2604.02793v2, 2026. https://arxiv.org/abs/2604.02793.

[20] Boyan Xu and Lvzhou Li. *Fanout Complexity of Symmetric Boolean Functions in QAC^0*. arXiv:2609.05153v1, 2026. https://arxiv.org/abs/2609.05153.

[21] Yangjing Dong, Fengning Ou, and Penghui Yao. *Linear-Size QAC^0 Channels: Learning, Testing and Hardness*. arXiv:2510.00593v2, 2025. https://arxiv.org/abs/2510.00593.

[22] Ramamohan Paturi. *On the Degree of Polynomials that Approximate Symmetric Boolean Functions*. Proceedings of the 24th ACM Symposium on Theory of Computing, 468–474, 1992. https://cseweb.ucsd.edu/~paturi/myPapers/pubs/Paturi_1992_stoc.pdf.

[23] Roman Smolensky. *Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity*. Proceedings of the 19th ACM Symposium on Theory of Computing, 77–82, 1987. https://people.cs.umass.edu/~immerman/cs601/Smolensky.pdf.