

Parity in Shallow QAC Circuits: Correlation Decay and Exact Lower Bounds

Shiva Kintali*

September 5, 2026

Abstract

We prove two exponential bounds for single-output parity computation by shallow QAC circuits. At entangling depth at most three, uniform-input parity correlation is at most $Ke^{-\kappa n}$, independently of circuit size and ancillary width. At entangling depth at most four, exact parity requires at least e^{cn} multiqubit gates for all sufficiently large n . The constants are absolute. The model allows arbitrary one-qubit gates, arbitrary-arity generalized Toffoli gates, any finite number of input-independent product-initialized ancillas, including mixed product states, and unrestricted final garbage outside one measured qubit.

1 Introduction

Parity asks whether an odd number of input bits are one. Its full dependence on the input makes it a basic test of how information passes through a shallow circuit. Classical constant-depth circuits with unbounded-fan-in AND and OR gates and NOT gates cannot compute parity with polynomial size; switching lemmas and Fourier concentration explain this limitation [9, 13]. For quantum circuits, interference and unbounded-arity gates make an ordinary bounded-fan-in light-cone argument insufficient.

We study circuits with one computational-basis copy of each input bit, generalized Toffoli gates, and arbitrary one-qubit unitaries. Depth counts layers of multiqubit gates with pairwise disjoint supports, including controls; one-qubit gates contribute neither to depth nor to size. Throughout, depth d means entangling depth at most d . Ancillas start in a fixed product state independent of the input. The numbers of gates and ancillary wires may be arbitrarily large but finite. Only one measured output qubit must report the answer: the inputs may be overwritten and all other outputs may contain garbage. There are no intermediate measurements, postselection, additional supplied input copies, or free fan-out gates.

Moore introduced the QAC^0 model and related coherent parity, fan-out, and cat-state preparation [15]; see also [7, 3]. With fan-out gates, constant-depth quantum circuits admit approximate counting and arithmetic constructions [11] and exact threshold constructions [18]. Here QAC denotes the gate-set model without a polynomial resource restriction, while QAC^0 denotes constant-depth families with polynomially many wires and gates. The distinction matters: one of our conclusions rules out fixed parity advantage at depth three regardless of finite circuit size, whereas the other gives an exponential gate requirement at depth four.

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

1.1 Two main results

Let $\chi(x) = (-1)^{x_1 + \dots + x_n}$ for $x \in \{0, 1\}^n$, $n \geq 1$. A fixed binary measurement on the designated output qubit has effects $E_+, E_- \geq 0$ with $E_+ + E_- = I$. Label its outcomes $+1$ for even parity and -1 for odd parity, and set $A_o = E_+ - E_-$. For a circuit C with full output state ρ_x , define

$$c(C, A_o) = \mathbb{E}_{x \in \{0, 1\}^n} \chi(x) \text{Tr}(\rho_x A_o).$$

Here the average is uniform and A_o is tensored with the identity on all other wires. Since $-I \leq A_o \leq I$, it is a Hermitian contraction. The correct-outcome probability on input x is $[1 + \chi(x) \text{Tr}(\rho_x A_o)]/2$, so the uniform-input success probability is $(1 + c(C, A_o))/2$. Our main results are:

- (i) *Exponential correlation decay at depth three.* There are absolute $K_0 \geq 1$ and $\kappa > 0$ such that

$$|c(C, A_o)| \leq 6K_0 e^{-\kappa n}$$

for every depth-three circuit, independently of gate count, ancillary width, and gate arity (Theorem 1). Thus no fixed positive uniform-input advantage, and hence no fixed worst-case advantage, is possible for sufficiently large n . A zero-ancilla construction has correlation $\exp(-\Theta(n))$, showing that the exponential order is attainable (Proposition 14); this does not assert a matching constant in the exponent or a fixed worst-case advantage.

- (ii) *Exponential gate growth for exact depth four.* There are absolute $c_* > 0$ and n_0 such that every exact depth-four parity circuit on $n \geq n_0$ inputs has at least $e^{c_* n}$ multiqubit gates (Theorem 2). The ancillary width and gate arities remain unrestricted and finite.

Both statements permit general binary measurements on any designated qubit, including an input wire, and arbitrary final garbage. They also hold for mixed product ancillary states. They do not allow arbitrary entangled initial advice, and they do not resolve parity at every constant depth. The depth-four theorem concerns exact computation; it does not exclude all bounded-error depth-four families.

Explicit constants for the principal bounds are recorded in Appendix A.

1.2 A common estimate and two circuit applications

For the proof overview, take pure product initialization; the mixed-product case follows by Lemma 3. A *one-input source* is a first-layer gate component, or an untouched input singleton, containing exactly one variable input. All its other wires start in fixed pure product states. Its *ports* are its nonempty intersections with second-layer gates, together with one additional port of idle wires if any remain. Let ρ_0, ρ_1 be the source states after the first layer. Writing $F = (\rho_0 - \rho_1)/2$ and $\bar{\rho} = (\rho_0 + \rho_1)/2$, we assign weight $h_j = 1 - \lambda_{\max}(\bar{\rho}_j)$ to port j , where $\bar{\rho}_j$ is the reduced average state on that port. Lemma 4 gives the budget $\sum_j h_j \leq 1/2 + 2/e$, independently of physical width.

For r independent one-input sources, Theorem 11 bounds the absolute full-parity coefficient of any *port-product effect* $E = \bigotimes_j E_j$, $0 \leq E_j \leq I_j$, after the second layer by $K_0 e^{-ar}$, for absolute constants $K_0 \geq 1$ and $a > 0$. An input-independent resource operator of arbitrary finite dimension and trace norm at most one may be tensored with the sources; its ports are included in the product defining E . In expansion branches this operator can be non-Hermitian, so the resulting functional is interpreted as a possibly complex trace pairing. Each effect E_j may act jointly on all physical qubits of its port. This generality is needed for the depth-four certificate.

The proof calls a second-layer gate *heavy* when the sum of the source weights incident to it exceeds a fixed threshold. Heavy gates are treated by an exact reflection expansion. Several corrections can touch different ports of one source; they must be treated as one joint insertion. A centered rank-one inequality transfers the source's upper bound on the squared trace norm of F

after random port retention to this insertion, using the original weights. The remaining light gates and private ports receive independent random choices; sources meeting the same light gate share its choice. Bounded differences controls the resulting dependence between sources. A convex estimate for the loss from a joint insertion controls terms with many heavy-gate corrections. Summing separately over terms with few and many corrections gives exponential decay.

Two arguments then connect this estimate to parity. For depth three, a rank-two correction handles the signed state of a first-layer gate containing many input bits. Combining this expansion with conditioning on the smaller components gives an exponential bound for events that factor over individual physical qubits, with arbitrary first layers (Theorem 12). Pulling back one output observable through the third layer expresses it as a linear combination of such events with total absolute coefficient mass at most six (Lemma 13) and proves the correlation theorem.

For exact depth four, the cleanup in Lemma 15 uses classical and quantum restrictions to retain $r \geq n/3$ variable inputs with one variable per source, without increasing depth or gate count. Output determinism yields a certificate involving the sector on which the final output gate acts as the identity. Label each control by its input source or the grouped input-free resource, and by its gate-or-idle incidence in layers two and three. The resulting cells number at most $(r+1)(m_2+1)(m_3+1)$, where m_2, m_3 count gates in those two layers (Lemma 18). Each cell lies in one two-layer port, so its effect may be treated jointly within that port. The number of cells is polynomial in the retained input count and gate count and is independent of ancillary width. Applying the common event bound to these cells forces exponentially many gates. Exactness is used in the depth-four argument.

1.3 Related work and attribution

Ancilla-restricted lower bounds at general depths. Fang, Fenner, Green, Homer, and Zhang [5] proved logarithmic depth lower bounds for clean exact implementations of coherent parity and fan-out with constantly many ancillas, together with a depth-ancilla tradeoff. Here clean means that the ancillary registers are returned to their initial zero state. Their convention counts one-qubit gates in the depth; the logarithmic conclusion is unchanged by the constant-factor conversion to our convention. Bera [2] gave a different lower-bound method excluding ancilla-free constant-depth exact parity circuits.

Nadimpalli, Parham, Vasconcelos, and Yuen [16] developed Pauli-spectrum concentration, in the spirit of the classical Fourier-analytic approach. For polynomial-size circuits of fixed depth d , their Theorem 33 implies uniform-input parity success at most $1/2 + 2^{-\Omega(n^{1/d})}$ when the ancilla count is a sufficiently small constant multiple of $n^{1/d}$. Anshu, Dong, Ou, and Yao [1] used spectral-norm approximation by low-degree operators to obtain the stated ancilla threshold $\tilde{\Omega}(n^{1+3^{-d}})$ for parity. Dong, Ou, and Yao [4] improved the corresponding parity threshold to $\tilde{\Omega}(n^{1+2^{-d}})$ using operator and unitary dilation; in the small-ancilla regime of their Corollary 3.3, uniform-input success is at most $1/2 + O(d/n)$. The tilde suppresses logarithmic factors in these resource thresholds. At depth four, the two stated superlinear thresholds have exponents $82/81$ and $17/16$. These three spectral methods apply across arbitrary constant depths and allow arbitrary input-independent ancillary states. Our removal of any ancilla-count restriction must instead be read together with the product-initialization assumption.

The quoted exponents are the papers' simplified corollary thresholds, not optimal inversions of their estimates. At fixed depth $d \geq 1$, the cited estimates give a polynomial that uniformly approximates the single-output acceptance probability and has degree $\tilde{O}(N^{1-\alpha})$, where N is the total number of wires and $\alpha = 3^{-d}$ or 2^{-d} , respectively. For any fixed positive success advantage, they therefore imply an ancillary requirement $\tilde{\Omega}(n^{1/(1-\alpha)})$; see [1, Theorem 4.1] and [4, Theorem 3.1]. Indeed, let $p(x)$ be the probability of outputting 1 (odd parity). Since $\mathbb{E}_x \chi(x) = 0$, uniform-input

success at least $1/2 + \delta$ gives $\mathbb{E}_x \chi(x)p(x) \leq -\delta$. Every polynomial of total degree less than n is orthogonal to χ on the uniform Boolean cube: each of its monomials omits some variable, and averaging over that variable cancels its contribution. Thus, for any such polynomial $\tilde{p}$,

$$\delta \leq |\mathbb{E}_x \chi(x)(p(x) - \tilde{p}(x))| \leq \|p - \tilde{p}\|_\infty,$$

where the norm is the maximum over $\{0, 1\}^n$. Consequently, a uniform approximation with error less than δ requires degree at least n . The parameter restriction in [4, Theorem 3.1] permits the required choice of its accuracy parameter of order $\log N$ at fixed depth and initial operator degree one, for sufficiently large N . The approximation error in [1, Theorem 4.1] also tends to zero at fixed depth. Inverting $n \leq \tilde{O}(N^{1-\alpha})$ yields $N \geq \tilde{\Omega}(n^{1/(1-\alpha)})$. Since $1/(1-\alpha) > 1$, subtracting the n input wires preserves this asymptotic lower bound on the ancillary count. Thus the directly derived depth-four exponents are $81/80$ and $16/15$, respectively.

Fixed-depth lower bounds and approximate constructions. Rosenthal [17, Corollary 1.6 and Theorem 1.7] studied the coherent parity operation, which adds parity into a target and preserves the other input wires, allowing final garbage. He obtained an $\Omega(n/(d+1))$ multiqubit-gate requirement for worst-case coherent approximation with squared-overlap fidelity exceeding $1/2$ by a fixed positive amount. At depth two, fidelity on a specified product quantum input is at most $1/2 + \exp(-\Omega(n))$, independently of size. These are coherent fidelity statements, not average-case Boolean-success bounds; allowing garbage does not remove input preservation.

His Corollary 1.2 also gives, for every $d \geq 7$ and $0 < \varepsilon < 1$, circuits of size and ancilla count $\exp(\text{poly}(n^{1/d} \log(n/\varepsilon)))$ whose output has squared Euclidean distance at most ε from the ideal coherent parity output tensored with an all-zero ancillary register, on every pure quantum input. This phase-sensitive state-vector guarantee permits a small deviation of the actual ancillas from zero. Such greater-depth approximate constructions are consistent with an exact depth-four lower bound.

Fenner, Grier, Padé, and Thierauf [6] proved that exact single-output parity on more than three non-target input bits is impossible at depth two, regardless of gate count or the number of product-initialized ancillas. Their proof uses irreducible polynomials and entanglement. Joshi, Tal, Vasconcelos, and Wright [12] proved exact depth-three impossibility for $n > 100$, again independently of size and the number of product-initialized ancillas, and a depth-two single-output parity-correlation bound $\exp(-\Omega(\sqrt{n}))$. Their exact analysis uses quantum restrictions and classical AC^0 simulation.

Gretta, Gupta, and Joshi [8, Theorems 1.1 and 1.3] relate parity lower bounds to Fourier concentration and show that inverse-polynomial average-case parity correlation suffices for exact parity computation under QAC^0 reductions. These results concern constant-depth classes with a change in depth; they do not preserve the specific depth-four restriction studied here.

For exact parity with unrestricted product ancillary width, Theorem 2 reaches depth four with an exponential gate lower bound. The depth-three result of [12] excludes circuits of every size at that smaller depth. Compared with general-depth ancilla bounds, it removes any numerical limit on product-initialized ancillary space at this fixed depth. Compared with Rosenthal’s general-depth gate bound, it gives an exponential lower bound for exact computation and applies to the more permissive single-output task. It neither excludes exponentially large depth-four circuits nor establishes a lower bound at every constant depth.

Theorem 1 supplies a different improvement: it bounds average-case parity correlation at depth three by $\exp(-\Omega(n))$ without a size or product-ancilla-width restriction. The exact depth-three obstruction and the depth-two correlation result cited above do not themselves give this conclusion.

Inherited ingredients. Product-reflection normalization is used in [17]. Our exact cleanup reproduces the mechanism of [12, Lemma 4.1], including the preparation of the fixed-parity pairs and

gate accounting. The final-gate active/identity-sector decomposition appears in [12, Section 2.3.1]. The additional depth-four step is the quantitative certificate and its grouping into control cells. Bounded differences and the bounded-range exponential-moment inequality are classical tools [14, 10]. We also use standard trace-norm duality, Schatten Hölder, convexity, and spectral decomposition. The contributions are the source and joint-insertion estimates, their network summation, and the circuit interfaces; no historical priority claim is made for general matrix or concentration inequalities.

Organization. Section 2 fixes the model and states the main theorems. Sections 3 and 4 prove the shared source and event estimates. Section 5 proves the depth-three theorem and gives the matching exponential-order construction. Section 6 proves the exact depth-four theorem. The appendices give explicit constants and a self-contained proof of the concentration inequality used in the main argument.

2 Circuit model, notation, and statements

Registers, gates, and resources. Fix integers $N \geq n \geq 1$. A circuit acts on N qubits, with one copy of the classical input $|x\rangle$, $x \in \{0, 1\}^n$, and a pure ancillary state

$$|\alpha\rangle = \bigotimes_{q=n+1}^N |\alpha_q\rangle,$$

where each α_q is a unit vector in $\mathbb{C}^2$. An absent register is the one-dimensional space $\mathbb{C}$, and an empty tensor product is the scalar 1; this includes $N = n$. We initially work with pure product initialization; Lemma 3 below extends the conclusions to mixed product states. Write $\rho_x^{\text{in}} = |x\rangle\langle x| \otimes |\alpha\rangle\langle\alpha|$. For mixed product initialization, replace $|\alpha\rangle\langle\alpha|$ by $\tau = \bigotimes_{q=n+1}^N \tau_q$, where each τ_q is a single-qubit density operator: $\tau_q \geq 0$ and $\text{Tr } \tau_q = 1$. The circuit, ancillary state, output wire, and measurement are fixed independently of x . There are no intermediate measurements or postselection. Pure product ancillas may be taken to start in $|0\rangle$ on every ancillary wire, because preparing their specified states requires only one-qubit unitaries.

A generalized controlled- Z gate on a set S of at least two qubits is

$$\text{CZ} * S = I_S - 2|1^S\rangle\langle 1^S|, \quad |1^S\rangle = \bigotimes_{q \in S} |1\rangle_q.$$

For any target $t \in S$, conjugating this gate by the Hadamard on t gives the generalized Toffoli with controls $S \setminus \{t\}$:

$$H_t \text{CZ} * S H_t : |z\rangle|b\rangle \mapsto |z\rangle|b \oplus \bigwedge_{q \in S \setminus \{t\}} z_q\rangle.$$

Here $z \in \{0, 1\}^{S \setminus \{t\}}$ labels the control register and $b \in \{0, 1\}$ labels the target; their displayed order is only a notation for the tensor factors. The same conjugation gives the reverse conversion. Thus these two gate sets are equivalent once one-qubit unitaries are free, as in the depth convention below.

For an integer $d \geq 0$, a circuit has *entangling depth at most d* if it can be written with d layers of controlled- Z gates, some possibly empty, with pairwise disjoint gate supports within each layer, including all control and target wires. Arbitrary products of one-qubit unitaries may occur before, between, and after these layers; they contribute neither to entangling depth nor to size. We use “depth d ” to mean entangling depth at most d . The *size* M is the total number of multiqubit gates. Every circuit has finitely many wires and gates, but we impose no a priori bound on N or on gate arities as functions of n . There is one designated output qubit o , which may be an original input wire. Every other wire may contain arbitrary final garbage; input preservation and ancillary uncomputation are not required.

Reflection normal form. The conversion to product reflections preserves the number and supports of the multiqubit gates, but retains a final local unitary. For $d = 0$ the circuit is already a product of one-qubit unitaries. For $d \geq 1$, write

$$C = L_d D_d L_{d-1} D_{d-1} \cdots L_1 D_1 L_0, \quad L_{\leq j} = L_j L_{j-1} \cdots L_0,$$

where each L_j is a product of one-qubit unitaries and each D_j is a disjoint controlled- Z layer. Then

$$\tilde{D}_j = L_{\leq j-1}^* D_j L_{\leq j-1}, \quad C = L_{\leq d} \tilde{D}_d \tilde{D}_{d-1} \cdots \tilde{D}_1.$$

Indeed, $L_{\leq j} \tilde{D}_j = L_j D_j L_{\leq j-1}$; iterating this identity from $j = d$ down to $j = 1$ gives the displayed factorization. Here and below, $*$ denotes the adjoint. Because $L_{\leq j-1}$ is local, each gate in $\tilde{D}_j$ has the same support S as its original gate and has the form

$$D_S = I_S - 2\Pi_S, \quad \Pi_S = \bigotimes_{q \in S} |\theta_q\rangle\langle\theta_q|, \quad (1)$$

for one-qubit unit vectors θ_q . Conversely, choose one-qubit unitaries V_q with $V_q|1\rangle = |\theta_q\rangle$. For $V_S = \bigotimes_{q \in S} V_q$,

$$D_S = V_S C_Z V_S^*.$$

Disjointness allows these local conjugations to be performed simultaneously within each layer, and neighboring local layers can be merged. This proves equivalence in both directions without changing entangling depth or multiqubit-gate count. The remaining $L_{\leq d}$ may be omitted for output statistics if the one-qubit measurement observable A_o is replaced by $L_{\leq d,o}^* A_o L_{\leq d,o}$, with $L_{\leq d,o}$ its output-qubit factor.

Parity coefficients, events, and norms. All logarithms are natural unless explicitly stated otherwise. Write $[k] = \{1, \dots, k\}$ for integers $k \geq 0$, with $[0] = \emptyset$, and put $\chi(x) = (-1)^{x_1 + \cdots + x_n}$. All input averages are uniform. For $q : \{0, 1\}^n \rightarrow \mathbb{C}$, its full-parity coefficient is

$$\hat{q}([n]) = \mathbb{E}_x \chi(x) q(x) = 2^{-n} \sum_{x \in \{0,1\}^n} \chi(x) q(x).$$

In auxiliary estimates we also use this definition for zero variables: $\{0, 1\}^0$ has one element, its parity sign is 1, and the coefficient is the value of q at that element. For a state ρ_x at a specified circuit stage, write $\langle B \rangle_x = \text{Tr}(\rho_x B)$. In such expectations, an operator specified on only some wires is tensored with identity on the remaining wires.

A *physical-qubit product event*, or simply a *product event* when no other partition is specified, is $E = \bigotimes_q E_q$, where each E_q is an identity or a one-qubit rank-one orthogonal projector. It represents the joint event specified by these local projectors, with no condition on identity factors. Where a later statement explicitly permits $0 \leq E_q \leq I_q$, the term also includes these more general product measurement effects. For a specified partition into blocks, a *port-product effect* is instead a tensor product $\bigotimes_j E_j$ with $0 \leq E_j \leq I_j$ on each block. Such a factor need not be a product over its physical qubits. Section 4 specifies the source-second-layer ports for its event estimate. This must be distinguished from the physical-qubit product-event statement in Section 5.

For a linear map $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ between specified finite-dimensional Hilbert spaces, we use the unnormalized trace, operator, and Hilbert–Schmidt norms, respectively:

$$\|T\|_1 = \text{Tr} \sqrt{T^* T}, \quad \|T\|_\infty = \sup_{\|v\| \leq 1} \|Tv\|, \quad \|T\|_2 = \sqrt{\text{Tr}(T^* T)}.$$

A *contraction* has operator norm at most one. For an operator on a single Hilbert space, a *positive contraction*, or measurement effect, satisfies $0 \leq T \leq I$. In particular, $\|I_{\mathcal{H}}\|_1 = \dim \mathcal{H}$ and $\|I_{\mathcal{H}}\|_2 = \sqrt{\dim \mathcal{H}}$. The identity-extension convention for expectations does not extend the space on which a standalone trace norm or Hilbert–Schmidt norm is evaluated.

Output conventions and the main theorems. Let $\rho_x = C\rho_x^{\text{in}}C^*$ be the full output state. For pure initialization, write $|\Psi_x\rangle = C(|x\rangle \otimes |\alpha\rangle)$, so that $\rho_x = |\Psi_x\rangle\langle\Psi_x|$. A binary measurement on o consists of effects $E_+, E_- \geq 0$ with $E_+ + E_- = I_o$. Label its outcomes $+1$ for even parity and -1 for odd parity, and put

$$A_o = E_+ - E_- = 2E_+ - I_o.$$

Thus A_o is a one-qubit Hermitian contraction and $\text{Tr}(\rho_x A_o)$ is the mean of the signed measurement outcome. Conversely, every one-qubit Hermitian contraction defines such a measurement by $E_{\pm} = (I_o \pm A_o)/2$. For a general binary measurement the outcomes are these labels, not necessarily the eigenvalues of A_o .

We say that C , with its specified measurement A_o , *computes parity exactly* if

$$\text{Tr}(\rho_x A_o) = \chi(x) \quad \text{for every } x \in \{0, 1\}^n. \quad (2)$$

Because $n \geq 1$, both parity values occur. Since $\text{Tr}(\rho_x A_o)$ is a convex combination of the eigenvalues of A_o , weighted by the reduced state on o , attaining each of the extreme expectations $+1$ and -1 forces both values to occur in the spectrum of A_o . Since o is a qubit, A_o is therefore a Hermitian unitary with one eigenvalue of each sign. Its measurement effects are consequently the rank-one spectral projectors $E_{\pm} = (I_o \pm A_o)/2$. A one-qubit change of basis turns this into computational-basis measurement, so allowing general binary measurements does not enlarge the exact computation model. This argument also applies to mixed output states.

For pure initialization in the exact case, let $|a_{\pm}\rangle$ be fixed unit eigenvectors of A_o . For $y = \chi(x)$, the projector $P_y = \frac{I_o + yA_o}{2} \otimes I_{\text{rest}}$ satisfies

$$\|(I - P_y)|\Psi_x\rangle\|^2 = \frac{1 - y \text{Tr}(\rho_x A_o)}{2}.$$

For this Hermitian unitary A_o , exactness is consequently equivalent to

$$|\Psi_x\rangle = |a_{\chi(x)}\rangle * o \otimes |\gamma_x\rangle * \text{rest}$$

for a unit vector γ_x that may depend arbitrarily on x . The purity used here follows from pure initialization and unitary evolution. No preservation of the other registers is asserted.

For a general binary measurement, define the uniform-input parity correlation by

$$c(C, A_o) = \mathbb{E}_x \chi(x) \text{Tr}(\rho_x A_o). \quad (3)$$

This is real and belongs to $[-1, 1]$, since ρ_x is a density operator and A_o is a Hermitian contraction. On input x , the probability of the correct outcome is $[1 + \chi(x) \text{Tr}(\rho_x A_o)]/2$. Hence the average success probability is exactly

$$p_{\text{succ}}(C, A_o) = \frac{1 + c(C, A_o)}{2}.$$

Because every input has positive uniform probability, average success one is equivalent to success on every input. Thus $c(C, A_o) = 1$ if and only if C with measurement A_o computes parity exactly. In particular, uniform-input success at least $1/2 + \varepsilon$ implies $c(C, A_o) \geq 2\varepsilon$. A worst-case success guarantee implies the same uniform-input guarantee, but the converse is not assumed.

Theorem 1 (Exponential correlation at depth three). *There are absolute constants $K_0 \geq 1$ and $\kappa > 0$ such that every circuit in the model above of entangling depth at most three satisfies*

$$|c(C, A_o)| \leq 6K_0 e^{-\kappa n}, \quad p_{\text{succ}}(C, A_o) \leq \frac{1}{2} + 3K_0 e^{-\kappa n}.$$

The constants are independent of the finite gate count, ancillary width, gate arities, product ancillary state, and binary one-qubit measurement. The same statement holds for mixed product ancillary states. In particular, one may use

$$|c(C, A_o)| \leq \min\{1, 1800e^{-2^{-27}n}\}.$$

The qualitative bound is proved in Section 5; Appendix A derives the displayed numerical specialization.

Theorem 2 (Exact depth-four lower bound). *There exist an absolute constant $c_* > 0$ and an absolute integer $n_0 \geq 1$ such that, for every $n \geq n_0$ and every circuit in the model above of entangling depth at most four that computes n -bit parity exactly, its multiqubit-gate count satisfies*

$$M \geq e^{c_* n}.$$

The constants are independent of the finite number of product-initialized ancillary wires, their initial state, all gate arities, and the binary one-qubit measurement. The same assertion holds for mixed product ancillary states by Lemma 3. In particular, exact parity has no polynomial-size depth-four QAC circuit family, even when only multiqubit gates are counted.

The proof is completed in Section 6, using the exponential event estimate and parity-preserving cleanup from Sections 4 and 6.1. It permits arbitrary garbage and also covers an output qubit that is an original input wire.

Lemma 3 (Mixed product initialization). *For a fixed circuit and measurement, any bound on $|c(C, A_o)|$ uniform over pure product ancillary states extends to input-independent mixed product ancillary states. The same holds for the absolute full-parity coefficient of any fixed output operator, including a non-Hermitian operator. If a fixed circuit and measurement compute parity exactly with mixed product ancillas, they do so with a pure product ancillary state on the same wires. Thus the exact lower bound also extends without changing depth, gate count, or arity.*

Proof. Expand each τ_q in an orthonormal eigenbasis. Their tensor product is a finite convex combination $\tau = \sum_z p_z |\alpha_z\rangle\langle\alpha_z|$ of pure product states, where $p_z \geq 0$, $\sum_z p_z = 1$, and the weights are independent of x . Consequently $\rho_x = \sum_z p_z \rho_{x,z}$, where $\rho_{x,z}$ is the output state for ancillary input $|\alpha_z\rangle$. Linearity gives $c(C, A_o) = \sum_z p_z c_z(C, A_o)$, so the triangle inequality proves the first claim; the same calculation applies to any fixed signed trace pairing.

For the exact assertion, let $s_z(x) \in [0, 1]$ be the correct-output probability for input x and ancillary state $|\alpha_z\rangle$. Exactness of the mixture implies, for every x ,

$$0 = 1 - \sum_z p_z s_z(x) = \sum_z p_z (1 - s_z(x)).$$

Every term is nonnegative. Thus, for every fixed z with $p_z > 0$, we have $s_z(x) = 1$ for every input x . At least one such z exists, and choosing it once gives a single pure product ancillary state that works for all inputs. The circuit, measurement, and physical wires are unchanged. The argument includes an absent ancillary register, whose decomposition has one scalar component of weight one. $\square$

3 A local source lemma

For each fixed classical input, pure product initialization and disjointness imply that the state after the first multiqubit layer is a tensor product over gate components: the gate supports and untouched singleton wires. A *source* in this section is one such component with exactly one variable input and any finite number of fixed pure product ancillas. Denote its variable-input wire by $\mathbf{d}$. Partition the source wires into nonempty *ports*, indexed by $\mathcal{J} = \{0, 1, \dots, m\}$ for an integer $m \geq 0$. The *owner* port 0 contains $\mathbf{d}$; all other ports contain only ancillary wires. This partition is arbitrary and will later be chosen from second-layer incidences. For $A \subseteq \mathcal{J}$, the notation Tr_{A^c} means tracing all wires in ports indexed by $\mathcal{J} \setminus A$. In particular, if A contains only ancillary-port indices, its complement includes the owner.

Using the reflection normal form and a local change of basis on $\mathbf{d}$, write the source gate as

$$G = I - 2|1\rangle\langle 1|_{\mathbf{d}} \otimes \Pi_{\text{anc}}, \quad \Pi_{\text{anc}} = \bigotimes_{q \in \text{Anc}} |v_q\rangle\langle v_q|,$$

where Anc is the set of ancillary wires in the source. All vectors v_q and α_q below are unit vectors. The two possible data inputs u_0, u_1 form an orthonormal basis: they are the images of the two actual data states under the one-qubit basis change that sends the gate's trigger vector on $\mathbf{d}$ to $|1\rangle$. The ancillary input is $\alpha = \bigotimes_{q \in \text{Anc}} \alpha_q$. One-qubit unitaries after G act within ports; they conjugate each port marginal and each retained marginal of F , preserving eigenvalues and trace norms. In the insertion lemma below they can also be absorbed into the inserted operator without changing its rank or norm, so we omit them here.

An untouched singleton is covered by taking no ancillas and using the formal one-qubit reflection $G = Z_{\mathbf{d}}$. If its actual output vectors are φ_0, φ_1 , set $u_x = Z_{\mathbf{d}}\varphi_x$, so $Gu_x = \varphi_x$. This is an analytic representation, not an added multiqubit gate. Empty tensor products are scalars on a one-dimensional space.

For $x \in \{0, 1\}$, put

$$\psi_x = G(u_x \otimes \alpha), \quad F = \frac{|\psi_0\rangle\langle\psi_0| - |\psi_1\rangle\langle\psi_1|}{2}, \quad \bar{\rho} = \frac{|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|}{2}.$$

The vectors ψ_0, ψ_1 are orthonormal because G is unitary. Thus $F = F^*$, $\text{Tr } F = 0$, and $\|F\|_1 = 1$. Here $\bar{\rho}$ is the uniform average over the source's input bit, and F is its signed average. For each port j , define

$$\bar{\rho}_j = \text{Tr}_{\mathcal{J} \setminus \{j\}} \bar{\rho}, \quad \lambda_j = \lambda_{\max}(\bar{\rho}_j), \quad h_j = 1 - \lambda_j.$$

Set $C_s = 1/2 + 2/e$.

Lemma 4 (Mixedness and random retention). *For a source as above:*

- (i) $h_0 = 1/2$ and $\sum_{j \in \mathcal{J}} h_j \leq C_s$.
- (ii) There is a nondecreasing function $f : 2^{\mathcal{J}} \rightarrow [0, 1]$ such that

$$\|\text{Tr}_{A^c} F\|_1^2 \leq f(A) \quad (A \subseteq \mathcal{J}).$$

If A is chosen by retaining each port independently with probability $1/2$, then $\mathbb{E}_A f(A) \leq 3/4$.

- (iii) For every $A \subseteq \mathcal{J}$ and every $j \in \mathcal{J}$,

$$|f(A \triangle \{j\}) - f(A)| \leq 4h_j,$$

where $\triangle$ denotes symmetric difference.

These assertions allow arbitrary finite port widths and arbitrarily many fixed pure product ancillas in the owner port.

Proof. Define

$$\xi = |\langle 1|u_1\rangle|^2 - |\langle 1|u_0\rangle|^2, \quad |\xi| \leq 1.$$

For each port j , let Anc_j be its ancillary wires and set

$$\alpha^{(j)} = \bigotimes_{q \in \text{Anc} * j} \alpha_q, \quad \Pi^{(j)} = \bigotimes_{q \in \text{Anc} * j} |v_q\rangle\langle v_q|.$$

The data wire is excluded from both products when $j = 0$. For every $j \in \mathcal{J}$, put

$$p_j = \langle \alpha^{(j)} | \Pi^{(j)} | \alpha^{(j)} \rangle = \prod_{q \in \text{Anc} * j} |\langle v_q | \alpha_q \rangle|^2, \quad \Theta = \prod_{j \in \mathcal{J}} p_j.$$

All these overlap probabilities lie in $[0, 1]$; empty products are one.

Since the uniform mixture of u_0, u_1 is $I_d/2$, averaging gives

$$\bar{\rho} = \frac{1}{2} |0\rangle\langle 0| \otimes |\alpha\rangle\langle \alpha| + \frac{1}{2} |1\rangle\langle 1| \otimes R_{\text{anc}} |\alpha\rangle\langle \alpha| R_{\text{anc}}, \quad R_{\text{anc}} = I - 2\Pi_{\text{anc}}. \quad (4)$$

On the owner port, the data-zero block is a pure ancillary projector of weight $1/2$. The orthogonal data-one block is $1/2$ times a density operator. Every eigenvalue is therefore at most $1/2$, and the data-zero block attains $1/2$. Hence $\lambda_0 = 1/2$, even when the owner contains ancillas.

If $\Theta = 0$, then $\Pi_{\text{anc}}\alpha = 0$, so G acts as the identity on the two-dimensional input subspace spanned by $u_x \otimes \alpha$. All ancillary-only ports remain pure. Retaining the owner preserves the orthogonal data states, whereas erasing it traces out the data wire and gives a zero signed marginal. Taking $f(A) = \mathbf{1}_{\{0 \in A\}}$ gives the exact squared retained norm: its mean is $1/2$, ancillary-port changes are zero, and the owner change is $1 = 2h_0 \leq 4h_0$. This proves all assertions in this case. Henceforth assume $\Theta > 0$, so every p_j is positive.

For an ancillary-only port j , put $r_j = \Theta/p_j \in (0, 1]$ and $R_j = I_j - 2\Pi^{(j)}$. Tracing (4) gives

$$\bar{\rho}_j = (1 - r_j/2) |\alpha^{(j)}\rangle\langle \alpha^{(j)}| + (r_j/2) R_j |\alpha^{(j)}\rangle\langle \alpha^{(j)}| R_j.$$

Indeed, the reduced reflected pure state is $(1 - r_j) |\alpha^{(j)}\rangle\langle \alpha^{(j)}| + r_j R_j |\alpha^{(j)}\rangle\langle \alpha^{(j)}| R_j$, as follows by expanding the reflection and tracing the other ancillas; the same expansion is written out below for a general retained set A , and the present case is $A = \{j\}$ with $\Theta_I = \Theta/p_j = r_j$. The marginal has rank at most two, so its two possibly nonzero eigenvalues are $1 - h_j, h_j$, with $0 \leq h_j \leq 1/2$. For a mixture $(1 - t)|u\rangle\langle u| + t|v\rangle\langle v|$ of unit vectors, the product of these two eigenvalues is $t(1 - t)(1 - |\langle u|v\rangle|^2)$: restrict to their span and use the orthonormal basis consisting of u and the normalized component of v orthogonal to u ; if the span is one-dimensional, both sides are zero. Here $\langle \alpha^{(j)} | R_j | \alpha^{(j)} \rangle = 1 - 2p_j$, giving

$$h_j(1 - h_j) = \frac{r_j}{2} \left(1 - \frac{r_j}{2}\right) [1 - (1 - 2p_j)^2] = \Theta(1 - p_j)(2 - r_j). \quad (5)$$

Since $h_j(1 - h_j) \leq h_j$ and $2 - r_j \geq 1$, this implies $h_j \geq \Theta(1 - p_j)$. For the other direction, the variational characterization of the largest eigenvalue gives

$$1 - h_j = \lambda_j \geq \langle \alpha^{(j)} | \bar{\rho}_j | \alpha^{(j)} \rangle = 1 - 2\Theta(1 - p_j).$$

Thus

$$\Theta(1 - p_j) \leq h_j \leq 2\Theta(1 - p_j). \quad (6)$$

Writing $\Theta_a = \prod_{j \neq 0} p_j$, the inequality $1 - p_j \leq -\log p_j$ and $\Theta = p_0 \Theta_a \leq \Theta_a$ yield

$$\sum_{j \neq 0} h_j \leq 2\Theta \sum_{j \neq 0} (1 - p_j) \leq -2\Theta \log \Theta_a \leq -2\Theta \log \Theta \leq \frac{2}{e}.$$

The last step maximizes $-2\Theta \log \Theta$ on $(0, 1]$. With $h_0 = 1/2$, this proves (i), including the empty ancillary-port collection.

We next compute the signed marginal when the owner is erased. For retained ancillary-only ports $A \subseteq \mathcal{J} \setminus \{0\}$, put

$$\Theta_A = \prod_{j \in A} p_j, \quad \Theta_I = p_0 \prod_{\substack{j \notin A \\ j \neq 0}} p_j, \quad \Theta = \Theta_A \Theta_I.$$

Here Θ_I includes the overlap probabilities of the owner ancillas. Let $\alpha_A = \bigotimes_{j \in A} \alpha^{(j)}$, $\rho_A = |\alpha_A\rangle\langle\alpha_A|$, $\Pi_A = \bigotimes_{j \in A} \Pi^{(j)}$, and $R_A = I_A - 2\Pi_A$. The data-zero and data-one diagonal coefficients of the signed input are $\xi/2$ and $-\xi/2$, respectively. The off-diagonal data blocks vanish under the partial trace over $\mathbf{d}$, so

$$\text{Tr}_{\mathbf{d}} F = \frac{\xi}{2} (|\alpha\rangle\langle\alpha| - R_{\text{anc}}|\alpha\rangle\langle\alpha|R_{\text{anc}}).$$

Expanding the reflection and tracing all erased ancillas gives

$$\begin{aligned} \text{Tr}_{\text{erased ancillas}}(R_{\text{anc}}|\alpha\rangle\langle\alpha|R_{\text{anc}}) &= \rho_A - 2\Theta_I \Pi_A \rho_A - 2\Theta_I \rho_A \Pi_A + 4\Theta_I \Pi_A \rho_A \Pi_A \\ &= \rho_A + \Theta_I (R_A \rho_A R_A - \rho_A). \end{aligned}$$

Consequently,

$$\text{Tr}_{A^c} F = \frac{\xi \Theta_I}{2} (\rho_A - R_A \rho_A R_A).$$

For unit vectors u, v , the trace norm of the difference of their pure-state projectors is

$$\| |u\rangle\langle u| - |v\rangle\langle v| \|_* = 2\sqrt{1 - |\langle u|v\rangle|^2}.$$

Indeed, on their span the difference has trace zero and eigenvalues $\pm\sqrt{1 - |\langle u|v\rangle|^2}$; the coincident-state case gives zero. Using $\langle\alpha_A|R_A|\alpha_A\rangle = 1 - 2\Theta_A$, we obtain

$$\begin{aligned} D(A) &:= \|\text{Tr} * A^c F\|_1 = 2|\xi| \Theta_I \sqrt{\Theta_A(1 - \Theta_A)}, \\ D(A)^2 &= 4\xi^2 \Theta^2 (\Theta_A^{-1} - 1). \end{aligned} \tag{7}$$

Division by Θ_A is legitimate because $\Theta > 0$. If $A = \emptyset$, its complement is every source port, $\Theta_A = 1$, and $D(A) = \|\text{Tr} F\| = 0$, consistently with this formula.

Define

$$f(A) = \begin{cases} 1, & 0 \in A, \\ D(A)^2, & 0 \notin A. \end{cases}$$

Every reduced signed state is half the difference of two density operators, so its trace norm is at most one by the triangle inequality. Thus $0 \leq \|\text{Tr}_{A^c} F\|_1^2 \leq f(A) \leq 1$. In the owner-erased case f is the exact squared norm; in the owner-retained case it is only an upper bound.

Conditioned on owner erasure, independent half-retention of the other ports gives

$$\mathbb{E}_A D(A)^2 = 4\xi^2 p_0^2 \Theta_a \left[\prod_{j \neq 0} \frac{1 + p_j}{2} - \Theta_a \right].$$

For example, this follows directly from $\mathbb{E}_A \Theta_A^{-1} = \prod_{j \neq 0} \frac{1+p_j^{-1}}{2}$. For numbers in $[0, 1]$,

$$\prod_j \frac{1+p_j}{2} \leq \frac{1 + \prod_j p_j}{2}.$$

The empty-product case is equality, and the induction step uses

$$\frac{1+ab}{2} - \frac{(1+a)(1+b)}{4} = \frac{(1-a)(1-b)}{4} \geq 0.$$

It follows that $\mathbb{E}_A D(A)^2 \leq 2\xi^2 p_0^2 \Theta_a(1 - \Theta_a) \leq 1/2$. The owner's independent retention coin therefore gives

$$\mathbb{E}f = \frac{1}{2} + \frac{1}{2} \mathbb{E}_A D(A)^2 \leq \frac{3}{4},$$

proving the mean bound in (ii).

Finally, suppose the owner is erased and $j \neq 0$ is not yet in A . By (7), adding j changes the majorant by

$$D(A \cup \{j\})^2 - D(A)^2 = 4\xi^2 \frac{\Theta^2}{\Theta_{Ap_j}} (1 - p_j) \leq 4\Theta(1 - p_j) \leq 4h_j.$$

The change is nonnegative; the first inequality uses $\xi^2 \leq 1$ and $\Theta/(\Theta_{Ap_j}) \leq 1$, the latter being the product of overlaps of the still-erased ancillas, and the last uses (6). The same absolute bound holds when j is erased instead. If the owner is already retained, an ancillary-port change leaves $f = 1$ unchanged. Adding or erasing the owner changes f by at most $1 = 2h_0 \leq 4h_0$. This proves (iii). Each single-port addition is nonnegative: for ancillary ports this was just proved, and adding the owner sets f to one. Thus f is nondecreasing under inclusion, completing (ii). $\square$

The partition into ports was arbitrary, so the lemma also applies after merging ports. The port containing d is designated the owner, and the mixedness of each merged port is computed from its joint marginal.

Remark 5. The constant $3/4$ in (ii) is sharp for majorants of the true squared retained norm. Take $u_x = |x\rangle$ and one ancilla in $|0\rangle$, apply CNOT from the input to the ancilla, and place the two wires in separate ports. This is directly a permitted source reflection:

$$\text{CNOT} = I - 2|1\rangle\langle 1|_d \otimes |-\rangle\langle -|, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

It has $\xi = 1$, $p_0 = 1$, and $p_1 = 1/2$, and its signed output is

$$F = \frac{|00\rangle\langle 00| - |11\rangle\langle 11|}{2}.$$

For $A = \emptyset, \{0\}, \{1\}, \{0, 1\}$, respectively, the squared retained trace norms are $0, 1, 1, 1$. The majorant above has these same values, so its mean is exactly $3/4$. Indeed, any majorant of the true squared retained norm must have mean at least $3/4$ in this example.

The example also shows why separate insertion penalties cannot be multiplied. Here $\lambda_0 = \lambda_1 = 1/2$, but

$$|\text{Tr}[(|0\rangle\langle 0| \otimes |0\rangle\langle 0|)F]| = \frac{1}{2} > \lambda_0 \lambda_1.$$

The joint marginal, rather than a product of single-port bounds, is needed for simultaneous insertions.

Lemma 6 (Rank-one insertion). *For a one-input source and any operator T of rank at most one on port j , put $K = \text{Tr}_j[(T \otimes I)F]$. Then*

$$\|K\|_1 \leq \|T\|_\infty \lambda_j.$$

In particular, a rank-one contraction gives $\|K\|_1 \leq \lambda_j$. For a union $J \subseteq \mathcal{J}$ of ports, take T of rank at most one on that union and put $K = \text{Tr}_J[(T \otimes I)F]$; Tr_J traces all wires in the ports indexed by J . The same statement holds with the joint marginal $\bar{\rho}_J = \text{Tr}_{J^c} \bar{\rho}$ and $\lambda_J = \lambda_{\max}(\bar{\rho}_J)$.

If Y_1, Y_2, Y_3 are contractions on the remaining wires, then $\|Y_1 K Y_2\|_1 \leq \|T\|_\infty \lambda_j$ and $|\text{Tr}(Y_3 K)| \leq \|T\|_\infty \lambda_j$. Further partial traces cannot increase the trace-norm bound. For insertion on a union J , these consequences hold with λ_J in place of λ_j . All these assertions concern unnormalized operators.

Proof. Define the isometry $W : \mathbb{C}^2 \rightarrow \mathcal{H}_{\text{source}}$ by $W|x\rangle = \psi_x$. Let $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$ on its logical input. Then

$$F = WZW^*/2, \quad \bar{\rho} = WW^*/2.$$

For a unit vector u on the inserted port define $W_u = (\langle u| \otimes I)W$. Cyclicity of trace gives

$$\|W_u\|_2^2 = \text{Tr}[(|u\rangle\langle u| \otimes I)WW^*] = 2\langle u|\bar{\rho}_j|u\rangle \leq 2\lambda_j.$$

For $T = |u\rangle\langle v|$ with u, v unit, direct partial tracing gives

$$K = W_v Z W_u^*/2.$$

Schatten Hölder and $\|W_v Z\|_2 = \|W_v\|_2$ yield

$$\|K\|_1 \leq \frac{1}{2} \|W_v\|_2 \|W_u\|_2 \leq \lambda_j.$$

Every nonzero rank-one T has the form $s|u\rangle\langle v|$ with $s = \|T\|_\infty$ and u, v unit; scaling proves the stated bound. For $T = 0$ it is immediate. Nothing requires T or K to be Hermitian.

The same calculation treats a union of ports as one tensor factor. It includes $J = \emptyset$, when the inserted register is scalar and $\lambda_J = \text{Tr} \bar{\rho} = 1$, and $J = \mathcal{J}$, when the remaining register is scalar and $\lambda_J = 1/2$. The remaining-wire assertions follow from $\|Y_1 K Y_2\|_1 \leq \|Y_1\|_\infty \|K\|_1 \|Y_2\|_\infty$ and trace-norm duality. For completeness, partial-trace contractivity also holds without Hermiticity:

$$\|\text{Tr}_B X\|_1 = \sup_{\|Y\|_\infty \leq 1} |\text{Tr}[(I_B \otimes Y)X]| \leq \|X\|_1.$$

Here Y acts on the untraced register, and $\|I_B \otimes Y\|_\infty = \|Y\|_\infty$, so no dimension factor occurs. The insertion argument uses only that W is an isometry. The special mixedness values and budget proved above still require the one-reflection, product-initialized source structure. $\square$

4 Exponential event bounds from the shared source estimate

4.1 Ports, joint resources, and random activation

Take r independent one-input sources and a fixed, input-independent resource operator σ with $\|\sigma\|_1 \leq 1$, in tensor product with the sources. Group all resource wires before defining ports. For each group, its ports are its nonempty intersections with second-layer gate supports and, if nonempty, its remaining idle wires. Let E be a fixed tensor product of positive contractions over these ports. A factor may be entangled within one port. One-qubit gates before and after the second layer can

be absorbed into source states and event factors, respectively; they act within ports and preserve the source weights and norms. The signed input to the second layer is

$$\Sigma = \sigma \otimes \bigotimes_{i=1}^r F_i, \quad F_i = (\rho_{i,0} - \rho_{i,1})/2.$$

The resource is not assumed to factor over its ports and need not be Hermitian. For operators X on $\mathbf{B} \otimes \mathbf{A}$ and T on $\mathbf{B}$, trace-norm duality gives

$$\|\mathrm{Tr}_{\mathbf{B}}[(T \otimes I)X]\|_1 \leq \|T\|_\infty \|X\|_1. \quad (8)$$

Indeed, pair the partial trace with an operator Y of norm at most one; then $|\mathrm{Tr}[(T \otimes Y)X]| \leq \|T\|_\infty \|X\|_1$. This includes ordinary partial-trace contractivity and requires no Hermiticity of X .

Represent each second-layer gate by a real vertex and each idle port by a private vertex. Every wire belongs to exactly one vertex and each source has at most one port at that vertex. Write $i \sim g$ when source i has a port at gate g . Write the effect and gate at a real vertex as

$$P_g = P_{\mathrm{res},g} \otimes \bigotimes_{i:i \sim g} P_{ig}, \quad D_g = I - 2\Pi_g, \quad \Pi_g = \Pi_{\mathrm{res},g} \otimes \bigotimes_{i:i \sim g} \Pi_{ig},$$

omitting absent factors. Each present Π factor is a rank-one orthogonal projector; each P factor is a positive contraction. The pulled-back event is the tensor product of $D_g P_g D_g$ over real vertices and the idle effects. The exact identity

$$D_g P_g D_g = P_g - 2\Pi_g P_g - 2P_g \Pi_g + 4\Pi_g P_g \Pi_g \quad (9)$$

has correction coefficient mass eight. Each coefficient-free correction factors into rank-at-most-one contractions on all its incident ports.

Lemma 7 (Joint random activation). *Let a tensor-product operator consist of arbitrary operators $K_1, \dots, K_r$ on disjoint source groups and an independent resource operator σ with $\|\sigma\|_1 \leq 1$. None of these operators need be Hermitian. Partition its wires into vertices so that each vertex meets each source in either one whole port or no wires. For any tensor product of positive contractions E_g over vertices,*

$$\left| \mathrm{Tr} \left[\left(\bigotimes_g E_g \right) \left(\sigma \otimes \bigotimes_i K_i \right) \right] \right| \leq \mathbb{E}_A \prod_i \|\mathrm{Tr}_{\mathcal{V}_i \setminus A} K_i\|_1, \quad (10)$$

where each vertex belongs to A independently with probability $1/2$ and $\mathcal{V}_i$ is the vertex set incident to source i . The partial trace removes the ports of source i at vertices outside A . Empty registers and products have the scalar conventions of Section 2.

Proof. Put $\Gamma_g = 2E_g - I_g$. It is a Hermitian contraction, and

$$\bigotimes_g E_g = \mathbb{E}_A \left[\left(\bigotimes_{g \in A} \Gamma_g \right) \otimes \left(\bigotimes_{g \notin A} I_g \right) \right].$$

For fixed A , trace the inactive wires in each source and in the joint resource. The resource marginal has trace norm at most one by (8). The active operator has norm at most one. Trace-norm duality and multiplicativity on tensor products give the asserted product bound, and averaging proves the claim. Different sources may share a coin; no expectation is factored over sources, and no tensor factorization of the resource's ports is used. $\square$

4.2 Retention after a joint insertion

Lemma 8 (Rank-one contraction bound). *Let ρ_0, ρ_1 be any density operators on $\mathbf{B} \otimes \mathbf{A}$, and put $F = (\rho_0 - \rho_1)/2$. For a rank-at-most-one contraction T on $\mathbf{B}$, put*

$$d = \|\mathrm{Tr}_{\mathbf{B}} F\|_1, \quad \nu = \|\mathrm{Tr}_{\mathbf{B}}[(T \otimes I)F]\|_1.$$

Then

$$\nu \leq \frac{1+d}{2}, \quad \nu^2 \leq \frac{1}{2} + \frac{1}{2}d^2. \quad (11)$$

The operator T and the contracted output need not be Hermitian.

Proof. We prove a stronger centered inequality for every Hermitian F . Write $K = \mathrm{Tr}_{\mathbf{B}}[(T \otimes I)F]$, $F_{\mathbf{A}} = \mathrm{Tr}_{\mathbf{B}} F$, $t = \|T\|_{\infty}$, and $\tau = \mathrm{Tr} T$. Then

$$\|2K - \tau F_{\mathbf{A}}\|_1 \leq t \|F\|_1. \quad (12)$$

The case $T = 0$ is immediate. Otherwise factor $T = |u\rangle\langle v|$, so $\|u\|\|v\| = t$. For a unit vector z on $\mathbf{B}$, put $P_z = |z\rangle\langle z|$ and $T_{zz} = \langle z|T|z\rangle$. Cauchy–Schwarz, applied to each term and then to the sum, gives

$$|T_{zz}| + |T_{zz} - \tau| \leq \|P_z u\| \|P_z v\| + \|(I - P_z)u\| \|(I - P_z)v\| \leq t.$$

Thus $T_c = 2T - \tau I$ satisfies $|\langle z|T_c|z\rangle| \leq t$ for every unit z . Consequently $-tI \leq \mathrm{Re}(e^{i\theta}T_c) \leq tI$ for every real θ , by evaluating its quadratic form on unit vectors. For any unitary U on $\mathbf{A}$, diagonalizing U shows that $\mathrm{Re}(T_c \otimes U)$ is a direct sum of blocks $\mathrm{Re}(e^{i\theta}T_c)$, each of operator norm at most t . Here $\mathrm{Re} X = (X + X^*)/2$. For Hermitian F , $\mathrm{Re} \mathrm{Tr}(XF) = \mathrm{Tr}[(\mathrm{Re} X)F]$. Trace-norm duality therefore gives

$$\|2K - \tau F_{\mathbf{A}}\|_1 = \max_{U \text{ unitary}} \mathrm{Re} \mathrm{Tr}[(T_c \otimes U)F] \leq t \|F\|_1.$$

No Hermiticity of K or T is used. Since $\|F\|_1 \leq 1$ and $|\tau| \leq t \leq 1$, the triangle inequality yields $2\nu \leq 1 + d$. Finally, for $0 \leq d \leq 1$,

$$1 - \nu^2 \geq \frac{(1-d)(3+d)}{4} \geq \frac{1}{2}(1 - d^2).$$

This proves (11). □

Lemma 9 (Transfer of the original retention majorant). *For one source, insert a rank-at-most-one contraction T on a union $\mathcal{B}$ of its original ports. Put $K = \mathrm{Tr}_{\mathcal{B}}[(T \otimes I)F]$ and let f_0 be the nondecreasing majorant from Lemma 4. For any subset A of the remaining ports, define*

$$f_{\mathcal{B}}(A) = \min\{\|K\|_1^2, (1 + f_0(A))/2\}.$$

This majorizes the squared retained norm of K , is at most one, has mean at most $7/8$ under independent fair retention, and has oscillation at most $2h_j$ when a remaining port j is toggled. The weights are the original source weights, including when the owner or every port is inserted.

Proof. Equation (8) gives $\|K\|_1 \leq 1$. Write K_A and F_A for the respective marginals on A . Trace all other remaining ports of both density operators defining F before applying Lemma 8 to the inserted union and A . This gives $\|K_A\|_1^2 \leq (1 + \|F_A\|_1^2)/2 \leq (1 + f_0(A))/2$. Partial-trace contractivity also bounds it by $\|K\|_1^2$. Couple fair choices on the remaining ports with independent fair choices $A_{\mathcal{B}}$ on the inserted ports. Monotonicity gives

$$\mathbb{E}_A f_0(A) \leq \mathbb{E}_{A, A_{\mathcal{B}}} f_0(A \cup A_{\mathcal{B}}) \leq 3/4.$$

Hence the new mean is at most $7/8$. Halving the original oscillation bound gives $2h_j$; clipping at the fixed number $\|K\|_1^2$ cannot increase oscillations. Empty registers are scalar spaces, so the same argument includes empty inserted or remaining sets and zero insertions. No inserted operator is renormalized. $\square$

4.3 One penalty for a union of ports

Lemma 10 (Convex joint deficit). *For one source from Section 3 and a union $\mathcal{B}$ of its original ports, put*

$$W(\mathcal{B}) = \sum_{j \in \mathcal{B}} h_j, \quad h^\cup(\mathcal{B}) = 1 - \lambda_{\max}(\bar{\rho}_{\mathcal{B}}).$$

Define

$$\Phi(0) = 0, \quad \Phi(t) = \frac{t}{4 \log(4/t)} \quad (0 < t \leq C_s).$$

Then Φ is increasing and convex on $[0, C_s]$, and $h^\cup(\mathcal{B}) \geq \Phi(W(\mathcal{B}))$ for every $\mathcal{B}$.

Proof. Write $W = W(\mathcal{B})$ and $h^\cup = h^\cup(\mathcal{B})$. The empty union and $W = 0$ are immediate. If $\mathcal{B}$ contains the owner, then $h^\cup = 1/2$ and $W \leq C_s < 5/4$. Since $4/W > 16/5 > e$, we have $\Phi(W) < W/4 < 5/16 < 1/2$.

Suppose $\mathcal{B}$ is ancillary-only. Let Θ be the full ancillary trigger probability and $p_{\mathcal{B}}$ its probability on $\mathcal{B}$. Since $W > 0$, the case $\Theta = 0$ is excluded: all ancillary-only weights would vanish. Also $p_{\mathcal{B}} = 1$ would make every selected port weight zero. Thus $0 < \Theta \leq p_{\mathcal{B}} < 1$ before any logarithm or division below is taken. Summing (6) over $j \in \mathcal{B}$ gives $W \leq 2\Theta \sum_{j \in \mathcal{B}} (1 - p_j) \leq 2\Theta \log(1/p_{\mathcal{B}})$; applying (6) to the coarser partition in which $\mathcal{B}$ is a single ancillary-only port gives $h^\cup \geq \Theta(1 - p_{\mathcal{B}})$; and $\Theta \leq p_{\mathcal{B}}$ because Θ is $p_{\mathcal{B}}$ times further trigger probabilities. Thus

$$W \leq 2\Theta \log(1/p_{\mathcal{B}}), \quad h^\cup \geq \Theta(1 - p_{\mathcal{B}}), \quad \Theta \leq p_{\mathcal{B}}.$$

Put $x = \log(1/p_{\mathcal{B}}) > 0$ and obtain

$$W \leq 2xe^{-x} \leq 2/e, \quad h^\cup \geq \frac{W(1 - e^{-x})}{2x}. \quad (13)$$

For $0 < x \leq 1$, the decreasing function $(1 - e^{-x})/x$ is at least $1 - e^{-1} > 5/8$, so $h^\cup > 5W/16$. Meanwhile $\log(4/W) \geq \log(2e) > 1$, giving $\Phi(W) < W/4 < h^\cup$.

For $x \geq 1$, the first inequality in (13) implies $\log(4/W) \geq x - \log(x/2)$. Differentiation shows that $\log(x/2)/x$ has global maximum $1/(2e) < 3/16$ on $x > 0$. Therefore

$$\frac{(1 - e^{-x}) \log(4/W)}{x} > \frac{5}{8} \frac{13}{16} = \frac{65}{128} > \frac{1}{2}.$$

Substitution into (13) proves $h^\cup \geq W/[4 \log(4/W)]$. All divisions were made only after $\Theta, W, x > 0$ was established. $\square$

Finally, writing $\theta = \log(4/t) > 0$, direct differentiation gives

$$\Phi'(t) = \frac{\theta + 1}{4\theta^2} > 0, \quad \Phi''(t) = \frac{\theta + 2}{4t\theta^3} > 0.$$

The value and right derivative at zero are both zero. This proves the asserted monotonicity and convexity on the closed interval. $\square$

4.4 The exponential source-event theorem

Theorem 11 (Exponential two-layer source bound). *There are absolute constants $K_0 \geq 1$ and $a > 0$ with the following property. Let $r \geq 0$ be an integer. Consider r independent one-input first-layer sources as in Section 3 and a tensor-independent resource operator σ with $\|\sigma\|_1 \leq 1$, independent of the input x . Apply a second layer of pairwise wire-disjoint generalized controlled-Z gates, with arbitrary one-qubit gates before and after it. Use the nonempty source-by-second-gate ports and private idle ports defined above, including ports on the resource.*

Let E be a fixed tensor product of positive contractions over these ports. After absorbing the adjacent one-qubit gates as above, set

$$q(x) = \text{Tr} \left[EU_2 \left(\sigma \otimes \bigotimes_{i=1}^r \rho_{i,x_i} \right) U_2^* \right], \quad x \in \{0,1\}^r,$$

where ρ_{i,x_i} is the state of input source i and U_2 is the second multiqubit layer. All gates and effects are independent of x . The full signed coefficient of q satisfies

$$|\hat{q}([r])| \leq K_0 e^{-ar}.$$

The functional $q(x)$ is an event probability if σ is a density operator, and may be complex otherwise. The bound is independent of gate count, port widths, and finite resource dimension. For $r = 0$ the coefficient is the single trace pairing, of magnitude at most one. An absent resource is the scalar 1.

Proof. The $r = 0$ assertion follows from trace-norm duality. Assume $r \geq 1$. Linearity of uniform signed averaging gives $\Sigma = \sigma \otimes \bigotimes_{i=1}^r F_i$ and $\hat{q}([r]) = \text{Tr}(U_2^* E U_2 \Sigma)$. Use the original port weights h_{ig} , with $\sum_g h_{ig} \leq C_s$ and $h_{ig} = 0$ for absent ports. All rows refer only to input sources, not the resource. Fix a threshold $w \geq 1$ for now. Call a real gate heavy if its input-source load $\sum_i h_{ig}$ exceeds w . Let $\mathcal{H}$ be the set of heavy gates and $N_{\mathcal{H}} = |\mathcal{H}|$. The row budgets imply

$$N_{\mathcal{H}} \leq C_s r / w. \tag{14}$$

Expand all heavy gates. For every $g \in \mathcal{H}$, expand its pulled-back event exactly:

$$D_g P_g D_g = P_g - 2\Pi_g P_g - 2P_g \Pi_g + 4\Pi_g P_g \Pi_g.$$

After removing the scalar coefficients $-2, -2, 4$, the three correction operators factor over source ports into contractions of rank at most one, as in (9). A zero correction can be omitted. The base term P_g is a product of positive port effects and is split into private vertices.

Choose a correction set $\mathcal{S} \subseteq \mathcal{H}$ and, for each $g \in \mathcal{S}$, a type $\ell_g \in \{1, 2, 3\}$. Denote its *coefficient-free* scalar term by $t_{\mathcal{S},\ell}$. With $(\theta_1, \theta_2, \theta_3) = (-2, -2, 4)$,

$$\hat{q}([r]) = \sum_{\mathcal{S} \subseteq \mathcal{H}} \sum_{\ell \in \{1,2,3\}^{\mathcal{S}}} \left(\prod_{g \in \mathcal{S}} \theta_{\ell_g} \right) t_{\mathcal{S},\ell}.$$

Explicitly, writing $B_g^{(\ell)}$ for the coefficient-free correction operator of type ℓ at gate g ,

$$t_{\mathcal{S},\ell} = \text{Tr} \left[\left(\bigotimes_{g \in \mathcal{S}} B_g^{(\ell_g)} \otimes \bigotimes_{g \in \mathcal{H} \setminus \mathcal{S}} P_g \otimes \bigotimes_{g \text{ light}} D_g P_g D_g \otimes \bigotimes_{v \text{ private}} E_v \right) \Sigma \right],$$

where the last factor collects the effects on the private idle vertices. For $k = |\mathcal{S}|$, the absolute coefficient mass of the inner sum is at most $(2 + 2 + 4)^k = 8^k$, with equality before zero terms are omitted. The empty correction set has one term with coefficient one.

Fix $(\mathcal{S}, \ell)$. On source i , the corrections form a contraction T_i of rank at most one on the union J_i of its selected ports. Insert it once on that whole union and trace:

$$K_i = \text{Tr}_{J_i}[(T_i \otimes I)F_i], \quad \nu_i = \|K_i\|_1.$$

For $J_i = \emptyset$ take the scalar $T_i = 1$, so $K_i = F_i$. Likewise, let J_{res} be the union of selected resource ports, let T_{res} be their tensor-product insertion, and put $\sigma' = \text{Tr}_{J_{\text{res}}}[(T_{\text{res}} \otimes I)\sigma]$. For an empty union take $T_{\text{res}} = 1$. Equation (8) gives $\|\sigma'\|_1 \leq 1$. No operator is renormalized. The remaining real gates are light; unselected heavy gates contribute private positive effects. These surviving effects act on wires disjoint from the inserted ports and remain positive contractions. Writing their tensor product as E_{rem} , partial tracing the inserted ports gives the exact identity

$$t_{\mathcal{S}, \ell} = \text{Tr} \left[E_{\text{rem}} \left(\sigma' \otimes \bigotimes_i K_i \right) \right].$$

The resource remains one joint operator, even when its remaining ports meet different vertices. All remaining ports keep their original weights.

Dense correction sets: a joint deficit. For a selected set $\mathcal{S}$ of k heavy gates, let W_i be the selected original port weight of source i , and let $h_i^\cup$ be its joint deficit on the selected union. Then $0 \leq W_i \leq C_s$ and $\sum_i W_i \geq kw$. By Lemma 6, every source's single joint rank-one insertion leaves trace norm at most $1 - h_i^\cup$. The independent resource is kept jointly and has trace norm at most one, including when it is non-Hermitian. The surviving effect is a contraction, so trace-norm duality bounds the term by $\prod_i \nu_i$. Using $1 - u \leq e^{-u}$, the bound $h_i^\cup \geq \Phi(W_i)$, Jensen's inequality, and monotonicity of Φ gives

$$|t_{\mathcal{S}, \ell}| \leq \prod_i (1 - h_i^\cup) \leq \exp \left[-r \Phi \left(\frac{kw}{r} \right) \right] = \exp \left[-\frac{kw}{4 \log(4r/(kw))} \right] \quad (k > 0). \quad (15)$$

Here $kw/r \leq C_s < 4$, so the denominator is positive. For $k = 0$ the middle expression equals one because $\Phi(0) = 0$; the last expression is used only for $k > 0$. No low-trigger source or selected weight is discarded.

Every light remainder: shared-coin concentration. Apply Lemma 7 to the surviving positive effects. Distinct remaining ports of one source have distinct fair coins, even after unselected heavy base terms have split into private vertices. Let $\mathcal{V}_i$ be the surviving vertices incident to source i , and write h_{iv} for the original weight of its port at vertex v , or zero if there is no such port. Each original port contributes to at most one surviving vertex. Light real vertices have load at most w ; private vertices have load at most $1/2$, since each belongs to one source or to the resource and every source port has $h_j \leq 1/2$. For a retained vertex set A , write $D_i(A) = \|\text{Tr}_{\mathcal{V}_i \setminus A} K_i\|_1$ and let $f_i(A \cap \mathcal{V}_i)$ be the majorant supplied by Lemma 9. For $F_* = \sum_i f_i$, the coordinate oscillations are bounded by $c_v = 2 \sum_i h_{iv}$. Thus $c_v \leq 2w$ for every surviving vertex and $\sum_v c_v \leq 2C_s r$, which give $\sum_v c_v^2 \leq 4C_s r w$. The majorants satisfy $D_i^2 \leq f_i \leq 1$ and $\mathbb{E} F_* \leq 7r/8$. The latter follows by linearity of expectation; no independence between the f_i is needed. For $0 < t \leq 1/2$, the inequality $\sqrt{f} \leq e^{-t(1-f)}$ holds on $[0, 1]$ by $\log f \leq f - 1$, with $f = 0$ handled directly. The exponential-moment bound (23), applied to the independent vertex coins, gives

$$|t_{\mathcal{S}, \ell}| \leq \mathbb{E} \prod_i D_i \leq e^{-tr} \mathbb{E} e^{tF_*} \leq \exp \left[-\frac{tr}{8} + \frac{t^2 C_s r w}{2} \right].$$

This does not factor an expectation over sources; only vertex coins are independent. Taking $t = 1/(8C_s w) < 1/2$ yields

$$|t_{\mathcal{S},\ell}| \leq e^{-c_1 r/w}, \quad c_1 = (128C_s)^{-1}. \quad (16)$$

The same moment bound includes constant majorants and vanishing oscillations without an exceptional-event argument.

Sum sparse and dense correction sets. Choose $0 < \eta < C_s$ such that $A_1 = \eta \log(8eC_s/\eta) < c_1$, and put $b_1 = c_1 - A_1 > 0$. Such a choice exists because $\eta \log(8eC_s/\eta) \rightarrow 0$ as $\eta \downarrow 0$. Choose $w \geq 1$ so that

$$\log(8eC_s/\eta) - c_\eta w \leq -1, \quad c_\eta = \frac{1}{4 \log(4/\eta)}.$$

This is possible because $c_\eta > 0$; both choices are absolute constants. With $\mu = r/w$, there are at most $C_s \mu$ heavy gates; write $N_{\mathcal{H}} \leq C_s \mu$ for their number. The total absolute correction-type coefficient for a selected set of size k is 8^k .

For $1 \leq k \leq \eta \mu$, the binomial bound is

$$\binom{N_{\mathcal{H}}}{k} 8^k \leq \exp\left(k \log \frac{8eC_s \mu}{k}\right) \leq e^{A_1 \mu}.$$

The last inequality follows because $s \log(8eC_s \mu/s)$ is increasing for $0 < s \leq \eta \mu$: its derivative is $\log(8eC_s \mu/s) > 0$. Including $k = 0$ and at most $C_s \mu + 1$ possible sizes, the sparse part is bounded by $(C_s \mu + 1)e^{-b_1 \mu}$. For $k > \eta \mu$, equation (15) gives $|t_{\mathcal{S},\ell}| \leq e^{-c_\eta k w}$, and hence the contribution of size k is at most

$$\left(\frac{8eC_s}{\eta} e^{-c_\eta w}\right)^k \leq e^{-k}.$$

The dense sum is at most $e^{-\eta \mu}/(1 - e^{-1})$. These estimates include empty summation ranges and the case of no heavy gates. Thus the source theorem holds with

$$a = \frac{\min\{b_1/2, \eta\}}{w}, \quad K_0 = 1 + \frac{2C_s}{eb_1} + \frac{1}{1 - e^{-1}}. \quad (17)$$

Indeed, use $\sup_{\mu \geq 0} \mu e^{-b_1 \mu/2} = 2/(eb_1)$ to absorb the sparse prefactor. The case $r = 0$ is bounded by one directly and is covered by $K_0 > 1$. $\square$

5 Depth-three correlation and its exponential order

5.1 Removing the first-layer input-arity restriction

The source theorem assumes one variable per input component. We now remove that assumption by expanding large first-layer components before applying the theorem. The event must factor over physical qubits at this interface, because the expansion changes the source partition.

Theorem 12 (Arbitrary first-layer inputs). *With K_0 as in Theorem 11, there is an absolute $\kappa > 0$ with the following property. Let $C^{(2)} = L_2 D_2 L_1 D_1 L_0$ be the two-layer prefix of a circuit in the model of Section 2, including its final one-qubit layer, and let $E = \bigotimes_q E_q$ with $0 \leq E_q \leq I_q$ be a physical-qubit product event. Put $q(x) = \text{Tr}[E C^{(2)} \rho_x^{\text{in}} (C^{(2)})^*]$ and write $\hat{q}([n]) = \mathbb{E}_x \chi(x) q(x)$ for its full n -input parity coefficient. Then*

$$|\hat{q}([n])| \leq K_0 e^{-\kappa n},$$

with arbitrary finite product ancillary initialization and circuit width. The constants are independent of the event and the circuit.

Proof. By Lemma 3, it suffices to prove a bound uniform over pure product ancillary initializations. Assume this from now on and use the constants in Theorem 11. Put

$$b = \min\{a, (\log 2)/2\}, \quad L = \max \left\{ 2, \left\lceil \frac{2 \log(8/b)}{\log 2} \right\rceil \right\}, \quad \kappa = b/L > 0.$$

The source theorem remains valid with the smaller rate b . Call a first-layer gate large if it contains more than L inputs. Let $\mathcal{G}$ be the set of large gates and let k_g be the number of inputs in $g \in \mathcal{G}$.

For a first-layer gate with k inputs, the signed pre-gate operator is

$$F_0 = \bigotimes_{j=1}^k \frac{\Xi_j}{2} \otimes \bigotimes_{q \in \text{anc}} |\alpha_q\rangle\langle\alpha_q|,$$

where each Ξ_j is a traceless Hermitian one-qubit unitary. These are the signed computational-basis operators after the initial local unitaries; the ancillary vectors also include those unitaries. The ancillary register may be absent. Tensor multiplicativity gives $\|F_0\|_\infty = 2^{-k}$ and $\|F_0\|_1 = 1$. Write the gate as $D = I - 2\Pi$, with $\Pi = |\vartheta\rangle\langle\vartheta|$ and $\|\vartheta\| = 1$, and put $v_\perp = (I - \Pi)F_0|\vartheta\rangle$. The signed post-gate operator has the binary expansion

$$DF_0D = F_0 + \Upsilon, \quad \Upsilon = -2(|\vartheta\rangle\langle v_\perp| + |v_\perp\rangle\langle\vartheta|).$$

Indeed, expand $DF_0D - F_0 = -2\Pi F_0 - 2F_0\Pi + 4\Pi F_0\Pi$ and cancel its Π -diagonal terms. Since $v_\perp \perp \vartheta$, the nonzero eigenvalues of Υ are $2\|v_\perp\|$ and $-2\|v_\perp\|$ when $v_\perp \neq 0$. Consequently

$$\|\Upsilon\|_1 = 4\|v_\perp\| \leq 4\|F_0\|_\infty = 4 \cdot 2^{-k}.$$

Thus $\Upsilon = 4 \cdot 2^{-k} \sigma_g$ for a Hermitian resource operator σ_g of trace norm at most one, with $\sigma_g = 0$ if $\Upsilon = 0$. This resource may be entangled within the component. It is independent of the remaining input variables: all k variables of its component have already been signed and averaged into F_0 . The baseline F_0 is the signed average for the component with its gate deleted: its k input factors are signed singleton sources. Its ancillary factors remain pure and product and can be included in the independent resource.

Expand only the large first-layer gates. In each original input-containing component with at most L inputs, keep one variable and condition the others classically. Explicitly, for such a component with k inputs and output states $\rho_{x,z}$, where x is the retained bit and $z \in \{0,1\}^{k-1}$, its signed operator is

$$2^{-(k-1)} \sum_z (-1)^{\sum_j z_j} \frac{\rho_{0,z} - \rho_{1,z}}{2}.$$

For fixed z , the conditioned wires are fixed pure product inputs, so $(\rho_{0,z} - \rho_{1,z})/2$ is the signed operator of a one-input source. The absolute averaging weights sum to one. Include untouched input singletons among these small components. Let r_0 be the number of small components and m_0 their total number of inputs, so $r_0 \leq m_0 \leq Lr_0$. For a fixed correction subset $\mathcal{S}$ of large gates, the remaining number of variable sources is exactly

$$r(\mathcal{S}) = r_0 + \sum_{g \in \mathcal{G} \setminus \mathcal{S}} k_g.$$

Group all corrected components, all input-free components, and the baseline ancillary factors into one resource. After extracting the factors $4 \cdot 2^{-k_g}$, this resource has trace norm at most one and

is tensor-independent of all remaining input sources. This remains true when the resource is not positive. Local unitaries after the first layer preserve the norms and all tensor separations used here.

The event factors over physical wires, so it factors over every new source-by-second-gate port partition, including after a large baseline component has split into singleton sources. This physical factorization is preserved when L_2 is absorbed into the event and is essential at this interface.

Let z collect all $m_0 - r_0$ conditioned bits, and let $q_{\mathcal{S},z}$ be the event functional of the resulting $r(\mathcal{S})$ one-input sources and their joint resource, after the factors 42^{-k_g} have been extracted. The circuit, resource, and event are fixed independently of these surviving variables. Linearity and the componentwise signed averages give the exact identity

$$\hat{q}([n]) = 2^{-(m_0-r_0)} \sum_{z \in \{0,1\}^{m_0-r_0}} (-1)^{\sum_j z_j} \sum_{S \subseteq \mathcal{G}} \left(\prod_{g \in S} 42^{-k_g} \right) \hat{q}_{\mathcal{S},z}([r(\mathcal{S})]).$$

Theorem 11 bounds each branch coefficient in absolute value by $K_0 e^{-br(\mathcal{S})}$. The absolute weights of the z -average sum to one. Taking absolute values and summing over correction subsets therefore gives

$$|\hat{q}([n])| \leq K_0 e^{-br_0} \prod_{g \in \mathcal{G}} (e^{-bk_g} + 42^{-k_g}). \quad (18)$$

This also covers branches with no surviving sources, by the $r = 0$ case of the source theorem.

For $k > L$, the chosen cutoff and $b \leq (\log 2)/2$ give

$$8e^{-(\log 2-b)k} \leq 82^{-k/2} \leq b \leq bk.$$

Consequently, using $1 + t \leq e^t$,

$$e^{-bk} + 42^{-k} \leq e^{-bk}(1 + bk/2) \leq e^{-bk/2}.$$

Moreover $e^{-br_0} \leq e^{-bm_0/L}$ and $b/L \leq b/2$. Since $n = m_0 + \sum_{g \in \mathcal{G}} k_g$, equation (18) is at most $K_0 e^{-\kappa n}$. Empty products, no large gates, and no small components are included without a separate normalization or limiting argument. $\square$

5.2 One measured output

Lemma 13 (One-output pullback). *Conjugating a one-qubit Hermitian contraction by a disjoint layer of product reflections expresses it as a linear combination of positive physical-qubit product events with total absolute coefficient mass at most six. Local unitaries before or after the layer preserve this claim.*

Proof. Absorb the output's final local unitary into A_o . Gates disjoint from o cancel, leaving at most one gate. If it is present, write

$$D = (I - P) \otimes I_o + P \otimes R_o, \quad P = \bigotimes_{j \in S \setminus \{o\}} |v_j\rangle\langle v_j|, \quad R_o = I_o - 2|v_o\rangle\langle v_o|.$$

Since $P(I - P) = 0$, direct multiplication gives

$$D(I \otimes A_o)D = I \otimes A_o - 2P \otimes B_o, \quad B_o = (A_o - R_o A_o R_o)/2. \quad (19)$$

Both A_o and B_o are Hermitian, and $\|B_o\|_\infty \leq (\|A_o\|_\infty + \|R_o A_o R_o\|_\infty)/2 \leq 1$. Their spectral decompositions into two rank-one qubit projectors have coefficient absolute sums at most two each,

yielding total mass at most $2 + 2 \cdot 2 = 6$. Every resulting factor is identity or a rank-one qubit projector, with identity on wires outside the gate. Degenerate eigenspaces may be resolved in any orthonormal basis. If no gate meets o , only the first spectral decomposition is needed. Finally, conjugation by a local layer preserves the physical-product form and positivity. $\square$

Proof of Theorem 1. Pad with empty layers if necessary. Pull the output observable through the final local layer and then the third multiqubit layer by Lemma 13. For $C = L_3 D_3 L_2 D_2 L_1 D_1 L_0$, the remaining prefix is $C^{(2)} = L_2 D_2 L_1 D_1 L_0$. Each resulting physical-product event has full-parity coefficient of absolute value at most $K_0 e^{-\kappa n}$ by Theorem 12. The triangle inequality gives $|c(C, A_o)| \leq 6K_0 e^{-\kappa n}$, and the success identity gives the stated advantage bound. Lemma 3 includes mixed product initialization. Appendix A provides the explicit constants. $\square$

5.3 Attaining exponential-order correlation

The following construction shows that the exponential-in- n order in Theorem 1 is attainable. Write $Z_o = |0\rangle\langle 0| - |1\rangle\langle 1|$ for computational-basis measurement on the output qubit.

Proposition 14. *For every $n \geq 1$, put $r = \lceil n/4 \rceil$. There is a zero-ancilla circuit C_n with at most three multiqubit layers and at most $n - r + 1$ multiqubit gates such that*

$$c(C_n, Z_o) = \begin{cases} 1, & r = 1, \\ 2^{2-r}, & r \geq 2. \end{cases}$$

Consequently, $c(C_n, Z_o) = \exp[-(\log 2)n/4 + O(1)]$ as $n \rightarrow \infty$.

Proof. Partition the inputs into r nonempty blocks of size at most four. In a block of size b , label the wires $q_1, \dots, q_b$ and store its parity on q_b using the following two-layer CNOT schedule; $i \rightarrow j$ denotes control q_i and target q_j , and a dash means no gate:

b	first layer	second layer
1	—	—
2	$1 \rightarrow 2$	—
3	$1 \rightarrow 2$	$2 \rightarrow 3$
4	$1 \rightarrow 2, 3 \rightarrow 4$	$2 \rightarrow 4$

The schedules use disjoint supports within each layer and run in parallel across blocks. Each CNOT, and the generalized Toffoli below, is one allowed multiqubit gate up to local Hadamards, by the equivalence in Section 2. The two preprocessing layers use $n - r$ gates in total and leave the block parities $z_1, \dots, z_r$ on the representative wires. If $r = 1$, measuring the representative already gives parity exactly.

For $r \geq 2$, apply a generalized Toffoli in the third layer with controls $z_1, \dots, z_{r-1}$ and target z_r . Thus the total gate count is $n - r + 1$. The target becomes $y = z_r \oplus \prod_{j < r} z_j$. Under uniform input, the block parities are independent uniform bits and their parity equals the input parity. Thus, before a possible output NOT,

$$\begin{aligned} \mathbb{E}_x(-1)^{\sum_i x_i + y} &= \mathbb{E}_z(-1)^{\sum_{j < r} z_j} \left(1 - 2 \prod_{j < r} z_j \right) \\ &= -2(-1)^{r-1} 2^{-(r-1)} = (-1)^r 2^{2-r}. \end{aligned}$$

The term without the product has mean zero, and the product is nonzero only when all $r - 1$ controls are one. An output NOT when r is odd makes the correlation positive; this one-qubit gate is free in the model. Measuring the target proves the formula and the gate count. Finally, $r = n/4 + O(1)$ gives the stated exponential order. $\square$

The construction matches the order of decay, without optimizing the exponential constant. It concerns uniform-input correlation: for $n \geq 9$ its correlation is less than one, so its deterministic output is incorrect on some inputs and its worst-case success probability is zero.

6 The exact depth-four gate lower bound

The shared event estimate also constrains exact computation through a fourth layer. We first retain one variable per source without increasing gate count, then use output determinism to obtain a certificate whose complexity depends on the gate count rather than on physical width.

6.1 Parity-preserving first-layer cleanup

We use the parity-preserving quantum restriction of [12, Lemma 4.1]. Its preparation of the fixed two-qubit states within the original first layer is also part of that proof. We give the parity-only argument with explicit circuit and resource accounting. Here a *restriction* may fix a pair of input wires to a superposition of computational-basis strings of the same parity; it is not necessarily an ordinary classical bit restriction.

Lemma 15 (Parity-preserving cleanup). *Let C be an exact n -input parity circuit in the model of Section 2, on N physical wires, with $n \geq 1$ and pure product initialization. There is a circuit C' on the same wires, with $\lceil n/3 \rceil \leq r \leq n$ variable inputs forming a subset of the original input wires, that computes r -bit parity or its complement exactly on the same designated output wire, with the same observable A_o , and does not increase entangling depth or multiqubit-gate count. Each first-layer component of C' contains at most one variable input. Components here are gate supports and individual idle wires, including untouched input singletons. All $N - r$ ancillary wires of C' are initialized in a fixed pure product state. Their number may exceed the original $N - n$: it is physical wire count, not ancillary wire count, that is preserved.*

Proof. Fix the first multiqubit layer D_1 , padding with an empty layer if the circuit has none. If D_1 is empty, every input wire is an idle singleton; leave the circuit unchanged and take $r = n$, which proves the lemma in that case. Assume henceforth that D_1 is nonempty.

Exactness permits fixed-parity substitutions. Use the output eigenvectors $|a_{\pm}\rangle$ from Section 2. For the original product ancillary state $|\alpha\rangle$, exactness gives

$$C(|x\rangle \otimes |\alpha\rangle) = |a_{\chi(x)}\rangle * o \otimes |\gamma_x\rangle.$$

Let $\mathcal{I} \subseteq [n]$ be a set of retained input positions. For a string t indexed by any subset of $[n]$, use $\chi(t) = (-1)^{\sum_j t_j}$, the sum running over the positions of t ; in particular the empty string has parity sign one. Suppose the other input positions are fixed to a unit vector $\Phi = \sum_t c_t |t\rangle$ supported on strings with $\chi(t) = \varsigma \in \{\pm 1\}$. All tensor factors are placed on their original physical wires. For $z \in \{0, 1\}^{\mathcal{I}}$, let $x(z, t)$ combine the retained bits z and fixed-position bits t , and set

$$\iota|z\rangle = \sum_t c_t |x(z, t)\rangle \otimes |\alpha\rangle.$$

This map is an isometry. Since $\chi(x(z, t)) = \chi(z)\chi(t) = \varsigma\chi(z)$ on the support of Φ , linearity gives

$$C\iota|z\rangle = |a * \varsigma\chi(z)\rangle * o \otimes \left(\sum_t c_t |\gamma * x(z, t)\rangle \right).$$

The parenthesized garbage vector has norm one because $C\iota$ is an isometry and the output factor is a unit vector. In particular, interference among the garbage vectors cannot spoil the deterministic output. Thus such a substitution preserves parity up to the fixed sign ς . The proof uses exactness, not merely a bounded-error guarantee on computational-basis inputs.

Deactivate gates without rotating the input parity basis. Write $C = C_{>1}D_1L_0$, where L_0 is the initial product of one-qubit unitaries. Set

$$\tilde{D}_1 = L_0^*D_1L_0, \quad C_{\text{suf}} = C_{>1}L_0, \quad C = C_{\text{suf}}\tilde{D}_1.$$

The gates of $\tilde{D}_1$ are disjoint product reflections with the same supports as those of D_1 . The factor L_0 has moved to the following local layer and can be merged with it; the computational input basis and the parity classes above have not been changed.

Consider a gate $G_g = I - 2|\theta_g\rangle\langle\theta_g|$ of $\tilde{D}_1$, where $|\theta_g\rangle$ is a product vector, and suppose its support contains $k \geq 3$ original input wires. Choose two, denoted i, j , and write their trigger factor as $|\theta_i\theta_j\rangle$. Put

$$\vartheta_{01} = \langle\theta_i\theta_j|01\rangle, \quad \vartheta_{10} = \langle\theta_i\theta_j|10\rangle.$$

Choose a unit vector $\phi_g = \alpha_g|01\rangle + \beta_g|10\rangle$ by

$$(\alpha_g, \beta_g) = \begin{cases} (\vartheta_{10}, -\vartheta_{01})/\sqrt{|\vartheta_{01}|^2 + |\vartheta_{10}|^2}, & (\vartheta_{01}, \vartheta_{10}) \neq (0, 0), \\ (1, 0), & (\vartheta_{01}, \vartheta_{10}) = (0, 0). \end{cases}$$

The coefficients are normalized, and $\vartheta_{01}\alpha_g + \vartheta_{10}\beta_g = 0$, including for complex $\vartheta_{01}, \vartheta_{10}$. The inner products defining the ϑ 's already contain the required conjugates. Thus ϕ_g is a unit vector of odd computational parity. The tensor-product form of the trigger projector implies

$$(|\theta_g\rangle\langle\theta_g|)(\phi_g \otimes \psi) = 0, \quad G_g(\phi_g \otimes \psi) = \phi_g \otimes \psi$$

for every vector ψ on the other wires in the gate. Thus fixing this pair to ϕ_g deactivates the entire gate, while all other $k - 2$ input wires remain variable. The selected pairs are disjoint because the original gate supports are disjoint.

For a first-layer gate with exactly two input wires, fix one of them to $|0\rangle$ and retain the gate on its original support; deactivating it would leave no variable input, whereas fixing one wire keeps one. It then has only one variable input; the fixed wire becomes an ancilla. Leave gates with one or zero input wires unchanged, and leave every idle input wire variable. In particular, an input-free gate contributes nothing to the count of variable inputs.

Prepare the pairs from product initialization in the same layer. For each deactivated gate define

$$|\omega_g\rangle = (\alpha_g|0\rangle + \beta_g|1\rangle) \otimes |1\rangle, \quad C_g = \text{CNOT}_{i \rightarrow j}.$$

The direct identity

$$C_g|\omega_g\rangle = \alpha_g|01\rangle + \beta_g|10\rangle = \phi_g$$

prepares the required pair from a product state. Moreover $C_g = H_j C Z_{\{i,j\}} H_j$, so it costs one allowed multiqubit gate and only free one-qubit unitaries.

Let $\mathcal{I}$ now be the surviving input positions. Let C be the product of these pair preparations, and let $\tilde{D}_{\text{keep}}$ be the product of the first-layer reflections that were not deactivated. Their supports are mutually disjoint and are disjoint from the prepared pairs. Let $|\Omega_z\rangle$ be the product input state with retained bits z , each selected pair in ω_g , each classically fixed input in $|0\rangle$, and all original ancillas still in α . If Φ is the tensor product of the ϕ_g and the classically fixed zero bits, then

$$C|\Omega_z\rangle = \iota|z\rangle, \quad \tilde{D}_1 \iota|z\rangle = \tilde{D}_{\text{keep}} \iota|z\rangle.$$

Consequently the circuit

$$C' = C_{\text{suf}} \tilde{D}_{\text{keep}} C \quad \text{satisfies} \quad C'|\Omega_z\rangle = C \iota|z\rangle$$

for every z . The state Φ has fixed parity $\varsigma = (-1)^{N_{\text{pair}}}$, where N_{pair} is the number of prepared pairs, so the preceding exactness argument applies. With the original observable A_o , the output sign is $\varsigma \chi(z)$, as required by the lemma. If desired, when $\varsigma = -1$, interchanging the two output labels gives parity itself without adding any gate.

The layer $\tilde{D}_{\text{keep}} C$ replaces each deactivated gate by one two-qubit gate on a subset of its original support and retains all other gates. It is therefore one entangling layer, with no increase in multiqubit-gate count. The local conjugations that implement its product reflections and CNOTs can be merged into the adjacent local layers. The suffix C_{suf} uses the original later multiqubit layers, so the full entangling depth does not increase. All physical wires, including the output wire even if it was one of the fixed inputs, are retained.

Each prepared pair is now an input-free first-layer component. All other wires of a deactivated gate are singleton components; in particular its surviving variable inputs are distinct sources. The undeactivated input-containing gates have at most one variable input each. Since Ω_z uses only fixed single-qubit ancillary states, initialization remains a pure product. This does not assert that the ancillary state after the first layer is product.

Count the retained inputs. Partition the original wires into first-layer gate supports and idle singletons. A component containing k input wires retains

$$r(k) = \begin{cases} 0, & k = 0, \\ 1, & k = 1 \text{ or } k = 2, \\ k - 2, & k \geq 3. \end{cases}$$

For every integer $k \geq 0$, $r(k) \geq k/3$; equality in the nontrivial $k \geq 3$ case occurs at $k = 3$. Since the component input counts sum to n , their retained counts sum to an integer $r \geq n/3$, hence $r \geq \lceil n/3 \rceil$. This also covers $n = 1, 2$ and arbitrary input-free or idle components. $\square$

6.2 Small-support observables after three layers

We convert a product of one-qubit test operators at the output of three layers into a linear combination of product events after two layers. The expansion is algebraic; the cleaned first-layer assumption is needed only when applying the event bounds.

Lemma 16 (Small-support pullback). *Consider a three-layer prefix $C^{(3)}$ in the model of Section 2, on $N \geq r \geq 1$ qubits with variable input $x \in \{0, 1\}^r$ and fixed pure product ancillas. Empty*

multiqubit layers and arbitrary intervening and boundary one-qubit gates are allowed. Suppose every input-containing first-layer component, including an untouched input singleton, contains exactly one variable input.

Let $s \geq 1$ be an integer and let S be a fixed set of at most s physical qubits. For each $q \in S$, let B_q be a fixed one-qubit contraction, independent of x , and put

$$B = \left(\bigotimes_{q \in S} B_q \right) \otimes I_{S^c}.$$

Write $\langle B \rangle_x = \text{Tr}(\rho_x^{(3)} B)$, where $\rho_x^{(3)}$ is the state after the entire prefix, and put $\chi(x) = (-1)^{x_1 + \dots + x_r}$. Then, with uniform averages over x ,

$$|\mathbb{E}_x \chi(x) \langle B \rangle_x| \leq 36^s K_0 e^{-ar},$$

with the absolute constants from Theorem 11. Neither Hermiticity of B nor exact parity computation by the prefix is assumed; the expectation may be complex.

Proof. Step 1: local gates and affected supports. Write

$$C^{(3)} = L_3 D_3 C^{(2)}, \quad C^{(2)} = L_2 D_2 L_1 D_1 L_0,$$

where the L_j are products of one-qubit unitaries and the D_j are disjoint controlled- Z layers. Let $\rho_x^{(2)}$ be the output state of $C^{(2)}$ and put $\tilde{B} = L_3^* B L_3$. Conjugation by L_3 preserves the one-qubit tensor form, the support bound, and the contraction bounds. Cyclicity of trace gives

$$\langle B \rangle_x = \text{Tr} \left[\rho_x^{(2)} D_3 \tilde{B} D_3 \right].$$

Set $t = |S| \leq s$. A third-layer gate disjoint from S conjugates identity and has no effect. If h_3 gates meet S , then $h_3 \leq t$, because the gate supports are pairwise disjoint. Supported qubits idle in this layer remain as individual tensor factors.

Step 2: expansion of the affected reflections. For an affected gate g with support $\text{supp}(g)$, write $D_g = I - 2\Pi_g$ with $\Pi_g = \bigotimes_{q \in \text{supp}(g)} \pi_q$, where each π_q is a rank-one orthogonal projector. For the restriction $\tilde{B}_g = \bigotimes_{q \in \text{supp}(g)} \tilde{B}_q$, with $\tilde{B}_q = I$ for $q \notin S$, the exact identity is

$$D_g \tilde{B}_g D_g = \tilde{B}_g - 2\Pi_g \tilde{B}_g - 2\tilde{B}_g \Pi_g + 4\Pi_g \tilde{B}_g \Pi_g.$$

After separating the scalar coefficients 1, -2, -2, 4, the four branches have one-qubit factors

$$\tilde{B}_q, \quad \pi_q \tilde{B}_q, \quad \tilde{B}_q \pi_q, \quad \pi_q \tilde{B}_q \pi_q,$$

respectively. Each factor has operator norm at most one by submultiplicativity. When $q \notin S$, these factors are I in the base branch and π_q in each correction branch. Thus the simultaneous expansion over all h_3 affected gates is a finite sum of physical-qubit product tensors with total coefficient absolute sum at most

$$(1 + 2 + 2 + 4)^{h_3} = 9^{h_3}.$$

The factors at the original t sites may be non-Hermitian contractions; every other factor is already an identity or a rank-one orthogonal projector. Zero terms may be omitted.

Step 3: decomposition of the remaining qubit factors. For any one-qubit contraction Y , write

$$Y = Y_{\Re} + iY_{\Im}, \quad Y_{\Re} = (Y + Y^*)/2, \quad Y_{\Im} = (Y - Y^*)/(2i).$$

Both $Y_{\Re}$ and $Y_{\Im}$ are Hermitian contractions, since $\|Y^*\|_{\infty} = \|Y\|_{\infty} \leq 1$ and the operator norm satisfies the triangle inequality. Choose orthonormal eigenbases separately to write

$$Y_{\Re} = \sum_{a=1}^2 \lambda_a P_a, \quad Y_{\Im} = \sum_{b=1}^2 \mu_b Q_b, \quad |\lambda_a|, |\mu_b| \leq 1,$$

where the P_a, Q_b are rank-one orthogonal projectors. Repeated eigenvalues cause no difficulty: choose an orthonormal basis of the eigenspace; zero coefficients may be discarded. Consequently Y is a linear combination of such projectors with coefficient absolute sum at most four. This need not be a convex combination, and the coefficients for $iY_{\Im}$ may be imaginary.

Apply this decomposition at each of the original t sites in each expanded tensor. Tensoring the decompositions multiplies their coefficient absolute sums, so

$$D_3 \tilde{B} D_3 = \sum_{\nu} c_{\nu} E_{\nu}, \quad \sum_{\nu} |c_{\nu}| \leq 9^{h_3} 4^t \leq 36^s,$$

where every E_{ν} is a physical-qubit product event: each factor is an identity or a rank-one orthogonal projector. The events and coefficients are fixed independently of x . For $t = 0$, there is just the identity event with coefficient one; the same estimates hold using empty-product conventions. No first-layer hypothesis was used in this algebraic expansion.

Step 4: application of the two-layer event bounds. For $q_{\nu}(x) = \text{Tr}(\rho_x^{(2)} E_{\nu})$, absorb the final local layer L_2 of $C^{(2)}$ into E_{ν} . This again preserves its physical-qubit product-event form. Grouping its one-qubit factors within each nonempty source-second-layer port produces a positive contraction on that port, exactly as required by Theorem 11.

The first-layer assumption supplies r independent one-input sources. All input-free first-layer components can be grouped into a density-operator resource tensor-independent of those sources and independent of x ; this follows from product initialization and first-layer disjointness. The intermediate local layer L_1 acts within these groups, so it preserves this separation. Thus every q_{ν} satisfies $|\hat{q}_{\nu}([r])| \leq K_0 e^{-ar}$. Finally, linearity of the uniform signed average gives

$$\mathbb{E}_x \chi(x) \langle B \rangle_x = \sum_{\nu} c_{\nu} \hat{q}_{\nu}([r]).$$

The triangle inequality and $\sum_{\nu} |c_{\nu}| \leq 36^s$ prove the claim. If S is empty, the coefficient is in fact zero because $r \geq 1$ and $\langle I \rangle_x = 1$. $\square$

6.3 The exact last-layer certificate

The following argument uses only the final layer and exact output determinism, not the depth or internal structure of the prefix. We use the identity/active-sector decomposition from [12, Section 2.3.1] to derive an explicit quantitative certificate on one- and two-qubit product observables.

Lemma 17 (Exact last-layer certificate). *Let a circuit in the model of Section 2 compute parity exactly on $n \geq 1$ input bits. Suppose the final multiqubit layer contains a gate incident to the output qubit o , with $b \geq 1$ other wires. Regard those wires as its controls in the product-reflection*

representation. Let ρ_x be the state at the output of the prefix preceding that layer, with any preceding local gates included, and put

$$\alpha(B) = 2^{-n} \sum_{x \in \{0,1\}^n} \chi(x) \text{Tr}(\rho_x B), \quad \chi(x) = (-1)^{x_1 + \dots + x_n}.$$

There is a fixed Hermitian product contraction B , independent of x and supported on at most two physical prefix-output qubits, such that

$$\alpha(B) \geq \frac{1}{2b+1}.$$

Here a product contraction means a tensor product of one-qubit contractions, with identity on every unspecified wire. If no gate in the final multiqubit layer touches o , there is instead a one-qubit Hermitian unitary B with $\alpha(B) = 1$.

Proof. Step 1: isolate the output gate. Absorb the terminal local rotation on o into its measurement observable A_o . Exactness and $n \geq 1$ make A_o a Hermitian unitary, by Section 2. All other terminal local gates act outside o and do not change its statistics. Within the final multiqubit layer, gate supports are disjoint. Writing that layer as UG , where G is the gate incident to o , the other gates U commute with both G and A_o . Hence

$$(UG)^* A_o (UG) = G A_o G.$$

Thus removing U does not change the output expectation, even when ρ_x is entangled across their supports.

Label the control wires by $1, \dots, b$, separately from the input indices. In the gate's product basis, write

$$P = \bigotimes_{j=1}^b P_j, \quad Q = I - P, \quad R_o = I_o - 2\pi_o, \quad G = Q \otimes I_o + P \otimes R_o,$$

where P_j and π_o are one-qubit rank-one orthogonal projectors. This representation is valid with any incident wire chosen as the output; no distinguished target of a Toffoli representation is needed. Suppressing identity tensor factors, set $A_G = G A_o G$. For every input, $y = \chi(x)$ satisfies

$$\text{Tr}(\rho_x A_G) = y, \quad H_x := I - y A_G \geq 0, \quad \text{Tr}(\rho_x H_x) = 0.$$

Indeed, A_G is a Hermitian unitary and $\text{Tr} \rho_x = 1$.

Step 2: exactness on each inactive control sector. For $Q_j = I - P_j$, extended by identity off control j , we have $Q_j P = 0$, $Q_j G = G Q_j = Q_j$, and $[Q_j, A_o] = 0$. Consequently Q_j commutes with both A_G and H_x and $Q_j A_G = Q_j A_o$. Since Q_j is an orthogonal projector,

$$0 \leq Q_j H_x \leq H_x.$$

Indeed, block decomposition with respect to Q_j writes H_x as the sum of the positive operators $Q_j H_x Q_j$ and $(I - Q_j) H_x (I - Q_j)$, which proves this order inequality. Positivity of ρ_x and $\text{Tr}(\rho_x H_x) = 0$ now give

$$0 = \text{Tr}(\rho_x Q_j H_x) = \text{Tr}(\rho_x Q_j) - y \text{Tr}(\rho_x A_o Q_j).$$

No projection of the state is performed or normalized here. Averaging the identity yields

$$q_j := \mathbb{E}_x \text{Tr}(\rho_x Q_j) = \alpha(A_o Q_j) \geq 0.$$

The commuting control projectors satisfy the operator union bound

$$Q = I - \prod_{j=1}^b (I - Q_j) \leq \sum_{j=1}^b Q_j.$$

To see this, simultaneously diagonalize the Q_j ; each eigenvalue is 0 or 1, and the indicator that at least one eigenvalue is 1 is at most the number of 1's. It follows that

$$q := \mathbb{E}_x \text{Tr}(\rho_x Q) \leq \sum_{j=1}^b q_j.$$

No statistical independence of the control outcomes is assumed.

Step 3: obtain the quantitative certificate. Put $A'_o = R_o A_o R_o$, a one-qubit Hermitian unitary. The orthogonal P, Q sectors give the exact identities

$$A_G = Q A_o + P A'_o, \quad A_G - A'_o = Q(A_o - A'_o).$$

Because $A_o - A'_o$ is Hermitian of norm at most two and acts only on o , it commutes with Q . Therefore

$$-2Q \leq A_G - A'_o \leq 2Q, \quad |\text{Tr}(\rho_x (A_G - A'_o))| \leq 2 \text{Tr}(\rho_x Q).$$

Since $\alpha(A_G) = \mathbb{E}_x y^2 = 1$, averaging with the parity sign gives

$$|1 - \alpha(A'_o)| \leq 2q.$$

All the coefficients here are real. Keeping their signs and using the bound on q , we obtain

$$1 \leq \alpha(A'_o) + 2q \leq \alpha(A'_o) + 2 \sum_{j=1}^b \alpha(A_o Q_j).$$

In particular, the following absolute-value form will also be useful:

$$1 \leq |\alpha(R_o A_o R_o)| + 2 \sum_{j=1}^b |\alpha(A_o Q_j)|. \tag{20}$$

The right-hand side of the signed inequality is at most

$$(2b + 1) \max\{\alpha(A'_o), \alpha(A_o Q_1), \dots, \alpha(A_o Q_b)\}.$$

Thus at least one of these coefficients is at least $1/(2b + 1)$, with its positive sign as claimed. The corresponding operator is either the one-qubit Hermitian unitary A'_o or a product $A_o Q_j$ of one-qubit Hermitian contractions on two distinct wires. All choices are fixed independently of the input.

If the final multiqubit layer has no gate on o , the same conjugation argument removes all its gates from the output expectation. Exactness then gives $\text{Tr}(\rho_x A_o) = \chi(x)$ and hence $\alpha(A_o) = 1$. This also covers an empty final layer; a circuit with no multiqubit layer can be padded with an empty layer. The proof uses neither purity of ρ_x nor a cleaned first layer; its positivity argument does require the pointwise exactness assumption. $\square$

Only the no-gate case and the identity-sector argument of Steps 2–3 are used below; Lemma 18 replaces the $1/(2b + 1)$ certificate by a gate-count inequality.

6.4 Control cells and completion of the lower bound

The preceding certificate has a prefactor depending on the number of controls of the final gate. We replace individual controls by blocks, called *cells*, determined by their first-layer source and their second- and third-layer incidences. Each cell lies in a single two-layer port and meets at most one third-layer gate. These two properties allow its entire control projector to be treated as one rank-one operator, with a cost independent of the number of physical qubits in the cell.

Lemma 18 (Cell certificate). *Consider an exact parity circuit in the model of Section 2, on $N \geq r \geq 1$ qubits with r variable inputs and fixed pure product ancillas. Write it with four multiqubit layers, allowing empty layers and arbitrary intervening and boundary one-qubit gates. Suppose every first-layer component that contains a variable input, including an idle variable-input singleton, contains exactly one such input; wires carrying fixed original inputs count as ancillas. Let $m_2, m_3 \geq 0$ be the numbers of multiqubit gates in layers two and three.*

There is a nonnegative integer N_{cell} , depending only on the circuit, such that

$$1 \leq [36 + 2664N_{\text{cell}}]K_0e^{-ar}, \quad N_{\text{cell}} \leq (r+1)(m_2+1)(m_3+1). \quad (21)$$

The constants are those of Theorem 11. If the fourth multiqubit layer does not touch the output, one may take $N_{\text{cell}} = 0$.

Proof. Step 1: fix the circuit stages and the event bound. Write

$$C = L_4 D_4 L_3 D_3 C^{(2)}, \quad C^{(2)} = L_2 D_2 L_1 D_1 L_0,$$

where the L_j are products of one-qubit unitaries and the D_j are disjoint controlled- Z layers. Let $\rho_x^{(2)}$ be the state after $C^{(2)}$, and let ρ_x be the state after $L_3 D_3 C^{(2)}$, immediately before D_4 . Absorb L_4 into the final output measurement, denoting its one-qubit Hermitian unitary by A_o , and set

$$\alpha(B) = \mathbb{E}_{x \in \{0,1\}^r} \chi(x) \text{Tr}(\rho_x B), \quad \chi(x) = (-1)^{x_1 + \dots + x_r},$$

with a uniform average.

Put $\beta = K_0 e^{-ar}$. Theorem 11 bounds the absolute full-parity coefficient of every admissible two-layer port-product event by β . Product initialization and first-layer disjointness give r independent one-input sources. Group all input-free first-layer components into one density-operator resource independent of these sources. Define its ports after grouping, as required by the theorem.

If D_4 does not touch o , Lemma 17 gives $\alpha(A_o) = 1$, and Lemma 16 with $s = 1$ gives $1 \leq 36\beta$. This proves the assertion with $N_{\text{cell}} = 0$. Henceforth let G be the gate of D_4 incident to o . All other gates in D_4 cancel from its output expectation, as in Lemma 17. Let $\mathcal{T}$ be the set of control wires of G , excluding o . Write

$$P = \bigotimes_{q \in \mathcal{T}} P_q, \quad Q = I - P, \quad G = Q \otimes I_o + P \otimes R_o,$$

where P_q are one-qubit rank-one orthogonal projectors and R_o is the induced one-qubit reflection.

Step 2: partition the controls into cells. For $1 \leq i \leq r$, let $V_{1,i}$ be the wire set of input source i , and let $V_{1,0}$ be the union of the input-free first-layer components; $V_{1,0}$ may be empty. For $\ell = 2, 3$, let $V_{\ell,j}$ be the support of gate j in layer ℓ , for $1 \leq j \leq m_\ell$, and let $V_{\ell,0}$ be the set of wires idle in that layer. Define the cells to be the nonempty sets

$$\mathfrak{c}_{ijk} = \mathcal{T} \cap V_{1,i} \cap V_{2,j} \cap V_{3,k}, \quad 0 \leq i \leq r, \quad 0 \leq j \leq m_2, \quad 0 \leq k \leq m_3.$$

Each wire has exactly one source label and one gate-or-idle label in each layer, so these cells partition $\mathcal{T}$. Their number satisfies

$$N_{\text{cell}} \leq (r+1)(m_2+1)(m_3+1).$$

For $j > 0$, the cell lies in the port $V_{1,i} \cap V_{2,j}$. For $j = 0$, it lies in the single private idle port $V_{1,i} \cap V_{2,0}$. The global idle label therefore does not merge different source ports. Likewise, the cell lies entirely in one third-layer gate if $k > 0$, and is entirely idle if $k = 0$. The output is excluded from every cell, even if it has the same three labels. Tensor-factor reorderings used below are not physical SWAP gates.

Step 3: apply the exact identity-sector argument to cells. For a cell $\mathfrak{c}$, put

$$P_{\mathfrak{c}} = \bigotimes_{q \in \mathfrak{c}} P_q, \quad Q_{\mathfrak{c}} = I_{\mathfrak{c}} - P_{\mathfrak{c}},$$

with identity extensions off $\mathfrak{c}$. The cells partition the controls, so $P = \prod_{\mathfrak{c}} P_{\mathfrak{c}}$ and $Q \leq \sum_{\mathfrak{c}} Q_{\mathfrak{c}}$. Put $A_G = GA_oG$ and $H_x = I - \chi(x)A_G$. Exactness gives $H_x \geq 0$ and $\text{Tr}(\rho_x H_x) = 0$. Each $Q_{\mathfrak{c}}$ commutes with A_o and G , annihilates P , and satisfies $Q_{\mathfrak{c}}G = GQ_{\mathfrak{c}} = Q_{\mathfrak{c}}$. The argument of Step 2 in the proof of Lemma 17, the block decomposition of H_x with respect to a commuting orthogonal projector, therefore applies verbatim to $Q_{\mathfrak{c}}$, regardless of the dimension of the cell. It gives

$$\alpha(A_o Q_{\mathfrak{c}}) = \mathbb{E}_x \text{Tr}(\rho_x Q_{\mathfrak{c}}) \geq 0.$$

Combining the operator union bound with the comparison $GA_oG - R_oA_oR_o = Q(A_o - R_oA_oR_o)$, exactly as in that lemma, yields

$$1 \leq |\alpha(R_oA_oR_o)| + 2 \sum_{\mathfrak{c}} |\alpha(A_o Q_{\mathfrak{c}})|. \quad (22)$$

This uses neither independence of the cells nor commutation of ρ_x with their projectors. It is a reuse of the proof's identity-sector argument, not an assertion that a cell is one physical qubit.

Step 4: pull back one whole-cell projector. We prove, for each fixed cell $\mathfrak{c}$,

$$|\alpha(A_o P_{\mathfrak{c}})| \leq 1296\beta.$$

Write $L_{3,q}$ for the one-qubit factor of L_3 on wire q , and put $L_{3,\mathfrak{c}} = \bigotimes_{q \in \mathfrak{c}} L_{3,q}$. Define

$$\tilde{A}_o = L_{3,o}^* A_o L_{3,o}, \quad \tilde{P}_{\mathfrak{c}} = L_{3,\mathfrak{c}}^* P_{\mathfrak{c}} L_{3,\mathfrak{c}}.$$

Then $L_3^*(A_o P_{\mathfrak{c}})L_3 = \tilde{A}_o \tilde{P}_{\mathfrak{c}}$, where $\tilde{A}_o$ is a one-qubit contraction and $\tilde{P}_{\mathfrak{c}}$ is still a product of one-qubit rank-one orthogonal projectors on $\mathfrak{c}$. Cyclicity of trace gives

$$\text{Tr}(\rho_x A_o P_{\mathfrak{c}}) = \text{Tr} \left[\rho_x^{(2)} D_3 (\tilde{A}_o \tilde{P}_{\mathfrak{c}}) D_3 \right].$$

The observable meets at most two gates in D_3 : one at o and one containing $\mathfrak{c}$. These gates may coincide, or either may be absent. All other third-layer gates conjugate identity.

For an affected gate $D_g = I - 2\Pi_g$, expand its conjugation as

$$D_g B D_g = B - 2\Pi_g B - 2B\Pi_g + 4\Pi_g B\Pi_g.$$

Keep the coefficients 1, -2, -2, 4 separate from the operator factors. If $h_3 \leq 2$ gates are affected, their simultaneous expansion has coefficient absolute sum at most $9^{h_3} \leq 81$. In each scalar-free

branch, the output factor is a one-qubit contraction. If $\mathfrak{c}$ is in gate g , let $\Pi_{\mathfrak{c}}$ be the tensor product of the trigger projectors of Π_g on $\mathfrak{c}$. The cell factor is one of

$$\tilde{P}_{\mathfrak{c}}, \quad \Pi_{\mathfrak{c}}\tilde{P}_{\mathfrak{c}}, \quad \tilde{P}_{\mathfrak{c}}\Pi_{\mathfrak{c}}, \quad \Pi_{\mathfrak{c}}\tilde{P}_{\mathfrak{c}}\Pi_{\mathfrak{c}}.$$

Each has rank at most one and norm at most one, since $\tilde{P}_{\mathfrak{c}}$, a tensor product of one-qubit rank-one orthogonal projectors, is a rank-one orthogonal projector on the cell register, and so is $\Pi_{\mathfrak{c}}$. If $\mathfrak{c}$ is idle, its factor is just $\tilde{P}_{\mathfrak{c}}$. Every physical-qubit factor outside $\{o\} \cup \mathfrak{c}$ is identity or a rank-one orthogonal projector. When the output and cell meet the same gate, the branch choices on them are synchronized, but each branch still factors over the disjoint blocks $\{o\}$, $\mathfrak{c}$, and the remaining physical wires.

For completeness, let Y be a rank-at-most-one contraction on an arbitrary finite-dimensional space. If $Y = 0$, omit the term. Otherwise write $Y = uv^*$ and put

$$Y_{\Re} = (Y + Y^*)/2, \quad Y_{\Im} = (Y - Y^*)/(2i).$$

Both are Hermitian contractions whose ranges are contained in $\text{span}\{u, v\}$. Hence each has at most two nonzero eigenvalues, all of absolute value at most one. Their separate spectral decompositions express $Y = Y_{\Re} + iY_{\Im}$ as a linear combination of rank-one orthogonal projectors with coefficient absolute sum at most four. Repeated eigenvalues admit an orthonormal eigenbasis; zero eigenvalues contribute nothing. The coefficients need not be positive or real. The one-qubit output contraction also admits coefficient mass at most four, by the decomposition in Lemma 16.

Applying these decompositions separately on the output and cell blocks gives a finite, input-independent expansion

$$D_3(\tilde{A}_o\tilde{P}_{\mathfrak{c}})D_3 = \sum_{\nu} d_{\nu}E_{\nu}, \quad \sum_{\nu} |d_{\nu}| \leq 81 \cdot 4 \cdot 4 = 1296.$$

A spectral projector on $\mathfrak{c}$ may be entangled across its physical wires. Nevertheless, $\mathfrak{c}$ lies in one source-second-layer port. For that port V , its factor in E_{ν} has the form $\Pi_{\nu, \mathfrak{c}} \otimes E_{\nu, V \setminus \mathfrak{c}}$, where $\Pi_{\nu, \mathfrak{c}}$ is the cell's rank-one spectral projector and $E_{\nu, V \setminus \mathfrak{c}}$ is a product of positive qubit contractions. This includes any output factor in the same port. The two factors act on disjoint registers, so their tensor product is a positive contraction. Every other port contains only positive qubit factors. Thus regrouping E_{ν} by the ports gives a tensor product of positive contractions, even when the cell projector is entangled. This also applies to ports on the grouped resource source.

To apply the two-layer bound, let τ_x be the state immediately before D_2 . Up to tensor-factor order, the first-layer separation gives

$$\tau_x = \sigma \otimes \bigotimes_{i=1}^r \rho_{i, x_i},$$

where σ is a fixed density operator and each source state ρ_{i, x_i} includes the local layer L_1 . Thus $\rho_x^{(2)} = L_2 D_2 \tau_x D_2^* L_2^*$. Set $F_{\nu} = L_2^* E_{\nu} L_2$. Local conjugation preserves positivity and port factorization, including any entangled cell projector, and

$$\text{Tr}(\rho_x^{(2)} E_{\nu}) = \text{Tr}(F_{\nu} D_2 \tau_x D_2).$$

The event theorem applied to F_{ν} therefore gives

$$\left| \mathbb{E}_x \chi(x) \text{Tr}(\rho_x^{(2)} E_{\nu}) \right| \leq \beta.$$

Linearity and the triangle inequality prove $|\alpha(A_o P_c)| \leq 1296\beta$. No physical-product-only event bound has been substituted for the port-product bound in this step.

Step 5: sum the cell certificate. Lemma 16 with $s = 1$ gives

$$|\alpha(A_o)|, |\alpha(R_o A_o R_o)| \leq 36\beta.$$

Since $\alpha(A_o Q_c) = \alpha(A_o) - \alpha(A_o P_c)$,

$$|\alpha(A_o Q_c)| \leq (36 + 1296)\beta = 1332\beta.$$

Substitution into (22) yields

$$1 \leq 36\beta + 2N_{\text{cell}}(1332\beta) = [36 + 2664N_{\text{cell}}]\beta.$$

This proves (21). $\square$

Proof of Theorem 2. By Lemma 3, it suffices to use pure product initialization. Let C be an exact n -input circuit as in the theorem, with M multiqubit gates. Apply Lemma 15, obtaining a product-initialized circuit on the same physical wires with $\lceil n/3 \rceil \leq r \leq n$ variable inputs and at most M multiqubit gates. If its task is the complement of parity, reverse the output labels, replacing A_o by $-A_o$. This changes neither depth nor gate count. Pad with empty layers at the end if necessary to obtain four designated multiqubit layers. The first-layer source condition is preserved, and the second- and third-layer gate counts satisfy $m_2, m_3 \leq M$.

The exponential cell certificate gives

$$1 \leq [36 + 2664(r+1)(M+1)^2]K_0 e^{-ar} \leq (M+1)^2 K_0 [36 + 2664(r+1)]e^{-ar}.$$

The second inequality uses $(M+1)^2 \geq 1$. All quantities are positive, so taking natural logarithms yields

$$\begin{aligned} 2 \log(M+1) &\geq ar - \log(K_0 [36 + 2664(r+1)]) \\ &\geq \frac{an}{3} - \log(K_0 [36 + 2664(n+1)]). \end{aligned}$$

The last bound uses $r \geq n/3$ in the linear term and $r \leq n$ in the logarithmic term, so it is uniform over all permitted r .

We now exhibit an admissible explicit value of n_0 . Put $C_0 = 2700K_0$ and choose

$$n_0 = \left\lceil \max \left\{ 1, \frac{12 \log C_0}{a}, \frac{144}{a^2}, \frac{24 \log 2}{a} \right\} \right\rceil.$$

For $n \geq 1$, $36 + 2664(n+1) \leq 2700(n+1)$ and $\log(n+1) \leq \sqrt{n}$. The latter follows since $\sqrt{t} - \log(1+t)$ vanishes at $t = 0$ and has derivative $(\sqrt{t} - 1)^2 / [2\sqrt{t}(1+t)] \geq 0$ for $t > 0$. For $n \geq n_0$, its definition ensures $\log C_0 \leq an/12$ and $\sqrt{n} \leq an/12$. Consequently,

$$\log(K_0 [36 + 2664(n+1)]) \leq \log C_0 + \sqrt{n} \leq \frac{an}{6}.$$

It follows that $2 \log(M+1) \geq an/6$, hence $M+1 \geq e^{an/12}$. Finally, $e^{an/24} \geq 2$ for $n \geq n_0$, and $\theta^2 - 1 \geq \theta$ for every real $\theta \geq 2$; hence

$$M \geq e^{an/12} - 1 = (e^{an/24})^2 - 1 \geq e^{an/24}.$$

Thus Theorem 2 holds with $c_* = a/24$ and the displayed n_0 . These constants depend only on the absolute constants a, K_0 , not on physical wire count or gate arities. No bound on the number of ancillary wires was used. $\square$

A Explicit constants for both main theorems

The shared source theorem admits explicit constants without changing either circuit interface. The following choices are conservative.

Proposition 19. *Theorem 11 holds with $a = 2^{-21}$ and an absolute constant K_0 satisfying $1 \leq K_0 < 300$. Consequently every circuit in Theorem 1 satisfies*

$$|c(C, A_o)| \leq \min\{1, 1800e^{-2^{-27}n}\}.$$

Theorem 2 holds with $c_ = 2^{-21}/24$ and $n_0 = 144 \cdot 2^{42} = 633318697598976$.*

Proof. Recall $C_s = 1/2 + 2/e$ and the quantities used in the proof of Theorem 11:

$$c_1 = \frac{1}{128C_s}, \quad b_1 = c_1 - \eta \log(8eC_s/\eta), \quad c_\eta = \frac{1}{4 \log(4/\eta)}.$$

Take $\eta = 2^{-12} < C_s$ and $w = 512 \geq 1$. The elementary bounds $27/10 < e < 3$ and $\log 2 < 7/10$ give $C_s < 5/4$ and $8eC_s/\eta < 30 \cdot 4096 < (27/10)^{12} < e^{12}$. Hence

$$c_1 > \frac{1}{160}, \quad \log(8eC_s/\eta) < 12, \quad b_1 > \frac{1}{160} - 12 \cdot 2^{-12} > \frac{33}{10000}.$$

In particular $b_1 > 0$ and $b_1/2 > \eta$. Also $c_\eta = 1/(56 \log 2) > 5/196$, so

$$\log(8eC_s/\eta) - c_\eta w < 12 - 640/49 = -52/49 < -1.$$

Thus both parameter conditions in the source theorem's proof hold. Equation (17) gives

$$a = 2^{-21}, \quad K_0 < 3 + \frac{5/2}{(27/10)(33/10000)} < 300.$$

Here $1 + 1/(1 - e^{-1}) < 3$, since $e > 2$, which justifies the prefactor estimate. In Theorem 12, $b = \min\{a, (\log 2)/2\} = a$, and the integer cutoff and rate are

$$L = \max\{2, \lceil 2 \log_2(8/a) \rceil\} = 48, \quad \kappa = a/L = 2^{-21}/48 > 2^{-27}.$$

The one-output pullback contributes coefficient mass at most six, so $6K_0e^{-\kappa n} \leq 1800e^{-2^{-27}n}$. Combining this with $|c(C, A_o)| \leq 1$ proves the correlation bound.

For depth four, Section 6 gives $c_* = a/24$ and

$$n_0 = \left\lceil \max \left\{ 1, \frac{12 \log(2700K_0)}{a}, \frac{144}{a^2}, \frac{24 \log 2}{a} \right\} \right\rceil.$$

Since $2700K_0 < 810000 < (27/10)^{14} < e^{14}$ and $a = 2^{-21} < 1/2$, each other entry in the maximum is strictly smaller than $144/a^2$. This gives the stated integer n_0 . These conservative constants are uniform over the finite widths and gate arities allowed by the model. $\square$

B The bounded-differences inequality

We record both the tail and exponential-moment forms of the bounded-differences inequality [14]. The proof uses the centered bounded-variable estimate of Hoeffding [10] and is included to specify the constants used in the network argument.

Lemma 20 (Bounded differences and exponential moments). *Let $m \geq 0$ be an integer, and let $X_1, \dots, X_m$ be independent random variables, where X_j takes values in a finite set $\mathcal{X}_j$, every point of which has positive probability. Let $\Lambda : \prod_{j=1}^m \mathcal{X}_j \rightarrow \mathbb{R}$ satisfy $|\Lambda(x) - \Lambda(x')| \leq c_j$ whenever x, x' differ only in coordinate j , where $c_j \geq 0$. Put $V = \sum_{j=1}^m c_j^2$. Write Λ also for the random variable $\Lambda(X_1, \dots, X_m)$. Then, for every real t ,*

$$\mathbb{E}e^{t(\Lambda - \mathbb{E}\Lambda)} \leq e^{t^2 V/8}. \quad (23)$$

If $V > 0$, then for every $u > 0$,

$$\mathbb{P}[\Lambda - \mathbb{E}\Lambda \geq u] \leq \exp\left(-\frac{2u^2}{V}\right).$$

If $V = 0$, then Λ is constant on the product support and this upper-tail probability is zero. These statements include $m = 0$, when the product support is a singleton and $V = 0$.

Proof. First let Y be a centered finite random variable in an interval of length c_Y . Its log moment generating function $\psi_Y(t) = \log \mathbb{E}e^{tY}$ satisfies $\psi_Y(0) = \psi'_Y(0) = 0$. Differentiating the finite sum gives $\psi''_Y(t) = \text{Var}_t(Y)$, where the subscript denotes the probability distribution tilted by the positive factor e^{tY} . Its support remains in the same interval. If y_{mid} is the interval midpoint, then $\text{Var}_t(Y) \leq \mathbb{E}_t(Y - y_{\text{mid}})^2 \leq c_Y^2/4$. Integrating twice gives $\psi_Y(t) \leq t^2 c_Y^2/8$ for $t \geq 0$; applying this to $-Y$ gives the same bound for $t < 0$.

Let $\mathcal{F}_j = \sigma(X_1, \dots, X_j)$, with $\mathcal{F}_0$ trivial, and set $\Lambda_j = \mathbb{E}[\Lambda \mid \mathcal{F}_j]$. Then $\Lambda_0 = \mathbb{E}\Lambda$ and $\Lambda_m = \Lambda$. Fix preceding values $x_{<j} = (x_1, \dots, x_{j-1})$ and put $g_j(z) = \mathbb{E}_{X_{>j}} \Lambda(x_{<j}, z, X_{>j})$, where $X_{>j} = (X_{j+1}, \dots, X_m)$; an empty tuple is allowed. Independence makes this the same future law for every z , so

$$|g_j(z) - g_j(z')| \leq \mathbb{E}_{X_{>j}} |\Lambda(x_{<j}, z, X_{>j}) - \Lambda(x_{<j}, z', X_{>j})| \leq c_j.$$

On this conditioning event, $\Lambda_j = g_j(X_j)$ and $\Lambda_{j-1} = \mathbb{E}_{X_j} g_j(X_j)$ is fixed. Thus $Y_j = \Lambda_j - \Lambda_{j-1}$ has conditional mean zero and lies in an interval of length at most c_j . The centered-variable estimate, applied separately to each conditional law, gives for every real t

$$\mathbb{E}[e^{tY_j} \mid \mathcal{F}_{j-1}] \leq e^{t^2 c_j^2/8}.$$

For $S_j = \sum_{k=1}^j Y_k$, S_{j-1} is $\mathcal{F}_{j-1}$ -measurable. The tower property therefore gives

$$\mathbb{E}e^{tS_j} = \mathbb{E}\left[e^{tS_{j-1}} \mathbb{E}(e^{tY_j} \mid \mathcal{F}_{j-1})\right] \leq e^{t^2 c_j^2/8} \mathbb{E}e^{tS_{j-1}}.$$

Iterating from $S_0 = 0$ and using $S_m = \Lambda - \mathbb{E}\Lambda$ proves (23). For $V > 0$, exponential Markov gives $\mathbb{P}[\Lambda - \mathbb{E}\Lambda \geq u] \leq \exp(-tu + t^2 V/8)$ for $t > 0$; choose $t = 4u/V$. If $V = 0$, all coordinate oscillations vanish. Any two points of the product support can be joined by successive single-coordinate changes, so Λ is constant there. The empty product case has the same conclusion directly. $\square$

Use of AI: The author prepared an eight-page draft without AI assistance, outlining previous QAC lower bounds and several high-level strategies to extend them for the exact and bounded-error parity QAC lower bounds at several depths. The author then fed this draft to multiple AI tools (Claude, ChatGPT and two open-weight models with open-source harnesses) and prompted them to work out the details of the proofs and identify missing gaps. The author fixed those missing gaps and iteratively verified and expanded the proofs using AI tools. The author is solely responsible for the correctness of the proofs.

References

- [1] Anurag Anshu, Yangjing Dong, Fengning Ou, and Penghui Yao. On the computational power of QAC^0 with barely superlinear ancillae. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1476–1487, 2025. doi:10.1145/3717823.3718189; full version: <https://arxiv.org/abs/2410.06499v4>.
- [2] Debajyoti Bera. A lower bound method for quantum circuits. *Information Processing Letters*, 111(15):723–726, 2011. doi:10.1016/j.ipl.2011.05.002.
- [3] Debajyoti Bera, Frederic Green, and Steven Homer. Small depth quantum circuits. *ACM SIGACT News*, 38(2):35–50, 2007. doi:10.1145/1272729.1272739.
- [4] Yangjing Dong, Fengning Ou, and Penghui Yao. Linear-size QAC^0 channels: learning, testing and hardness. arXiv:2510.00593, 2025; version 2. <https://arxiv.org/abs/2510.00593v2>.
- [5] Maosen Fang, Stephen A. 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] 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>.
- [7] 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. doi:10.26421/QIC2.1-3.
- [8] Lucas Gretta, Meghal Gupta, and Malvika Raj Joshi. Parity $\notin \text{QAC}^0 \iff \text{QAC}^0$ is Fourier-concentrated. arXiv:2604.02793, 2026; version 2. <https://arxiv.org/abs/2604.02793v2>.
- [9] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, pages 6–20, 1986. doi:10.1145/12130.12132.
- [10] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:10.1080/01621459.1963.10500830.
- [11] Peter Høyer and Robert Špalek. Quantum fan-out is powerful. *Theory of Computing*, 1(5):81–103, 2005. doi:10.4086/toc.2005.v001a005.
- [12] Malvika Raj Joshi, Avishay Tal, Francisca Vasconcelos, and John Wright. Improved lower bounds for QAC^0 . In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing*, pages 2199–2209, 2026. doi:10.1145/3798129.3800922; full version, arXiv:2512.14643, version 4: <https://arxiv.org/abs/2512.14643v4>.
- [13] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, Fourier transform, and learnability. *Journal of the ACM*, 40(3):607–620, 1993. doi:10.1145/174130.174138.
- [14] Colin McDiarmid. On the method of bounded differences. In J. Siemons, editor, *Surveys in Combinatorics, 1989*, London Mathematical Society Lecture Note Series 141, pages 148–188. Cambridge University Press, 1989. doi:10.1017/CBO9781107359949.008.
- [15] Cristopher Moore. Quantum circuits: fanout, parity, and counting. ECCC Report TR99-032, 1999. <https://eccc.weizmann.ac.il/report/1999/032/>.
- [16] Shivam Nadimpalli, Natalie Parham, Francisca Vasconcelos, and Henry Yuen. On the Pauli spectrum of QAC^0 . In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1498–1506, 2024. doi:10.1145/3618260.3649662; full version: <https://arxiv.org/abs/2311.09631v4>.

- [17] Gregory Rosenthal. Bounds on the QAC^0 complexity of approximating parity. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, LIPIcs 185, Article 32, 2021. [doi:10.4230/LIPIcs.ITCS.2021.32](https://doi.org/10.4230/LIPIcs.ITCS.2021.32).
- [18] Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. *Computational Complexity*, 25(4):849–881, 2016. [doi:10.1007/s00037-016-0140-0](https://doi.org/10.1007/s00037-016-0140-0).