

A simplified proof of $\text{PARITY} \notin \text{QAC}^0$

Minbo Gao^{*1,2}, Zhengfeng Ji^{†3}, and Chenghua Liu^{‡1,2}

¹Institute of Software, Chinese Academy of Sciences, Beijing, China

²University of Chinese Academy of Sciences, Beijing, China

³Department of Computer Science and Technology, Tsinghua University, Beijing, China

Abstract

We give a simplified, self-contained proof of OpenAI’s parity lower bound for constant-depth quantum circuits with arbitrary one-qubit gates, unbounded-arity Toffoli gates, and polynomially many ancillas. The proof establishes product-projection localization and uses it to bound the parity Fourier coefficient of the acceptance probability. We remove one circuit layer directly, avoiding intermediate low-count propagation. This yields a leakage recurrence that closes by induction over depth and threshold scales.

1 Introduction

The classical lower bound $\text{PARITY} \notin \text{AC}^0$ of Furst, Saxe, and Sipser [FSS84] and, independently, Ajtai [Ajt83] established parity as a central test for shallow circuits. Here AC^0 consists of polynomial-size, constant-depth circuits with unbounded-fan-in AND and OR gates and NOT gates. Håstad [Hås86] strengthened this separation using the switching lemma, obtaining the almost optimal size lower bound $\exp(\Omega(n^{1/(d-1)}))$ for parity at every fixed depth $d \geq 2$. Moore [Moo99] asked whether the analogous separation holds for QAC^0 , whose circuits use arbitrary one-qubit unitaries and unbounded-arity Toffoli gates. Let $\mathcal{U}_d(N)$ denote circuits on N qubits with at most d layers of pairwise disjoint Toffoli gates, interleaved with arbitrary one-qubit unitaries. Counting the one-qubit layers as well gives depth at most $2d + 1$, so this convention preserves the class of constant-depth circuits. We emphasize that the model has no fanout primitive. Unbounded fanout gives a powerful extension of this model. Høyer and Špalek [HŠ05] showed that polynomial-size, constant-depth quantum circuits with unbounded fanout can approximate parity, threshold, and counting operations with polynomially small error.

Early progress on Moore’s question quantified the tradeoff between depth and ancillary space. Fang, Fenner, Green, Homer, and Zhang [FFGHZ06] obtained a depth bound $\Omega(\log(n/(a+1)))$ for clean, exact parity with a ancillas, counting one-qubit layers. Rosenthal [Ros21] developed reductions between coherent parity approximation and generalized cat-state preparation, connecting circuit lower bounds to the preparation of quantum states.

Spectral and polynomial methods gave lower bounds for circuits whose answer is supplied by a single measured output. Nadimpalli, Parham, Vasconcelos, and Yuen [NPVY24] established low-degree concentration for quantum circuit channels and derived parity lower bounds. Anshu, Dong, Ou, and Yao [ADOY25] used operator-norm approximation to prove lower bounds with barely-superlinear numbers of ancillas. Dong, Ou, and Yao [DOY25] further improved these bounds through sharper polynomial approximations.

^{*}gmb17@tsinghua.org.cn

[†]jizhengfeng@tsinghua.edu.cn

[‡]liuch.russell@gmail.com

At small depths, a complementary line of work established bounds independent of ancillary space. Fenner, Grier, Padé, and Thierauf [FGPT25] exclude exact single-output parity at depth two. Joshi, Tal, Vasconcelos, and Wright [JTVW26, Theorem 1.2 and Corollary 5.9] extend the exact lower bound to depth three and bound the depth-two parity correlation by $\exp(-\Omega(\sqrt{n}))$, regardless of size and ancilla count. Kintali’s preprint [Kin26, Theorems 1–2] reports exponential parity correlation decay at depth three, independently of size and ancilla count, and an exponential lower bound on the number of multi-qubit gates for exact parity at depth four.

Recently, OpenAI [Ope26a, Ope26b] used its internal AI models to resolve Moore’s question, proving a parity lower bound for QAC^0 circuits of arbitrary constant depth with polynomially many ancillas. We state this result using the depth convention above.

Theorem 1 (Parity lower bound [Ope26a, Theorem 1]). *Fix an integer $d \geq 0$, a real $c \geq 1$, and $0 < \varepsilon \leq 1/2$. For all sufficiently large n , every circuit in $\mathcal{U}_d(N)$ with n input qubits and $n \leq N \leq n^c$ total qubits has an input $x \in \{0, 1\}^n$ on which its probability of outputting parity is less than $1/2 + \varepsilon$. The ancillas start in $|0\rangle$, and the answer is obtained by measuring one designated output qubit in the computational basis. The remaining qubits may be discarded in arbitrary final states.*

This paper gives a self-contained simplification of OpenAI’s localization proof. Our aim is to better understand the AI-generated proofs and make the argument more accessible. The technical proof occupies about five pages. We replace the low-count enumeration and three-factor propagation in [Ope26a, Lemmas 5–6] with a direct layer-removal argument. Together with the original inverse-circuit argument, this gives a single leakage recurrence. We use the original zero-count polynomial, parity reduction, and induction over threshold scales.

For a circuit U , let $p(x)$ be its probability of outputting one on input x . Its parity success probability on a uniform input is $1/2 - \hat{p}([n])$, where $\hat{p}([n])$ is the full-parity Fourier coefficient. After a Hadamard change of input basis, this coefficient is an average of matrix elements of $U^\dagger B U$ between product states that are orthogonal on every input wire, with B the output measurement projection. To bound these matrix elements, we study *product projections*, which fix chosen wires to chosen one-qubit states and act as identity elsewhere. These projections also arise naturally from the gates. A Toffoli gate with control set S and target o is

$$I - 2 \left(\bigotimes_{i \in S} |1\rangle\langle 1|_i \right) \otimes |-\rangle\langle -|_o, \quad |\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}.$$

Local basis changes, absorbed into adjacent one-qubit layers, give reflections $I - 2|v\rangle\langle v|$ for any unit product vector $|v\rangle$ on the gate’s wires. For a product projection B and a mismatch-count operator M , we study the leakage amplitude

$$\|M_{\geq r} U^\dagger B U M_0\|, \tag{1}$$

where M counts mismatches with a chosen product state, M_0 projects onto the product state, and $M_{\geq r}$ projects onto states with at least r mismatches. The localization theorem bounds this amplitude uniformly over M, B . Since the matrix elements defining $\hat{p}([n])$ involve n mismatches, localization makes this coefficient small and bounds parity success by $1/2 + o(1)$.

The layer-removal step uses the product structure of these reflections. A product test A on k wires meets at most k gates in the last disjoint layer. Expanding those gates on both sides of A costs at most 9^k in total absolute coefficient. Each term is already a product projection outside the support of A . Each of the at most k remaining one-qubit factors is a linear combination of projections with total absolute coefficient at most two. Thus $U^\dagger A U$ becomes a linear combination of conjugated product tests at depth $d - 1$, with total absolute coefficient at most $9^k 2^k = 18^k$, independently of the gate widths. This uses the outside-support factorization of [Ope26b, Proposition 4.1] to remove a layer without intermediate low-count propagation.

To obtain the leakage recurrence, write $B = D_0$ for another mismatch-count operator D , and approximate B by a polynomial $f(D)$ that equals one at zero and is at most ε in magnitude at positive integer counts below a threshold R . A monomial of degree j expands into product tests on at most j wires. Layer removal therefore bounds the polynomial contribution by its expansion cost times the leakage at depth $d - 1$. To control the approximation error, we first square the leakage norm and replace only one copy of $Q = U^\dagger B U$ by $U^\dagger f(D) U$. The remaining exact copy of Q lets us bound the error at counts at least R by the same expansion cost times the leakage of the inverse circuit, with M_0 and B exchanged. This gives [Eq. \(10\)](#), combining the error ε , leakage at depth $d - 1$, and leakage at depth d at the larger threshold R . The zero-count polynomial has sufficiently small expansion cost for both leakage bounds to absorb it. At depth zero, the product structure gives the required bound directly. At each subsequent depth, leakage is zero above the maximum count N . This starts a descending induction over threshold scales and completes the proof of localization.

2 Reducing parity to localization

For a set S of wires and normalized one-qubit states $|u_i\rangle$, define the operator

$$M = \sum_{i \in S} \Pi_i, \quad \Pi_i = I - |u_i\rangle\langle u_i|. \quad (2)$$

Identities on all other wires are implicit. The projections Π_i act on distinct wires and therefore commute. In their joint eigenbasis, each contributes zero when its local constraint is satisfied and one when it is violated. Thus the eigenvalues of M are the integers from 0 to $|S| \leq N$, recording the number of mismatches; we call M the *mismatch-count operator*.

For a real number γ , let M_γ be the orthogonal projection onto the eigenspace of M with eigenvalue γ , with $M_\gamma = 0$ if γ is not an eigenvalue. Thus M_γ selects states with exactly γ mismatches. We also write

$$M_{\geq \gamma} = \sum_{j \geq \gamma} M_j, \quad M_0 = \bigotimes_{i \in S} |u_i\rangle\langle u_i|.$$

Here $M_{\geq \gamma}$ selects states with at least γ mismatches. Every product projection is M_0 for some mismatch-count operator M . We use the same subscript notation for spectral projections of other Hermitian operators. Throughout, $\|\cdot\|$ is the operator norm and logarithms are natural.

Theorem 2 (Product-projection localization [[Ope26a](#), Theorem 2]). *For every fixed integer $d \geq 0$ and real exponents $0 \leq s < t$, there is $N_0(d, s, t)$ such that, for all $N \geq N_0$, every $U \in \mathcal{U}_d(N)$, product projection B , and mismatch-count operator M satisfy*

$$\|M_{\geq N^t} U^\dagger B U M_0\| \leq e^{-N^s}. \quad (3)$$

The estimate acts on the whole zero-count subspace. In particular, it allows arbitrary entanglement on the uncounted register. This is what permits the two product constraints to exchange roles in the proof.

Proof of Theorem 1 from Theorem 2. Let $C \in \mathcal{U}_d(N)$, write $a = N - n$, and absorb the output measurement basis into the final local layer. For input $x \in \{0, 1\}^n$, the probability of measuring one on output qubit o is

$$p(x) = \|\lvert 1 \rangle \langle 1 \rvert_o C(\lvert x \rangle \otimes \lvert 0^a \rangle)\|^2 = \langle x | Q | x \rangle,$$

where the acceptance operator on the input register is

$$Q = (I \otimes \langle 0^a |) C^\dagger \lvert 1 \rangle \langle 1 \rvert_o C (I \otimes \lvert 0^a \rangle).$$

For a uniform input with $n \geq 1$, the parity success probability is

$$2^{-n} \left(\sum_{|x| \text{ even}} (1 - p(x)) + \sum_{|x| \text{ odd}} p(x) \right) = \frac{1}{2} - \hat{p}([n]),$$

where $\hat{p}([n]) = 2^{-n} \sum_x (-1)^{|x|} p(x)$ is the full-parity Fourier coefficient. It therefore suffices to bound $|\hat{p}([n])|$. Let H be the Hadamard gate and X, Z the Pauli matrices. Since $HZH = X$ and $H^2 = I$, we have

$$\begin{aligned} \hat{p}([n]) &= 2^{-n} \text{Tr}(Z^{\otimes n} Q) \\ &= 2^{-n} \text{Tr}(X^{\otimes n} H^{\otimes n} Q H^{\otimes n}) \\ &= 2^{-n} \sum_{x \in \{0,1\}^n} \langle \bar{x} | H^{\otimes n} Q H^{\otimes n} | x \rangle, \end{aligned} \tag{4}$$

where $\bar{x}$ is the bitwise complement of x . The second line uses cyclicity of the trace, and the last uses $X^{\otimes n} |x\rangle = |\bar{x}\rangle$.

For each x , let $|\psi_x\rangle = H^{\otimes n} |x\rangle \otimes |0^a\rangle$. Expanding the definition of Q gives

$$\langle \bar{x} | H^{\otimes n} Q H^{\otimes n} | x \rangle = \langle \psi_{\bar{x}} | C^\dagger | 1 \rangle \langle 1 |_o C | \psi_x \rangle.$$

Count mismatches with the input states $H|x_i\rangle$ and leave the ancillas uncounted. The states $H|x_i\rangle$ and $H|1 - x_i\rangle$ are orthogonal on every input wire, so $|\psi_x\rangle$ lies in the range of M_0 and $|\psi_{\bar{x}}\rangle$ lies in the range of M_n . Choose fixed $0 < s < t < 1/c$. Since $N \leq n^c$, eventually $N^t < n$, and [Eq. \(3\)](#) bounds every summand in [Eq. \(4\)](#) by e^{-N^s} . The normalized average incurs no dimension factor. Thus $|\hat{p}([n])| \leq e^{-N^s}$. The uniform-average success probability is at most $1/2 + e^{-N^s} = 1/2 + o(1)$. For sufficiently large n , this is less than $1/2 + \varepsilon$, so some input has success probability below $1/2 + \varepsilon$. $\square$

3 Deriving the leakage recurrence

Fix the width N . The operator $M_{\geq r} U^\dagger B U M_0$ describes leakage into states with at least r mismatches; its norm is the *leakage amplitude*. Define the worst-case leakage profile by

$$L_d(r) = \sup_{U, B, M} \|M_{\geq r} U^\dagger B U M_0\|, \quad U \in \mathcal{U}_d(N), \tag{5}$$

where B ranges over product projections and M over mismatch-count operators. Always $0 \leq L_d(r) \leq 1$, and $L_d(r) = 0$ for $r > N$. We suppress the dependence on N .

Lemma 3 (Direct layer removal). *Let $d \geq 1$, $U \in \mathcal{U}_d(N)$, and let A be a product projection supported on at most k wires. There are a circuit $V \in \mathcal{U}_{d-1}(N)$, product projections B_α , and complex numbers c_α such that*

$$U^\dagger A U = \sum_\alpha c_\alpha V^\dagger B_\alpha V, \quad \sum_\alpha |c_\alpha| \leq 18^k. \tag{6}$$

Consequently, for every mismatch-count operator M and $r > 0$,

$$\|M_{\geq r} U^\dagger A U M_0\| \leq 18^k L_{d-1}(r). \tag{7}$$

Proof. Write $U = LGV$, with L local and G the last reflection layer. Replace A by $L^\dagger A L$, preserving its product form and support. Put $S_A = \text{supp } A$ and list the gates of G whose supports meet S_A as $R_j = I - 2P_j$ for $1 \leq j \leq \ell$, where each P_j is a product projection. Their supports are pairwise disjoint, and each contains a distinct wire of S_A , so $\ell \leq |S_A| \leq k$. Write

$G = G_{\text{hit}}G_{\text{miss}}$, where $G_{\text{hit}} = \prod_{j=1}^{\ell} R_j$ and G_{miss} contains all remaining gates. The two factors commute, G_{miss} commutes with A , and $G_{\text{miss}}^2 = I$. Consequently,

$$GAG = G_{\text{hit}}AG_{\text{hit}}.$$

For $I \subseteq \{1, \dots, \ell\}$, put $P_I = \prod_{j \in I} P_j$, with $P_{\emptyset} = I$. Expanding both copies of G_{hit} gives

$$\begin{aligned} G_{\text{hit}} &= \sum_{I \subseteq \{1, \dots, \ell\}} (-2)^{|I|} P_I, \\ GAG &= \sum_{I, J \subseteq \{1, \dots, \ell\}} (-2)^{|I|+|J|} P_I A P_J. \end{aligned}$$

The total absolute coefficient in this expansion is

$$\sum_{I, J \subseteq \{1, \dots, \ell\}} 2^{|I|+|J|} = \left(\sum_{I \subseteq \{1, \dots, \ell\}} 2^{|I|} \right)^2 = (1 + 2)^{2\ell} \leq 9^k.$$

Each resulting term $P_I A P_J$ factors over the wires. Outside $\text{supp } A$, every factor is a projection or identity. If both sides act on a wire, their projections come from the same gate of G and have the same axis. On a wire in $\text{supp } A$, the factor T is a contraction of rank at most one. Its trace norm satisfies $\|T\|_1 \leq 1$, and

$$T = \frac{T + T^\dagger}{2} + i \frac{T - T^\dagger}{2i}.$$

Both Hermitian parts have trace norm at most one. Their spectral decompositions express T as a linear combination of rank-one orthogonal projections with total absolute coefficient at most two. Tensoring these decompositions on the at most k affected wires costs 2^k . This proves Eq. (6); applying Eq. (5) termwise proves Eq. (7). $\square$

The projections B_α can be supported on many wires. The induction still applies because it controls *every* product test. Only the support of the inserted A is charged.

For a polynomial $f(x) = \sum_j c_j x^j$, package the expansion and layer-removal costs into

$$\mathcal{C}_N(f) = \sum_j |c_j| (18N)^j. \quad (8)$$

Let $d \geq 1$, $U \in \mathcal{U}_d(N)$, $r > 0$, and let M, D be mismatch-count operators. Each monomial D^j expands into at most N^j products of one-wire projections, each supported on at most j wires. Repeated indices collapse. Lemma 3 and the triangle inequality therefore give

$$\begin{aligned} \|M_{\geq r} U^\dagger f(D) U M_0\| &\leq \mathcal{C}_N(f) L_{d-1}(r), \\ \|f(D)\| &\leq \mathcal{C}_N(f). \end{aligned} \quad (9)$$

Lemma 4 (Leakage recurrence). *Let $d \geq 1$, $r, R > 0$, and $\varepsilon \geq 0$. Suppose a real polynomial f satisfies $f(0) = 1$ and $|f(j)| \leq \varepsilon$ at every integer $1 \leq j < R$. Then*

$$L_d(r)^2 \leq \varepsilon + \mathcal{C}_N(f)(L_{d-1}(r) + L_d(R)). \quad (10)$$

Proof. Fix U, B, M in the supremum defining Eq. (5). Choose a mismatch-count operator D with $D_0 = B$, and set $K = U^\dagger D U$. Then $K_0 = U^\dagger B U$. Since M_0 and K_0 are projections, we can square the norm and approximate only the left copy of K_0 . Using that projections are contractions gives

$$\begin{aligned} \|M_{\geq r} K_0 M_0\|^2 &= \|M_{\geq r} K_0 M_0 K_0 M_{\geq r}\| \\ &\leq \|M_{\geq r} f(K) M_0\| + \|(K_0 - f(K)) M_0 K_0\|. \end{aligned} \quad (11)$$

No noncommuting factors have been exchanged. The first term is at most $\mathcal{C}_N(f) L_{d-1}(r)$ by Eq. (9). For the second term, $K_0 - f(K)$ vanishes at count zero, is bounded by ε on $0 < K < R$, and equals $-f(K)$ on $K \geq R$. Hence

$$\|(K_0 - f(K))M_0K_0\| \leq \varepsilon + \mathcal{C}_N(f) \|K_{\geq R}M_0K_0\|.$$

Conjugating the remaining factor by U gives the leakage amplitude for the inverse circuit,

$$\begin{aligned} \|K_{\geq R}M_0K_0\| &= \|D_{\geq R}UM_0U^\dagger D_0\| \\ &\leq L_d(R). \end{aligned} \tag{12}$$

Taking the supremum proves Eq. (10). $\square$

The rightmost K_0 in Eq. (11) is essential. Replacing K_0 directly in $M_{\geq r}K_0M_0$ would leave an error on an arbitrary vector, not the leakage amplitude in Eq. (12). The polynomial need not be bounded by one outside its approximation interval, since its full cost is explicitly paid by the coarser tail.

4 Completing the induction

We first construct the scalar polynomial needed in Eq. (10). The construction is the integer zero filter used in [Ope26a, Lemma 3]. Handling small positive counts exactly and suppressing the error at larger counts is a standard technique, also used in small-error quantum query algorithms for OR; see de Wolf [Wol08].

Lemma 5 (Separating zero from positive counts). *Fix $0 < q < b$ and $\gamma > (q + b)/2$. For all sufficiently large N , there is a real polynomial f with*

$$\begin{aligned} f(0) &= 1, & |f(j)| &\leq e^{-N^q} \quad (1 \leq j \leq \lceil N^b \rceil, j \in \mathbb{Z}), \\ \deg f &\leq N^\gamma, & \log \mathcal{C}_N(f) &\leq N^\gamma. \end{aligned} \tag{13}$$

The exponent b may exceed one.

Proof. Put $a = \lceil N^q \rceil$, $B = \lceil N^b \rceil$, and $m = \lceil 2\sqrt{aB} \log B \rceil$. Eventually $B \geq 2a$. With T_m the Chebyshev polynomial of the first kind, set

$$F(x) = \prod_{j=1}^a \left(1 - \frac{x}{j}\right), \quad y(x) = \frac{B + a - 2x}{B - a}, \quad f(x) = F(x) \frac{T_m(y(x))}{T_m(y(0))}. \tag{14}$$

The first a positive integers are exact roots, and $f(0) = 1$. On $[a, B]$, $|F(x)| \leq B^a$ and $|T_m(y(x))| \leq 1$. Writing $z = \sqrt{a/B}$ gives $y(0) = \cosh(2 \operatorname{artanh} z)$, and therefore

$$T_m(y(0)) \geq \frac{1}{2} \exp\left(2m\sqrt{a/B}\right) \geq \frac{1}{2} B^{4a}.$$

Thus $|f(x)| \leq 2B^{-3a} \leq e^{-N^q}$ on $[a, B]$ for large N .

The norm defined in Eq. (8) is submultiplicative, with

$$\mathcal{C}_N(F) \leq (1 + 18N)^a, \quad \mathcal{C}_N(y) \leq 3 + 36N.$$

The defining recurrence $T_{j+1} = 2yT_j - T_{j-1}$, with $T_0 = 1$ and $T_1 = y$, gives $\mathcal{C}_N(T_m(y)) \leq (2\mathcal{C}_N(y) + 1)^m$. Since $T_m(y(0)) \geq 1$, we obtain

$$\deg f = O\left(N^{(q+b)/2} \log N\right), \quad \log \mathcal{C}_N(f) = O\left(N^{(q+b)/2} (\log N)^2\right).$$

The strict gap in γ absorbs the logarithmic factors. $\square$

Proof of Theorem 2. We induct on d . At depth zero, $Q = U^\dagger B U$ is a product projection. On a counted wire with source $|u\rangle$ and test $|v\rangle\langle v|$, let $\beta = |\langle u|v\rangle|^2$ and $\Pi = I - |u\rangle\langle u|$. Since Π is a projection, $2^\Pi = I + \Pi$, and hence

$$\|2^{\Pi/2}|v\rangle\langle v|u\rangle\|^2 = \beta\langle v|(I + \Pi)|v\rangle = \beta(2 - \beta) \leq 1.$$

Identity tests contribute one, and uncounted factors are contractions. Tensoring, and then applying $M_{\geq r}$, gives

$$\|2^{M/2} Q M_0\| \leq 1, \quad L_0(r)^2 \leq 2^{-r}. \quad (15)$$

This proves Eq. (3) at depth zero. The factor β accounts for the probability that the test succeeds.

Assume the theorem at depth $d - 1$. We first prove the implication

$$L_d(N^b) \leq e^{-N^t} \implies L_d(N^t) \leq e^{-N^s}, \quad 0 \leq s < t < b < 2t - s. \quad (16)$$

Choose

$$s < q < 2t - b, \quad \frac{q + b}{2} < \gamma < t,$$

and use Lemma 5. Apply the preceding-depth theorem at threshold exponent t and precision exponent $(\gamma + t)/2$. In Lemma 4, set $r = N^t$, $R = N^b$, and $\varepsilon = e^{-N^q}$. This gives

$$\begin{aligned} L_d(N^t)^2 &\leq e^{-N^q} + e^{N^\gamma - N^{(\gamma+t)/2}} + e^{N^\gamma - N^t} \\ &\leq e^{-2N^s} \end{aligned} \quad (17)$$

for all sufficiently large N . This proves Eq. (16). The square-root degree scale puts the cost exponent $(q + b)/2$ below t , leaving room for both tail estimates to absorb that cost.

It remains to start the same-depth implications without circularity. Fix $0 \leq s < t \leq 1$, let $g = t - s$, and define

$$t_{-1} = s, \quad t_0 = t, \quad t_{j+1} = t_j + \frac{g}{j+2} \quad (j \geq 0). \quad (18)$$

Successive gaps decrease strictly, so

$$0 \leq t_{j-1} < t_j < t_{j+1} < 2t_j - t_{j-1}.$$

Their harmonic sum diverges; let $J \geq 1$ be the first index with $t_J > 1$. This index depends only on s, t . For $N > 1$, the threshold N^{t_J} exceeds the largest possible mismatch count N , so

$$L_d(N^{t_J}) = 0 \leq e^{-N^{t_{J-1}}}.$$

This supplies the initial hypothesis for a descending induction. Suppose, for some $j \in \{0, \dots, J - 1\}$, that $L_d(N^{t_{j+1}}) \leq e^{-N^{t_j}}$ has already been established. The strict gap inequalities above allow us to apply Eq. (16) with $(s, t, b) = (t_{j-1}, t_j, t_{j+1})$, obtaining

$$L_d(N^{t_j}) \leq e^{-N^{t_{j-1}}}.$$

This is the hypothesis needed at the next index $j - 1$. Starting at $j = J - 1$ and continuing down to $j = 0$ therefore gives $L_d(N^{t_0}) \leq e^{-N^{t_{-1}}}$, which is the desired bound $L_d(N^t) \leq e^{-N^s}$.

Only finitely many fixed exponent choices were used, so one threshold $N_0(d, s, t)$ works for all of them, uniformly over the circuits and product constraints. The argument includes $s = 0$ and $t = 1$; for $t > 1$ the tail is already zero. This completes the induction. $\square$

References

- [Ajt83] Miklós Ajtai. Σ_1^1 -formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983. doi:10.1016/0168-0072(83)90038-6.
- [ADOY25] Anurag Anshu, Yangjing Dong, Fengning Ou, and Penghui Yao. On the computational power of QAC^0 with barely superlinear ancillae. In *STOC*, pages 1476–1487, 2025. doi:10.1145/3717823.3718189. Full version: arXiv:2410.06499v4.
- [DOY25] Yangjing Dong, Fengning Ou, and Penghui Yao. Linear-size QAC^0 channels: Learning, testing and hardness. arXiv:2510.00593v2, 2025.
- [FFGHZ06] Maosen Fang, Stephen Fenner, Frederic Green, Steven Homer, and Yong Zhang. Quantum lower bounds for fanout. *Quantum Information and Computation*, 6(1):46–57, 2006. doi:10.26421/QIC6.1-3.
- [FGPT25] Stephen Fenner, Daniel Grier, Daniel Padé, and Thomas Thierauf. Tight bounds on depth-2 QAC -circuits computing parity. arXiv:2504.06433v1, 2025.
- [FSS84] Merrick Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984. doi:10.1007/BF01744431.
- [Hås86] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *STOC*, pages 6–20, 1986. doi:10.1145/12130.12132.
- [HŠ05] Peter Høyer and Robert Špalek. Quantum fan-out is powerful. *Theory of Computing*, 1(5):81–103, 2005. doi:10.4086/toc.2005.v001a005.
- [JTVW26] Malvika Raj Joshi, Avishay Tal, Francisca Vasconcelos, and John Wright. Improved lower bounds for QAC^0 . In *STOC*, pages 2199–2209, 2026. doi:10.1145/3798129.3800922. Full version: arXiv:2512.14643v4.
- [Kin26] Shiva Kintali. Parity in shallow QAC circuits: Correlation decay and exact lower bounds. Preprint, September 5, 2026. https://shivakintali.github.io/papers/QAC.pdf.
- [Moo99] Cristopher Moore. Quantum circuits: Fanout, parity, and counting. arXiv:quant-ph/9903046v3, 1999.
- [NPVY24] Shivam Nadimpalli, Natalie Parham, Francisca Vasconcelos, and Henry Yuen. On the Pauli spectrum of QAC^0 . In *STOC*, pages 1498–1506, 2024. doi:10.1145/3618260.3649662. Full version: arXiv:2311.09631v4.
- [Ope26a] OpenAI. Product-projection localization and the QAC^0 parity lower bound. Preprint, September 24, 2026. Repository version adc7f124.
- [Ope26b] OpenAI. Regular trajectories, pruning and quantum parity. Preprint, September 24, 2026. Repository version adc7f124.
- [Ros21] Gregory Rosenthal. Bounds on the QAC^0 complexity of approximating parity. In *ITCS*, volume 185 of *LIPIcs*, pages 32:1–32:20, 2021. doi:10.4230/LIPIcs.ITCS.2021.32. Full version: arXiv:2008.07470v3.
- [Wol08] Ronald de Wolf. A note on quantum algorithms and the minimal degree of ε -error polynomials for symmetric functions. *Quantum Information and Computation*, 8(10):943–950, 2008. arXiv:0802.1816.