On the Computational Power of QAC0 with Barely Superlinear Ancillae

Published version
https://doi.org/10.48550/arXiv.2410.06499

On the Computational Power of $\mathbf{QAC}^0$ with Barely Superlinear Ancillae

Anurag Anshu* Yangjing Dong† Fengning Ou‡ Penghui Yao§¶

November 1, 2024

Abstract

$\mathbf{QAC}^0$ is the family of constant-depth polynomial-size quantum circuits consisting of arbitrary single qubit unitaries and multi-qubit Toffoli gates. It was introduced by Moore [Moo99] as a quantum counterpart of $\mathbf{AC}^0$, along with the conjecture that $\mathbf{QAC}^0$ circuits cannot compute PARITY. In this work, we make progress on this long-standing conjecture: we show that any depth-d $\mathbf{QAC}^0$ circuit requires $n^{1+3^{-d}}$ ancillae to compute a function with approximate degree $\Theta(n)$, which includes PARITY, MAJORITY and MOD_k. We further establish superlinear lower bounds on quantum state synthesis and quantum channel synthesis. This is the first lower bound on the super-linear sized $\mathbf{QAC}^0$. Regarding PARITY, we show that any further improvement on the size of ancillae to $n^{1+\exp(-o(d))}$ would imply that PARITY $\notin \mathbf{QAC}^0$.

These lower bounds are derived by giving low-degree approximations to $\mathbf{QAC}^0$ circuits. We show that a depth-d $\mathbf{QAC}^0$ circuit with a ancillae, when applied to low-degree operators, has a degree $(n+a)^{1-3^{-d}}$ polynomial approximation in the spectral norm. This implies that the class $\mathbf{QLC}^0$, corresponding to linear size $\mathbf{QAC}^0$ circuits, has an approximate degree $o(n)$. This is a quantum generalization of the result that $\mathbf{LC}^0$ circuits have an approximate degree $o(n)$ by Bun, Kothari, and Thaler [BKT19]. Our result also implies that $\mathbf{QLC}^0 \neq \mathbf{NC}^1$.

*School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. Email: anuraganshu@fas.harvard.edu.
†State Key Laboratory for Novel Software Technology, Nanjing University, China. Email: dongmassimo@gmail.com.
‡State Key Laboratory for Novel Software Technology, Nanjing University, China. Email: reverymoon@gmail.com.
§State Key Laboratory for Novel Software Technology, Nanjing University, China. Email: phyao1985@gmail.com.
¶Hefei National Laboratory, Hefei 230088, China.

Contents

1 Introduction

Shallow quantum circuits (quantum circuits with constant depths) stand as a fundamental construction within the realm of quantum computing, arising naturally from the practical constraints and design of quantum hardware. These circuits, while seemingly simple in structure, represent an exciting area of research within quantum information theory and quantum complexity theory. One of the notable achievements in recent years has been the discovery of the quantum advantages of constant-depth quantum circuits [BGK18]. It is now well recognized that their computational reach is particularly constrained in their ability to generate long-range quantum entanglement, a cornerstone quantum phenomenon. In essence, the flow of information from a qubit is limited to the "light cone" of the qubit in such circuits, severely restricting the computational power of constant-depth quantum circuits, especially when solving decision problems. This raises a fundamental question: What is the simplest architecture beyond shallow quantum circuits that is beyond the light-cone constraint?

Inspired by the study of $\mathbf{AC}^0$ circuits as extensions of $\mathbf{NC}^0$ circuits, Green, Homer, Moore and Pollett [Moo99, GHMP01] introduce the notion of $\mathbf{QAC}^0$ circuits, where they augmented shallow quantum circuits with generalized Toffoli gates $U_F |x_1, \ldots, x_n, b\rangle = |x_1, \ldots, x_n, b \oplus \wedge_i x_i\rangle$. $\mathbf{QAC}^0$ gives a simple theoretical model beyond light-cone constraints, with which to study the power of many-qubit operations in quantum computing. Researchers have also investigated the realization of many-qubit operations on different platforms [GKH+21, YWL+22, BLS+22, GDC+22].

Despite being a quantum generalization of $\mathbf{AC}^0$, our understanding of the computational power of $\mathbf{QAC}^0$ is still very limited. We even do not know whether $\mathbf{QAC}^0$ contains $\mathbf{AC}^0$. To see it, arbitrary fan-out is allowed in $\mathbf{AC}^0$ circuits, which means that the output of a gate can be fed to arbitrary many other gates as input. However, for quantum circuits, due to the quantum no-cloning theorem, the output of a quantum gate can only be used once as the input for another quantum gate, so fan-out is restricted to 1. Another difference that arises from the quantumness of $\mathbf{QAC}^0$ circuits is that quantum computations are reversible, while classical circuits can perform the **AND** and **OR** gates, which are nonreversible. To mitigate this, $\mathbf{QAC}^0$ circuits allow other auxiliary quantum qubits in the circuit, referred to as ancillae. The ancillae act as the memory or internal nodes in classical $\mathbf{AC}^0$ circuits. The size of the ancillae together with the inputs is analogous to the size of classical circuits.

Does $\mathbf{QAC}^0$ contain a decision problem that is not in $\mathbf{AC}^0$? This problem remains open. It is well known that PARITY and MAJORITY are not in $\mathbf{AC}^0$ [Hås86]. The question of whether these functions are contained in $\mathbf{QAC}^0$ was immediately raised by Moore [Moo99] when the class $\mathbf{QAC}^0$ was first proposed. Since then, many works have been devoted to investigating this problem, which are largely summarized in Figure 1. However, progress on this problem is still very slow. Before this work, it was still unknown whether any linear-sized $\mathbf{QAC}^0$circuit can compute PARITY.

Understanding the computational power of bounded-depth circuits is not only interesting in quantum computing but is also a core topic in classical complexity theory. The celebrated results of Håstad showed that PARITY is not in $\mathbf{AC}^0$ via the well-known switching lemma [Hås86], which nowadays is still one of the most powerful tools to prove the circuit lower bounds. Despite decades of effort, constant-depth circuits still represent the frontier of our understanding of circuit lower bounds.

A natural approach to proving the lower bounds in $\mathbf{QAC}^0$ is to generalize the existing techniques for the lower bounds of classical circuits [Hås86, Raz87, Smo87, Ros16]. However, it seems that all of the techniques face certain barriers when generalizing to the quantum world. As mentioned above, a crucial difference between $\mathbf{AC}^0$ and $\mathbf{QAC}^0$ is that all the circuits in $\mathbf{QAC}^0$ are reversible and, thus the number of output qubits is the same as the number of input qubits. The ancilla qubits act as intermediate nodes in $\mathbf{AC}^0$ circuits, which play a crucial role in quantum computing. Indeed, Rosenthal proved that an exponential-sized and constant-depth quantum circuit allowing many-qubits Toffoli gates can compute parity [Ros21]. To our knowledge, all the lower bound techniques for classical

PARITY	$a \geq n2^{-d} - 1$	exact	[FFG+06]
	$a \geq \infty$ when $d = 2$	exact	[PFGT20]
	$a \geq \infty$ when $d = 2$	average case	[Ros21]
	$a \leq \exp(O(n \log n/\varepsilon))$ when $d = 7$	worst case	[Ros21]
	$a \geq n^{\Omega(1/d)}$	average case	[NPVY24]
	$a \geq n^{1+3^{-d}}$	average/worst case	**This work**
MAJORITY	$a \geq n^{\Omega(1/d)}$	average case†	[NPVY24]
	$a \geq n^{1+3^{-d}}$	average/worst case	**This work**
MOD_k	$a \geq n^{1+3^{-d}}$	worst case	**This work**

Figure 1: Hardness of Boolean functions

† For the MAJORITY function, the ancillae lower bound for average case works if the average case error is small then $1/\sqrt{n}$. This also applies to our average case lower bound for MAJORITY.

circuits, roughly speaking, substitute part of the inputs with simpler ones, which eliminate intermediate nodes. However, the input qubits and the ancilla qubits are fed into the circuit at the same time. Therefore, it is unclear how to eliminate ancilla qubits via simplifying the inputs.

1.1 Our Results

Our work extends the Pauli analysis framework on quantum circuits by [NPVY24]. The main technical result is a low-degree approximation for $\mathbf{QAC}^0$ circuits. The approximation shows that $\mathbf{QAC}^0$ circuits will map approximately low-degree operators to approximately low-degree operators. Here an operator having degree k means that it can be represented as a summation of k-local operators. An operator is "low-degree" means its degree is of order $o(n)$, and the approximation is with respect to the spectral norm. The following is the key technical theorem in this paper.

Theorem 1.1 (informal of Corollary 3.6)**.** *For any $2^n \times 2^n$ operator A with degree ℓ, and any unitary U implemented by a depth-d $\mathbf{QAC}^0$ circuit, the approximate degree of $UAU^\dagger$ is upper bounded by $\widetilde{O}\left(n^{1-3^{-d}}\ell^{3^{-d}}\right)$.*

Since we are working with the spectral norm, this degree upper bound also holds when we post-select on ancillae. This is because for an operator A and ancillae in the state $|\varphi\rangle$, it holds that

$$\|(\mathbb{1} \otimes \langle\varphi|)\, A\, (\mathbb{1} \otimes |\varphi\rangle)\| \leq \|A\|.$$

In this case, the upper bound of degrees is with respect to the total number of qubits. For linear ancillae $\mathbf{QAC}^0$ circuits, which form the class $\mathbf{QLC}^0$, Theorem 1.1 shows that $\mathbf{QLC}^0$ circuits have approximate degree $o(n)$. See Corollary 4.2.

With Theorem 1.1, we study the power of $\mathbf{QAC}^0$ circuits in different quantum computational tasks.

Compute Boolean functions Our first results are about the hardness for computing Boolean functions. We investigate both the worst case and the average case. In the worst case, the circuit computes all inputs correctly with high probability. For the average case, the circuit computes the function correctly with high probability when the input is drawn uniformly. The results along with previous works are summarized in Figure 1.

We prove that $\mathbf{QAC}^0$ circuits with ancillae only slightly more than linear can not compute Boolean functions with approximate degree $\Omega(n)$ in the worst case with any probability strictly larger than $1/2$. To the best of our knowledge, these are the first hardness results that allow superlinear pure state ancillae.

Theorem 1.2 (informal of Theorem 4.3)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function with approximate degree $\Omega(n)$. Suppose U is a depth d $\mathbf{QAC}^0$ circuit with n input qubits and $a = \widetilde{O}\left(n^{1+3^{-d}}\right)$ ancillae initialized in any quantum state. Then U cannot compute f with the worst-case error strictly below $1/2$. In particular, this includes PARITY, MAJORITY, and also MOD_k when $k \le cn$ for some $c < 1$.*

We further adapt the techniques from [NPVY24] to obtain average-case hardness results. For Boolean functions that remain high degree even with Frobenius norm approximations, we prove that it is hard for $\mathbf{QAC}^0$ circuits to approximate them within a certain error regime. In particular, we state our average-case hardness results for PARITY and MAJORITY.

Theorem 1.3 (informal of Theorem 4.6)**.** *Suppose U is a depth d $\mathbf{QAC}^0$ circuit with n input qubits and $a = \widetilde{O}\left(n^{1+3^{-d}}\right)$ ancillae initialized in any quantum state. Let $C_U(x) \in \{0,1\}$ denote the classical output of the circuit on input x. Then it hold that:*

- *U can not approximate* Parity_n *over uniform inputs, i.e.,*

$$\mathop{\mathbb{E}}_{\boldsymbol{x}\in\{0,1\}^n}\left[\Pr\left[C_U(\boldsymbol{x}) = \text{Parity}_n(\boldsymbol{x})\right]\right] \le \frac{1}{2} + O(d/n).$$

- *U can not approximate* Majority_n *over uniform inputs, i.e.,*

$$\mathop{\mathbb{E}}_{\boldsymbol{x}\in\{0,1\}^n}\left[\Pr\left[C_U(\boldsymbol{x}) = \text{Majority}_n(\boldsymbol{x})\right]\right] \le 1 - \Omega\left(\frac{1}{\sqrt{n}}\right) + O(d/n).$$

Our result is incomparable to the results of [NPVY24]: On the one hand, if the ancillae are initialized into pure states, we have a super-linear lower bound on the ancillae, while their lower bound is $n^{\Omega(1/d)}$. On the other hand, if the ancillae are initialized as mixed states, our bound remains $n^{1+3^{-d}}$, while [NPVY24] proved that $\mathbf{QAC}^0$ with an arbitrarily dimensional maximally mixed state as ancillae cannot compute PARITY.

Our lower bound $\widetilde{\Omega}\left(n^{1+3^{-d}}\right)$ seems far away from an arbitrarily polynomial lower bound. Surprisingly, we show that any improvement of the exponent d to $o(d)$ would lead to PARITY $\notin \mathbf{QAC}^0$, a complete resolution of the problem.

Theorem 1.4 (informal of Corollary 5.3)**.** *If any $\mathbf{QAC}^0$ circuit with $n^{1+\exp(-o(d))}$ ancillae, where d is the depth of this circuit family, can not compute* Parity_n *with the worst-case error* $\text{negl}(n)$*, then any $\mathbf{QAC}^0$ circuit family with arbitrary polynomial ancillae can not compute* Parity_n *with the worst-case error* $\text{negl}(n)$.

Quantum State Synthesis We further investigate the hardness of quantum state synthesis in $\mathbf{QAC}^0$ circuits.

Theorem 1.5 (informal of Theorem 6.4)**.** *Let $\varphi = |\varphi\rangle\langle\varphi|$ be a pure state on n qubits. The approximate degree of φ is $\Omega(n)$. Suppose that there exists a depth-d $\mathbf{QAC}^0$ circuit working on a qubits such that the first n qubits of $U|0^a\rangle$ measure to φ with constant probability, then we have*

$$a = \widetilde{\Omega}(n^{1+3^{-d}/2}).$$

In particular, the n-nekomata state, and the low energy states of the code Hamiltonian in [ABN22] satisfies $\Omega(n)$ approximate degree. Hence $\mathbf{QAC}^0$ circuits with linear ancillae cannot synthesize these states.

Quantum Channels Synthesis Our last result is a general hardness result for quantum channel synthesis.

Theorem 1.6 (informal of Theorem 7.2)**.** *Suppose $\mathcal{E}_{U,\psi}$ is a quantum channel from n qubits to k qubits, implemented by a depth-d $\mathbf{QAC}^0$ circuit U with n input qubits and a ancillae. The upper bound of approximate degree of the Choi representation $\Phi_{U,\psi}$ of $\mathcal{E}_{U,\psi}$ is then given by $\widetilde{O}\left((n+a)^{1-3^{-d}}k^{3^{-d}/2}\right)$.*

1.2 Related Works

The study of quantum circuit complexity was initiated in large part by Green, Homer, Moore and Pollett [Moo99, GHMP01], who introduced quantum counterparts of a number of classical circuit classes, including $\mathbf{QAC}^0$. The problem whether the parity function can be computed by $\mathbf{QAC}^0$ circuits was also raised in their papers. However, this problem has made slow progress over the years.

To compute PARITY , we need to implement unitary

$$|b, x_1, \ldots, x_n\rangle \mapsto \left|b \oplus \bigoplus_i x_i, x_1, \ldots, x_n\right\rangle.$$

Green, Homer, Moore and Pollett [GHMP01] have shown that PARITYis equivalent to fan-out until conjugation with Hadamard gates, where fan-out is unitary $|b, x_1, \ldots, x_n\rangle \mapsto |b, x_1 \oplus b, \ldots, x_n \oplus b\rangle$.

Fang, Fenner, Green, Homer, and Zhan [FFG+06] gave the first lower bound on this question. They showed that $\mathbf{QAC}^0$circuits with sublinear ancillae cannot compute parity exactly, where "exactly" means that the output qubit should be precisely in the state $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ for any input. In particular, they showed that for any depth-d $\mathbf{QAC}^0$ circuit with a ancillae, to exactly compute parity, we should have $d \geq 2\log(n/(a+1))$. This lower bound is derived in the following way: They are able to find an input state on $(a+1)2^{d/2}$ qubits such that the circuit output is fixed to be 1 irrelevant to the other input qubits. Then if $(a+1)2^{d/2} < n$, such circuits clearly cannot compute parity exactly, because any flip on another input qubit changes the output. The case of 0 ancillae is also implied by a result of Bera [Ber11]. Padé, Fenner, Grier, and Thierau [PFGT20] showed that no depth-2 circuits can compute the Parity_n function exactly for $n \geq 4$, even with any number of ancillae. This is achieved by a careful analysis on the structure of depth-2 quantum circuits.

These techniques heavily rely on the condition of exact computation, and thus are difficult to extend to the approximate case. Actually, Rosenthal [Ros21] has showed a depth-7 quantum circuit with many-qubit Toffoli gates and exponentially ancillae that approximately computes the PARITY function. This implies that when proving the hardness of computing PARITY, we need to take into account the number of ancilla. Rosenthal has also established a series of equivalent problems to PARITY. In particular, a circuit preparing the cat state defined as $\left|🐱_n\right\rangle = \frac{1}{\sqrt{2}}(|0^n\rangle + |1^n\rangle)$ is equivalent to parity up to constant-degree circuit reductions. In the same work, Rosenthal showed that any depth-2 $\mathbf{QAC}^0$circuits with arbitrary ancilla qubits cannot compute parity even in the approximate case, generalizing the results of [PFGT20] to the average case. Very recently, Nadimpalli, Parham, Vasconcelos, and Yuen [NPVY24] have initiated the study of Pauli spectrum of $\mathbf{QAC}^0$ circuits and proved the first lower bound on $\mathbf{QAC}^0$ circuit with an arbitrary depth. More specifically, they proved that a $\mathbf{QAC}^0$ circuit with depth d that computes PARITY needs at least $n^{\Omega(1/d)}$ ancillae.

Fourier Analysis and Pauli Analysis. Fourier analysis on the space of Boolean functions seeks to understand Boolean functions via their Fourier transformations. The very first application of Fourier analysis on Boolean functions is the well-known KKL theorem proved by Kahn, Kalai, and Linial [KKL88] in 1988. Today, Fourier analysis on Boolean functions has played a crucial role in various areas. Readers may refer to O'Donnell's excellent book [O'D14] for a thorough treatment.

Pauli analysis is a quantum generalization of the Fourier analysis of Boolean functions. Recall that the set of n-qubit Pauli operators $\mathcal{P}_n = \{I, X, Y, Z\}^{\otimes n}$ forms an orthonormal basis for the space of all operators acting on n-qubits. Given an n-qubit operation M, its Pauli expansion is $M = \sum_{P \in \mathcal{P}_n} \widehat{M}(P)\, P$, analogous to the Fourier expansion of a Boolean function, where the coefficients $\widehat{M}(P)$'s are the Pauli spectrum of M. Such expansions can be further generalized to quantum channels [BY23, NPVY24]. To see that Pauli analysis is indeed a generalization of Fourier analysis, for a Boolean function $f : \{0,1\}^n \to \mathbb{R}$, the Pauli expansion of the diagonal matrix $M_f = \sum_{x \in \{0,1\}^n} f(x) |x\rangle\langle x|$ is exactly the same as the Fourier expansion of f.

Pauli analysis was first studied by Montanaro and Obsorne in the context of the so-called quantum Boolean functions [MO10] and has received increasing attention in the past couple of years, finding applications in testing and learning quantum operations [CdDPH+24, BY23, CNY23, ADEGP24, RWZ24], classical simulations of noisy quantum circuits [AGL+23], multiprover quantum interactive proof systems in the noisy world [Yao19, QY21, QY23, DFN+24]. Very recently, Nadimpalli, Parham, Vasconcelos, and Yuen [NPVY24] initiated the study on the Pauli spectrum of $\mathbf{QAC}^0$circuits, which has led to a series of lower bounds and learning algorithms for $\mathbf{QAC}^0$.

Comparison with [NPVY24] Similarly to the work [NPVY24], our work is also built on the theory of Pauli analysis. Instead of analyzing the Pauli spectrum of the channels induced by $\mathbf{QAC}^0$ circuits as in [NPVY24], we study the Pauli spectrum of the projectors induced by a $\mathbf{QAC}^0$ circuits, namely $U^\dagger (|0\rangle\langle 0| \otimes \mathbb{1})\, U$, where U is a $\mathbf{QAC}^0$ circuit.

[NPVY24] removed the large Toffoli gates in a circuit and argued that the Frobenius norm does not change significantly. In this way, they obtain a low-degree approximation of $\mathbf{QAC}^0$ circuits within the Frobenius norm. However, with this approach, the ancillary qubits are restricted to $n^{O(1/d)}$, because the normalized Frobenius norm grows exponentially after operation $A \mapsto (\mathbb{1} \otimes \langle 0^a|)\, A\, (\mathbb{1} \otimes |0^a\rangle)$, which projects the ancillary registers to the initial state $|0^a\rangle$.

In our work, we consider the approximation with respect to the spectral norm. We use Chebyshev polynomials to approximate the high-degree gates, and thus are able to get $\sqrt{n}$-degree approximations with respect to the spectral norm. Meanwhile, the projection of the ancilla registers to the initial state does not increase the spectral norm. One more advantage of using the spectral norm is that we can prove the hardness for a much broader class of Boolean functions. Since the Frobenius norm can be exponentially smaller than the spectral norm, many Boolean functions that have low-degree approximations with small Frobenius norm actually have an $\Omega(n)$ approximate degree if we consider the spectral norm. Hence in our work we are also able to prove optimal $1/2$ worst case hardness results for the MAJORITY function and the MOD_k function. For comparison, [NPVY24] shows that it is hard for $\mathbf{QAC}^0$ circuits to approximate the MAJORITY function with a success probability greater than $1 - \frac{1}{\sqrt{n}}$.

In this paper, we further establish lower bounds on quantum state synthesis. Notice that the normalized Frobenius norms of quantum states are exponentially small, while the spectral norm can be as large as 1. Hence, the results of [NPVY24] cannot be applied directly to quantum state synthesis.

1.3 Discussion

It is still open whether $\mathbf{QAC}^0$ circuits can compute the PARITY function. Despite being hard in the quantum case, the classical counterpart, that is, $\mathbf{AC}^0$ circuits cannot compute the PARITY function, has admitted several different proofs. Here we briefly explain some well-known techniques for classical circuit lower bounds and explain why it seems to be difficult to generalize to quantum circuits.

The most well studied technique goes to Håstad's switching lemma [Hås86], which describes the effect of random restrictions on depth-2 circuits. Random restriction fixes part of the input in a random manner, which effectively kills large **AND** or **OR** gates. Moreover, **AND**-**OR** circuits and **OR**-**AND** circuits switch with each other via decision tree representation, which decreases the depth of circuits. For quantum circuits, however, there is not an obvious way to apply random restrictions.

Fang, Fenner, Green, Homer, and Zhang [FFG+06] proved that for any depth d $\mathbf{QAC}^0$ circuit with a ancillae, there exists a quantum state φ on $(a+1)2^{d/2}$ qubits such that when part of the input is restricted to the state φ, then the output of the $\mathbf{QAC}^0$ circuit is fixed to be 1. This can be seen as a restriction-based method. However, in their work the state φ is chosen in a deterministic manner. Also, their method only works if the circuit computes PARITY exactly. The difficulty for quantum random restrictions arises from the fact that large Toffoli gates might interact with many ancilla qubits but only a few input qubits, such as Rosenthal's circuit for nekomata states synthesis [Ros21]. Then random restricting inputs won't kill many large Toffoli gates.

Rossman [Ros16] provided a simple proof of PARITY $\notin$ $\mathbf{AC}^0$ via the DNF sparsification introduced by Gopalan, Meka and Reingold [GMR12]. An m-junta is a Boolean function $f : \{0,1\}^n \to \mathbb{R}$ that actually only looks at the input in the m indices. The DNF sparsification theorem from [GMR12] allows one to approximate any k-DNF or k-CNF to an m-junta, where m only depends on k and the approximation error. Then, through a random restriction, one can convert a large junta to a smaller junta. These two steps together can convert an depth-3 circuit, which are essentially k-DNFs or k-CNFs, into a k-junta, which can be represented by a constant sized depth-2 circuit, decreasing the depth by 1. By repeating the process above, we can reduce an arbitrary $\mathbf{AC}^0$ circuit into a trivial depth-2 k-junta. On the other hand, the PARITY function is robust against random restrictions, thus we can prove that $\mathbf{AC}^0$ circuits cannot compute the PARITY function.

This methods appears to be easier to adapt to the quantum case, compared to the switching lemma of [Hås86]. This is because the notion of "quantum junta" is more natural and has been studied in [MO10, Wan11, CNY23, BY23]. However, it is still built on random restrictions, and it is not clear how to extend it to quantum circuits.

The polynomial method is a different approach to prove the hardness results of $\mathbf{AC}^0$. Razborov [Raz87] and Smolensky [Smo87] to prove that MAJORITY and MOD_3 cannot be efficiently computed by $\mathbf{AC}^0[\oplus]$. Refer to the excellent survey of polynomial methods by Williams [Wil14] for the polynomial method in circuit complexity. For classical circuits, it is first shown that the elementary gates (**AND**, **OR**, and **NOR**) can be approximated by low-degree polynomials. And then by combining these polynomial together, we can represent the whole $\mathbf{AC}^0$ circuit by a low-degree polynomial. Finally, since PARITY, MAJORITY, and MOD_k have an $\Omega(n)$ approximate degree, we can show that these problems are hard for $\mathbf{AC}^0$ circuits.

Our techniques resemble the polynomial method in spirit, where low-degree polynomials are replaced by low-degree operators. The elementary gates in a $\mathbf{QAC}^0$ circuit are the multi-qubit CZ-gates, along with all single-qubit unitaries. The single-qubit unitaries are already of degree 1 so we do not need further actions on them. For an n qubit CZ-gate, we use Corollary 3.3 to represent it by a degree $\widetilde{O}(\sqrt{n})$ operator. Using a layer-by-layer argument, we prove Corollary 3.6, which states that the whole effect of a $\mathbf{QAC}^0$ circuit is low degree. A crucial difference between our proof and the proofs for $\mathbf{AC}^0$ is that the approximation for $\mathbf{AC}^0$ is over a certain finite field while our approximation for $\mathbf{QAC}^0$ is over complex number fields. It is not clear how to represent the computation of $\mathbf{QAC}^0$ in a finite field.

Our work uses spectral norm approximations, which do not increase when the number of ancillae increase. However, this requires us to approximate $\mathbf{QAC}^0$ circuits in the spectral norm using low-degree operators. Even in the classical case, proving that $\mathbf{AC}^0$ circuits have $o(n)$ approximate degree is a notorious open problem [BT22], and if solved we may get several consequences in complexity theory; see [Slo24] and [BT22] for more discussion. Our work can be seen as a quantum generalization of the work of Bun, Kothari, and Thaler [BKT19, Theorem 5], who proved that $\mathbf{AC}^0$ circuits with linear ancillae and depth d, aka $\mathbf{LC}^0_d$, have an approximate degree $\widetilde{O}(n^{1-2^{-d}})$. However, in both the quantum case and the classical case, proving approximate degree upper bounds for $\mathbf{QAC}^0$ circuits with superlinear ancillae or $\mathbf{AC}^0$ circuits of superlinear size requires novel techniques.

Acknowledgement

We thank Zongbo Bao for his contribution in the early stage of this project. PY thanks Xun Gao for the helpful discussion. This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing, supported by DOE QSA grant number FP00010905. AA acknowledges support through the NSF Award No. 2238836 and the NSF award QCIS-FF: Quantum Computing & Information Science Faculty Fellow at Harvard University (NSF 2013303). YD, FO, and PY were supported by National Natural Science Foundation of China (Grant No. 62332009, 12347104) and Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302901).

2 Preliminaries

2.1 Analysis of Boolean Functions

In this subsection, we briefly introduce the theory of analysis of Boolean functions. Readers may refer to O'Donnell's excellent book [O'D14] for a thorough treatment.

Given a Boolean function $f : \{0,1\}^n \to \mathbb{R}$, its p-norm is defined to be $\|f\|_p = \left(\mathbb{E}_{\mathbf{x}}\left[|f(\mathbf{x})|^p\right]\right)^{1/p}$ for $p \geq 1$. Its infinity norm is defined to be $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p = \max_x |f(x)|$. We will use the notation $\|f\| = \|f\|_\infty$.

Given Boolean functions $f, g : \{0,1\}^n \to \mathbb{R}$, the inner product of f and g is $\langle f, g\rangle = \mathbb{E}_{\mathbf{x}}[f(\mathbf{x})g(\mathbf{x})]$, where $\mathbf{x}$ is uniformly distributed over $\{0,1\}^n$. For any $S \subseteq [n]$, define the Fourier basis χ_S as $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$. It is not hard to see that $\{\chi_S\}_{S\subseteq[n]}$ is an orthonormal basis. The Fourier expansion of f is $f = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S$.

The following is the well-known Parseval theorem, which relates the 2-norm and Fourier coefficients of a Boolean function.

Theorem 2.1 (Parseval's theorem). *Let $f : \{0,1\}^n \to \mathbb{R}$ be a Boolean function. Then*

$$\|f\|_2^2 = \sum_{S\subseteq[n]} \widehat{f}(S)^2.$$

Definition 2.2. Let $f : \{0,1\}^n \to \mathbb{R}$ be a Boolean function with Fourier expansion

$$f = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S.$$

Then the degree of f is defined as

$$\deg(f) = \max_{S:\widehat{f}(S)\neq 0} |S|.$$

Definition 2.3 (Approximate Degree). Let $f : \{0,1\}^n \to \mathbb{R}$ be a Boolean function. For $\varepsilon \in [0,1]$, the approximate degree of f is defined as

$$\widetilde{\deg}_\varepsilon(f) = \min_{g:\|f-g\|\leq\varepsilon} \deg(g).$$

It is worth noticing that the approximation is with respect to the infinity norm. The notion of approximate degrees has played a crucial role in quantum query complexity and quantum communication complexity [BBC+01, BdW01].

A Boolean function $f : \{0,1\}^n \to \mathbb{R}$ is symmetric if $f(x_1,\ldots,x_n) = f\left(x_{\sigma(1)},\ldots,x_{\sigma(n)}\right)$ for any $\sigma \in S_n$.

Fact 2.4 ([BT22, Theorem 3]). Let $f : \{0,1\}^n \to \{0,1\}$ be a symmetric function and t be the number function such that f is constant on all inputs of Hamming weight between t and $n-t$. Then for $\varepsilon \in (2^{-n}, 1/3)$, we have

$$\widetilde{\deg}_\varepsilon(f) = \Theta\left(\sqrt{nt} + \sqrt{n\log(1/\varepsilon)}\right).$$

Example 2.5. With Fact 2.4, we can show that the following functions have an $\Omega(n)$ approximate degree:

- For any n define the function $\text{Parity}_n : \{0,1\}^n \to \{0,1\}$ as
$$\text{Parity}_n(x) = \bigoplus_i x_i.$$
Then $\widetilde{\deg}_{1/3}\left(\text{Parity}_n\right) = \Theta(n)$.
- For any odd n define the function $\text{Majority}_n : \{0,1\}^n \to \{0,1\}$ as
$$\text{Majority}_n(x) = \begin{cases} 1 & \text{if } \sum_i x_i \geq n/2 \\ 0 & \text{if } \sum_i x_i < n/2 \end{cases}$$
Then $\widetilde{\deg}_{1/3}\left(\text{Majority}_n\right) = \Theta(n)$.
- For any n and k, define the function $\text{Mod}_{n,k} : \{0,1\}^n \to \{0,1\}$ as
$$\text{Mod}_{n,k}(x) = \begin{cases} 1 & \text{if } \sum_i x_i \bmod k \neq 0 \\ 0 & \text{if } \sum_i x_i \bmod k = 0 \end{cases}$$
Then $\widetilde{\deg}_{1/3}\left(\text{Mod}_{n,k}\right) = \Omega(n-k)$. Note that $\text{Mod}_{n,2} = \text{Parity}_n$.

2.2 Quantum Information and Pauli analysis

A quantum system A is associated with a finite-dimensional Hilbert space, which we also denote by A. The quantum registers in the quantum system A are represented by *density operators*, which are trace-one positive semidefinite operators, in the Hilbert space A. We also use the Dirac notation $|\varphi\rangle$ to represent a pure state. In this case, we have the convention that $\varphi = |\varphi\rangle\langle\varphi|$, where here φ is a rank-one density operator. For two separate quantum registers φ and σ from quantum systems A and B, the compound register is the Kronecker product $\varphi \otimes \sigma$. A *positive operator-valued measure* (POVM) is a quantum measurement described by a set of positive semidefinite operators that sum up to identity. Let $\{P_a\}_a$ be a POVM applied on a quantum register φ, then the probability that the measurement outcome is a is $\text{Tr}\left[P_a\varphi\right]$.

For any integer $n \geq 2$, let $\mathcal{M}_n$ be the set of $n \times n$ matrix. For any matrix $M \in \mathcal{M}_n$, we let $|M| = \sqrt{M^\dagger M}$. For any $M, N \in \mathcal{M}_n$, the inner product of M, N is $\langle M, N\rangle = \text{Tr}\left[M^\dagger N\right]/n$. It is evident that $(\mathcal{M}_n, \langle\cdot,\cdot\rangle)$ forms a Hilbert space.

For $p \geq 1$, the *normalized* Schatten p-norm of M is defined to be
$$\|M\|_p = \left(\frac{1}{n}\text{Tr}\left[|M|^p\right]\right)^{1/p}.$$
It is not hard to see that $\langle M, M\rangle = \|M\|_2^2$. Moreover, $\|\cdot\|_p$ is monotone non-decreasing with respect to p and $\|\cdot\|_\infty = \lim_{p\to\infty}\|\cdot\|_p$ is the spectral norm. The fidelity between two quantum states ρ and φ is defined as
$$F(\rho,\sigma) = \text{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right].$$
The above definition is symmetric: $F(\rho,\sigma) = F(\sigma,\rho)$. When one of the input states is pure, say $\rho = |\rho\rangle\langle\rho|$, then we have
$$F(|\rho\rangle\langle\rho|, \sigma) = \sqrt{\langle\rho|\sigma|\rho\rangle}.$$
The Fuchs–van de Graaf inequalities give a relation between the norms and fidelity:

Lemma 2.6 ([Wat18, Theorem 3.33]). *Let ρ, σ be positive semi-definite operators of size $2^n \times 2^n$. Let $\|\rho\|_{TD} = 2^n \|\rho\|_1$ denote the unnormalized trace norm of an operator. It holds that*

$$1 - \frac{1}{2}\|\rho - \sigma\|_{TD} \leq F(\rho, \sigma) \leq \sqrt{1 - \frac{1}{4}\|\rho - \sigma\|_{TD}^2}.$$

Equivalently,

$$2 - 2F(\rho, \sigma) \leq \|\rho - \sigma\|_{TD} \leq 2\sqrt{1 - F(\rho, \sigma)^2}.$$

Also, for any operator ρ, we have $\|\rho\|_{TD} \geq \|\rho\|_p$ for any $p \geq 1$ or $p = \infty$.

The following lemma will also be used throughout this work. The proof is deferred to Appendix A.

Lemma 2.7. *Let $A, B, \widetilde{A}, \widetilde{B}$ be operators satisfying*

- $\|A\| \leq 1$ *and* $\|B\| \leq 1$.
- $\left\|A - \widetilde{A}\right\| \leq \varepsilon_0$.
- $\left\|B - \widetilde{B}\right\| \leq \varepsilon_1$.

Then $\|AB\| \leq 1$ and $\left\|AB - \widetilde{A}\widetilde{B}\right\| \leq \varepsilon = \varepsilon_0 + \varepsilon_1 + \varepsilon_0\varepsilon_1 = (1+\varepsilon_0)(1+\varepsilon_1) - 1$.

The Pauli matrices $\mathcal{B}_0, \ldots, \mathcal{B}_3$ are

$$\mathcal{B}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathcal{B}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \mathcal{B}_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \mathcal{B}_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

which form an orthonormal basis in $\mathcal{M}_2$. For integer $n \geq 1$ and $\sigma \in \{0, 1, 2, 3\}^n$, we define

$$\mathcal{B}_\sigma = \mathcal{B}_{\sigma_1} \otimes \cdots \otimes \mathcal{B}_{\sigma_n}.$$

The set of Pauli matrices $\{\mathcal{B}_\sigma\}_{\sigma \in \{0,1,2,3\}^n}$ forms an orthonormal basis in $\mathcal{M}_{2^n}$. For any $2^n \times 2^n$ matrix A, the Pauli expansion of A is

$$A = \sum_{\sigma \in \{0,1,2,3\}^n} \widehat{A}(\sigma) \cdot \mathcal{B}_\sigma.$$

The coefficients $\widehat{A}(\sigma)$'s are called the Pauli coefficients of A. We can then define the degree and the approximate degree of a matrix.

Definition 2.8. Let n be an integer and A be a $2^n \times 2^n$ matrix. The degree of A is defined as

$$\deg(A) = \max_{\sigma : \widehat{A}(\sigma) \neq 0} |\sigma|,$$

where $|\sigma| = |\{i : \sigma_i \neq 0\}|$. For $\varepsilon \in [0, 1]$, the approximate degree of A is defined as

$$\widetilde{\deg}_\varepsilon(A) = \min_{B : \|A - B\| \leq \varepsilon} \deg(B),$$

where $\|\cdot\|$ is the spectral norm.

Lemma 2.9. *Let A be a $2^n \times 2^n$ matrix satisfying* $\deg(A) = \ell$. *Let U be a unitary of the form $U = \bigotimes_i U_i$, where each U_i is a local unitary acting on at most t qubits. Then*

$$\deg\left(UAU^\dagger\right) \leq \ell t.$$

Proof. Since A has degree at most ℓ, we have the Pauli expansion

$$A = \sum_{\substack{\sigma \in \{0,1,2,3\}^n \\ |\sigma| \le \ell}} \widehat{A}(\sigma) \cdot \mathcal{B}_\sigma.$$

Then

$$UAU^\dagger = \sum_{\substack{\sigma \in \{0,1,2,3\}^n \\ |\sigma| \le \ell}} \widehat{A}(\sigma) \cdot U\mathcal{B}_\sigma U^\dagger.$$

Note that each $U\mathcal{B}_\sigma U^\dagger$ acts non-trivially on at most ℓt qubits, because for the U_j that do not act on the non-trivial part of $\mathcal{B}_\sigma$, we have $U_j U_j^\dagger = \mathbb{1}$. This proves that $\deg\left(UAU^\dagger\right) \le \ell t$. □

Lemma 2.10. *Let M be a $2^n \times 2^n$ matrix and* $\mathrm{diag}(M)$ *be the diagonal matrix containing the diagonal entries of M. If M has a Pauli expansion*

$$M = \sum_{\sigma \in \{0,1,2,3\}^n} \widehat{M}(\sigma)\mathcal{B}_\sigma,$$

then

$$\mathrm{diag}(M) = \sum_{\sigma \in \{0,3\}^n} \widehat{M}(\sigma)\mathcal{B}_\sigma.$$

Proof.

$$\mathrm{diag}(M) = \sum_x \langle x| M |x\rangle \cdot |x\rangle\langle x| = \sum_{x,\sigma} \widehat{M}(\sigma) \cdot \langle x| \mathcal{B}_\sigma |x\rangle \cdot |x\rangle\langle x|.$$

Since $\mathcal{B}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, we have $\langle b| \mathcal{B}_1 |b\rangle = 0$ for any $b \in \{0,1\}$. This also holds for $\mathcal{B}_2$. So $\langle x| \mathcal{B}_\sigma |x\rangle \ne 0$ only if $\sigma \in \{0,3\}^n$. □

One consequence is that the approximate degree of the diagonal part does not exceed the approximate degree of the original matrix.

Lemma 2.11. *Given a $2^n \times 2^n$ matrix M, it holds that, for any ε,*

$$\widetilde{\deg}_\varepsilon\left(\mathrm{diag}(M)\right) \le \widetilde{\deg}_\varepsilon(M).$$

Proof. Let $\widetilde{M}$ be the operator satisfying

- $\left\|M - \widetilde{M}\right\| \le \varepsilon$.
- $\deg(\widetilde{M}) = \widetilde{\deg}_\varepsilon(M)$.

By Lemma 2.10, $\deg\left(\mathrm{diag}\left(\widetilde{M}\right)\right) \le \deg\left(\widetilde{M}\right)$. Also

$$\begin{aligned}\left\|\mathrm{diag}(M) - \mathrm{diag}\left(\widetilde{M}\right)\right\| &= \max_x \left|\langle x| \left(\mathrm{diag}(M) - \mathrm{diag}\left(\widetilde{M}\right)\right) |x\rangle\right| \\ &= \max_x \left|\langle x| \left(M - \widetilde{M}\right) |x\rangle\right| \\ &\le \left\|M - \widetilde{M}\right\|.\end{aligned}$$

□

Lemma 2.12. *Let M be a $2^n \times 2^n$ Hermitian matrix. For any $k \le n$ and $2^k \times 2^k$ density operator φ, let*

$$M_\varphi = \mathrm{Tr}_{n-k+1,\dots,n}\left[(\mathbb{1} \otimes \varphi)\, M\right].$$

For any $\varepsilon \in [0,1]$, we have

$$\widetilde{\deg}_\varepsilon\left(M_\varphi\right) \le \widetilde{\deg}_\varepsilon\left(M\right).$$

Proof. We first consider the exact degree, i.e., $\varepsilon = 0$. Suppose M has Fourier expansion

$$M = \sum_{\sigma \in \{0,1,2,3\}^n} \widehat{M}(\sigma)\mathcal{B}_\sigma.$$

Then

$$\begin{aligned} M_\varphi &= \mathrm{Tr}_{n-k+1,\dots,n}\left[(\mathbb{1} \otimes \varphi)\, M\right] \\ &= \sum_{\sigma \in \{0,1,2,3\}^n} \widehat{M}(\sigma)\, \mathrm{Tr}_{n-k+1,\dots,n}\left[(\mathbb{1} \otimes \varphi)\, \mathcal{B}_\sigma\right] \\ &= \sum_{\substack{\sigma_1 \in \{0,1,2,3\}^{n-k} \\ \sigma_2 \in \{0,1,2,3\}^k}} \widehat{M}(\sigma_1\sigma_2)\, \mathrm{Tr}\left[\varphi \mathcal{B}_{\sigma_2}\right] \mathcal{B}_{\sigma_1} \end{aligned}$$

and we can easily see that $\deg(M_\varphi) \le \deg(M)$.

Then for $\varepsilon > 0$, Let $\widetilde{M}$ be the matrix satisfying $\deg(\widetilde{M}) = \widetilde{\deg}_\varepsilon(M)$ and $\left\|M - \widetilde{M}\right\| \le \varepsilon$. Then let $\widetilde{M}_\varphi = \mathrm{Tr}_{k+1,\dots,n}\left[(\mathbb{1} \otimes \varphi)\, \widetilde{M}\right]$. From the $\varepsilon = 0$ result We have $\deg(\widetilde{M}_\varphi) \le \deg(\widetilde{M})$. Let $S_n = \{\phi \in \mathcal{M}_{2^n} : \mathrm{Tr}\,|\phi| = 1\}$. Then

$$\begin{aligned} \left\|M_\varphi - \widetilde{M}_\varphi\right\| &= \max_{\phi \in S_{n-k}} \mathrm{Tr}\left[\left(M_\varphi - \widetilde{M}_\varphi\right)\phi\right] \\ &= \max_{\phi \in S_{n-k}} \mathrm{Tr}\left[\left(M - \widetilde{M}\right)(\phi \otimes \varphi)\right] \\ &\le \max_{\phi \in S_n} \mathrm{Tr}\left[\left(M - \widetilde{M}\right)\phi\right] \\ &= \left\|M - \widetilde{M}\right\| \\ &\le \varepsilon. \end{aligned}$$

□

2.3 Embedding Boolean functions in the space of operators

A natural approach to embedding Boolean functions in the space of matrices is viewing a Boolean function as a diagonal matrix. More specifically, let $f : \{0,1\}^n \to \mathbb{R}$ be a function, we use M_f to denote the $2^n \times 2^n$ square matrix

$$M_f = \sum_x f(x) \cdot |x\rangle\langle x| = \begin{bmatrix} f(0^n) & & \\ & \ddots & \\ & & f(1^n) \end{bmatrix}. \tag{1}$$

The notion of Fourier expansion of f coincides with the Pauli expansion of M_f:

Fact 2.13. Suppose f has a Fourier expansion

$$f = \sum_{S \subseteq [n]} \widehat{f}(S) \chi_S.$$

For each $\sigma \in \{0,3\}^n$, define $S_\sigma = \{i : \sigma_i = 3\}$. Then M_f has Pauli expansion

$$M_f = \sum_{\sigma \in \{0,3\}^n} \widehat{f}(S_\sigma) \cdot \mathcal{B}_\sigma.$$

One consequence is that the degrees and the approximate degrees of f and M_f are the same, respectively.

Fact 2.14. $\widetilde{\deg}_\varepsilon(f) = \widetilde{\deg}_\varepsilon(M_f)$ for any $\varepsilon \in [0,1]$.

Proof. Note that since M_f is diagonal, by Lemma 2.10, the operator B that satisfies $\|M_f - B\| \leq \varepsilon$ and minimizes the degree is also diagonal. □

2.4 Quantum Circuits

A quantum circuit is a model of quantum computation. The computation involves an initial input quantum register with the state $|\varphi\rangle$, an ancillary quantum register with the state $|\psi\rangle$, and a series of quantum gates $U_s, \ldots, U_1$, where each U_i is a unitary operator drawn from a predefined gate set $\mathcal{U}$, and acts on a subset of working qubits. After the computation, the working quantum register contains the state

$$U(|\varphi\rangle \otimes |\psi\rangle) = U_s \ldots U_1 (|\varphi\rangle \otimes |\psi\rangle).$$

We can implement an n-qubit to k-qubit quantum channel by tracing out all but the first k qubits after implementing all unitaries. This channel with ancillae in the state $|\psi\rangle$ is denoted by $\Phi_{k,U,|\psi\rangle}$. The subscript k is omitted whenever it is clear from the context.

We may get a classical output by applying a computational basis measurement on the first qubit. That is, we apply the measurement $\{M_0 = |0\rangle\langle 0| \otimes \mathbb{1}, M_1 = |1\rangle\langle 1| \otimes \mathbb{1}\}$, and the probability that we output 1 is

$$\mathrm{Tr}\left[(|0\rangle\langle 0| \otimes \mathbb{1})\, U (|\varphi\rangle\langle\varphi| \otimes |\psi\rangle\langle\psi|)\, U^\dagger\right].$$

We use $C_{U,|\psi\rangle}$ to denote the above classical output of a quantum channel. When $|\psi\rangle = |0\rangle^a$ or there are no ancillae, we may simply write C_U.

We say a quantum circuit C computes f with the worst-case probability $1-\varepsilon$ (with the worst-case error ε) if for any $x \in \{0,1\}^n$,

$$\Pr[C(x) \neq f(x)] \leq \varepsilon,$$

where $C(x)$ is the output of the circuit on input x. Similarly, a quantum circuit computes f with the average-case probability $1 - \varepsilon$ (with the average-case error ε) if

$$\mathop{\mathbb{E}}_{\mathbf{x} \in \{0,1\}} [\Pr[C(\mathbf{x}) \neq f(\mathbf{x})]] \leq \varepsilon.$$

In this work we are concerned with **QAC** circuits. The gate set for **QAC** circuits includes all single-qubit unitaries, along with the multi-qubit CZ-gates[1]. An n-qubit CZ-gate is defined as

$$\mathrm{CZ} = \mathbb{1} - 2\,|1\rangle\langle 1|^n. \tag{2}$$

The gates can be written in the form $U = L_d M_d \ldots M_1 L_0$, where each L_i is a tensor product of single qubit unitaries, and each M_i is a tensor product of CZ-gates. The depth of this circuit is d. A sample **QAC** circuit with input state $|x\rangle$ and ancillae $|\psi\rangle$ is depicted in Fig. 2.

[1] Some definitions use Generalized Toffoli gates. They are equivalent in our case.

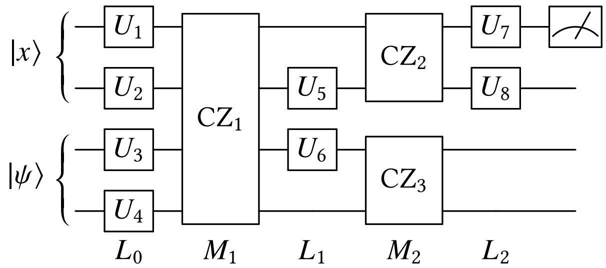


Figure 2: $\mathbf{QAC}^0$ Circuit Example

The complexity class $\mathbf{QAC}^0$ consists of all languages that can be decided by constant-deph and polynomial-sized **QAC** quantum circuits. Formally, a language L is in $\mathbf{QAC}^0$ if there exists a family of constant-deph and polynomial-sized **QAC** quantum circuits $\{C_n\}_{n\in\mathbb{N}}$ such that for any $n \in \mathbb{N}$ and $x \in \{0,1\}^n$, if $x \in L$ then $\Pr[C_n(x) = 1] \geq 2/3$, and if $x \notin L$, then $\Pr[C_n(x) = 0] \geq 2/3$ where $C_n(x)$ is the measurement outcome on the output qubits of the circuit C_n on input x. We also introduce the class of $\mathbf{QLC}^0$ circuits, which consists of $\mathbf{QAC}^0$ circuits with linear-sized ancillae. $\mathbf{QLC}^0$ is a quantum counterpart of the classical circuit family $\mathbf{LC}^0$ introduced in [CR96], which is one of the most interesting subclasses of $\mathbf{AC}^0$ and has received significant attention from various perspectives [KPLT06, CIP09, SS12].

2.5 State Synthesis

In this paper, we also investigate the circuit complexity of quantum state synthesis. In particular, we are interested in the complexity of preparing "cat states"[2] and also "nekomata" states introduced in [Ros21].

Definition 2.15. For $n \geq 1$, the cat state, denoted by $|🐱_n\rangle$, is defined as

$$|🐱_n\rangle = \frac{1}{\sqrt{2}}(|0^n\rangle + |1^n\rangle).$$

Definition 2.16. For $n \geq 1$, a quantum state $|\nu\rangle$ is an n-nekomata, if $|\nu\rangle = \frac{1}{\sqrt{2}}(|0^n, \psi_0\rangle + |1^n, \psi_1\rangle)$ for some arbitrary states $|\psi_0\rangle, |\psi_1\rangle$.

Rosenthal [Ros21] showed that the syntheses of cat states and n-nekomatas are equivalent up to constant depth reductions. In addition, they are equivalent to computing PARITY up to constant depth reductions.

Definition 2.17. We say an n-qubit pure state $\psi = |\psi\rangle\langle\psi|$ is synthesized by a $\mathbf{QAC}^0$ circuit U with a ancillae and fidelity $1-\delta$, if the fidelity between ψ and the first n qubits of $U|0\rangle^{n+a}$ is above $1-\delta$, or formally,

$$F\left(|\psi\rangle\langle\psi|, \mathrm{Tr}_{n+1,\dots,n+a}\left[U|0\rangle\langle 0|^{n+a}U^\dagger\right]\right) = \sqrt{\langle\psi|\,\mathrm{Tr}_{n+1,\dots,n+a}\left[U|0\rangle\langle 0|^{n+a}U^\dagger\right]|\psi\rangle} \geq 1-\delta.$$

Definition 2.18. A family of quantum states $\{\psi_n\}_{n\in\mathbb{N}}$, where ψ_n is an n-qubit state, is in $\mathbf{stateQAC}^0\,[\delta]$ if there exists a family of $\mathbf{QAC}^0$ circuits $\{U_n\}_{n\in\mathbb{N}}$, such that for each n, U_n synthesizes ψ_n with fidelity $1-\delta$. We use $\mathbf{stateQAC}^0$ to represent $\mathbf{stateQAC}^0\,[1/3]$. The class $\mathbf{stateQLC}^0\,[\delta]$ and $\mathbf{stateQLC}^0$ are defined analogously, but with linear ancillae.

[2]In the literature, the name "cat states" appeared in [Sho96, DS96, GHMP01]. They are also known as Greenberger–Horne–Zeilinger states (GHZ states) [GHZ89].

3 Low Degree Approximation of $\mathbf{QAC}^0$ Circuits

In this section, we present our main technical result. In Corollary 3.6, we give a sublinear upper bound on the approximate degree of the operator $UAU^\dagger$, where U is the unitary implemented by a $\mathbf{QAC}^0$ circuit and A is a low-degree operator.

The main ingredient in our proof is the following low-degree approximation of high-degree tensor product states, adapted from [KAAV17a, AM23]. It states that tensor product states such as $|0\rangle\langle 0|^{\otimes n}$ can be approximated by a $\sqrt{n}$-local operator.

Lemma 3.1 ([AM23, Lemma 3.1], see also [KAAV17a]). *Let $H = \sum_{i=1}^n H_i$ be a sum of n commuting projectors each acting on ℓ qubits, and $|\psi\rangle$ be the maximum-energy eigenstate of H. Then, for any $r \in \left(\sqrt{n}, n\right)$, let $\varepsilon = 2^{-\frac{r^2}{2^8 n}}$,*

$$\widetilde{\deg}_\varepsilon \left(|\psi\rangle\langle\psi|\right) \le \ell r.$$

Corollary 3.2. *Let $|\psi\rangle$ be an ℓ-qubit pure state. Then for any $r \in \left(\sqrt{n}, n\right)$, let $\varepsilon = 2^{-\frac{r^2}{2^8 n}}$. It holds that*

$$\widetilde{\deg}_\varepsilon \left(|\psi\rangle\langle\psi|^{\otimes n}\right) \le \ell r.$$

Proof. For $i = 1, \ldots, n$, define $H_i = \mathbb{1}_{i-1} \otimes |\psi\rangle\langle\psi| \otimes \mathbb{1}_{n-i}$. Then clearly H_i are commuting projectors, each acting on ℓ qubits. In addition, the maximum energy eigenstate of $H = \sum_i H_i$ is precisely $|\psi\rangle^{\otimes n}$. So we can apply Lemma 3.1. □

This approximation also works for CZ-gates, as an n-qubit CZ-gate can be written as

$$\mathrm{CZ}_n = \mathbb{1}_n - 2\,|1\rangle\langle 1|^{\otimes n}.$$

Corollary 3.3. *For any CZ-gate CZ acting on n qubits and real number $1 < r < n$, there exists an operator $\widetilde{\mathrm{CZ}}$ such that*

$$\left\|\mathrm{CZ} - \widetilde{\mathrm{CZ}}\right\| \le 2^{1-2^{-8} r}$$

and

$$\deg\left(\widetilde{\mathrm{CZ}}\right) \le \sqrt{nr}.$$

Proof. We invoke Corollary 3.2 with $\ell \leftarrow 1$ and $r \leftarrow \sqrt{nr}$, to get a degree $\sqrt{nr}$ operator Z satisfying $\left\| |1\rangle\langle 1|^n - Z\right\| \le 2^{-2^{-8} r}$. Then we let $\widetilde{\mathrm{CZ}} = \mathbb{1} - 2Z$. □

The above approximation will be applied to a $\mathbf{QAC}^0$ circuit in a layer-by-layer fashion.

3.1 Approximating a Single Layer

The following lemma captures the increase of approximate degrees, when we apply a layer of CZ-gates to a low-degree operator.

Lemma 3.4. *Let $n \ge 1$ be an integer and $U = \bigotimes_i \mathrm{CZ}_i$ be a layer of CZ-gates acting on totally n qubits. For any integer ℓ and $r \in (1, n)$, there exists an operator $\widetilde{U}$ such that*

$$\left\|U - \widetilde{U}\right\| \le \varepsilon = n \cdot 2^{1-2^{-8} r} \log e$$

and for any $2^n \times 2^n$ operator A with degree at most ℓ,

$$\deg\left(\widetilde{U} A \widetilde{U}^\dagger\right) \le 3 n^{\frac{2}{3}} \ell^{\frac{1}{3}} r^{\frac{1}{3}}.$$

Proof. Suppose for each i, CZ_i acts on s_i qubits. Let $t = n^{\frac{2}{3}}\ell^{-\frac{2}{3}}r^{\frac{1}{3}}$ be a threshold and divide the CZ-gates into the set of small CZ-gates $S = \{i : s_i \le t\}$ and the set of large CZ-gates gates $T = \{i : s_i > t\}$. Let $U_S = \bigotimes_{i\in S}\mathrm{CZ}_i$ and $U_T = \bigotimes_{i\in T}\mathrm{CZ}_i$. Clearly $U = U_S \otimes U_T$.

Note that we can assume without loss of generality that $r < t$ because otherwise we have $r \ge t = n^{\frac{2}{3}}\ell^{-\frac{2}{3}}r^{\frac{1}{3}}$, implying $n \le \ell r$, and then $\deg\left(\widetilde{U}A\widetilde{U}^\dagger\right) \le n \le 3n^{\frac{2}{3}}\ell^{\frac{1}{3}}r^{\frac{1}{3}}$ trivially holds by setting $\widetilde{U} = U$.

Then for any $i \in T$, we have $r < t < s_i$ and we can approximate CZ_i using Corollary 3.3 with parameter r. For each $i \in T$, we get the approximation operator $\widetilde{\mathrm{CZ}}_i$ satisfying

- $\left\|\mathrm{CZ}_i - \widetilde{\mathrm{CZ}}_i\right\| \le 2^{1-2^{-8r}}$ and
- $\deg\left(\widetilde{\mathrm{CZ}}_i\right) \le \sqrt{s_i r}$.

We define $\widetilde{U}_T = \bigotimes_{i\in T}\widetilde{\mathrm{CZ}}_i$, and our approximation for U will be $\widetilde{U} = U_S \otimes \widetilde{U}_T$. We first argue that U and $\widetilde{U}$ are close in spectral norm: By Lemma 2.7,

$$\left\|U - \widetilde{U}\right\| = \left\|U_S \otimes U_T - U_S \otimes \widetilde{U}_T\right\| = \left\|U_T - \widetilde{U}_T\right\| = \left\|\bigotimes_{i\in T}\mathrm{CZ}_i - \bigotimes_{i\in T}\widetilde{\mathrm{CZ}}_i\right\| \le \left(1 + 2^{1-2^{-8}r}\right)^n - 1 \le \varepsilon.$$

Then we upper bound the degree of $\widetilde{U}A\widetilde{U}^\dagger$. By Lemma 2.9, we have $\deg\left(U_S A U_S^\dagger\right) \le \ell t$. For the large CZ-gates $\widetilde{U}_T = \bigotimes_{i\in T}\widetilde{\mathrm{CZ}}_i$ part we have

$$\deg\left(\widetilde{U}_T\right) \le \sum_{i\in T}\deg\left(\widetilde{\mathrm{CZ}}_i\right) \le \sum_{i\in T}\sqrt{s_i r}.$$

Note that $s_i > t$ for each $i \in T$ and $\sum_i s_i \le n$. We have

$$\sum_{i\in T}\sqrt{s_i r} = \sum_{i\in T} s_i \cdot \sqrt{\frac{r}{s_i}} < \sum_{i\in T} s_i \cdot \sqrt{\frac{r}{t}} \le n\sqrt{\frac{r}{t}}.$$

Thus by plugging $t = n^{\frac{2}{3}}\ell^{-\frac{2}{3}}r^{\frac{1}{3}}$,

$$\deg\left(\widetilde{U}A\widetilde{U}^\dagger\right) = \deg\left(\widetilde{U}_T U_S A U_S^\dagger \widetilde{U}_T^\dagger\right) \le \deg\left(U_S A U_S^\dagger\right) + 2\deg\left(\widetilde{U}_T\right) \le \ell t + 2n\sqrt{\frac{r}{t}} \le 3n^{\frac{2}{3}}\ell^{\frac{1}{3}}r^{\frac{1}{3}}.$$

□

3.2 Approximating Multiple Layers

A $\mathbf{QAC}^0$ circuit consists of multiple layers of CZ-gates and single-qubit unitaries. Note that by Lemma 2.9, single-qubit unitaries do not change the degree of an operator. Hence, only the CZ-gates concern in our case. Thus, we upper bound the degree of a $\mathbf{QAC}^0$ circuit by applying Lemma 3.4 inductively.

Lemma 3.5. *Let $n \ge 1$ be an integer and $U = L_d M_d \dots L_1 M_1 L_0$ be a $\mathbf{QAC}^0$ circuit of depth d, where L_i consists of only single-qubit unitaries and M_i is the i-th layer of CZ-gates. For any $r \in (1, n)$ there exists an operator $\widetilde{U}$ satisfying*

$$\left\|U - \widetilde{U}\right\| \le d \cdot n \cdot 2^{1-2^{-8}r}\log^2 e$$

and for any $2^n \times 2^n$ operator A with degree at most ℓ,

$$\deg\left(\widetilde{U}A\widetilde{U}^\dagger\right) = O\left(n^{1-3^{-d}} \cdot \ell^{3^{-d}} \cdot r^{1/2}\right).$$

For each different layer i, we will carefully choose the parameter ℓ for Lemma 3.4, to get an approximation $\widetilde{M_i}$ for M_i such that the overall degree will not exceed $\Theta(n)$. Also, Lemma 3.4 ensures that the approximation operator $\widetilde{M_i}$ does not depend on each local monomials of A. Instead of bounding the degree for each monomial individually, we get a unified unitary $\widetilde{M_i}$. This ensures that the error does not depend on the number of local monomials. Hence, we do not need to worry about the norm of each monomial as concerned in [AM23].

Proof of Lemma 3.5. We are only concerned about the M_i gates, because by Lemma 2.9 the single qubit unitaries do not change the degree. The idea is to apply Lemma 3.4 to each M_i layer by layer, totally d times, and get the approximation operators $\widetilde{M}_i$ for each M_i. Then our overall approximation operator will be

$$\widetilde{U} = L_d \widetilde{M}_d \dots L_1 \widetilde{M}_1 L_0.$$

For clarity, for $i = 1, 2, \dots, d$, we define

$$U_i = L_i M_i L_{i-1} M_{i-1} \dots L_1 M_1 L_0,$$

$$\widetilde{U}_i = L_i \widetilde{M}_i L_{i-1} \widetilde{M}_{i-1} \dots L_1 M_1 L_0.$$

We can check that $U_d = U, \widetilde{U}_d = \widetilde{U}$ and $U_0 = \widetilde{U}_0 = L_0$.

We can now start our induction argument. Initially, we let $\ell_0 = \ell$ and $\varepsilon_0 = 0$. For the base case we have

$$\left\| U_0 - \widetilde{U}_0 \right\| = \| L_0 - L_0 \| = 0 = \varepsilon_0$$

and by Lemma 2.9,

$$\deg\left(\widetilde{U}_0 A \widetilde{U}_0^\dagger\right) = \deg\left(L_0 A L_0^\dagger\right) = \deg(A) = \ell = \ell_0.$$

Then, for each i'th iteration where $i = 1, 2, \dots, d$:

- Before applying Lemma 3.4, we have

$$\deg\left(\widetilde{U}_{i-1} A \widetilde{U}_{i-1}^\dagger\right) = \deg\left(L_{i-1} \widetilde{M}_{i-1} \cdots L_0 A L_0^\dagger \cdots \widetilde{M}_{i-1}^\dagger L_{i-1}^\dagger\right) \leq \ell_{i-1}$$

  and

$$\left\| U_{i-1} - \widetilde{U}_{i-1} \right\| = \left\| L_{i-1} M_{i-1} \cdots L_0 - L_{i-1} \widetilde{M}_{i-1} \cdots L_0 \right\| \leq \varepsilon_{i-1}.$$

- After applying Lemma 3.4 with parameter $\ell \leftarrow \ell_{i-1}$, the degree changes from ℓ_{i-1} to

$$\deg\left(\widetilde{U}_i A \widetilde{U}_i^\dagger\right) = \deg\left(L_i \widetilde{M}_i \widetilde{U}_{i-1} A \widetilde{U}_{i-1}^\dagger \widetilde{M}_i^\dagger L_i^\dagger\right) = \deg\left(\widetilde{M}_i \widetilde{U}_{i-1} A \widetilde{U}_{i-1}^\dagger \widetilde{M}_i^\dagger\right) \leq \ell_i = 3n^{2/3} \cdot \ell_{i-1}^{1/3} \cdot r^{1/3}.$$

- By Lemma 2.7, the distance changes from ε_{i-1} to

$$\left\| U_i - \widetilde{U}_i \right\| = \left\| L_i M_i U_{i-1} - L_i \widetilde{M}_i \widetilde{U}_{i-1} \right\| \leq \varepsilon_i = (1 + \varepsilon_{i-1}) \left(1 + n \cdot 2^{1-2^{-8} r} \log e\right) - 1.$$

Now we analyze the degree of $\widetilde{U} A \widetilde{U}^\dagger$: By direct calculation,

$$\ell_i = 3^{\frac{3}{2}\left(1-3^{-i}\right)} n^{1-3^{-i}} \cdot \ell^{3^{-i}} \cdot r^{\frac{1}{2}\left(1-3^{-i}\right)}.$$

So the degree of $\widetilde{U} A \widetilde{U}^\dagger$ is

$$\deg\left(\widetilde{U} A \widetilde{U}^\dagger\right) \leq \ell_d = 3^{\frac{3}{2}\left(1-3^{-d}\right)} n^{1-3^{-d}} \cdot \ell^{3^{-d}} \cdot r^{\frac{1}{2}\left(1-3^{-d}\right)} \leq 3^{\frac{3}{2}} n^{1-3^{-d}} \cdot \ell^{3^{-d}} \cdot r^{1/2}.$$

The distance to such a local operator:

$$\varepsilon_i = \left(1 + n \cdot 2^{1-2^{-8}r} \log e\right)^i - 1$$

So by using the identity $(1+x)^d \leq 1 + xd \log e$ when x is small enough, we have

$$\varepsilon_d = \left(1 + n \cdot 2^{1-2^{-8}r} \log e\right)^d - 1 \leq d \cdot n \cdot 2^{1-2^{-8}r} \log^2 e.$$

□

We can combine Lemma 3.5 and Lemma 2.12 to get the following result

Corollary 3.6. *Let A be a $2^n \times 2^n$ operator acting on n qubits with $\|A\| \leq 1$, and satisfies $\widetilde{\deg}_\varepsilon(A) = \ell$. Let U be a depth d $\mathbf{QAC}^0$ circuit working on a qubits. Then for $\varepsilon' = (1+\varepsilon)(1 + O(d/n)) - 1$, we have*

$$\widetilde{\deg}_{\varepsilon'}\left(UAU^\dagger\right) = \widetilde{O}\left(n^{1-3^{-d}} \cdot \ell^{3^{-d}}\right).$$

Moreover, let $k \leq n$ and φ be a $2^k \times 2^k$ density operator. It holds that

$$\widetilde{\deg}_\varepsilon\left(\mathrm{Tr}_{n-k+1,\ldots,n}\left[UAU^\dagger\left(\mathbb{1} \otimes \varphi\right)\right]\right) = \widetilde{O}\left(n^{1-3^{-d}} \cdot \ell^{3^{-d}}\right).$$

Proof. Let $\widetilde{A}$ be a degree-ℓ operator satisfying $\left\|A - \widetilde{A}\right\| \leq \varepsilon$. Let $\widetilde{U}$ be the operator obtained from Lemma 3.5 with $r = 2^8(1 + 2\log n)$. Then

$$\deg\left(\widetilde{U}\widetilde{A}\widetilde{U}^\dagger\right) \leq 3^{3/2} n^{1-3^{-d}} \ell^{3^{-d}} r^{1/2} = \widetilde{O}\left(n^{1-3^{-d}} \cdot \ell^{3^{-d}}\right).$$

Moreover, by Lemma 2.7,

$$\left\|UAU^\dagger - \widetilde{U}\widetilde{A}\widetilde{U}^\dagger\right\| \leq (1+\varepsilon)\left(1 + dn2^{1-2^{-8}r} \log^2 e\right)^2 - 1 = (1+\varepsilon)(1 + O(d/n)) - 1.$$

□

4 Boolean Functions

In this section we present our hardness results for computing Boolean functions using $\mathbf{QAC}^0$ circuits. The starting point is the following theorem, which states that if a depth-d quantum circuit with a ancillae computes a function $p : \{0,1\}^n \to \mathbb{R}$, then the approximate degree of p is upper bounded by $\widetilde{O}((n+a)^{1-3^{-d}})$. This implies that when the number of ancillae is only $n^{1+3^{-d}}$, which is slightly more than linear, then $(n+a)^{1-3^{-d}} = o(n)$, and this $\mathbf{QAC}^0$ circuit cannot compute Boolean functions of an approximate degree $\Omega(n)$. These include the PARITY, MAJORITY, and MOD_k functions that we define in Example 2.5.

Let U be the unitary implemented by a $\mathbf{QAC}^0$ circuit on $n+a$ qubits, where the first n qubits serve as input, and the last a input qubits are initialized to the state φ. We will upper bound the approximate degree of the operator

$$\mathrm{Tr}_{n+1,\ldots,n+a}\left[U^\dagger\left(|1\rangle\langle 1| \otimes \mathbb{1}\right) U \left(\mathbb{1} \otimes \varphi\right)\right],$$

which is actually the projector onto the input space which outputs 1. The diagonal entries of this operator are actually the probabilities of outputting 1 on every classical input, and corresponds the the matrix M_p defined in Eq. (1).

Theorem 4.1. *Let $n \geq 1$ and U be a depth-d $\mathbf{QAC}^0$ circuit with n input qubits and a ancillae initialized in state φ. Let $p : \{0,1\}^n \to \mathbb{R}$ be the probability that U outputs 1. That is, for an input $x \in \{0,1\}^n$,*

$$p(x) = \mathrm{Tr}\left[(|1\rangle\langle 1| \otimes \mathbb{1})\, U\, (|x\rangle\langle x| \otimes \varphi)\, U^\dagger\right].$$

Then $\widetilde{\deg}_\varepsilon(p) = \widetilde{O}\left((n+a)^{1-3^{-d}}\right)$ for $\varepsilon = O(d/n)$.

Setting $a = O(n)$, we have the following corollary, which extends the analogous result for $\mathbf{LC}^0$ proved in [BT22].

Corollary 4.2. *For any $\varepsilon > 0$, it holds that*

$$\widetilde{\deg}_\varepsilon\left(\mathbf{QLC}^0\right) = o(n).$$

Proof of Theorem 4.1. Let $A = U^\dagger(|1\rangle\langle 1| \otimes \mathbb{1})\, U$. Then we have

$$p(x) = \mathrm{Tr}\left[(|x\rangle\langle x| \otimes \varphi)\, A\right] = \langle x|\, \mathrm{Tr}_{n+1,\ldots,n+a}\left[A\,(\mathbb{1} \otimes \varphi)\right] |x\rangle.$$

By Corollary 3.6, we have $\widetilde{\deg}_\varepsilon\left(\mathrm{Tr}_{n+1,\ldots,n+a}\left[A\,(\mathbb{1}\otimes\varphi)\right]\right) \leq \widetilde{O}\left((n+a)^{1-3^{-d}}\right)$. Then the diagonal matrix M_p as defined in Equation (1) can be obtained by zeroing out all the non-diagonal entries of the matrix

$$\mathrm{Tr}_{n+1,\ldots,n+a}\left[A\,(\mathbb{1}\otimes\varphi)\right].$$

So by Lemma 2.11 we have

$$\widetilde{\deg}_\varepsilon\left(M_p\right) \leq \widetilde{\deg}_\varepsilon\left(\mathrm{Tr}_{n+1,\ldots,n+a}\left[A\,(\mathbb{1}\otimes\varphi)\right]\right) \leq \widetilde{O}\left((n+a)^{1-3^{-d}}\right).$$

Finally by Fact 2.14 we have $\widetilde{\deg}_\varepsilon(p) = \widetilde{\deg}_\varepsilon\left(M_p\right)$ and we prove our theorem. □

With Theorem 4.1 we are able to prove that $\mathbf{QAC}^0$ circuits with linear ancillae cannot compute high-degree functions. In the following subsections, we give two flavors of hardness results. In Section 4.1 we prove the worst-case hardness and in Section 4.2 we prove the average-case hardness. The average-case hardness results follow from the framework of [NPVY24]. Although worst-case hardness results are weaker than average-case hardness results, they can be applied to a much broader class of Boolean functions. For example, for the MAJORITY function, we prove an optimal $1/2$ worst-case hardness result, while for the average case, we can only prove that approximating within a success probability of $1 - \frac{1}{\sqrt{n}}$ in the average case is impossible. For the MOD_k function, we also prove an optimal $1/2$ worst-case hardness result. However, in the average case, since MOD_k is a biased function, we can even achieve a success probability of $1 - 1/k$ simply by always outputting 1.

The reason for this distinction between worst-case hardness and average-case hardness comes from the fact that the normalized Frobenius norm can be exponentially smaller than the spectral norm. So a Boolean function f is much more easily approximated using low-degree functions in the Frobenius norm. But for the average-case hardness framework from [NPVY24] that we are using, we need a function that is of $\Omega(n)$ degree even if approximated within the Frobenius norm.

In Section 5, we also show that our worst-case ancillae lower bound of $n^{1+3^{-d}}$ for the PARITY function is just one step from a complete resolution of the PARITY $\notin \mathbf{QAC}^0$ conjecture: We show that any improvement of our ancillae lower bound to $n^{1+\exp(-o(d))}$, will imply that $\mathbf{QAC}^0$ circuits of arbitrary polynomial ancillae can not compute PARITY.

4.1 Worst-Case Hardness

With Theorem 4.1, we are able to prove lower bounds on Boolean functions with approximate degrees $\Omega(n)$.

Theorem 4.3. *Let $n \geq 1$. Suppose that we have a depth d QAC^0 circuit U that computes a function $f : \{0,1\}^n \to \{0,1\}$ with the worst-case probability $1-\delta$. U has n input qubits and a ancillae initialized to the state φ. Then for any $\varepsilon = \delta + O(d/n)$, we have*

$$\widetilde{\deg}_{\varepsilon}(f) = \widetilde{O}\left((n+a)^{1-3^{-d}}\right).$$

Moreover, if there exists a QAC^0 circuit that computes a function f satisfying $\widetilde{\deg}_{1/3}(f) = \Omega(n)$, with a constant error strictly below $1/2$, then we have $a = \widetilde{\Omega}\left(n^{1+3^{-d}}\right)$.

Corollary 4.4. *Suppose a QAC^0 circuit with depth d and a ancillae computes a Boolean function $f : \{0,1\}^n \to \{0,1\}$ where $f \in \left\{\mathrm{Parity}_n, \mathrm{Majority}_n, \mathrm{Mod}_{n,k}\right\}$ with $2 \leq k \leq cn$ for some $c < 1$, and the worst-case error is a constant below $1/2$, then*

$$a = \widetilde{\Omega}\left(n^{1+3^{-d}}\right).$$

Proof of Theorem 4.3. By Theorem 4.1, there exists two Boolean functions $p, \widetilde{p} : \{0,1\} \to \mathbb{R}$ satisfying

- $\|p - \widetilde{p}\| \leq O(d/n)$.
- $\deg(\widetilde{p}) = \widetilde{O}((n+a)^{1-3^{-d}})$.

Also, since the circuit U computes f with the worst-case error δ, we have $\|f - p\| \leq \delta$. This implies $\|f - \widetilde{p}\| \leq \|f - p\| + \|p - \widetilde{p}\| \leq \delta + O(d/n)$, and this proves the first part of the theorem.

Given a QAC^0 circuit that computes a Boolean function $f : \{0,1\}^n \to \mathbb{R}$, with a constant error strictly less than $1/2$ in the worst case, we may reduce the error below $1/3$ using the standard error reduction, i.e., repeating the computation constant times and taking the majority output, which multiplies the depth of the circuit and the size of ancillae by a constant. Thus, we still have $(n + a)^{1-3^{-d}} = \widetilde{\deg}_{1/3}(f) = \widetilde{\Omega}(n)$, which concludes the proof. □

For QAC^0 circuits where the depth d is a constant, the theorem above implies that if we wish to compute a function with $\Omega(n)$ approximate degree, and with any nontrivial constant worst-case success probability, the number of ancilla qubits needs to be superlinear.

4.2 Average-Case Hardness

In this subsection, we will prove the average-case hardness result for high-degree Boolean functions using arguments similar to those of [NPVY24]. To prove the average-case hardness, we require that the Boolean function be of approximately high degree in terms of Frobenius norm.

The level-k Fourier weight of f, denoted by $\mathrm{W}^{=k}(f)$, is the sum of the squares of the Fourier coefficients of degree exactly k. That is

$$\mathrm{W}^{=k}(f) = \sum_{\substack{S \subseteq [n] \\ |S| = k}} \widehat{f}(S)^2.$$

The quantites $\mathrm{W}^{\leq k}(f), \mathrm{W}^{<k}(f), \mathrm{W}^{\geq k}(f)$ and $\mathrm{W}^{>k}(f)$ are defined analogously. For a Boolean function that has large high-degree Fourier coefficients, that is, $\mathrm{W}^{>k}(f)$ is large, it cannot be approximated by a low-degree Boolean function with respect to the Frobenius norm. This also means that its approximate degree is large as the Frobenius norm is an lower bound of the spectral norm.

Lemma 4.5. *Let $f : \{0,1\}^n \to \mathbb{R}$ be a Boolean function. If $\mathrm{W}^{\geq k}(f) \geq \delta$, then for any $\varepsilon < \sqrt{\delta}$, we have $\widetilde{\deg}_\varepsilon(f) \geq k$.*

Proof. For any function $g : \{0,1\}^n \to \mathbb{R}$ with $\|f - g\| \leq \varepsilon$, we have $\|f-g\|_2 \leq \|f-g\| \leq \varepsilon$. Thus, $\mathrm{W}^{\geq k}(g) \geq \mathrm{W}^{\geq k}(f) - \|f-g\|_2^2 \geq \delta - \varepsilon^2 > 0$. This implies $\deg(g) \geq k$. □

By Theorem 4.1, we can prove average-case hardness results for Boolean functions with large high-degree Pauli weights.

Theorem 4.6. *Suppose U is a QAC^0 circuit with depth d. U has n input qubits and a ancillae initialized in state φ. Let $f : \{0,1\}^n \to \{0,1\}$ be any Boolean function. It holds that*

$$\Pr_{x\sim_U\{0,1\}^n}[C_U(x) = f(x)] \leq \frac{1}{2} + \frac{1}{2}\cdot\sqrt{1 - 4\mathrm{W}^{>k}(f)} + O(d/n),$$

for $k \geq \widetilde{\Theta}\left((n+a)^{1-3^{-d}}\right)$. In particular, for $a = \widetilde{O}(n^{1+3^{-d}})$,

$$\Pr_{x\sim_U\{0,1\}^n}\left[C_U(x) = \mathrm{Parity}_n(x)\right] \leq \frac{1}{2} + O(d/n)$$

and

$$\Pr_{x\sim_U\{0,1\}^n}\left[C_U(x) = \mathrm{Majority}_n(x)\right] \leq 1 - \Omega\left(\frac{1}{\sqrt{n}}\right) + O(d/n).$$

Remark 4.7. This theorem also holds in the case where we want to compute a Boolean function with randomized output. In this case, the function f now has the form $f : \{0,1\}^n \to \mathbb{R}$, where for each input $x \in \{0,1\}^n$, the value $f(x)$ denotes the probability of output 1. The proof is literally the same.

Proof of Theorem 4.6. For any $x \in \{0,1\}^n$, let $p(x)$ be the probability that the circuit outputs 1 on input x. By Theorem 4.1, $\widetilde{\deg}_\varepsilon(p) = \widetilde{O}\left((n+a)^{1-3^{-d}}\right)$. Then there exists a function $\widetilde{p} : \{0,1\}^n \to \mathbb{R}$ satisfying $\|p - \widetilde{p}\| \leq \varepsilon$ and $\deg(\widetilde{p}) \leq \widetilde{O}\left((n+a)^{1-3^{-d}}\right) \leq k$. Let $q = 2p-1$, $\widetilde{q} = 2\widetilde{p}-1$ and $g = 2f-1$. Then $\|q - \widetilde{q}\|_2 \leq \|q-\widetilde{q}\| \leq 2\varepsilon$. Let $\mathbf{x}$ be a random variable uniformly distributed over $\{0,1\}^n$.

$$\begin{aligned}
2\cdot\Pr_{x\sim_U\{0,1\}^n}[C_U(x) = f(x)] - 1 &= \mathop{\mathbb{E}}_{x\sim_U\{0,1\}^n}[q(x)g(x)] \\
&= \sum_{S\subseteq[n]}\widehat{q}(S)\widehat{g}(S) \\
&= \sum_{|S|\leq k}\widehat{q}(S)\widehat{g}(S) + \sum_{|S|>k}\widehat{q}(S)\widehat{g}(S) \\
&\leq \sqrt{\mathrm{W}^{\leq k}(q)\,\mathrm{W}^{\leq k}(g)} + \sqrt{\mathrm{W}^{>k}(q)\,\mathrm{W}^{>k}(g)} \\
&\leq \sqrt{\mathrm{W}^{\leq k}(g)} + \|q-\widetilde{q}\|_2\sqrt{\mathrm{W}^{>k}(g)} \\
&\leq \sqrt{\mathrm{W}^{\leq k}(g)} + 2\varepsilon\sqrt{\mathrm{W}^{>k}(g)}.
\end{aligned}$$

The second inequality follows since $\widetilde{q}$ has degree below k, so $\sqrt{\mathrm{W}^{>k}(q)} = \sqrt{\mathrm{W}^{>k}(q-\widetilde{q})} \leq \|q-\widetilde{q}\|_2$. Tidying up, we have

$$\begin{aligned}
\Pr_{x\sim_U\{0,1\}^n}[C_U(x) = f(x)] &\leq \frac{1}{2} + \frac{1}{2}\cdot\sqrt{\mathrm{W}^{\leq k}(g)} + \varepsilon\cdot\sqrt{\mathrm{W}^{>k}(g)} \\
&\leq \frac{1}{2} + \frac{1}{2}\cdot\sqrt{1 - \mathrm{W}^{>k}(g)} + \varepsilon\cdot\sqrt{\mathrm{W}^{>k}(g)} \\
&\leq \frac{1}{2} + \frac{1}{2}\cdot\sqrt{1 - 4\mathrm{W}^{>k}(f)} + \varepsilon.
\end{aligned}$$

Notice that $\text{Parity}_n = \frac{1}{2} + \frac{1}{2}\chi_{[n]}$. So we have $W^{>k}(\text{Parity}_n) = \frac{1}{4}$ for $k < n$. For the Majority function, we have $W^{>k}(\text{Majority}_n) \geq \Omega\left(\frac{1}{\sqrt{k}}\right)$ for odd k [O'D14, Equation 5.11]. □

This theorem is an analog of Proposition 32 and Theorem 33 of [NPVY24]. One main difference is that we use Theorem 4.1 to directly upper bound the degree of the Boolean function, instead of considering the Choi representation of quantum channels. This turns out to be more straightforward and efficient in our applications.

The result of [NPVY24] only allows $n^{\frac{1}{d}}$-qubit ancillary pure state, the barrier being the increase of the Frobenius norm when applying the operation $P \mapsto \text{Tr}_{n+1,\ldots,n+a}[P(\mathbb{1} \otimes \varphi)]$. We overcome this barrier by using spectral norm approximations. However, our method creates new barriers. Since we are not "removing" the large CZ-gates as in [NPVY24], in the case where there are too many, say $\omega(n)$ large CZ-gates, our method immediately fails. Also, it is still open whether we can approximate $\mathbf{AC}^0$ circuits with degree $o(n)$ polynomials [BT22]. It would be even more challenging to do so with quantum circuits.

Rosenthal [Ros21] proved that computing Parity_n is equivalent to synthesizing the n-qubit cat state. It is interesting to note that our hardness results Theorem 4.3 and Theorem 4.6 hold for any linear-sized ancillae, which implies that only providing a cat state in the ancillae is not enough to compute PARITY. Instead, a quantum circuit that synthesizes the cat state is needed for the reduction.

4.3 Application: Agnostic Learning for $\mathbf{QLC}^0$

Kearns, Schapire and Sllie [KSS92] proposed the agnostic PAC learning model, which is a more general model than the standard PAC learning model. Let $\mathcal{D}$ be a distribution on $\{0,1\}^n \times \{0,1\}$ and let $\mathcal{C}$ be a concept class. For any $h : \{0,1\}^n \to \{0,1\}$, the error is defined to be $\text{err}_{\mathcal{D}}(h) = \Pr_{(x,b)\sim\mathcal{D}}[h(x) \neq y]$ and one defines $\text{opt} = \min_{c\in\mathcal{C}} \text{err}_{\mathcal{D}}(c)$. We say that $\mathcal{C}$ is agnostically learnable in time $T(n,\varepsilon,\delta)$, if there exists an algorithm that takes as input n,δ and has access to an example oracle of $\mathcal{D}$, and satisfies the following properties: It runs in time at most $T(n,\varepsilon,\delta)$ and with probability at least $1-\delta$, it outputs a hypothesis h that satisfies $\text{err}_{\mathcal{D}}(h) \leq \text{opt} + \varepsilon$. The agnostic model is believed to be closer to the realistic scenario than the standard PAC model. However, designing efficient agnostic learning algorithms is challenging in general. Even very few concept classes are known to be agnostically learnable in subexponential time. Bun, Kothari, and Thaler [BKT19] gave a subexponential-time agnostic learning algorithm for the class of functions with approximate degree n^c for $c < 1$, built on [KKMS08].

Fact 4.8 ([BKT19, Corollary 24])**.** Let C be a set of Boolean functions on $\{0,1\}^n$. Suppose that for every $c \in C$, the ε-approximate degree of c is at most d. Then for every $\delta > 0$, there is an algorithm running in time $\text{poly}(n^d, 1/\varepsilon, \log(1/\delta))$ that agnostically learns C to error ε with respect to any (unknown) distribution D over $\{0,1\}^n \times \{0,1\}$.

As a simple corollary of Theorem 4.1, we know that the classical functions that a depth-d $\mathbf{QLC}^0$ circuit computes has approximate degree $\widetilde{O}(n^{1-3^{-d}})$. As an immediate corollary, we have a subexponential learning algorithm for the functions in $\mathbf{QLC}^0$ in the distribution-free agnostic PAC learning model, generalizing the classical counterpart [BKT19, Theorem 7].

Theorem 4.9. *The concept class of n-bit functions computed by $\mathbf{QLC}^0$ circuits of depth d can be learned in the distribution-free agnostic PAC model in time $2^{\widetilde{O}\left(n^{1-3^{-d}}\right)}$.*

4.4 $\mathbf{QAC}^0$ with low-degree classical post-processing

For a quantum channel, we measure the output qubits in the computational basis. Theorem 4.1 relies on the fact that we only look at the first qubit of the output of a quantum circuit. That is, we consider the projector $|0\rangle\langle 0| \otimes \mathbb{1}$, which is of degree 1. We can allow more complicated classical post-processing of the quantum output, e.g. linear-size $\mathbf{AC}^0$ circuits, as long as the post-processing is low degree. By slightly modifying the proof of Theorem 4.1, we obtain the following theorem.

Theorem 4.10. *Let $n \geq 1$. Let U be a depth-d $\mathbf{QAC}^0$ circuit with n input qubits and a ancillae initialized in the state φ. Let $f : \{0,1\}^{n+a} \to \mathbb{R}$ be a Boolean function satisfying $\widetilde{\deg}_\delta(f) = \ell$. Let $p : \{0,1\}^n \to \mathbb{R}$ be the function describing the probability that $f \circ U$ outputs 1. That is, for an input $x \in \{0,1\}^n$,*

$$p(x) = \mathrm{Tr}\left[M_f \cdot U\left(|x\rangle\langle x| \otimes \varphi\right) U^\dagger\right].$$

Then $\widetilde{\deg}_\varepsilon(p) = \widetilde{O}\left((n+a)^{1-3^{-d}} \cdot \ell^{3^{-d}}\right)$ for $\varepsilon = (1+\delta)(1+O(d/n)) - 1$. In particular, it implies that $L \notin \mathbf{LC}^0 \circ \mathbf{QLC}^0$ for $L \in \{\text{PARITY}, \text{MAJORITY}, \text{MOD}_k\}$, where $0 < c < 1$ and $2 \leq k \leq cn$.

The proof is almost identical to the proof of Theorem 4.1, we provide a proof for completeness in Appendix A. This improves upon the results of Slote [Slo24], where he proved a similar result for QNC^0 pre-processing.

5 Towards PARITY $\notin$ QAC^0

In this section, we show that our result Corollary 4.4 is just one step away from the conjecture that PARITY $\notin$ $\mathbf{QAC}^0$, despite that our "barely superlinear" lower bound of $\widetilde{\Omega}\left(n^{1+3^{-d}}\right)$ seems far away from the arbitrary polynomial ancillae allowed in $\mathbf{QAC}^0$ circuits. In particular, we show that $n^{1+\exp(-\Theta(d))}$ ancillae is a threshold. Any ancillae lower bound of the form $n^{1+\exp(-o(d))}$, e.g., $n^{1+\exp(d/\log d)}$, will lead to a lower bound for any polynomial ancillae. To see this, we show that given a $\mathbf{QAC}^0$ circuit family with depth d and ancillae n^c, we can get a $\mathbf{QAC}^0$ circuit computing Parity_n with depth $O(cdt)$ and ancillae $n^{1+2^{-t}}$, for any constant t.

Definition 5.1. A $[d \,|\, n \,|\, a]$ circuit is a $\mathbf{QAC}^0$ circuit with depth d, input size n and ancillae size a.

We first present the main theorem of this section, which we will prove by directly combining Lemma 5.6 and Lemma 5.7 at the end of this section.

Theorem 5.2. *Suppose there exist constants $d \in \mathbb{Z}^+, c \in \mathbb{Z}^+, N_0 \in \mathbb{Z}^+$, and a $\mathbf{QAC}^0$ circuit family $\{U_n\}_{n \geq N_0}$ such that for each $n \geq N_0$, the circuit U_n is a $[d \,|\, n \,|\, n^c]$ circuit, and computes Parity_n with the worst-case error $\text{negl}(n)$. Then there exists $D = O(cd)$, for any $K \geq 1$ and infinitely many n, we can construct a $\left[KD \,\middle|\, n \,\middle|\, n^{1+2^{-K}}\right]$ circuit, which computes Parity_n with the worst-case error $\text{negl}(n)$.*

Corollary 5.3. *For any function $\delta : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{x\to\infty} \delta(x) = \infty$, if any $\mathbf{QAC}^0$ circuit with $n^{1+\exp(-d/\delta(d))}$ ancillae, where d is the depth of this circuit family, can not compute Parity_n with the worst-case error $\text{negl}(n)$, then any $\mathbf{QAC}^0$ circuit family, where arbitrary polynomial ancillae is allowed, can not compute Parity_n with the worst-case error $\text{negl}(n)$.*

Proof. Suppose on the contrary, there exists a $\mathbf{QAC}^0$ circuit family that computes Parity_n with the worst-case error $\text{negl}(n)$. Then let D be the constant in Theorem 5.2, there exists a large enough constant K such that $D/\delta(KD) \leq 1/2$, and then

$$n^{1+\exp(-KD/\delta(KD))} \geq n^{1+\exp(-K/2)} \geq n^{1+2^{-K}}.$$

By Theorem 5.2, we can construct a $\mathbf{QAC}^0$ circuit with depth KD and $n^{1+2^{-K}} \leq n^{1+\exp(-(KD)/\delta(KD))}$ ancillae, a contradiction. □

5.1 Proof Overview of Theorem 5.2

For any $c > 1$, given a $[d \,|\, n \,|\, n^c]$ circuit U_n, we wish to construct a $\left[KD \middle| N \middle| N^{1+2^{-K}}\right]$ circuit for some constant D, any K, and infinitely many N. We use the idea that Parity_m can be computed recursively. Thus, we can construct a circuit to compute Parity_m for some $m > n$ using U_n as a building block. In this way of computing parity, ancillae size grows slower than m^c. We first give a very simple example. For $m = 2n$, we can construct a 2-layer circuit that computes Parity_m with depth $d + 1$ and $2n^c$ ancillae, as shown in Fig. 3: In the first layer, we use one instance of U_n to

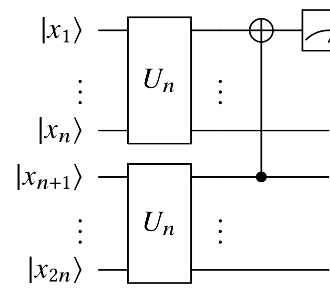


Figure 3: Circuit computing Parity_m for $m = 2n$

compute the parity of the first half of the input bits, and use another instance of U_n to compute the parity of the second half of the input bits. In the second layer, we get the parity of the whole input by computing the XOR of these two outputs.

The advantage of this construction is that now the ancillae size is $2n^c$, which is smaller than $m^c = 2^c n^c$ since $c > 1$. That is, the ancillae size is decreased relative to the input size. We can now amplify this advantage by constructing larger circuits. In the second layer, we can compute the parity of n bits instead of 2 bits, by replacing the CNOT gate into an instance of U_n. Then we can put n instances of U_n instead of 2 instances in the first layer. This yields a circuit computing Parity_m for $m = n^2$, with depth $2d$ and $(n+1) \cdot n^c \approx n^{c+1}$ ancillae. Moreover, by repeating this construction and further considering k-layer circuits as in Fig. 5, where the bottom layer has n^{k-1} instances of U_n, we con construct circuits that compute Parity_m for $m = n^k$, that have depth kd, and use approximately n^{c+k-1} ancillae. Note that

$$n^{c+k-1} = m^{\frac{c+k-1}{k}} = m^{1+\frac{c-1}{k}},$$

so now for any k, we have a $\left[kd \middle| m \middle| m^{1+\frac{c-1}{k}}\right]$ circuit computing Parity_m. For any constant $\varepsilon > 0$ that is independent of the depth of the circuit, we can get a $\left[O(1) \middle| m \middle| m^{1+\varepsilon}\right]$ circuit by choosing $k = (c-1)/\varepsilon$.

In the k-layer construction above, the depth of the circuit grows linearly with respect to k. The best we could expect with this construction is a $\left[d \middle| n \middle| n^{1+\varepsilon}\right]$ circuit for $\varepsilon = d^{-c}$ $(0 < c < 1)$. For instance, if $\varepsilon = d^{-0.99}$, then we can choose a constant k satisfying

$$1 + \frac{c-1}{k} \leq 1 + \varepsilon = 1 + (kd)^{-0.99}.$$

To further compress the size of ancillae, we adapt a more efficient construction. In particular, since the construction above alone can not prove Theorem 5.2, for the sake of clean parameters, we will choose $k = c$, and get a $\left[O(1) \middle| m \middle| m^2\right]$ circuit.

Now the goal becomes reducing the ancillae size from n^2 to $n^{1+\exp(-o(d))}$. Looking back to the k-layer construction in Fig. 5, we observe that we are using the ancillae very inefficiently when the number of layers grows: In the first (bottom) layer, we use n^{k-1} instances of U_n and thus there are n^{c+k-1} ancillae. However, in the second layer, we use n^{k-2} instances of U_n and thus the ancillae used is only $n^{c+k-2} = \frac{1}{n} \cdot n^{c+k-1}$. The top layer is the worst, which uses only n^c ancillae. This indicates that an improvement may be possible if we can adjust the ancillae size in each layer to make them equal. This would not change the order of magnitude of the total number of ancillae. However, we might be able to compute the parity function for larger input sizes, thus effectively reducing the ancilla size with respect to the input size.

This improvement can be easily implemented when we have a $\mathbf{QAC}^0$ circuit family $\{U_n\}$, where each U_n is a $\left[d\,\middle|\,n\,\middle|\,n^2\right]$ circuit. We also know that such a circuit family is promised by the construction in Fig. 5. We illustrate the improved k-layer construction in Fig. 6. In the first (bottom) layer, we fill it with instances of the circuit U_n. In the second layer, we fill it with instances of the circuit U_{n^2}. In the ith layer, we fill it with instances of the circuit $U_{n^{2^{i-1}}}$. Now this construction computes Parity_m for

$$m = \prod_{i=1}^{k} n^{2^{i-1}} = n^{2^k-1}.$$

And the number of instances of $U_{n^{2^{i-1}}}$ in the ith layer can be calculated to be

$$\prod_{j=i+1}^{k} n^{2^{j-1}} = n^{2^k-2^i}.$$

So the ancillae size in the ith layer is

$$n^{2^k-2^i} \cdot \left(n^{2^{i-1}}\right)^2 = n^{2^k-2^i} \cdot n^{2^i} = n^{2^k}.$$

And the total ancillae size is $k \cdot n^{2^k} \approx n^{2^k}$. The depth of this construction is kd. In Conclusion, we get a $\left[kd\,\middle|\,n^{2^k-1}\,\middle|\,n^{2^k}\right]$ circuit computing $\text{Parity}_{2^{2^k-1}}$. With $m = n^{2^k-1}$, we have

$$n^{2^k} = m^{\frac{2^k}{2^k-1}} = m^{1+\frac{1}{2^k-1}} \approx m^{1+2^{-k}}.$$

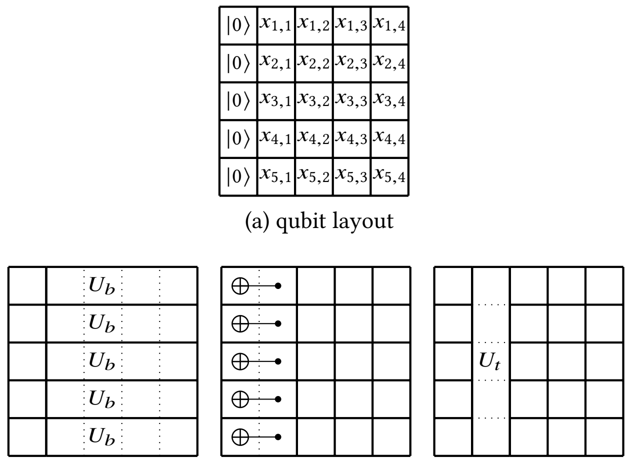
So we have a $\left[kd\,\middle|\,m\,\middle|\,m^{1+2^{-k}}\right]$ circuit computing Parity_m, proving Theorem 5.2.

There are still two caveats. The first one is regarded to the error. The above construction seems to be correct only when each U_n computes Parity_n without any error. When each U_n computes Parity_n with worst case probability $(1+\delta)/2$, and there are totally t instances of these circuits, we show that the total correct probability is $(1+\delta^t)/2$. Since t will always be a polynomial in n, the above construction is stable as long as $\delta = 1 - \text{negl}(n)$. Second, for this accept probability to hold, we need to measure the first output qubit of each U_n before feeding them to the next layer. Although measurements are in general not allowed in $\mathbf{QAC}^0$ circuits, we can achieve the same effect by applying a CNOT gate controlled on the first output qubit of each U_n, targeted to an ancilla initialized to the state $|0\rangle$. These coherent operations slightly increase the overall depth and acillae size of our constructions. See Lemma 5.4 for details.

Therefore, the overall proof of Theorem 5.2 consists of two steps. The first step is to reduce an arbitrary $\mathbf{QAC}^0$circuit computing PARITY to an $\left[O(1)\,\middle|\,n\,\middle|\,n^2\right]$ circuit. The second step further reduces the circuit to $\left[d\,\middle|\,n\,\middle|\,n^{1+\exp(-o(d))}\right]$. It is worth noticing that if we optimize the usage of ancillae in the first step, as we did in the second step, we could also get $\left[d\,\middle|\,n\,\middle|\,n^{1+\exp(-o(d))}\right]$. However, the parameters would be complicated. To keep the clearness of the parameters, we maintain this two-step reduction in our proof.

5.2 Proof of Theorem 5.2

The rest of this section is devoted to proving Theorem 5.2. The building block of our construction is the following lemma, which states that given two $\mathbf{QAC}^0$ circuits computing Parity_{n_1} and Parity_{n_2} respectively, we can construct a $\mathbf{QAC}^0$ circuit computing $\text{Parity}_{n_1 n_2}$.



(a) qubit layout

(b) first layer (c) second layer (d) third layer

Figure 4: Circuit for $\mathrm{Parity}_{n_1 n_2}$

Lemma 5.4. *Suppose there exist two* QAC^0 *circuits* U_t *and* U_b[3]*. such that the circuit* U_t *(resp.* U_b*) is a* $[d_t \mid n_t \mid a_t]$ *(resp.* $[d_b \mid n_b \mid a_b]$*) circuit, and computes* Parity_{n_t} *(resp.* Parity_{n_b}*) with the worst-case probability at least* $(1+\delta_t)/2$ *(resp.* $(1+\delta_b)/2$*). Then we can construct a* QAC^0 *circuit* U*, such that* U *is a* $[d_t + d_b + 1 \mid n_t n_b \mid n_t(a_b + 1) + a_t]$ *circuit, and computes* $\mathrm{Parity}_{n_t n_b}$ *with the worst-case probability at least* $(1 + \delta_b^{n_t}\delta_t)/2$*.*

Remark 5.5. Note that the term $\delta_b^{n_t}$ in the worst-case probability may decrease exponentially fast if δ_b is smaller than $1 - 1/n_t$, e.g., when δ_b is a constant. Also, the input size n_t can be as large as any polynomial in the constructions below. So to keep the worst-case probability of correctly computing $\mathrm{Parity}_{n_t n_b}$ non-trivial, We will always assume $\delta_b = 1 - \mathrm{negl}(n)$ for some negligible function.

Proof of Lemma 5.4. Let $n = n_t n_b$. Observe that for $x \in \{0,1\}^n$, we can write it as $x = x_1 \dots x_{n_t}$ where each $x_i \in \{0,1\}^{n_b}$, and we can implement the function Parity_n by n_t instances of Parity_{n_b}, followed by one instance of Parity_{n_t}:

$$\mathrm{Parity}_n(x) = \mathrm{Parity}_{n_t}\left(\mathrm{Parity}_{n_b}(x_1), \dots, \mathrm{Parity}_{n_b}(x_{n_t})\right).$$

Also, the evaluations of $\mathrm{Parity}_{n_b}(x_1), \dots, \mathrm{Parity}_{n_b}(x_n)$ are independent to each other. So they can be carried out in parallel. This suggests a three-layer circuit for Parity_n, by using the circuits U_t and U_b as subroutines: In the first layer, we compute each $\mathrm{Parity}_{n_b}(x_i)$ in parallel using n_t instances of the circuit U_b. In the second layer, we need to measure each output of U_b in the computational basis. This can be achieved by applying a CNOT gate, controlled on the qubit to be measured, to an ancilla initialized to the state $|0\rangle$. Finally in the third layer, we compute the parity of these outputs, by an instance of the circuit U_t.

The circuit of this construction for $n_t = 5, n_b = 4$ can then be depicted in Fig. 4. The ancillae used by the circuits U_t and U_b are omitted for clarity. We partition the n inputs qubits into an $n_t \times n_b$ grid as in Fig. 4a. In the first layer (Fig. 4b), we apply the circuit U_b to each row, and the output qubit is the left-most qubit. In the second layer (Fig. 4c), we "measure" the output of each row. Each "measurement" is implemented by a CNOT gate plus an ancilla initialized in the state $|0\rangle$. In the third and final layer (Fig. 4d), we apply an instance of the circuit U_t, on the input qubits in the left-most column. Now the result of $\mathrm{Parity}_n(x)$ should be stored in the top-left qubit, originally containing $x_{1,1}$, the first bit of the input. To simplify calculation, we assume that the circuits U_i for $i \in \{b, t\}$ computes Parity_{n_i} for any input with probability exactly $(1 + \delta_i)/2$. This is a reasonable assumption

[3]Here 't' stands for top and 'b' stands for bottom.

because otherwise, the circuit U_i does better for some inputs, and the overall correct probability with be even larger. Now fix any input. Let $|\varphi_i\rangle = \alpha_{i,0}|0\rangle|\mu_i\rangle + \alpha_{i,1}|1\rangle|\nu_i\rangle$ be the output of the circuit computing $\text{Parity}_{n_b}(x_i)$. Let $b_i = \text{Parity}_{n_b}(x_i)$. Then $|\alpha_{i,b_i}|^2 = (1+\delta)/2$. After the second layer consisting of CNOT gates, the output state becomes $|\alpha_{i,0}|^2 |0\rangle\langle 0| + |\alpha_{i,1}|^2 |1\rangle\langle 1|$.

For $i \in \{t, b\}$, let π_i be a distribution on $\{-1, 1\}$ such that $\pi_i(1) = (1+\delta_i)/2$ and $\pi_i(-1) = (1-\delta_i)/2$. For $i = 1, \ldots, n$, let w_i be independent random variables, such that $w_i = 1$ if and only if the measurement outcome of the circuit computing $\text{Parity}_{n_b}(x_i)$ is correct. Let v be an independent random variable such that $v = 1$ if and only if the output of the U_t in the third layer is correct. Then w_i are all distributed according to the distribution π_b, and v is distributed according to the distribution π_t. Also, the circuit U computes $\text{Parity}_n(x)$ correctly if and only if

$$v \cdot \prod_{i=1}^{n_t} w_i = 1.$$

Then the circuit U correctly computes Parity_n with probability exactly

$$\begin{aligned}\Pr\left[v \cdot \prod_{i=1}^{n_t} w_i = 1\right] &= \frac{\Pr\left[v \cdot \prod_i w_i = 1\right] + \left(1 - \Pr\left[v \cdot \prod_i w_i = -1\right]\right)}{2} \\ &= \frac{1 + \mathbb{E}\left[v \cdot \prod_i w_i\right]}{2} \\ &= \frac{1 + \mathbb{E}[v] \cdot \prod_i \mathbb{E}[w_i]}{2} \\ &= \frac{1 + \delta_b^{n_t}\delta_t}{2}.\end{aligned}$$

□

The proof of Theorem 5.2 will be split in two steps:

- In step 1, we show how to construct a $\mathbf{QAC}^0$ circuit with n^2 ancillae, from any $\mathbf{QAC}^0$ circuit with arbitrary polynomial ancillae.
- In step 2, given a circuit family using at most n^2 ancillae, we show how to construct $\mathbf{QAC}^0$ circuits with $n^{1+\exp(-o(d))}$ ancillae.

Step 1 Given a $\mathbf{QAC}^0$ circuit U_n computing Parity_n with n^c ancillae, the goal of the first step is to construct a $\mathbf{QAC}^0$ circuit computing Parity_{n^c} with n^{2c} ancillae, using the construction in Fig. 5: We consider a c-layer $\mathbf{QAC}^0$ circuit computing Parity_{n^c}. In the top layer, there is a single circuit U_n that computes Parity_n. Following, from top to down, for $i = 2, \ldots, c$, each layer consists of n^{i-1} instances of the circuit U_n. These layers are connected by Lemma 5.4. So the overall depth will be $cd + c - 1$. The total number of instances of U_n is

$$1 + n + \cdots + n^{c-1} = \frac{n^c - 1}{n - 1} \le n^c.$$

Since each U_n uses at most n^c ancillae, this circuit has at most n^{2c} ancillae. And the worst case probability of computing Parity_{n^c} is $\left(1 + \delta_n^{n^c}\right)/2$, which is negligibly close to 1 if $\delta_n = 1 - \text{negl}(n)$. Formally, we prove the first step of our construction with the following lemma.

Lemma 5.6. *For $n \ge 3$, suppose we have a $\mathbf{QAC}^0$ circuit which is a $[d \,|\, n \,|\, n^c]$ circuit, and computes* Parity_n *with the worst-case probability $(1+\delta)/2$. Then we can construct a $\mathbf{QAC}^0$ circuit U' such that the circuit U' is a $\left[c(d+1) \,\middle|\, n^c \,\middle|\, n^{2c}\right]$ circuit, and computes* Parity_{n^c} *with the worst-case probability* $\left(1 + \delta^{cn^{c-1}}\right)/2$.

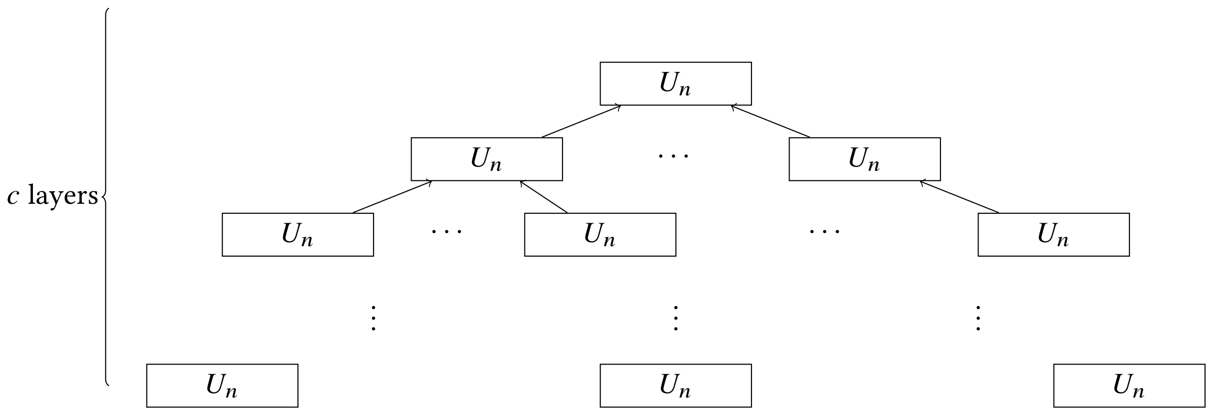


Figure 5: Step 1, $n^c \longrightarrow n^2$ ancillae

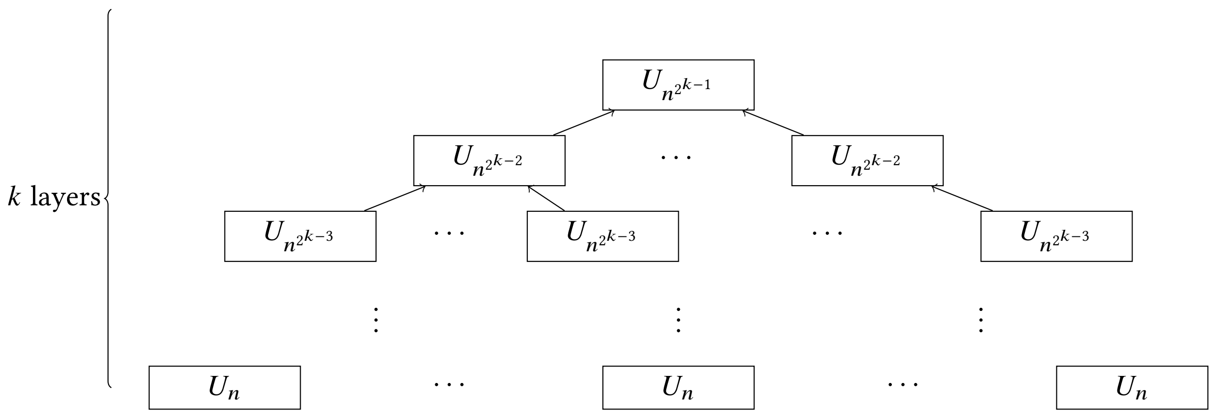


Figure 6: Step 2, $n^2 \longrightarrow n^{1+\exp(-o(d))}$ ancillae

Proof. We use Lemma 5.4 for $c-1$ times to construct the circuit in Fig. 5 in a top-to-down manner. Initially, we let the circuit $V_1 = U$. Then, for $i = 2, \ldots, c$, we apply Lemma 5.4 with $U_b \leftarrow U$ and $U_t \leftarrow V_{i-1}$, and let the resulting $\mathbf{QAC}^0$ circuit be V_i. Finally, we will let $U' = V_c$.

By induction, we prove that each V_i is a $\left[i(d+1) \,\middle|\, n^i \,\middle|\, n^{i+c}\right]$ circuit that computes Parity_{n^i} with probability $\left(1+\delta^{in^{(c-1)}}\right)/2$. For the base case, $V_1 = U$ is a $[d \,|\, n \,|\, n^c]$ circuit, and computes Parity_n with worst case probability $(1+\delta)/2 \geq \left(1+\delta^{c-1}\right)/2$. Then for each $i = 2, \ldots, c$, by Lemma 5.4, the circuit V_i satisfies:

- depth: $(i-1)(d+1) + d + 1 = i(d+1)$.
- input: $n^{i-1} \cdot n = n^i$.
- ancillae: $n^{i-1}\left(n^c+1\right) + n^{i-1+c} \leq 3n^{i-1+c} \leq n^{i+c}$.
- probability: $\left(1+\delta^{(i-1)n^{(c-1)}}\delta^{i-1}\right)/2 \geq \left(1+\delta^{in^{(c-1)}}\right)/2$.

The final circuit V_c will be a $\left[c(d+1) \,\middle|\, n^c \,\middle|\, n^{2c}\right]$ circuit, that computes Parity_{n^c} with probability $\left(1+\delta^{cn^{(c-1)}}\right)/2$. □

Step 2 In the second step, we push the ancillae rate from n^2 to $n^{1+\exp(-o(d))}$, using the construction in Fig. 6. The construction is similar to the construction Fig. 5 in step 1. The difference is that in step 2, we work with $\mathbf{QAC}^0$ circuit families instead of a single $\mathbf{QAC}^0$ circuit. The reason we need a circuit family is that we will use different U_n in different layers. Specifically, for each n, for a k layer construction, we let the top layer circuit have $n^{2^{k-1}}$ inputs, and then the next layer have $n^{2^{k-1}}$ instances of circuits with $n^{2^{k-2}}$ inputs, and so on. The bottom layer are circuits with $n^{2^0} = n$ inputs.

By using Lemma 5.4, the depth of this construction is $kd + k - 1$, and it computes the function Parity_N for

$$N = \prod_{i=0}^{k-1} n^{2^i} = n^{2^k-1}.$$

Also, we can check that the ancillae used in each layer is exactly $n^{2^k} \approx N^{1+2^{-k}}$. So the overall ancillae size is $kn^{2^k} \approx N^{1+2^{-k}}$, with the depth being $O(kd)$. We give this construction formally using the following lemma. To keep the constructed circuit compute Parity_N with non-trivial probability, we require the circuits to have worst-case correct probability $1 - \text{negl}(n)$.

Lemma 5.7. *Let $k \geq 2$ be any constant integer. Suppose there exists a constant $d \in \mathbb{Z}^+$ and an integer n such that for each $0 \leq i \leq k-1$, there exist a QAC^0 circuit U_i that is a $\left[d \,\middle|\, n^{2^i} \,\middle|\, n^{2^{i+1}}\right]$ circuit, and computes $\text{Parity}_{n^{2^i}}$ with worst case probability $1-\text{negl}(n)$. Then we can construct a $\left[k(d+1) \,\middle|\, n^{2^k-1} \,\middle|\, 2kn^{2^k}\right]$ circuit and computes $\text{Parity}_{n^{2^k-1}}$ with worst case probability $1 - \text{negl}(n)$.*

In particular, we have $2kn^{2^k} = n^{2^k+\log_n(2k)}$ and for $n \geq 4k^2$,

$$\frac{2^k + \log_n(2k)}{2^k - 1} = 1 + \frac{1 + \log_n(2k)}{2^k - 1} \leq 1 + \frac{2 - 2^{-k+1}}{2^k - 1} = 1 + 2^{-k+1}.$$

So we get a $\left[k(d+1) \,\middle|\, N \,\middle|\, N^{1+2^{-k+1}}\right]$ circuit for $N = n^{2^k-1}$, with depth $k(d+1)$.

Proof. We use Lemma 5.4 for $k-1$ times to construct the circuit in Fig. 6 in a down-to-top manner. We will use V_i to represent a sub-tree in Fig. 6 with i layers. Initially, we let $V_1 = U_0$. Then for each $i = 2, \ldots, k$, we apply Lemma 5.4 with $U_b \leftarrow V_{i-1}$ and $U_t \leftarrow U_{i-1}$, and let the resulting QAC^0 circuit be V_i. Remember by assumption, U_{i-1} is a $\left[d \,\middle|\, n^{2^{i-1}} \,\middle|\, n^{2^i}\right]$ circuit. We will prove by induction that each V_i is a $\left[i(d+1) \,\middle|\, n^{2^i-1} \,\middle|\, 2in^{2^i}\right]$ circuit, that computes $\text{Parity}_{n^{2^i-1}}$ with worst-case probability $1 - \text{negl}(n)$.

For the base case, $V_1 = U_0$ is a $\left[d \,\middle|\, n \,\middle|\, n^2\right]$ circuit. Now for each $i = 2, \ldots, k$, by Lemma 5.4, the circuit V_i satisfies

- depth: $(i-1)(d+1) + d + 1 \leq i(d+1)$.
- input: $n^{2^{i-1}-1} \cdot n^{2^{i-1}} = n^{2^i-1}$.
- ancillae: $n^{2^{i-1}}\left(2(i-1)n^{2^{i-1}} + 1\right) + n^{2^i} = 2(i-1)n^{2^i} + n^{2^{i-1}} + n^{2^i} \leq 2in^{2^i}$.

Finally, since k is a constant, the above construction only uses polynomial instances of the circuits U_i, so the circuit V_k computes $\text{Parity}_{n^{2^k-1}}$ with worst-case probability $1 - \text{negl}(n)$.

□

Putting together With Lemma 5.6 and Lemma 5.7, we can directly prove Theorem 5.2, which is restated below:

Theorem 5.2. *Suppose there exist constants $d \in \mathbb{Z}^+, c \in \mathbb{Z}^+, N_0 \in \mathbb{Z}^+$, and a QAC^0 circuit family $\{U_n\}_{n \geq N_0}$ such that for each $n \geq N_0$, the circuit U_n is a $[d \,|\, n \,|\, n^c]$ circuit, and computes Parity_n with the worst-case error $\text{negl}(n)$. Then there exists $D = O(cd)$, for any $K \geq 1$ and infinitely many n, we can construct a $\left[KD \,\middle|\, n \,\middle|\, n^{1+2^{-K}}\right]$ circuit, which computes Parity_n with the worst-case error $\text{negl}(n)$.*

Proof. Let $D = 3c(d+1)$. Let n be any integer satisfying

- $n = m^{(2^{K+1}-1)c}$ for some integer $m \geq N_0$.
- $m^c \geq 4K^2$.

For $i = 0, \ldots, K$, apply Lemma 5.6 to the circuit $U_{m^{2^i}}$ to get the $\left[c(d+1) \middle| m^{2^i c} \middle| m^{2^{i+1}c}\right]$ circuit U'_i. Then apply Lemma 5.7 with $k \leftarrow K+1$ and $n \leftarrow m^c$ to the circuits $\{U'_i\}_{i \in \{0,\ldots,K\}}$. We get a $\left[(K+1)(c(d+1)+1) \middle| n \middle| n^{1+2^{-K}}\right]$ circuit. Then we finish the proof by showing

$$(K+1)(c(d+1)+1) \leq KD.$$

□

6 Quantum State Synthesis

In this section, we investigate the hardness of quantum state synthesis. Rosenthal [Ros21] has shown that the cat state synthesis and PARITY are equivalent via $\mathbf{QAC}^0$ reduction. Thus, the hardness of cat state synthesis can be derived from Theorem 4.3. This section will provide a more generic method to prove the hardness of state synthesis via low-degree property of $\mathbf{QAC}^0$ circuits in Corollary 3.6.

For a quantum state φ, its normalized Frobenius norm is exponentially small.

$$\|\varphi\|_2 = \left(2^{-n}\operatorname{Tr}\left(\varphi^2\right)\right)^{1/2} \leq 2^{-n/2}.$$

Hence if we approximate quantum states with constant Frobenius distance, we can always get trivial approximation results. However, the spectral norm of a quantum state can be as large as 1 for pure states. This means that the approximation results of [NPVY24] are insufficient, and spectral approximate degree is essential in the task of quantum state synthesis hardness.

6.1 Lower Bound on the Degrees of Quantum States

We first study the approximate degree of quantum states. For a quantum state φ, we continue to use the approximate degree $\widetilde{\deg}_\varepsilon(\varphi)$ defined in Definition 2.8. For low-degree pure quantum states, we have the following concentration bound when we measure it in the computational basis:

Lemma 6.1 ([AM23, KAAV17b]). *Let $\varphi = |\varphi\rangle\langle\varphi|$ be a pure state satisfying $\widetilde{\deg}_\varepsilon(\varphi) \leq k$. We measure φ in the computational basis, and let W_φ denote the Hamming weight of the measurement outcome. That is, we let W_φ be the random variable satisfying $\Pr\left[W_\varphi = i\right] = \sum_{x\in\{0,1\}^n:|x|=i}\langle x|\varphi|x\rangle$. Let m be an integer median of W_φ that satisfies $\Pr\left[W_\varphi \leq m\right] \geq 1/2$, then*

$$\Pr\left[W_\varphi > m + k\right] \leq 4\varepsilon^2. \tag{3}$$

Similarly, if m is an integer median that satisfies $\Pr\left[W_\varphi \geq m\right] \geq 1/2$, then

$$\Pr\left[W_\varphi < m - k\right] \leq 4\varepsilon^2. \tag{4}$$

Lemma 6.1 implies that the quantum states that do not satisfy this concentration property actually have high approximate degree. This allows us to prove a tight bound on the approximate degree of a nekomata state.

Corollary 6.2. *For $\varepsilon < \frac{1}{4\sqrt{2}}$ and n-nekomata state $|\nu\rangle = \frac{1}{\sqrt{2}}(|0^n,\psi_0\rangle + |1^n,\psi_1\rangle)$, we have*

$$\widetilde{\deg}_\varepsilon(|\nu\rangle\langle\nu|) \geq n.$$

As a special case, for cat states, we have

$$\widetilde{\deg}_\varepsilon(😺_n) \geq n.$$

Proof. We first prove the special case for cat states. Let $W_{\text{🐱}}$ be the random variable defined in Lemma 6.1. It is easy to see that

$$W_{\text{🐱}_n} = \begin{cases} n \text{ with probability } 1/2. \\ 0 \text{ with probability } 1/2. \end{cases}$$

Applying Lemma 6.1 with $k \leftarrow n-1$ and $m \leftarrow 0$, we conclude the result $\widetilde{\deg}_{\varepsilon}\left(\text{🐱}_n\right) \geq n$.

Now for a general n-nekomata, assume $\langle \psi_0 | \psi_1 \rangle \geq 0$. Let

$$\widetilde{\nu} = \operatorname{Tr}_{\geq n+1}\left[\left(\mathbb{1} \otimes \frac{|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1| + |\psi_0\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_0|}{\operatorname{Tr}\left(|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1| + |\psi_0\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_0|\right)}\right)|\nu\rangle\langle\nu|\right].$$

With Lemma 2.12, $\widetilde{\deg}_{\varepsilon}(\widetilde{\nu}) \leq \widetilde{\deg}_{\varepsilon}(\nu)$.

We finish our proof with the fact

$$\widetilde{\nu} = \left(\frac{|1 + \langle\psi_0|\psi_1\rangle|^2}{\operatorname{Tr}\left(|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1| + |\psi_0\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_0|\right)}\right)\text{🐱}_n.$$

□

6.2 Quantum State Synthesis

Now we can prove our hardness results for quantum state synthesis. The formal definition of quantum state synthesis is Definition 2.17. This definition coincides with the approximate dirty state generation problem considered by Rosenthal [Ros21], and is the easiest one among his definitions.

We will be using the following lemma, with the proof deferred to Appendix A.

Lemma 6.3. *Let $\varphi = |\varphi\rangle\langle\varphi|$ be a pure state on n qubits. Let $\psi = |\psi\rangle\langle\psi|$ be a pure state on $n+a$ qubits. Suppose*

$$\left\| |\varphi\rangle\langle\varphi| - \operatorname{Tr}_{n+1,\ldots,n+a}\left[|\psi\rangle\langle\psi|\right]\right\| \leq \varepsilon,$$

then there exists a pure quantum state $|\nu\rangle$ on a qubits such that

$$\| |\varphi\rangle\langle\varphi| \otimes |\nu\rangle\langle\nu| - |\psi\rangle\langle\psi| \| \leq 5\sqrt{\varepsilon}.$$

Theorem 6.4. *Let $\varphi = |\varphi\rangle\langle\varphi|$ be a pure state on n qubits. Let U be a $\mathbf{QAC}^0$ circuit with a ancillae and depth d that synthesizes φ with fidelity $1-\delta$. Then for $\varepsilon = 10\delta^{1/4} + \Omega(d/n)$, we have*

$$(n+a)^{1-3^{-d}/2} = \widetilde{\Omega}\left(\widetilde{\deg}_{\varepsilon}(\varphi)\right).$$

In particular, for the n-nekomata state, and for some fidelity $1-\delta_0 < 0$, we have

$$(n+a)^{1-3^{-d}/2} = \widetilde{\Omega}(n).$$

Setting $a = O(n)$ and combing with Corollary 6.2, we have the following corollary.

Corollary 6.5. *For any $0 < \delta \leq 1$, it holds that*

$$\widetilde{\deg}_{\varepsilon}\left(\mathbf{stateQLC}^0[\delta]\right) = o(n).$$

for $\varepsilon = 10\delta^{1/4} + \Omega(d/n)$. In particular, for any family of n-nekomata state $\{|\nu_n\rangle\langle\nu_n|\}_{n\in\mathbb{N}}$, it holds that $\{|\nu_n\rangle\langle\nu_n|\}_{n\in\mathbb{N}} \notin \mathbf{stateQLC}^0$.

Remark 6.6. Our Definition 2.17 requires that the output of the $\mathbf{QAC}^0$ circuit have fidelity $1 - \delta$ with the target state ψ, which is a relatively strong assumption. In the proof of Theorem 6.4, to prove our circuit size lower bound, we actually only require the output state of the $\mathbf{QAC}^0$ circuit to be close to the target state ψ in spectral norm. This is a very weak assumption because, according to the Fuchs–van de Graaf inequalities [Wat18, Theorem 3.33], high fidelity is equivalent to low unnormalized trace norm, and the spectral norm can be exponentially smaller than the unnormalized trace norm. Hence, the first step of the proof would be acquiring a spectral norm upper bound given the fidelity requirement.

Proof of Theorem 6.4. By Definition 2.17, we have

$$F\left(|\varphi\rangle\langle\varphi|, \mathrm{Tr}_{n+1,\ldots,n+a}\left[U|0\rangle\langle 0|^{n+a}U^\dagger\right]\right) = \sqrt{\langle\varphi|\,\mathrm{Tr}_{n+1,\ldots,n+a}\left[U|0\rangle\langle 0|^{n+a}U^\dagger\right]|\varphi\rangle} \geq 1-\delta.$$

By the Fuchs–van de Graaf inequalities Lemma 2.6, along with the fact that the unnormalized trace norm is larger than the spectral norm, we have

$$\left\|\varphi - \mathrm{Tr}_{n+1,\ldots,n+a}\left[U|0\rangle\langle 0|^{n+a}U^\dagger\right]\right\| \leq 2\sqrt{\delta}.$$

Now by Lemma 6.3, there exists a pure state $\nu = |\nu\rangle\langle\nu|$ on a qubits such that

$$\left\|\varphi\otimes\nu - U|0\rangle\langle 0|^{n+a}U^\dagger\right\| \leq 10\delta^{1/4}.$$

We turn to prove a low-degree approximation of the operator $U|0\rangle\langle 0|^{n+a}U^\dagger$. First by Corollary 3.2, we have

$$\widetilde{\deg}_{O(1/(n+a))}\left(|0\rangle\langle 0|^{n+a}\right) = \widetilde{O}\left(\sqrt{n+a}\right).$$

Hence for $\tau = O(d/n)$, by Corollary 3.6, we have $\widetilde{\deg}_\tau\left(U|0\rangle\langle 0|^{n+a}U^\dagger\right) \leq \widetilde{O}\left((n+a)^{1-3^{-d}/2}\right)$. This implies that the state $\varphi\otimes\nu$ has a degree $\widetilde{O}\left((n+a)^{1-3^{-d}/2}\right)$ approximation with spectral distance at most $10\delta^{1/4} + O(d/n) \leq \varepsilon$. Finally, the lemma follows because

$$\widetilde{\deg}_\varepsilon(\varphi) \leq \widetilde{\deg}_\varepsilon(\varphi\otimes\nu).$$

Now we prove the circuit lower bound for synthesizing n-nekomata states. By Corollary 6.2, we need $\varepsilon < \frac{1}{4\sqrt{2}}$. Hence we only need to set δ_0 such that $10\delta_0^{1/4} < \varepsilon < \frac{1}{4\sqrt{2}}$, which implies $\delta_0 < \frac{1}{10240000}$. ☐

6.3 Synthesizing long range correlation using $\mathbf{QAC}^0$ circuits

It is well known that the output state θ of a shallow quantum circuit - with a product state input - has zero correlation length. That is, the "two-point correlation function" $\mathrm{Tr}(AB\theta) - \mathrm{Tr}(A\theta)\,\mathrm{Tr}(B\theta)$ is 0 for all local operators A, B that are supported out of each other's light cones. In this section, we give an example which shows that $\mathbf{QAC}^0$ circuits are more powerful and can produce long range entanglement in depth 1.

We start with the state

$$\left|\rho^0\right\rangle = \left(\sqrt{1-\frac{1}{n}}\,|0\rangle + \sqrt{\frac{1}{n}}\,|1\rangle\right)^{\otimes n} = \sum_{k=0}^{n} a_k\,|\psi_k\rangle,$$

where $a_k = \sqrt{\binom{n}{k}\left(1-\frac{1}{n}\right)^n \frac{1}{(n-1)^k}}$ and $|\psi_k\rangle = \frac{1}{\sqrt{\binom{n}{k}}}\sum_{x:|x|=k}|x\rangle$ is the uniform superposition over strings of Hamming weight k. The following state in $\mathbf{QAC}^0$:

$$\left|\rho^1\right\rangle = X^{\otimes n}\,\mathrm{CZ}\,X^{\otimes n}\left|\rho^0\right\rangle = -a_0\,|\psi_0\rangle + \sum_{k=1}^{n} a_k\,|\psi_k\rangle$$

will be shown to have a long range correlation. Towards this, define two projectors for a subset S of qubits:

$$\Pi_S^0 = |0\rangle\langle 0|^{\otimes |S|}, \quad \Pi_S^1 = \frac{1}{2}\left(|0\rangle^{\otimes |S|} + \frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}|x_S\rangle\right)\left(\langle 0|^{\otimes |S|} + \frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}\langle x_S|\right).$$

We have

$$\Pi_S^0 \left|\rho^0\right\rangle = a_0 |\psi_0\rangle + |0\rangle^{|S|} \otimes \sum_{k=1}^{n} a_k \left|\psi_k^{S_c}\right\rangle, \quad \Pi_S^0 \left|\rho^1\right\rangle = -a_0 |\psi_0\rangle + |0\rangle^{|S|} \otimes \sum_{k=1}^{n} a_k \left|\psi_k^{S_c}\right\rangle,$$

where $\left|\psi_k^{S_c}\right\rangle = \frac{1}{\sqrt{\binom{n}{k}}}\sum_{x_{S_c}:|x_{S_c}|=k}\left|x_{S_c}\right\rangle$ is a sub-normalized state on the qubits in S_c (the complement of S) with Hamming weight k. Note that

$$\|\Pi_S^0 \left|\rho^0\right\rangle\| = \|\Pi_S^0 \left|\rho^1\right\rangle\| = \left(1 - \frac{1}{n}\right)^{|S|},$$

since the minus sign in front of ψ_0 does not affect the norm and the last equality is the probability that $|S|$ sequences of 0 are seen. Next,

$$\begin{aligned}
\Pi_S^1 \left|\rho^0\right\rangle &= \frac{1}{2}\left(|0\rangle^{\otimes |S|} + \frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}|x_S\rangle\right)\left(a_0 |0\rangle^{\otimes |S_c|} + \sum_{k=1}^{n} a_k \left|\psi_k^{S_c}\right\rangle + a_1\sqrt{\frac{|S|}{n}}|0\rangle^{\otimes |S_c|} + |\mu\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle^{\otimes |S|} + \frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}|x_S\rangle\right)\left(\left(a_0 + a_1\sqrt{\frac{|S|}{n}}\right)|0\rangle^{\otimes |S_c|} + \sum_{k=1}^{n} a_k \left|\psi_k^{S_c}\right\rangle + |\mu\rangle\right) \\
\Pi_S^1 \left|\rho^1\right\rangle &= \frac{1}{2}\left(|0\rangle^{\otimes |S|} + \frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}|x_S\rangle\right)\left(-a_0 |0\rangle^{\otimes |S_c|} + \sum_{k=1}^{n} a_k \left|\psi_k^{S_c}\right\rangle + a_1\sqrt{\frac{|S|}{n}}|0\rangle^{\otimes |S_c|} + |\mu\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle^{\otimes |S|} + \frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}|x_S\rangle\right)\left(\left(-a_0 + a_1\sqrt{\frac{|S|}{n}}\right)|0\rangle^{\otimes |S_c|} + \sum_{k=1}^{n} a_k \left|\psi_k^{S_c}\right\rangle + |\mu\rangle\right)
\end{aligned}$$

where $|\mu\rangle := \left(\frac{1}{\sqrt{|S|}}\sum_{x_S:|x_S|=1}\langle x_S|\right)\left(\sum_{k=2}^{n} a_k |\psi_k\rangle\right)$ is a state with Hamming weight at least 1 on the qubits in S_c. Note that

$$\|\Pi_S^1 \left|\rho^0\right\rangle\|^2 - \|\Pi_S^1 \left|\rho^1\right\rangle\|^2 = \frac{\left(a_0 + a_1\sqrt{\frac{|S|}{n}}\right)^2 - \left(-a_0 + a_1\sqrt{\frac{|S|}{n}}\right)^2}{2} = 2a_0 a_1\sqrt{\frac{|S|}{n}}. \tag{5}$$

From here, we can compare the correlation functions

$$C_0 := \left\langle\rho^0\right|\Pi_S^0\Pi_T^1\left|\rho^0\right\rangle - \left\langle\rho^0\right|\Pi_S^0\left|\rho^0\right\rangle\left\langle\rho^0\right|\Pi_T^1\left|\rho^0\right\rangle, \quad C_1 := \left\langle\rho^1\right|\Pi_S^0\Pi_T^1\left|\rho^1\right\rangle - \left\langle\rho^1\right|\Pi_S^0\left|\rho^1\right\rangle\left\langle\rho^1\right|\Pi_T^1\left|\rho^1\right\rangle,$$

for two distinct sets of qubits S, T. The difference

$$\left\langle\rho^0\right|\Pi_S^0\Pi_T^1\left|\rho^0\right\rangle - \left\langle\rho^1\right|\Pi_S^0\Pi_T^1\left|\rho^1\right\rangle = \|\Pi_T^1\Pi_S^0\left|\rho^0\right\rangle\|^2 - \|\Pi_T^1\Pi_S^0\left|\rho^1\right\rangle\|^2,$$

can be evaluated similar to Equation 5 by replacing the coefficients a_k with $a_k\sqrt{\frac{\binom{n-|S|}{k}}{\binom{n}{k}}}$ (since, for each $b \in \{0,1\}$, the states $\left|\rho^b\right\rangle$ and $\Pi_S^0\left|\rho^b\right\rangle$ are very similar except that the former is defined on n qubits, latter is defined on $n - |S|$ qubits and the coefficients are rescaled). This shows that

$$\left\langle\rho^0\right|\Pi_S^0\Pi_T^1\left|\rho^0\right\rangle - \left\langle\rho^1\right|\Pi_S^0\Pi_T^1\left|\rho^1\right\rangle = 2a_0 a_1\sqrt{\frac{n-|S|}{n}}\sqrt{\frac{|T|}{n}}.$$

Collectively, using the fact that $C_0 = 0$ for the product state ρ^0, we have

$$\begin{aligned} -C_1 &= C_0 - C_1 = 2a_0a_1\sqrt{\frac{n-|S|}{n}}\sqrt{\frac{|T|}{n}} - 2a_0a_1\sqrt{\frac{|S|}{n}} \cdot \|\Pi_S^0 \left|\rho^0\right\rangle\|^2 \\ &= 2a_0a_1\left(\sqrt{\frac{n-|S|}{n}}\sqrt{\frac{|T|}{n}} - \sqrt{\frac{|S|}{n}}\left(1-\frac{1}{n}\right)^{|S|}\right). \end{aligned}$$

We can choose $|S| = |T| = \Theta(n)$, which gives $|C_1| = \Theta(1)$ using $a_0, a_1 = \Theta(1)$. This is near maximal correlation between two large regions and if we are considering lattices, one can arrange for these regions to be far apart. If we insist on considering operators of constant locality, we can choose $|S| = 1, |T| = 2$ and obtain $|C_1| = \Theta(1/\sqrt{n})$.

6.4 Bounds on low energy state preparation

Given that $\mathbf{QAC}^0$ circuits can produce long range correlations, the possibility of using them to probe low-energy regime of local Hamiltonians emerges. More concretely, we can ask if a sufficiently low-energy state of any local Hamiltonian can be prepared by a $\mathbf{QAC}^0$ circuit. Here we argue that this is not possible - Corollary 3.6 implies that the low-energy states of the local Hamiltonian considered in [ABN22] cannot be generated by $\mathbf{QAC}^0$ circuits. This holds since we can use a concentration bound similar to Lemma 6.1 to show approximate degree lower bounds on the low-energy states of [ABN22].

Lemma 6.7. *Consider a $[[n, k, d]]$ CSS code satisfying Property 1 from [ABN22] with parameters δ_0, c_1, c_2 as stated. Let $\mathbf{H} = \mathbf{H}_x + \mathbf{H}_z$ be the corresponding local Hamiltonian, with m_x and m_z local terms, respectively. Then for*

$$\varepsilon < \frac{1}{400c_1}\left(\frac{\min\{m_x, m_z\}}{n}\right) \cdot \min\left\{\left(\frac{k-1}{4n}\right)^2, \delta_0, \frac{c_2}{2}\right\},$$

and every pure state $\varphi = |\varphi\rangle\langle\varphi|$ such that $\mathrm{Tr}\left[\mathbf{H}\varphi\right] \le \varepsilon n$, we have for $\delta = 1/8000$,

$$\widetilde{\deg}_\delta(\varphi) = \Omega(n).$$

In particular, the quantum Tanner codes [LZ22] satisfy the above property.

Proof. Let W_z be the measurement outcome of φ on the computational basis, and W_x be the measurement outcome of φ on the Hadamard basis, which is equivalent to measuring $H^{\otimes n}\varphi H^{\otimes n}$ on the computational basis. By [ABN22, Lemma 2], there exist $b \in \{x, z\}$ and two sets B_0 and B_1 such that

$$W_b(B_0) \ge \frac{1}{400} \text{ and } W_b(B_1) \ge \frac{1}{400}.$$

Also, the distance between B_0 and B_1 is $\Omega(n)$.

Afterward, if $b = z$, then we will prove the approximate degree lower bounds for φ, and if $b = x$, we will prove the approximate degree lower bounds for $H^{\otimes n}\varphi H^{\otimes n}$. Since $\widetilde{\deg}_\delta(\varphi) = \widetilde{\deg}_\delta\left(H^{\otimes n}\varphi H^{\otimes n}\right)$, we can assume without loss of generality that $b = z$.

Now we can use an argument similar to Lemma 6.1, to prove that $\widetilde{\deg}_\delta(\varphi) = \Omega(n)$. Let Π_0 and Π_1 be the projectors on the strings in B_0 and B_1 respectively. For any operator R with degree smaller than $\mathrm{dist}(B_0, B_1) = \Omega(n)$, we have $\Pi_0 R \Pi_1 = 0$. Then

$$\begin{aligned}\|\varphi - R\| &\geq \|\Pi_0 (\varphi - R) \Pi_1\| \\ &= \|\Pi_0 \varphi \Pi_1\| \\ &\geq \|\Pi_0 |\varphi\rangle\langle\varphi| \Pi_1 |\varphi\rangle\|_2 \\ &= \langle\varphi| \Pi_1 |\varphi\rangle \|\Pi_0 |\varphi\rangle\|_2 \\ &= \mathrm{Tr}[\Pi_1 \varphi] \cdot (\mathrm{Tr}[\Pi_0 \varphi])^{1/2} \\ &= W_0(B_1) \cdot \sqrt{W_0(B_0)} \\ &\geq \frac{1}{8000}.\end{aligned}$$

Thus every operator R with degree lower than $\Omega(n)$ has spectral distance at least $\frac{1}{8000}$ to φ. Hence we have

$$\widetilde{\deg}_{1/8000}(\varphi) = \Omega(n).$$

□

With the approximate degree lower bound above, we can immediately invoke Theorem 6.4 to get the following hardness result for $\mathbf{QAC}^0$ circuits.

Corollary 6.8. *There exists a constant $\varepsilon_0 > 0$, such that for all depth-d $\mathbf{QAC}^0$ circuits with a ancillae that synthesize the low energy state φ with fidelity $1 - \delta = 1 - \varepsilon_0 + O(d^4/n^4)$, we have*

$$a = \widetilde{\Omega}\left(n^{1+3^{-d}/2}\right).$$

Proof. By Theorem 6.4, we have for

$$\varepsilon = 10\delta^{1/4} + \Omega(d/n) = 10\left(\varepsilon_0 - O(d^2/n^2)\right)^{1/4} + \Omega(d/n) = 10\varepsilon_0^{1/4},$$

$$(n + a)^{1-3^{-d}/2} = \widetilde{\Omega}\left(\widetilde{\deg}_{\varepsilon}(\varphi)\right).$$

For $\varepsilon = \frac{1}{8000}$, we have $\widetilde{\deg}_{\varepsilon}(\varphi) = \Omega(n)$. Thus for $\varepsilon_0 = \frac{\varepsilon^4}{10000}$, we have

$$(n + a)^{1-3^{-d}/2} = \widetilde{\Omega}(n).$$

□

7 Quantum Channels Synthesis

In this section, we prove $\mathbf{QAC}^0$ hardness results for general quantum channels. Recall for a unitary operator U, we use $\mathcal{E}_{k,U,\psi}$ to define the k qubit output quantum channel using ψ as ancilla as

$$\mathcal{E}_{k,U,\psi}(\rho) = \mathrm{Tr}_{[k]^c}\left[U(\rho \otimes \psi) U^\dagger\right].$$

The Choi representation of $\mathcal{E}_{k,U,\psi}$ is denoted by

$$\Phi_{k,U,\psi} = \left(\mathcal{E}_{k,U,\psi} \otimes \mathbb{1}\right)(\mathrm{EPR}_n),$$

where EPR_n is the density operator of unnormalized n-qubits EPR state $\sum_{x \in \{0,1\}^n} |x\rangle \otimes |x\rangle$. The subscript k may be omitted if it is clear from context. If there is no ancilla, we use the notation $\mathcal{E}_U$ and Φ_U. We adapt the following identity for the Choi state of quantum channels.

Fact 7.1 ([NPVY24]).

$$\Phi_U = \left(\mathbb{1} \otimes U^T\right) \left(\mathrm{EPR}_k \otimes \mathbb{1}_{n-k}\right) \left(\mathbb{1} \otimes \overline{U}\right).$$

If $\psi = |\psi\rangle\langle\psi|$ is a pure state,

$$\Phi_{U,\psi} = \langle\psi| \Phi_U |\psi\rangle .$$

Since we are working with spectral norm approximations of matrices, the spectral norms of the Choi states are important. In fact, by Fact 7.1 and [DY24, Lemma 4], we see that the spectral norms of Choi states are upper bounded by 2^k. We approximate the operator $2^{-k}\Phi_{U,\psi}$, which is achieved by using Corollary 3.6.

Theorem 7.2. *Let $n \geq 1$. Suppose U is a depth-d $\mathbf{QAC}^0$ circuit with n input qubits and a ancillae initialized in the state ψ. For $k \leq n$, we take the first k qubits as output and implement the quantum channel $\mathcal{E}_{U,\psi}$. For $\ell = \widetilde{O}\left((n+a)^{1-3^{-D}} k^{3^{-D}/2}\right)$, $\varepsilon = O(d/n)$, we have*

$$\widetilde{\deg}_{\varepsilon}\left(2^{-k}\Phi_{U,\psi}\right) \leq \ell. \tag{6}$$

Remark 7.3. This result is incomparable to the result of [NPVY24]. In our result, we approximate the channel by an $o(n)$-degree operator with respect to the spectral norm. In [NPVY24], they approximate the channel by a constantly local operator with respect to the Frobenius norm.

Proof of Theorem 7.2. By Fact 7.1,

$$2^{-k}\Phi_{U,\psi} = 2^{-k}\langle\psi| \Phi_U |\psi\rangle = \langle\psi| \left(\mathbb{1} \otimes U^T\right) \left(2^{-k}\mathrm{EPR}_k \otimes \mathbb{1}_{n-k}\right) \left(\mathbb{1} \otimes \overline{U}\right) |\psi\rangle .$$

So the proof idea would be to get a low-degree approximation of EPR_k and then invoke Corollary 3.6. However, note that $\|\mathrm{EPR}_k\| = 2^k$, so instead we approximate the operator $2^{-k}\mathrm{EPR}_k$. Note that $2^{-k}\mathrm{EPR}_k$ is the tensor product of k EPR pairs. Hence by Corollary 3.2,

$$\widetilde{\deg}_{O(1/n)}\left(2^{-k}\mathrm{EPR}_k\right) \leq \widetilde{O}\left(\sqrt{k}\right).$$

Then by Corollary 3.6, the theorem follows. □

Using this theorem, we can prove it is hard for $\mathbf{QAC}^0$ circuits with linear ancillae to implement high degree quantum channels. Here, we use the completely bounded spectral norm as a measure for quantum channels.

Definition 7.4 (Completely Bounded Spectral Norm). Let $\mathcal{E}$ be a quantum channel. The completely bounded spectral norm of $\mathcal{E}$ is defined as

$$\|\mathcal{E}\| = \max_{X:\|X\|\leq 1} \|(\mathcal{E} \otimes \mathbb{1})(X)\|.$$

We then have the following corollary of Theorem 7.2:

Corollary 7.5. *Let $\mathcal{E}$ be a quantum channel from n qubits to k qubits. Suppose there exists a $\mathbf{QAC}^0$ circuit U with a ancillae initialized in the state ψ that approximates $\mathcal{E}$. That is, for some $\varepsilon \leq 1$,*

$$\left\|\mathcal{E} - \mathcal{E}_{U,\psi}\right\| \leq 2^k \varepsilon,$$

then let Φ be the Choi state of $\mathcal{E}$, we have

$$\widetilde{\deg}_{\varepsilon+O(d/n)}\left(2^{-k}\Phi\right) \leq \widetilde{O}\left((n+a)^{1-3^{-D}} k^{3^{-D}/2}\right).$$

Proof. By the definition of completely bounded spectral norms,

$$2^k \varepsilon \geq \left\| \mathcal{E} - \mathcal{E}_{U,\psi} \right\| = \max_{X:\|X\|\leq 1} \left\| \left(\left(\mathcal{E} - \mathcal{E}_{U,\psi} \right) \otimes \mathbb{1} \right) (X) \right\| \geq \left\| \Phi - \Phi_{U,\psi} \right\|.$$

Then by Theorem 7.2, for $\varepsilon' = \varepsilon + O(d/n)$, we have

$$\widetilde{\deg}_{\varepsilon + O(d/n)} \left(2^{-k} \Phi \right) \leq \widetilde{O} \left((n+a)^{1-3^{-D}} k^{3^{-D}/2} \right).$$

□

References

[ABN22] A. Anshu, N.P. Breuckmann, and C. Nirkhe. NLTS Hamiltonians from good quantum codes. In *55th Annual ACM Symposium on Theory of Computing (STOC 2023)*, pages 1090–1096. ACM, 2022.

[ADEGP24] Srinivasan Arunachalam, Arkopal Dutt, Francisco Escudero Gutiérrez, and Carlos Palazuelos. Learning Low-Degree Quantum Objects. In Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson, editors, *51st International Colloquium on Automata, Languages, and Programming (ICALP 2024)*, volume 297 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:19, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[AGL+23] Dorit Aharonov, Xun Gao, Zeph Landau, Yunchao Liu, and Umesh Vazirani. A Polynomial-Time Classical Algorithm for Noisy Random Circuit Sampling. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 945–957, New York, NY, USA, 2023. Association for Computing Machinery.

[AM23] Anurag Anshu and Tony Metger. Concentration bounds for quantum states and limitations on the QAOA from polynomial approximations. *Quantum*, 7:999, May 2023.

[BBC+01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, July 2001.

[BdW01] Harry Buhrman and Ronald de Wolf. Communication Complexity Lower Bounds by Polynomials. In *Proceedings of Computational Complexity. Sixteenth Annual IEEE Conference*, page 0120, Los Alamitos, CA, USA, jun 2001. IEEE Computer Society.

[Ber11] Debajyoti Bera. A lower bound method for quantum circuits. *Inf. Process. Lett.*, 111(15):723–726, aug 2011.

[BGK18] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[BKT19] Mark Bun, Robin Kothari, and Justin Thaler. Quantum algorithms and approximating polynomials for composed functions with shared inputs. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, page 662–678, USA, 2019. Society for Industrial and Applied Mathematics.

[BLS+22] Dolev Bluvstein, Harry Levine, Giulia Semeghini, Tout T. Wang, Sepehr Ebadi, Marcin Kalinowski, Alexander Keesling, Nishad Maskara, Hannes Pichler, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. A quantum processor based on coherent transport of entangled atom arrays. *Nature*, 604(7906):451–456, Apr 2022.

[BT22] Mark Bun and Justin Thaler. Approximate degree in classical and quantum computing. *Found. Trends Theor. Comput. Sci.*, 15(3–4):229–423, December 2022.

[BY23] Zongbo Bao and Penghui Yao. On Testing and Learning Quantum Junta Channels. In Gergely Neu and Lorenzo Rosasco, editors, *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pages 1064–1094. PMLR, 12–15 Jul 2023.

[CdDPH+24] Sitan Chen, Jaume de Dios Pont, Jun-Ting Hsieh, Hsin-Yuan Huang, Jane Lange, and Jerry Li. Predicting quantum channels over general product distributions, 2024.

[CIP09] Chris Calabro, Russell Impagliazzo, and Ramamohan Paturi. The complexity of satisfiability of small depth circuits. In Jianer Chen and Fedor V. Fomin, editors, *Parameterized and Exact Computation*, pages 75–85, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[CNY23] Thomas Chen, Shivam Nadimpalli, and Henry Yuen. Testing and Learning Quantum Juntas Nearly Optimally. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1163–1185, 2023.

[CR96] Shiva Chaudhuri and Jaikumar Radhakrishnan. Deterministic restrictions in circuit complexity. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, page 30–36, New York, NY, USA, 1996. Association for Computing Machinery.

[DFN+24] Yangjing Dong, Honghao Fu, Anand Natarajan, Minglong Qin, Haochen Xu, and Penghui Yao. The Computational Advantage of MIP* Vanishes in the Presence of Noise. In Rahul Santhanam, editor, *39th Computational Complexity Conference (CCC 2024)*, volume 300 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 30:1–30:71, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[DS96] David P. DiVincenzo and Peter W. Shor. Fault tolerant error correction with efficient quantum codes. *Phys. Rev. Lett.*, 77:3260, 1996.

[DY24] Yangjing Dong and Penghui Yao. Communication Complexity of Common Randomness Generation With Isotropic States. *IEEE Transactions on Information Theory*, 70(8):5681–5691, 2024.

[FFG+06] M. Fang, S. Fenner, F. Green, S. Homer, and Y. Zhang. Quantum lower bounds for fanout. *Quantum Info. Comput.*, 6(1):46–57, jan 2006.

[GDC+22] Andrew Y. Guo, Abhinav Deshpande, Su-Kuan Chu, Zachary Eldredge, Przemyslaw Bienias, Dhruv Devulapalli, Yuan Su, Andrew M. Childs, and Alexey V. Gorshkov. Implementing a fast unbounded quantum fanout gate using power-law interactions. *Phys. Rev. Res.*, 4:L042016, Oct 2022.

[GHMP01] Frederic Green, Steve Homer, Cristopher Moore, and Christopher Pollett. Counting, Fanout, And The Complexity Of Quantum ACC. *Quantum Information and Computation*, 2, 12 2001.

[GHZ89] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. *Going Beyond Bell's Theorem*, pages 69–72. Springer Netherlands, Dordrecht, 1989.

[GKH⁺21] P. Gokhale, S. Koretsky, S. Huang, S. Majumder, A. Drucker, K. R. Brown, and F. T. Chong. Quantum fan-out: Circuit optimizations and technology modeling. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 276–290, Los Alamitos, CA, USA, oct 2021. IEEE Computer Society.

[GMR12] Parikshit Gopalan, Raghu Meka, and Omer Reingold. DNF Sparsification and a Faster Deterministic Counting Algorithm. In *2012 IEEE 27th Conference on Computational Complexity*, pages 126–135, 2012.

[Hås86] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, STOC '86, page 6–20, New York, NY, USA, 1986. Association for Computing Machinery.

[KAAV17a] Tomotaka Kuwahara, Itai Arad, Luigi Amico, and Vlatko Vedral. Local reversibility and entanglement structure of many-body ground states. *Quantum Science and Technology*, 2(1):015005, 2017.

[KAAV17b] Tomotaka Kuwahara, Itai Arad, Luigi Amico, and Vlatko Vedral. Local reversibility and entanglement structure of many-body ground states. *Quantum Science and Technology*, 2(1):015005, feb 2017.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on Boolean functions. In *29th Annual Symposium on Foundations of Computer Science*, pages 68–80, 1988.

[KKMS08] Adam Tauman Kalai, Adam R. Klivans, Yishay Mansour, and Rocco A. Servedio. Agnostically learning halfspaces. *SIAM Journal on Computing*, 37(6):1777–1805, 2008.

[KPLT06] M. Koucky, S. Poloczek, C. Lautemann, and D. Therien. Circuit lower bounds via Ehrenfeucht-Fraisse games. In *21st Annual IEEE Conference on Computational Complexity (CCC'06)*, pages 12 pp.–201, 2006.

[KSS92] Michael J. Kearns, Robert E. Schapire, and Linda M. Sellie. Toward efficient agnostic learning. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT '92, page 341–352, New York, NY, USA, 1992. Association for Computing Machinery.

[LZ22] A. Leverrier and G. Zemor. Quantum Tanner codes. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 872–883, Los Alamitos, CA, USA, nov 2022. IEEE Computer Society.

[MO10] Ashley Montanaro and Tobias J. Osborne. Quantum Boolean functions. *Chicago Journal of Theoretical Computer Science*, 2010(1), January 2010.

[Moo99] Cristopher Moore. Quantum circuits: Fanout, parity, and counting. *arXiv preprint quant-ph/9903046*, 1999.

[NPVY24] Shivam Nadimpalli, Natalie Parham, Francisca Vasconcelos, and Henry Yuen. On the Pauli spectrum of QAC0. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 1498–1506, New York, NY, USA, 2024. Association for Computing Machinery.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions.* Cambridge University Press, 2014.

[PFGT20] Daniel Padé, Stephen A. Fenner, Daniel Grier, and Thomas Thierauf. Depth-2 QAC circuits cannot simulate quantum parity. *CoRR*, abs/2005.12169, 2020.

[QY21] Minglong Qin and Penghui Yao. Nonlocal Games with Noisy Maximally Entangled States are Decidable. *SIAM Journal on Computing*, 50(6):1800–1891, 2021.

[QY23] Minglong Qin and Penghui Yao. Decidability of Fully Quantum Nonlocal Games with Noisy Maximally Entangled States. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*, volume 261 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 97:1–97:20, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[Raz87] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical notes of the Academy of Sciences of the USSR*, 41(4):333–338, April 1987.

[Ros16] Benjamin Rossman. PARITY $\notin$ AC^0 from DNF sparsification. 2016.

[Ros21] Gregory Rosenthal. Bounds on the QAC^0 Complexity of Approximating Parity. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[RWZ24] Cambyse Rouzé, Melchior Wirth, and Haonan Zhang. Quantum Talagrand, KKL and Friedgut's Theorems and the Learnability of Quantum Boolean Functions. *Communications in Mathematical Physics*, 405(4):95, Apr 2024.

[Sho96] P.W. Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65, 1996.

[Slo24] Joseph Slote. Parity vs. AC0 with Simple Quantum Preprocessing. In Venkatesan Guruswami, editor, *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, volume 287 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 92:1–92:21, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 77–82, New York, NY, USA, 1987. Association for Computing Machinery.

[SS12] Rahul Santhanam and Srikanth Srinivasan. On the limits of sparsification. In Artur Czumaj, Kurt Mehlhorn, Andrew Pitts, and Roger Wattenhofer, editors, *Automata, Languages, and Programming*, pages 774–785, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[Wan11] Guoming Wang. Property testing of unitary operators. *Physical Review A*, 84:052328, 2011.

[Wat18] John Watrous. *The theory of quantum information*. Cambridge university press, 2018.

[Wil14] Richard Ryan Williams. The Polynomial Method in Circuit Complexity Applied to Algorithm Design. In Venkatesh Raman and S. P. Suresh, editors, *34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS 2014)*, volume 29 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 47–60, Dagstuhl, Germany, 2014. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[Yao19] Penghui Yao. A doubly exponential upper bound on noisy EPR states for binary games. *arXiv preprint arXiv:1904.08832*, 2019.

[YWL+22] Dongmin Yu, Han Wang, Jin-Ming Liu, Shi-Lei Su, Jing Qian, and Weiping Zhang. Multiqubit Toffoli gates and optimal geometry with Rydberg atoms. *Phys. Rev. Appl.*, 18:034072, Sep 2022.

A Deferred Proofs

Proof of Lemma 2.7. The fact that $\|AB\| \leq 1$ follows from the submultiplicativity of the Schatten p-norm [Wat18, Eq. 1.176]. Then

$$\begin{aligned}\left\|AB - \widetilde{A}\widetilde{B}\right\| &= \left\|AB - A\widetilde{B} + A\widetilde{B} - \widetilde{A}\widetilde{B}\right\| \\ &\leq \left\|A\left(B - \widetilde{B}\right)\right\| + \left\|\left(A - \widetilde{A}\right)\widetilde{B}\right\| \\ &\leq \|A\|\left\|\left(B - \widetilde{B}\right)\right\| + \left\|\left(A - \widetilde{A}\right)\right\|\left\|\widetilde{B}\right\| \\ &\leq \varepsilon_1 + \varepsilon_0\left\|\widetilde{B} - B + B\right\| \\ &\leq \varepsilon_1 + \varepsilon_0\left(\left\|\widetilde{B} - B\right\| + \|B\|\right) \\ &\leq \varepsilon_0 + \varepsilon_1 + \varepsilon_0\varepsilon_1.\end{aligned}$$

□

Proof of Theorem 4.10. Let $A = U^\dagger M_f U$. Then we have

$$p(x) = \operatorname{Tr}\left[\left(|x\rangle\langle x| \otimes \varphi\right) A\right] = \langle x| \operatorname{Tr}_{n+1,\ldots,n+a}\left[A\left(\mathbb{1} \otimes \varphi\right)\right]|x\rangle .$$

Then the diagonal matrix M_p as defined in Equation (1) can be obtained by zeroing out all the non-diagonal entries of the matrix

$$\operatorname{Tr}_{n+1,\ldots,n+a}\left[A\left(\mathbb{1} \otimes \varphi\right)\right].$$

By Corollary 3.6, we have $\widetilde{\deg}_\varepsilon\left(\operatorname{Tr}_{n+1,\ldots,n+a}\left[A\left(\mathbb{1} \otimes \varphi\right)\right]\right) \leq \widetilde{O}\left((n+a)^{1-3^{-d}} \cdot \ell^{3^{-d}}\right)$. Finally, by Lemma 2.11 and Fact 2.14 we prove our theorem. □

Proof of Lemma 6.3. First we write down a Schmidt decomposition of ψ as

$$|\psi\rangle = \sum_i \sqrt{s_i}\,|\mu_i\rangle \otimes |\nu_i\rangle ,$$

where $\{\mu_i\}$ is a set of orthogonal basis on n qubits, and $\{\nu_i\}$ is a set of orthogonal basis on a qubits. Also, we assume $s_1 \geq s_2 \geq \cdots \geq s_{\mathrm{rank}(\psi)}$. Then

$$\operatorname{Tr}_{n+1,\ldots,n+a}\left[|\psi\rangle\langle\psi|\right] = \sum_i s_i\,|\mu_i\rangle\langle\mu_i| ,$$

and

$$\begin{aligned}
\varepsilon &\geq \left\|\varphi - \mathrm{Tr}_{n+1,\ldots,n+a}\left[|\psi\rangle\langle\psi|\right]\right\| \\
&= \left\|\varphi - \sum_i s_i\, |\mu_i\rangle\langle\mu_i|\right\| \\
&\geq \langle\varphi| \left(\varphi - \sum_i s_i\, |\mu_i\rangle\langle\mu_i|\right) |\varphi\rangle \\
&= 1 - \sum_i s_i |\langle\mu_i|\varphi\rangle|^2 \\
&\geq 1 - \sum_i s_1 |\langle\mu_i|\varphi\rangle|^2 \\
&= 1 - s_1.
\end{aligned}$$

Hence $s_1 \geq 1 - \varepsilon$, and then $\sum_{i\geq 2} s_i \leq 1 - s_1 \leq \varepsilon$.

$$\|\varphi - \mu_1\| \leq \left\|\varphi - \sum_i s_i\, |\mu_i\rangle\langle\mu_i|\right\| + \left\|(1 - s_1)\, |\mu_1\rangle\langle\mu_1| - \sum_{i\geq 2} s_i\, |\mu_i\rangle\langle\mu_i|\right\| \leq \varepsilon + \varepsilon = 2\varepsilon.$$

$$\begin{aligned}
\||\psi\rangle - |\mu_1\rangle \otimes |\nu_1\rangle\| &= \left\|-(1 - \sqrt{s_1})\, |\mu_1\rangle \otimes |\nu_1\rangle + \sum_{i\geq 2} \sqrt{s_i}\, |\mu_i\rangle \otimes |\nu_i\rangle\right\| \\
&= \sqrt{\left(1 - \sqrt{s_1}\right)^2 + \sum_{i\geq 2} s_i} \\
&\leq \sqrt{\left(1 - \sqrt{1 - \varepsilon}\right)^2 + \varepsilon} \\
&\leq \sqrt{2\varepsilon}.
\end{aligned}$$

So

$$\||\psi\rangle\langle\psi| - |\mu_1\rangle\langle\mu_1| \otimes |\nu_1\rangle\langle\nu_1|\| \leq \sqrt{8\varepsilon}$$

and

We let $|\nu\rangle = |\nu_1\rangle$, then

$$\|\varphi \otimes \nu - \psi\| \leq \|\varphi \otimes \nu - \mu_1 \otimes \nu\| + \|\mu_1 \otimes \nu - \psi\| \leq 2\varepsilon + \sqrt{8\varepsilon} \leq 5\sqrt{\varepsilon}.$$

□