Lower Bounds on Parity Using Analytic Techniques

Arjun Aggarwal

May 2024

1 Introduction

In this paper, we summarize recent attempts at proving $\mathsf{PAR} \not\in \mathsf{QAC}^0$. Specifically, we focus on [Nad+24], which we refer to as Yuen et al.'s work. In Section 2, we give a brief history of the problem and try to highlight its significance. In Section 3, we give a formal definition of the QAC^0 circuit family. We discuss the new notion of the Pauli spectrum of a quantum channel introduced in the aforementioned paper and their attempt at using it to prove lower bounds for parity. We analyze the shortcomings of Yuen et al.'s approach in Section 4 and present our ideas for improving their results.

2 Background and Motivation

QAC^0 circuits are the quantum generalization of AC^0 circuits, which are constant depth circuits with unbounded fan-in. Roughly speaking, AC^0 circuits are the most powerful constant-depth circuits that are composed only of the $\mathsf{AND}, \mathsf{OR}$, and NOT gates (one can get more powerful circuit classes if one allows other gates). From the perspective of complexity theory, AC^0 circuits are interesting because they sit in a Goldilocks zone, wherein they are complicated enough that proving lower bounds is not trivial, yet not so complicated that they are beyond the reach of our mathematical tools. Indeed, using his celebrated Switching Lemma, Håstad proved that $\mathsf{PAR} \not\in \mathsf{AC}^0$ [Has86]. To this day, the result remains one of the strongest unconditional lower bounds in for classical circuits. Thus, it is natural to try to extend this result for quantum circuits. As discussed in the next section, due to the no-cloning theorem, the computational power of AC^0 and QAC^0 are different. Besides, there is no natural extension of the Switching Lemma to the quantum case. Due to these obstacles, proving an analogous lower bound for quantum circuits has remained open since Moore introduced the problem in 1999 [Moo99].

Another line of motivation comes from the study of Fourier Analytic techniques for Boolean functions. Not only can the Switching Lemma be restated in this language, but analytic techniques have led to breakthroughs in several areas such as Cryptography, Social Choice Theory, and Learning Theory [ODo21]. Thus, there is a natural desire to extend these techniques to the quantum setting.

3 A Summary of Previous Results

In this section, we discuss the notion of the Pauli spectrum of a quantum channel, a quantum analog of the Fourier spectrum of a Boolean function, and its application in proving lower bounds on Parity. Roughly speaking, Yuen et al. show that a QAC^0 circuit cannot compute Parity if only a small number of auxiliary qubits is allowed. The result is established by showing an LMN-type concentration bound on the Pauli Spectrum of channels associated with QAC^0 circuits. The authors first prove low-degree concentration for circuits without auxiliary qubits via a lightcones-based argument. Then, they prove concentration for the general case by reducing it to the no-auxiliary qubit case. However, the bounds obtained via the reduction are much weaker with an exponential order dependency on the number of auxiliary qubits.

Presently, their techniques seem insufficient to establish the result in general, but the work presents a promising direction. Moreover, the new notion of Pauli Spectrum introduced in the paper may prove useful beyond just proving lower bounds on Parity.

We start a more detailed discussion of these ideas by first defining QAC^0 and the Pauli spectrum of quantum channels.

3.1 QAC^0 Circuits

In classical Complexity Theory, AC^0 is the class of polynomial size, constant depth circuits, composed of alternating layers of AND and OR gates (NOT gates only allowed in the input layer) with unbounded fan-in. Since OR gates can be implemented by applying the NOT gate to every input of an AND gate, we can alternatively define the class by allowing alternating layers of multi-input AND gates and single-input NOT gates.

The quantum generalization of this family, QAC^0, is the class of polynomial size, constant depth circuits that have alternating layers of multi-qubit generalized Tofolli gates and arbitrary single-qubit unitaries. It is important to note that in the classical case, we can also assume unbounded fanout since we can add polynomially many copies of any gate to a circuit by increasing its size by a polynomial factor. In the quantum setting, however, the unbounded fanout assumption cannot be made without loss of generality due to the no-cloning theorem. It turns out that allowing unbounded fanout is equivalent to being able to compute parity [Moo99]. Thus, we do not make this assumption for our definition.

Another technicality worth noting is that we can use multi-qubit CZ gates in our definition instead of generalized Tofolli gates, where the action of a CZ

gate is defined on a computational basis string $|x\rangle = |x_1 x_2 \dots x_n\rangle$

$$\mathsf{CZ}_n \, |x_1 x_2 \dots x_n\rangle = (-1)^{x_1 \cdot x_2 \dots x_n} \, |x_1 x_2 \dots x_n\rangle$$

If the CZ gate only acts on a subset of qubits $S \subseteq [n]$, we write

$$\mathsf{CZ}_S \, |x_1 x_2 \dots x_n\rangle = I_{[n]\setminus S} \otimes \mathsf{CZ} \, |x_i : i \in S\rangle$$

While Tofolli gates are a more natural generalization of AND gates, we adopt CZ in our definition since the gate has no designated output qubit and commutes with the SWAP gate. It is easy to observe that the two definitions are equivalent since a Tofolli gate can be obtained from a CZ gate by conjugating the desired output qubit with Hadamards.

3.2 Pauli Spectrum of Quantum Channels

As stated earlier, we want our notion of the Pauli spectrum to be the analog of the Fourier spectrum of Boolean functions. Since the Pauli Matrices on n-qubits, $\mathcal{P}_n := \{I, X, Y, Z\}^{\otimes n}$, form an orthonormal basis for the space of $N \times N$ matrices $\mathbb{M}_N$, where $N := 2^n$, a first attempt might be to decompose the unitary associated with the circuit in the Pauli basis. Continuing our analogy with the classical case, we want the spectrum of "simple" circuits to be supported only by low-degree terms. However, consider the following example: $U = X^{\otimes n}$. Arguably, U corresponds to a very simple circuit, yet the above notion of the Pauli spectrum assigns non-zero values to a degree-n term.

Any notion of a Pauli spectrum on circuits with n outputs faces the same issue. This suggests we must define the notion on circuits with a single designated output qubit, much like Boolean functions that have a single output. Thus, instead of looking at unitaries from n-qubits to n-qubits, we consider *channels* from n-qubits to 1-qubit. For a circuit U that acts on $n + a$ qubits, where n is the number of inputs and a the number of auxiliary qubits, and has a single output qubit, we define its associated channel as

$$\mathcal{E}_U(\rho) = \mathrm{Tr}_{\mathrm{in+aux}}(U \rho U^\dagger)$$

where in + aux refers to the input and auxiliary qubits.

A convenient way to represent a channel is via its *Choi Matrix.* Intuitively, the Choi matrix stores the result of applying the channel to $|r\rangle\langle s|$ in its $(r, s)^{th}$ entry. We give a more formal definition.

Definition 1 (Choi Matrix)**.** For a channel n-to-l qubit channel $\mathcal{E}$, the *Choi Matrix* $\Phi_{\mathcal{E}}$ is given by

$$\Phi_{\mathcal{E}} = I^{\otimes n} \otimes \mathcal{E}\left(|\mathrm{EPR}_n\rangle\langle \mathrm{EPR}_n|\right)$$

where $|\mathrm{EPR}_n\rangle = \sum_{x \in \{0,1\}^n} |x\rangle \, |x\rangle$ is the (unnormalized) maximally entangled state on n qubits.

We note that the Choi Matrix can equivalently be written as

$$\Phi_{\mathcal{E}} = \sum_{x,x' \in \{0,1\}^n} |x\rangle\langle x'| \otimes \mathcal{E}(|x\rangle\langle x'|)$$

We define the Pauli spectrum of a channel by the representation of its Choi Matrix in the Pauli basis.

Definition 2 (Pauli Coefficient)**.** For $A \in \mathbb{M}_N$ and $P \in \mathcal{P}_n$, where $N = 2^n$, we define the P^{th} *Pauli coefficient* of A as

$$\widehat{A}(P) = \frac{1}{N}\langle P, A\rangle$$

where $\langle \cdot, \cdot \rangle$ is the Hilbert-Schmidt innerproduct on $N \times N$ matrices. Thus, the *Pauli-basis representation* of A is given by

$$A = \sum_{P \in \mathcal{P}_n} \widehat{A}(P)P$$

The Pauli spectrum of a matrix is an element of the space $\mathbb{C}^{4^n}$. Thus, the natural inner product and the induced norm on the Pauli spectra of $A, B \in \mathbb{M}_N$ are

$$\langle \widehat{A}, \widehat{B}\rangle = \sum_{P \in \mathcal{P}_n} \widehat{A}(P)^* \widehat{B}(P)$$

$$\|\widehat{A}\|_2 = \sqrt{\langle \widehat{A}, \widehat{A}\rangle} = \sqrt{\sum_{P \in \mathcal{P}_n} |\widehat{A}(P))|^2}$$

Since the Pauli representation can be thought of as a change of basis, the inner product and l_2 norm on $\mathbb{M}_N$ are preserved. Thus, we obtain the Plancherel and Parseval identities.

Lemma 3 (Plancherel/Parseval)**.** *For $A, B \in \mathbb{M}_N$*

$$\frac{1}{N}\langle A, B\rangle = \langle \widehat{A}, \widehat{B}\rangle \quad \frac{1}{N}\|A\|_2^2 = \|\widehat{A}\|_2^2$$

In analogy to the classical case, where the Fourier representation of a function is a polynomial, we can also think of the Pauli representation of a channel as a polynomial. Under this analogy $P \in \mathcal{P}_n$ can be considered a monomial. We define the *degree* of P, denoted by $|P|$, as the number of qubits P acts non-trivially on. We say that P is in the *support* of the A's spectrum if $\widehat{A}(P) \neq 0$. We can also define the k^{th} *level weight* of a matrix A as follows

$$\mathbf{W}^{=k}[A] = \sum_{|P|=k} |\widehat{A}(P)|^2$$

3.3 Concentration for Circuits without Auxiliary Qubits

In this subsection, we prove LMN-like concentration for the Pauli spectra of QAC^0 circuits without auxiliary qubits. More precisely, we prove the following theorem

Theorem 4. *Let C be a depth d, size s QAC^0 circuit acting on n input qubits and no auxiliary qubits. Let Φ_C be the Choi Matrix associated with the circuit. Then,*

$$\mathbf{W}^{>k}\left[\Phi_C\right] \leq O(s^2 2^{-k^{1/d}})$$

As discussed earlier, the general strategy for proving this theorem is to show that C is "close" to a circuit $\widetilde{C}$ whose output depends on a constant number of inputs. We construct $\widetilde{C}$ via what we call "the lightcones transformation"; we delete all the gates in C with large fan-in. Since these gates act on many inputs, the gates behave as the identity operator for most inputs to the circuit. Intuitively, removing these gates should not affect the circuit for most inputs. Thus, we expect the $\widetilde{C}$'s spectrum to be close to that of C. A lightcones argument shows that $\widetilde{C}$ has no weight on high-degree terms.

Lemma 5. *Let C be QAC circuit with m CZ gates that have fan-in at least l. Let $\widetilde{C}$ be the circuit obtained by removing these m gates. For $P \in \mathcal{P}_{n+1}$, if $P > l^d$*

$$\widehat{\Phi_{\widetilde{C}}}(P) = 0$$

Proof. Consider the past lightcone of the output wire of $\widetilde{C}$, i.e. all the qubits that could have influenced the state of the output qubit. Since there are only d layers of multiqubit gates and every multiqubit gate acts on at most l qubits, there are at most l^d qubits in the lightcone of the output qubit. Let L be the set of qubits within the lightcone. We can write $\widetilde{C}$ as

$$\widetilde{C} = (I_L \otimes U_{\text{rest}})(U_L \otimes I_{\text{rest}})$$

where U_L contains all gates that act on L and U_{rest} is the rest of the circuit. Since U_{rest} does not affect the output at all, we can write the channel associated to $\widetilde{C}$ as

$$\mathcal{E}_{\widetilde{C}}(\rho) = \text{Tr}_L(U_L \rho U_L^\dagger) = \mathcal{E}_{U_L}(\rho)$$

Thus, the Choi Matrix of $\widetilde{C}$ can be written as

$$\begin{aligned}\Phi_{\widetilde{C}} &= I^{\otimes n} \otimes \mathcal{E}_{\widetilde{C}}(|\text{EPR}_n\rangle\langle\text{EPR}_n|) \\ &= I_{\text{rest}} \otimes I_L \otimes \mathcal{E}_{U_L}(|\text{EPR}_{|L|}\rangle\langle\text{EPR}_{|L|}|) \\ &= I_{\text{rest}} \otimes \Phi_{U_L}\end{aligned}$$

By the above equation, $P \in \mathcal{P}_{n+1}$ is in the support of $\widetilde{C}$ if and only if it is in the support of U_L. Since U_L acts on l^d qubits, any P in the support of the Φ_{U_L} must have degree at most l^d. This implies the result. □

Next, we argue that the l_2 distance between C and $\widetilde{C}$ must be small.

Lemma 6. *Let C and $\widetilde{C}$ be as described in Lemma 5*

$$\left\|\widehat{\Phi_C} - \widehat{\Phi_{\widetilde{C}}}\right\|_2^2 \leq O(m^2\, 2^{-l})$$

Since $m \leq s = \mathsf{poly}(n)$, the above bound is dominated by the 2^{-l} term as $l \to n$. Intuitively, this implies that removing gates with very large fan-in does not change the Pauli spectrum much.

To prove Lemma 6, we first prove that the distance between C and $\widetilde{C}$ is small.

This follows from a "hybrid worlds" argument. We remove the large fan-in gates one by one and argue that at each step the distance does not change significantly.

Lemma 7.

$$\frac{1}{2^n}\left\|C - \widetilde{C}\right\|_2^2 \leq O(m^2\, 2^{-l})$$

Proof. Let $S_1, S_2 \dots S_m \subseteq [n]$ be the subsets of qubits acted on by the large fan-in gates. We can write U and $\widetilde{U}$ as follows

$$\begin{aligned} C &= U_0\, \mathsf{CZ}_{S_1}\, U_1\, \mathsf{CZ}_{S_2} \dots U_{m-1}\, \mathsf{CZ}_{S_m}\, U_m \\ \widetilde{C} &= U_0 U_1 \dots U_{m-1} U_m \end{aligned}$$

The i^{th} intermediate unitary is obtained by removing the first i large fan-in gates.

$$\begin{aligned} U^{(i)} &:= U_0 U_1 \dots U_{i-1} U_i\, \mathsf{CZ}_{S_{i+1}}\, U_{i+1} \dots \mathsf{CZ}_{S_m}\, U_m \\ U^{(0)} &:= C \end{aligned}$$

We break up the distance between U and $\widetilde{U} = U^{(m)}$ in terms of the distance between subsequent $U^{(i)}$'s.

$$\begin{aligned} \left\|C - \widetilde{C}\right\|_2 &= \left\|\sum_{i=1}^{m} U^{(i-1)} - U^{(i)}\right\|_2 \\ &\leq \sum_{i=1}^{m} \left\|U^{(i-1)} - U^{(i)}\right\|_2 \end{aligned}$$

Each $U^{(i-1)}$ and $U^{(i)}$ differ only by a single CZ gate. Thus, for some unitaries V_1, V_2, their distance can be written as

$$\begin{aligned} \left\|U^{(i-1)} - U^{(i)}\right\|_2 &= \|V_1\, \mathsf{CZ}_{S_i}\, V_2 - V_1 V_2\|_2 \\ &= \|V_1(\mathsf{CZ}_{S_i} - I)V_2\|_2 \\ &= \|\mathsf{CZ}_{S_i} - I\|_2 \end{aligned}$$

where the last equality is obtained due to the unitary equivalence of the l_2 norm. Note that

$$\begin{aligned}\mathsf{CZ}_{S_i} - I &= I_{[n]\setminus S_i} \otimes (\mathsf{CZ} - I)\\ \mathsf{CZ} - I &= -2\,|1^{|S_i|}\rangle\langle 1^{|S_i|}|\end{aligned}$$

Thus,

$$\begin{aligned}\left\|\mathsf{CZ}_{S_i} - I\right\|_2 &= 2\left\|I_{[n]\setminus S_i}\right\|_2 \left\||1^{|S_i|}\rangle\langle 1^{|S_i|}|\right\|_2\\ &\le 2\cdot 2^{(n-l)/2}\end{aligned}$$

Finally, we get

$$\begin{aligned}\left\|C - \widetilde{C}\right\|_2 &\le 2m\cdot 2^{(n-l)/2}\\ \implies \frac{1}{2^n}\left\|C - \widetilde{C}\right\|_2^2 &\le 4m^2 2^{-l}\end{aligned}$$

□

Since C and $\widetilde{C}$ are close to each (in l_2 distance), it is natural to expect their Choi Matrices to be close as well. Via some (uninteresting) mathematical manipulations, it is possible to show the following lemma (see [Nad+24] for proof).

Lemma 8.

$$\left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2 \le 4\sqrt{2}\left\|U - \widetilde{U}\right\|_2$$

Combining Lemma 7, Lemma 8, and Parseval's identity gives us a proof for Lemma 6.

Proof (for Lemma 6).

$$\begin{aligned}\left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2 &\le 4\sqrt{2}\left\|U - \widetilde{U}\right\|_2\\ \implies \frac{1}{2^n}\left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2 &\le O(m^2\,2^{-l}) && \text{(by Lemma 7)}\\ \implies \left\|\widehat{\Phi_C} - \widehat{\Phi_{\widetilde{C}}}\right\|_2^2 &\le O(m^2\,2^{-l}) && \text{(by Parseval's Identity)}\end{aligned}$$

□

Through Lemma 5 and Lemma 6, we showed that the spectrum of $\widetilde{C}$ is supported only on low degree terms and through and that the spectra of C and $\widetilde{C}$ are close. Together these results imply the low degree concentration of Theorem 4.

Proof (for Theorem 4). Let $k = l^d + 1$.

$$\begin{aligned}
\mathbf{W}^{>k}\left[\Phi_C\right] &= \sum_{|P|>k} \left|\widehat{\Phi_C}(P)\right|^2 \\
&= \sum_{|P|>k} \left|\widehat{\Phi_C}(P) - \widehat{\Phi_{\widetilde{C}}}(P)\right|^2 \\
&= \left\|\widehat{\Phi_C} - \widehat{\Phi_{\widetilde{C}}}\right\|_2 \leq O(s^2 2^{-k^{1/d}})
\end{aligned}$$

□

3.4 Concentration for Circuits with Auxiliary Qubits

In this section, we extend our concentration results to circuits with auxiliary qubits. Let C be a circuit that acts on $n + a$ qubits, where n is the number of inputs and a is the number of auxiliary qubits. To ensure that the channel associated with the circuit is a function of only the input qubits, we require that the initial state of the input qubits and auxiliary qubits be a product state over the cut. Let ψ be a state on a qubits. We define $\mathcal{E}_{C,\psi}$ as the channel from the n to 1 qubit channel obtained by setting the state of the auxiliary qubits to ψ. Formally, let ρ be a state on n qubits,

$$\mathcal{E}_{C,\psi}(\rho) = \mathrm{Tr}_{\mathrm{in}\,+\,\mathrm{aux}}(C(\rho \otimes \psi)C^\dagger) = \mathcal{E}_C(\rho \otimes \psi)$$

Our primary interest is the case when the input qubits are "clean" i.e. $\psi = |0^a\rangle$. For convenience, we write

$$\mathcal{E}_{C,0} := \mathcal{E}_{C,|0^a\rangle\langle 0^a|}$$

Thus, we shift our focus to proving bounds on the Pauli Spectrum of this channel.

The key insight to extending results to circuits with auxiliary qubits is that the underlying circuit C makes no distinction between input and auxiliary qubits. If we expand the input to include the auxiliary qubits, the resulting circuit is a QAC^0 with no auxiliary qubits. In other words, the results from the previous section apply to the Pauli spectrum of Φ_C. If we can relate the spectrum of Φ_C and $\Phi_{C,0}$, we can get bounds on the latter's spectrum. Unfortunately, as we shall see shortly, the bounds obtained in this manner have an exponential dependency on a.

We first relate Φ_C and $\Phi_{C,0}$.

Lemma 9.

$$\Phi_{C,\psi} = \mathrm{Tr}_{\mathrm{aux}}((I_{\mathrm{in}} \otimes \psi^T \otimes I_{\mathrm{out}})\Phi_C) \tag{1}$$

Specifically,

$$\Phi_{C,0} = \langle 0^a| \, \Phi_C \, |0^a\rangle$$

where $|0^a\rangle$ *is short for* $I_{\mathrm{in}} \otimes |0^a\rangle \otimes I_{\mathrm{out}}$.

We omit the proof as it follows from a straightforward calculation. We recall that a channel $\mathcal{E}$ can be evaluated on an input ρ using its Choi Matrix as follows

$$\mathcal{E}(\rho) = \mathrm{Tr}_{\mathrm{in}}((\rho^T \otimes I_{\mathrm{out}})\Phi_{\mathcal{E}}))$$

Thus, Equation (1) can be loosely interpreted as a partially evaluating Φ_C with ψ in the auxiliary registers.

Using Equation we can relate the Pauli Spectra of Φ_C and $\Phi_{C,\psi}$

Lemma 10. *For all $P \in \mathcal{P}_{n+1}$*

$$\widehat{\Phi_{C,\psi}}(P) = \sum_{Q \in \mathcal{P}_a} \widehat{\Phi_C}(P \otimes Q)\,\mathrm{Tr}(\psi^T Q)$$

In the case $\psi = |0^a\rangle\langle 0^a|$,

$$\widehat{\Phi_{C,0}}(P) = \sum_{Q \in \mathcal{P}_a} \widehat{\Phi_C}(P \otimes Q)\,\langle 0^a|\,Q\,|0^a\rangle$$

We again omit the proof. Lemma 10 suggests a natural way to bound the spectrum of $\Phi_{C,0}$.

Theorem 11. *For $k \in [n+1]$,*

$$\mathbf{W}^{>k}\left[\Phi_{C,0}\right] \le 2^a \mathbf{W}^{>k}\left[\Phi_C\right]$$

Combining the above equation with Theorem 4 yields

$$\mathbf{W}^{>k}\left[\Phi_{C,0}\right] \le O(s^2 2^{-k^{1/d}+a})$$

Proof.

$$\begin{aligned}
\mathbf{W}^{>k}\left[\Phi_{C,0}\right] &= \sum_{|P|>k} \left|\widehat{\Phi_{C,0}}(P)\right|^2 \\
&= \sum_{|P|>k} \left| \sum_{Q \in \mathcal{P}_a} \widehat{\Phi_C}(P \otimes Q)\,\langle 0^a|\,Q\,|0^a\rangle \right|^2 \qquad (2)
\end{aligned}$$

We note that,

$$\langle 0^a|\,Q\,|0^a\rangle = \begin{cases} 1 & \text{if } Q \in \{I, Z\}^{\otimes a}, \\ 0 & \text{otherwise} \end{cases}$$

Thus, we can write (2) as

$$\begin{aligned}
\mathbf{W}^{>k}\left[\Phi_{C,0}\right] &= \sum_{|P|>k} \left| \sum_{Q \in \{I,Z\}^{\otimes a}} \widehat{\Phi_C}(P \otimes Q) \right|^2 \\
&\le 2^a \sum_{|P|>k} \sum_{Q \in \{I,Z\}^{\otimes a}} \left|\widehat{\Phi_C}(P \otimes Q)\right|^2 = 2^a \mathbf{W}^{>k}\left[\Phi_C\right]
\end{aligned}$$

where the inequality in the second line is obtained via Cauchy-Schwarz. □

Reading the above proof, it may seem that the 2^a factor arises due to Cauchy-Schwarz giving a loose estimate in this case, we will see later more fundamental reasons for the exponential dependence on a that point to the shortcomings of the lightcones transformation. However, Yuen et al. conjecture that, perhaps with new techniques, it should be possible to remove this factor [Nad+24].

3.5 Lower Bounds on Parity

The goal of this subsection is to use the low-degree bounds on the spectra of QAC^0 to show that they have a low correlation with Parity. The strategy is to prove a more general bound on the correlation of boolean functions and n to 1 qubit channel in terms of the Fourier spectrum of the boolean function and the Pauli spectrum of the channel. More specifically, we prove the following lemma

Lemma 12. *Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function and $\mathcal{E} : \left(\mathbb{C}^2\right)^{\otimes n} \to \mathbb{C}^2$ a channel. For $x \in \{0,1\}^n$, let $M(x)$ be the random variable representing the outcome of the measuring $\mathcal{E}(|x\rangle\langle x|)$ in the computational basis. Then,*

$$\mathbb{P}_{x\sim\{0,1\}^n}\left[M(x) = f(x)\right] \leq \frac{1}{2} + \sqrt{\mathbf{W}^{\leq k}\left[f\right]} + \sqrt{\mathbf{W}^{>k+1}\left[\Phi_{\mathcal{E}}\right]}$$

Since the entire Fourier weight of Parity is on a single n-degree term, choosing $k = n - 1$ readily yields the following bound,

Theorem 13. *Let C be a QAC^0 circuit and $M_C(x)$ be the result of measuring $\mathcal{E}_C(|x\rangle\langle x|)$ in the computational basis.*

$$\mathbb{P}_x\left[M_C(x) = \mathsf{PAR}(x)\right] \leq \frac{1}{2} + O(s \cdot \sqrt{2^{-n^{1/d}+a}})$$

In particular, when $a \leq \frac{1}{2} n^{1/d}$

$$\mathbb{P}_x\left[M_C(x) = \mathsf{PAR}(x)\right] \leq \frac{1}{2} + O(s \cdot 2^{\Theta(-n^{1/d})})$$

Intuitively, Theorem 13 says that when n is sufficiently large, a QAC^0 circuit with a small number of auxiliary qubits computes parity correctly on roughly half of the inputs. This means that the circuit does essentially no better than random guessing.

To prove Lemma 12, we consider first $\mathcal{E}_f$, the quantum channel that computes $f(x)$ perfectly. Formally,

$$\mathcal{E}_f(|x\rangle\langle x|) := |f(x)\rangle\langle f(x)| \qquad \Phi_f := \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes |f(x)\rangle\langle f(x)|$$

Notice that two spectra are associated with the boolean function f, the classical Fourier spectrum $\widehat{f}$ and the quantum Pauli spectrum $\widehat{\Phi_f}$. We start by relating the two spectra associated with $f(x)$.

Lemma 14. *Let $f : \{0,1\}^n \to \{0,1\}$ and $\mathcal{E}_f$ be as defined above.*

$$\Phi_f = \frac{1}{2} I^{\otimes n+1} + \frac{1}{2} \left(\sum_{S \subseteq [n]} \widehat{f}(S) Z_S \otimes Z \right)$$

For the sake of brevity, we omit the proof of this lemma.

To obtain the inequality in Lemma 12, we bound the correlation between $\mathcal{E}$ and f by the (l_2) inner-product between $\Phi_{\mathcal{E}}$ and Φ_f. This naturally gives a bound in terms of the Pauli spectrum of the two channels via Plancherel's identity. Since the Pauli spectrum of Φ_f is so intimately related to the Fourier spectrum of f as shown by Lemma 14, we express the inner product in terms of the Fourier coefficients. After that, a clever application of Cauchy-Schwarz yields the desired inequality.

Proof (for Lemma 12). Fix $x \in \{0,1\}^n$ and consider s_x the indicator variable that is 1 if we see $f(x)$ after measuring $\mathcal{E}(|x\rangle\langle x|)$. Note that the probability of succeeding on the string x is the same as the expected value of s_x. Since $|f(x)\rangle\langle f(x)|$ is a projection,

$$\begin{aligned}\mathbb{P}\left[M(x) = f(x)\right] &= \mathbb{E}\left[s_x\right] \\ &= \mathrm{Tr}(|f(x)\rangle\langle f(x)|\, \mathcal{E}(|x\rangle\langle x|)) \\ &= \mathrm{Tr}(\mathcal{E}_f(|x\rangle\langle x|)^{\dagger} \mathcal{E}(|x\rangle\langle x|))\end{aligned}$$

We observe that correlation with parity is the expected value of $\mathbb{P}\left[s_x = 1\right]$ over $x \in \{0,1\}^n$. Thus,

$$\begin{aligned}\mathbb{P}_{x \sim \{0,1\}^n}\left[M(x) = f(x)\right] &= \mathbb{E}_x[\mathbb{P}\left[s_x = 1\right]] \\ &= \mathbb{E}_x[\mathbb{E}\left[s_x\right]] \\ &= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \mathrm{Tr}(\mathcal{E}_f(|x\rangle\langle x|)\mathcal{E}(|x\rangle\langle x|)) \\ &= \frac{1}{2^n} \mathrm{Tr}\left(\sum_{x \in \{0,1\}^n} \mathcal{E}_f(|x\rangle\langle x|)\mathcal{E}(|x\rangle\langle x|) \right) \\ &= \frac{1}{2^n} \mathrm{Tr}\left(\Phi_f^{\dagger} \Phi_{\mathcal{E}} \right) \\ &= \frac{1}{2^n} \langle \Phi_f, \Phi_{\mathcal{E}} \rangle \\ &= 2 \langle \widehat{\Phi_f}, \widehat{\Phi_{\mathcal{E}}} \rangle\end{aligned}$$

where we get the last inequality via Plancherel. Expanding the inner product,

$$\langle \widehat{\Phi_f}, \widehat{\Phi_{\mathcal{E}}} \rangle = \sum_{P \in \mathcal{P}_{n+1}} \widehat{\Phi_f}(P)^* \widehat{\Phi_{\mathcal{E}}}(P) \tag{3}$$

By Lemma 14, we know that $\widehat{\Phi_f}(P) \neq 0$ if and only if $P = I^{\otimes n+1}$ or $P = Z_S \otimes Z$ for $S \subseteq [n]$. A simple calculation shows that $\widehat{\Phi_\mathcal{E}}(I) = \widehat{\Phi_f}(I) = \frac{1}{2}$. Thus, we can rewrite Equation (3) as

$$\langle \widehat{\Phi_f}, \widehat{\Phi_\mathcal{E}} \rangle = \frac{1}{4} + \frac{1}{2} \sum_{S \subseteq [n]} \widehat{f}(S) \widehat{\Phi_\mathcal{E}}(Z_S \otimes Z) \tag{4}$$

$$\tag{5}$$

where the first 1/4 term comes from the contribution of $P = I^{\otimes n+1}$. We bound the second term using Cauchy-Schwarz,

$$\begin{aligned} \sum_{S \subseteq [n]} \widehat{f}(S) \widehat{\Phi_\mathcal{E}}(Z_S \otimes Z) &= \sum_{|S| \leq k} \widehat{f}(S) \widehat{\Phi_\mathcal{E}}(Z_S \otimes Z) + \sum_{|S| > k} \widehat{f}(S) \widehat{\Phi_\mathcal{E}}(Z_S \otimes Z) \\ &= \|\widehat{f}\|_{\leq k} \|\widehat{\Phi_\mathcal{E}}(Z_S \otimes Z)\|_{\leq k} + \|\widehat{f}\|_{> k} \|\widehat{\Phi_\mathcal{E}}(Z_S \otimes Z)\|_{> k} \end{aligned}$$

where we define

$$\|\widehat{\Phi_\mathcal{E}}(Z_S \otimes Z)\|_{\leq k} := \sqrt{\sum_{|S| \leq k} \left| \widehat{\Phi_\mathcal{E}}(Z_S \otimes Z) \right|^2}$$

and $\|\widehat{f}\|_{\leq k}$ is defined similarly. Note that since $\|\widehat{f}\|_2, \|\widehat{\Phi_\mathcal{E}}\|_2 \leq 1$, implies

$$\begin{aligned} \|\widehat{f}\|_{\leq k} \|\widehat{\Phi_\mathcal{E}}(Z_S \otimes Z)\|_{\leq k} + \|\widehat{f}\|_{> k} \|\widehat{\Phi_\mathcal{E}}(Z_S \otimes Z)\|_{> k} &\leq \|\widehat{f}\|_{\leq k} + \|\widehat{\Phi_\mathcal{E}}(Z_S \otimes Z)\|_{> k} \\ &\leq \sqrt{\mathbf{W}^{\leq k}[f]} + \sqrt{\mathbf{W}^{> k}[\Phi_\mathcal{E}]} \end{aligned}$$

Combining this with Equation (3) and (4), we get the desired bound. □

4 Ideas for Improving Previous Results

In this section, we present ideas for improving the bounds presented in the previous section. In particular, we focus only on circuits that QAC^0 that compute "cleanly". Under this restriction, we examine the lightcones transformation and highlight its deficiencies. Based on this analysis, we propose a set of sufficient conditions for a transformation that would allow us to prove Yuen et al.'s conjecture.

4.1 Clean Computation

For the remainder of this section, we restrict our attention only to circuits that compute cleanly i.e. they revert the input and auxiliary qubits to their original state at the start of the computation. Note that any circuit that computes parity exactly can be converted into a circuit that computes parity cleanly using the "uncomputing garbage" trick from [Ben+97]. To use this trick, we need to add a "target" qubit to our circuit that is only used for copying over the output of the computation before the input and auxiliary qubits are uncomputed. Thus, the restriction is without loss of generality for this class of circuits. However, this is not the case for circuits that compute parity approximately.

Nevertheless, the restriction is useful as it offers a convenient interpretation of the Choi Matrix of circuits with auxiliary qubits. Let C be a cleanly computing circuit that acts on $n + a$ qubits. Let $C\,|x, y, 0\rangle = |x, y, \omega_{xy}\rangle$, where $\omega_{xy} \in \mathbb{C}^2$. The Choi Matrix of C can be written as

$$\begin{aligned}
\Phi_C &= \sum_{\substack{x,x' \in \{0,1\}^n \\ y,y' \in \{0,1\}^a}} |x\rangle\langle x'| \otimes |y\rangle\langle y'| \otimes \mathrm{Tr}_{\mathrm{in}\,+\,\mathrm{aux}}(C\,|x, y, 0\rangle\langle x', y', 0|\, C^\dagger) \\
&= \sum_{\substack{x,x' \in \{0,1\}^n \\ y,y' \in \{0,1\}^a}} |x\rangle\langle x'| \otimes |y\rangle\langle y'| \otimes \mathrm{Tr}_{\mathrm{in}\,+\,\mathrm{aux}}(|x, y, \omega_{xy}\rangle\langle x', y', \omega_{x'y'}|) \\
&= \sum_{\substack{x \in \{0,1\}^n \\ y \in \{0,1\}^a}} |x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \mathrm{Tr}_{\mathrm{in}\,+\,\mathrm{aux}}(C\,|x, y, 0\rangle\langle x, y, 0|\, C^\dagger)
\end{aligned}$$

We get the last equality since the partial trace over the input and auxiliary qubits evaluates to 0 when $x \neq x'$ and $y \neq y'$. Notice that by rearranging the order of the tensors, we get

$$\begin{aligned}
\Phi_C &= \sum_y |y\rangle\langle y| \otimes \left(\sum_x |x\rangle\langle x| \otimes \mathrm{Tr}_{\mathrm{in}\,+\,\mathrm{aux}}(C\,|x, y, 0\rangle\langle x, y, 0|\, C^\dagger) \right) \\
&= \sum_y |y\rangle\langle y| \otimes \Phi_{C,y}
\end{aligned}$$

Each term in the sum can be thought of as the projection of Φ_C on the subspace spanned by $I_{\mathrm{in}} \otimes |y\rangle\langle y| \otimes I_{\mathrm{out}}$. Alternatively, it can be thought of as restricting the auxiliary qubits of C to the string y. We make the following definition:

Definition 15. The y-auxiliary restriction of Φ_C is given by

$$\Phi_{C|_y} := |y\rangle\langle y| \otimes \Phi_{C,y}$$

Thus, normalizing Φ_C, we can we can write

$$\frac{1}{2^{n+a+1}} \Phi_C = \frac{1}{2^a} \sum_y \frac{1}{2^{n+1}} \Phi_{C|_y}$$

Note that $1/2^{n+1}$ is the correct normalization factor for $\Phi_{C,y}$. This suggests that we can interpret Φ_C as a uniform distribution over $\Phi_{C|_y}$, in the same way we interpret a mixed state as a classical distribution over its eigenvectors.

4.2 Unraveling the Lightcones Argument

As seen in the previous section, Yuen et al.'s approach to proving low-degree concentration for circuits with auxiliary qubits is to reduce it to the without auxiliary qubits case and then apply the lightcones argument. Under this approach, the 2^a factor in Theorem 11 seems rather arbitrary. The proof for Theorem 11 makes it seem we could remove the exponential factor if we used a sharper estimate than Cauchy-Schwarz.

However, unraveling the argument under the clean computation restriction reveals a more fundamental reason for the shortcomings of Theorem 11. Let C be a QAC^0 circuit that acts on $n+a$ qubits and $\widetilde{C}$ be the circuit obtained after the lightcones transformation (LT) i.e. after removing all the gates with fan-in greater than l. We know from Lemma 5 and Lemma 8 that

$$\left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2^2 \leq O(s^2\, 2^{n+a-l}) \tag{6}$$

Instead of bounding the Pauli weights of $\Phi_{C,0}$ in terms of Φ_C, we try to obtain a bound from the above inequality directly.

$$\left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2^2 = \left\| \sum_{y\in\{0,1\}^a} \Phi_{C|_y} - \Phi_{\widetilde{C}|_y} \right\|_2^2$$

Notice that for $y \neq y'$,

$$\langle \Phi_{C|_y} - \Phi_{\widetilde{C}|_y}, \Phi_{C|_{y'}} - \Phi_{\widetilde{C}|_{y'}} \rangle = 0$$

Thus, by the Pythagoras' Theorem,

$$\begin{aligned} \left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2^2 &= \sum_{y\in\{0,1\}^a} \left\| \Phi_{C|_y} - \Phi_{\widetilde{C}|_y} \right\|_2^2 \\ \implies \sum_{y\in\{0,1\}^a} \left\| \Phi_{C|_y} - \Phi_{\widetilde{C}|_y} \right\|_2^2 &\leq O(s^2\, 2^{n+a-l}) \\ \implies \sum_{y\in\{0,1\}^a} \left\| \widehat{\Phi_{C|_y}} - \widehat{\Phi_{\widetilde{C}|_y}} \right\|_2^2 &\leq O(s^2\, 2^{-l}) \end{aligned}$$

where the last inequality follows from Parseval. Furthermore,

$$\left\| \widehat{\Phi_{C|_0}} - \widehat{\Phi_{\widetilde{C}|_0}} \right\|_2^2 \leq \sum_{y\in\{0,1\}^a} \left\| \widehat{\Phi_{C|_y}} - \widehat{\Phi_{\widetilde{C}|_y}} \right\|_2^2 \leq O(s^2\, 2^{-l}) \tag{7}$$

A simple calculation shows that

$$\left\| \widehat{\Phi_{C,0}} - \widehat{\Phi_{\widetilde{C},0}} \right\|_2^2 = 2^a \left\| \widehat{\Phi_{C|_0}} - \widehat{\Phi_{\widetilde{C}|_0}} \right\|_2^2$$

Combining with Equation (7), we get

$$\left\| \widehat{\Phi_{C,0}} - \widehat{\Phi_{\widetilde{C},0}} \right\|_2^2 \leq O(s^2\, 2^{-l+a})$$

Following the rest of the argument from Section 3.3, we have

$$\mathbf{W}^{>k}\left[\Phi_{C,0}\right] \leq O(s^2\, 2^{-k^{1/d}+a})$$

Thus, we recover the bounds from Theorem 11. We note that the argument has equalities everywhere except Equation (6) and Equation (7). One might think that the first estimate in Equation (7) is very loose, which would explain the 2^a factor. However, as the following example illustrates, this bound can be sharp.

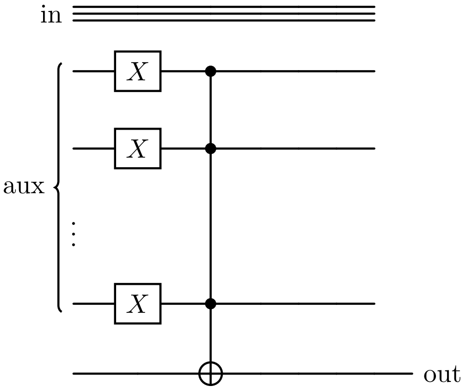

Figure 1: Example to Show Equation (7) is Sharp

Call the circuit in the above figure C (C acts on n input qubits and $n+1$ auxiliary qubits). Let $\widetilde{C}$ be the circuit obtained after applying LT to C for any $l \leq n$. The transformation only removes the CNOT gate that acts on the auxiliary qubits. Hence $\Phi_{C,y} = \Phi_{\widetilde{C},y}$ for all $y \neq 0^{n+1}$. Thus, the first inequality in Equation (7) is sharp in this case.

The above example highlights a deeper point about the LT approach. The Pauli spectrum of C obeys the low-degree concentration inequality without the 2^a factor (this is easy to check via direct calculation). However, if one follows the reconstruction of the lightcones argument presented in this subsection, all inequalities (expect (6)) are tight. Thus, the deficiency is not with our estimates but with LT itself.

4.3 In Search of a Better Transformation

Having observed the shortcomings of the LT, it is natural to ask what properties should a desirable transformation have. Said differently, what conditions should a transformation satisfy so that the argument in the previous subsection does not lead to the 2^a factor?

Note that LT possessed two properties that were crucial to the previous argument. Firstly, given a circuit C LT produced a circuit $\widetilde{C}$ such that the distance between the Pauli spectra decayed exponentially with l. This gave us Equation (6), which was the starting point of our analysis. Secondly, the Pauli spectrum of $\widetilde{C}$ exhibited low degree concentration. Bounds on the distance between Pauli spectra are useful only if this condition is satisfied.

As we argued in the previous section, the problem with LT lies in the first inequality of Equation (7). We note that if we could improve the bound in Equation (7) to

$$\left\|\widehat{\Phi_{C|_0}} - \widehat{\Phi_{\widetilde{C}|_0}}\right\|_2^2 \leq \frac{k}{2^a} \sum_{y\in\{0,1\}^a} \left\|\widehat{\Phi_{C|_y}} - \widehat{\Phi_{\widetilde{C}|_y}}\right\|_2^2 \tag{8}$$

where $k \leq \mathsf{poly}(n)$, then the 2^a factor would disappear from our final bound. Is there a natural interpretation for (8)? We define $\Delta\big|_y$ to be the distance between $\Phi_{C|_y}$ and $\Phi_{\widetilde{C}|_y}$ when y is chosen randomly. That is,

$$\Delta\big|_y := \left\|\widehat{\Phi_{C|_y}} - \widehat{\Phi_{\widetilde{C}|_y}}\right\|_2^2$$

We note that,

$$\mathbb{E}_y[\Delta\big|_y] = \frac{1}{2^a} \sum_{y\in\{0,1\}^a} \left\|\widehat{\Phi_{C|_y}} - \widehat{\Phi_{\widetilde{C}|_y}}\right\|_2^2$$

Thus, we can rewrite (8) as,

$$\left\|\widehat{\Phi_{C|_0}} - \widehat{\Phi_{\widetilde{C}|_0}}\right\|_2^2 \leq k\mathbb{E}_y[\Delta\big|_y]$$

Since there is nothing special about the string 0^a, we can strengthen the requirement to

$$\left\|\widehat{\Phi_{C|_z}} - \widehat{\Phi_{\widetilde{C}|_z}}\right\|_2^2 \leq k\mathbb{E}_y[\Delta\big|_y],\ \forall\, z \in \{0,1\}^a \tag{9}$$

We can interpret Equation 9 as follows: we require not only that the expected value of $\Delta\big|_y$ be small, but also that the distribution of Δ_y has small support. More precisely, we require the support to be at most polynomially large in n.

We distill our observation from this section into the following Proposition,

Proposition 16. *Let C be a QAC^0. If we find a transformation such that for any $l > 0$, the transformation produces $\widetilde{C}$ satisfying the following conditions, then we obtain the desired low-degree concentration bounds.*

1. $\left\|\Phi_C - \Phi_{\widetilde{C}}\right\|_2^2 \leq \mathsf{poly}(n) \cdot O(2^{n+a-l}) \iff \mathbb{E}_y[\Delta\big|_y] \leq \mathsf{poly}(n) \cdot O(2^{-l})$

2. $\mathbf{W}^{>l^d+1}\left[\Phi_{\widetilde{C}}\right] \leq \mathsf{poly}(n) \cdot O(2^{-l})$

3. *For all* $y \in \{0,1\}^a$, $\left\|\widehat{\Phi_{C|_z}} - \widehat{\Phi_{\widetilde{C}|_z}}\right\|_2^2 \leq \mathsf{poly}(n) \cdot \mathbb{E}_y[\Delta\big|_y]$

5 Conclusion

This work summarizes and analyzes Yuen et al.'s approach to proving lower bounds for parity using analytic techniques. Through our discussion, it is clear that Yuen et al.'s work presents a leap forward in obtaining lower bounds for general QAC^0 circuits and extending techniques from the analysis of boolean functions to the quantum setting. Yet, as we have seen much work remains to be done. By analyzing the shortcomings of Yuen et al.'s approach, we identified sufficient conditions for a transformation. The search for such a transformation continues!