Bounds on the QAC^0 Complexity of Approximating Parity

Gregory Rosenthal
University of Toronto, Canada
https://www.cs.toronto.edu/~rosenthal/
rosenthal@cs.toronto.edu

Abstract

QAC circuits are quantum circuits with one-qubit gates and Toffoli gates of arbitrary arity. QAC^0 circuits are QAC circuits of constant depth, and are quantum analogues of AC^0 circuits. We prove the following:

- For all $d \geq 7$ and $\varepsilon > 0$ there is a depth-d QAC circuit of size $\exp(\text{poly}(n^{1/d}) \log(n/\varepsilon))$ that approximates the n-qubit parity function to within error ε on worst-case quantum inputs. Previously it was unknown whether QAC circuits of sublogarithmic depth could approximate parity regardless of size.
- We introduce a class of "mostly classical" QAC circuits, including a major component of our circuit from the above upper bound, and prove a tight lower bound on the size of low-depth, mostly classical QAC circuits that approximate this component.
- Arbitrary depth-d QAC circuits require at least $\Omega(n/d)$ multi-qubit gates to achieve a $1/2 + \exp(-o(n/d))$ approximation of parity. When $d = \Theta(\log n)$ this nearly matches an easy $O(n)$ size upper bound for computing parity exactly.
- QAC circuits with at most two layers of multi-qubit gates cannot achieve a $1/2 + \exp(-o(n))$ approximation of parity, even non-cleanly. Previously it was known only that such circuits could not cleanly compute parity exactly for sufficiently large n.

The proofs use a new normal form for quantum circuits which may be of independent interest, and are based on reductions to the problem of constructing certain generalizations of the cat state which we name "nekomata" after an analogous cat yōkai.

2012 ACM Subject Classification Theory of computation → Circuit complexity; Theory of computation → Quantum complexity theory; Theory of computation → Quantum complexity theory

Keywords and phrases quantum circuit complexity, QAC^0, fanout, parity, nekomata

Digital Object Identifier 10.4230/LIPIcs.ITCS.2021.32

Related Version A full version of the paper is available at https://arxiv.org/abs/2008.07470. See "v3" for the full version that most closely corresponds to the ITCS proceedings version.

Funding Supported by NSERC (PGS D).

Acknowledgements Thanks to Benjamin Rossman and Henry Yuen for introducing me to this problem, and for having several helpful discussions throughout the research and writing processes. Thanks to Srinivasan Arunachalam, Daniel Grier, Ian Mertz, Eric Rosenthal, and Rahul Santhanam for helpful discussions as well. Part of this work was done while the author was visiting the Simons Institute for the Theory of Computing. Circuit diagrams were made using the Quantikz package [12].

1 Introduction

1.1 Background

A central problem in computational complexity theory is to prove lower bounds on the nonuniform circuit size required to compute explicit boolean functions. Since this appears to be out of reach given current techniques, research in circuit complexity has instead focused on

12th Innovations in Theoretical Computer Science Conference (ITCS 2021).
Editor: James R. Lee; Article No. 32; pp. 32:1–32:20
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

proving lower bounds in restricted circuit classes. There are now many known lower bounds in *classical* circuit complexity, as well as in quantum *query* complexity, but comparatively few lower bounds are known in quantum circuit complexity, which is the subject of the current paper.

The study of quantum circuit complexity was initiated in large part by Green, Homer, Moore and Pollett [8], who defined quantum analogues of a number of classical circuit classes. One of the seemingly most restrictive quantum circuit classes that they defined is the class of QAC^0 circuits, consisting of constant-depth QAC circuits, where QAC circuits are quantum circuits with arbitrary one-qubit gates and generalized Toffoli gates of arbitrary arity. (More precisely, $(n+1)$-ary generalized Toffoli gates are defined by $|x, b\rangle \mapsto |x, b \oplus \bigwedge_{j=1}^{n} x_j\rangle$ for $x = (x_1, \ldots, x_n) \in \{0,1\}^n, b \in \{0,1\}$.) This is analogous to the classical circuit class of AC^0 circuits, consisting of constant-depth AC circuits, where AC circuits are boolean circuits with NOT gates and unbounded-fanin AND and OR gates. Low-depth circuits are a model of fast parallel computation, and this is especially important for quantum circuits, because quantum computations need to be fast relative to the decoherence time of the qubits in order to avoid error.

One difference between AC and QAC circuits is that AC circuits are allowed fanout "for free", i.e. the input bits to the circuit and the outputs of gates may all be used as inputs to arbitrarily many gates. The quantum analogue of this would be to compute the unitary "fanout" transformation U_F, defined by $U_F|b, x_1, \ldots, x_{n-1}\rangle = |b, x_1 \oplus b, \ldots, x_{n-1} \oplus b\rangle$ for $b, x_1, \ldots, x_{n-1} \in \{0,1\}$, or at least to compute this in the case that we call "restricted fanout" in which $x_1 = \cdots = x_{n-1} = 0$. QAC^0 circuits with fanout gates are called QAC$^0_\mathrm{f}$ circuits, and can simulate arbitrary AC^0 circuits by using ancillae and restricted fanout to make as many copies as needed of the input bits and of the outputs of gates. In fact, QAC$^0_\mathrm{f}$ circuits are *strictly* more powerful than AC^0 circuits, because QAC$^0_\mathrm{f}$ circuits (even without generalized Toffoli gates) of polynomial size can also compute threshold functions [11, 14] whereas AC^0 circuits require exponential size to do so [10]. In contrast, little is known about the power of QAC^0 circuits and how it compares with that of AC^0 circuits.

Green et al. [8] observed that fanout can be computed by QAC circuits of logarithmic depth and linear size. This raises the question of whether QAC circuits of *sub*logarithmic depth can compute fanout, or at least restricted fanout, even if allowed *arbitrary* size. The same question can be asked about parity, which is a famous example of a function that requires exponential size to compute in AC^0 [10], and which is defined for quantum circuits as the unitary transformation $U_\oplus$ such that $U_\oplus|b, x\rangle = |b \oplus \bigoplus_{j=1}^{n-1} x_j, x\rangle$ for $b \in \{0,1\}, x = (x_1, \ldots, x_{n-1}) \in \{0,1\}^{n-1}$. In fact, all of these questions are equivalent: Green et al. [8] proved that parity and fanout are equivalent up to conjugation by Hadamard gates, and that they reduce to restricted fanout with negligible blowups in size and depth (see the full paper for illustrations).

Recent work [7, 9] suggests that QAC$^0_\mathrm{f}$ may be a physically realistic model of constant depth computation in certain quantum computing architectures (such as ion traps). As for QAC lower bounds, Fang, Fenner, Green, Homer and Zhang [5] proved that QAC circuits with a ancillae require depth at least $\Omega(\log(n/(a+1)))$ to compute the n-qubit parity and fanout functions, which is a nontrivial lower bound when a is $o(n)$. Bera [3] used a different approach to prove something slightly weaker than the $a = 0$ case of this result. Finally, Padé, Fenner, Grier and Thierauf [13] proved that QAC circuits with two layers of generalized Toffoli gates cannot cleanly[1] compute 4-qubit parity or fanout, regardless of the number of ancillae. A survey of Bera, Green and Homer [4] discusses some of the aforementioned QAC lower bounds and QAC$_\mathrm{f}$ upper bounds in greater detail.

[1] A clean computation is one in which the ancillae end in the all-zeros state.

1.2 Results and Selected Proof Overviews

1.2.1 Definitions of Complexity Measures

Call $|\langle\psi|\varphi\rangle|^2$ the *fidelity* of states $|\varphi\rangle$ and $|\psi\rangle$. We define the *size* of a QAC circuit to be the number of multi-qubit gates in it, and the *depth* of a QAC circuit to be the number of layers of multi-qubit gates in it. One motivation for not counting single-qubit gates, besides mathematical convenience, is that size and depth can be interpreted as measures of the reliability and computation time of a quantum circuit respectively, and in practice multi-qubit gates tend to be less reliable and take more time to apply as compared to single-qubit gates.

1.2.2 Reductions to and from Constructing Nekomata

Recall that Green et al. [8] proved that parity, fanout, and restricted fanout are all equivalent up to low-complexity QAC reductions. In Section 3 we make the more general observation that clean approximate and non-clean approximate versions of these problems are all equivalent in this sense. For brevity's sake, here in Section 1.2 we will only state immediate corollaries of these reductions insofar as they relate to our other results.

We also introduce another problem equivalent to parity, which all of our results about parity and fanout are proved via reductions to. The state $\frac{1}{\sqrt{2}}\sum_{b=0}^{1}|b^n\rangle$ is commonly called the *cat state* on n qubits, and we denote it by $|\text{😺}_n\rangle$. More generally, call a state $|\nu\rangle$ an n*-nekomata* if $|\nu\rangle = \frac{1}{\sqrt{2}}\sum_{b=0}^{1}|b^n,\psi_b\rangle$ for some states $|\psi_0\rangle, |\psi_1\rangle$ on any number of qubits (the word "nekomata" is also the name of two-tailed cats from Chinese and Japanese folklore), or equivalently if a standard-basis measurement of some n qubits of $|\nu\rangle$ outputs all-zeros and all-ones each with probability $1/2$.

Call a QAC circuit C acting on any number of qubits a solution to the "p-approximate n-nekomata problem" if there exists an n-nekomata $|\nu\rangle$ such that $C|0\dots0\rangle$ and $|\nu\rangle$ have fidelity at least p. (There is no need to allow "ancillae" in this problem, because if $|\nu\rangle$ is an n-nekomata then so is $|\nu,\psi\rangle$ for any state $|\psi\rangle$.) Note that the identity circuit on n or more qubits trivially solves the $1/2$-approximate n-nekomata problem. In informal discussions we will often say that a circuit "constructs an approximate n-nekomata" if it solves the p-approximate n-nekomata problem for some fixed $p \in (1/2, 1)$, say $p = 3/4$.

Constructing nekomata reduces to computing restricted fanout because $|\text{😺}_n\rangle = U_F(H \otimes I)|0^n\rangle$. Our reduction from parity to constructing nekomata is a variant of Green et al.'s [8] reduction from parity to restricted fanout.

1.2.3 Upper Bounds

▶ **Theorem 1.1.** *For all $\varepsilon > 0$ there exists a depth-2 QAC circuit C such that for some n-nekomata $|\nu\rangle$, the fidelity of $C|0\dots0\rangle$ and $|\nu\rangle$ is at least $1-\varepsilon$. Furthermore, the size of C and the number of qubits acted on by C are both $\exp(O(n\log(n/\varepsilon)))$.*

To state a stronger upper bound for approximating unitary transformations than can conveniently be done in terms of fidelity, call $1 - \||\varphi\rangle - |\psi\rangle\|_2^2$ the *phase-dependent fidelity* of states $|\varphi\rangle$ and $|\psi\rangle$. This quantity is at most the fidelity of $|\varphi\rangle$ and $|\psi\rangle$ (Equation (2)).

▶ **Corollary 1.2.** *For all $d \geq 7$ and $\varepsilon > 0$ there exist depth-d QAC circuits $C_\oplus, C_F, C_{\text{😺}}$ of size and number of ancillae $\exp(\mathrm{poly}(n^{1/d})\log(n/\varepsilon))$, where the $\mathrm{poly}(n^{1/d})$ term is at most $O(n)$, such that for all n-qubit states $|\phi\rangle$,*

- *the phase-dependent fidelity of $C_\oplus|\phi,0\dots0\rangle$ and $U_\oplus|\phi\rangle\otimes|0\dots0\rangle$ is at least $1-\varepsilon$;*
- *the phase-dependent fidelity of $C_F|\phi,0\dots0\rangle$ and $U_F|\phi\rangle\otimes|0\dots0\rangle$ is at least $1-\varepsilon$;*
- *the phase-dependent fidelity of $C_{\text{😺}}|0\dots0\rangle$ and $|\text{😺}_n,0\dots0\rangle$ is at least $1-\varepsilon$.*

The $d = 11$ case of Corollary 1.2 follows immediately from Theorem 1.1 and our reduction from parity to constructing nekomata. We decrease the minimum depth from 11 to 7 using an optimization specific to the circuit from our proof of Theorem 1.1. We prove Corollary 1.2 for higher depths using the fact that n-qubit restricted fanout can be computed by a circuit consisting of d layers of $n^{1/d}$-qubit restricted fanout gates.

If we were to also count one-qubit gates toward size and depth, then statements similar to Theorem 1.1 and Corollary 1.2 would still hold, because without loss of generality a depth-d QAC circuit acting on m qubits has at most $d+1$ layers of one-qubit gates and at most $(d+1)m$ one-qubit gates.

1.2.4 Tight Lower Bounds for Constructing Approximate Nekomata in "Mostly Classical" Circuits

Call a QAC circuit *mostly classical* if it can be written as $CLML^\dagger$ (i.e. C is applied last) such that C consists only of generalized Toffoli gates, L is a layer of one-qubit gates, and M is a layer of generalized Toffoli gates. The circuit C here is a close analogue of (classical) AC circuits with bounded fanout, since generalized Toffoli gates can simulate classical AND and NOT gates. The following is apparent from our proof of Theorem 1.1:

▶ Remark 1.3. Theorem 1.1 remains true even if "QAC circuit" is replaced by "mostly classical QAC circuit".

Motivated by Remark 1.3, we prove the following lower bound for constructing approximate nekomata in mostly classical circuits:

▶ **Theorem 1.4.** *Let C be a mostly classical circuit of size s and depth $o(\log n)$, acting on any number of qubits. Then for all n-nekomata $|\nu\rangle$, the fidelity of $C|0\dots0\rangle$ and $|\nu\rangle$ is at most* $1/2 + \exp\big(-n^{1-o(1)}/\max(\log s, \sqrt{n})\big)$.

(See Theorem 4.2 for a more precise tradeoff between depth and fidelity.) In particular, Theorem 1.4 implies that mostly classical circuits of depth $o(\log n)$ require size at least $\exp(n^{1-o(1)})$ to construct approximate n-nekomata, essentially matching the $\exp(\tilde{O}(n))$ size upper bound from Theorem 1.1 and Remark 1.3. This lower bound does not contradict the $\exp(n^{o(1)})$ size upper bounds of depth $\omega(1)$ from Corollary 1.2, because our reductions between parity, fanout, and constructing nekomata do not in general map mostly classical circuits to mostly classical circuits. Since the identity circuit is mostly classical, the upper bound on the fidelity of $C|0\dots0\rangle$ and $|\nu\rangle$ in Theorem 1.4 is tight up to the value being exponentiated. Finally, if we also allow r-qubit parity and fanout gates in mostly classical circuits – a natural model for small values of r, in light of the upper bounds from Corollary 1.2 – then a trivial generalization of our proof of Theorem 1.4 implies that an identical statement holds for circuits of depth $o(\log_{\max(r,2)} n)$.

To prove Theorem 1.4, it suffices to prove that the Hamming weight of a standard-basis measurement of any n qubits of $C|0\dots0\rangle$ is concentrated around some value. We use the fact that standard-basis measurements commute with generalized Toffoli gates, and, after some preparation, apply a concentration inequality of Gavinsky, Lovett, Saks and Srinivasan [6].

1.2.5 Lower Bounds for Arbitrary QAC Circuits of Low Size and Depth

Call the first n qubits of an n-nekomata $\frac{1}{\sqrt{2}}\sum_{b=0}^{1}|b^n, \psi_b\rangle$ the *targets* of that nekomata.

▶ **Theorem 1.5.** *There is a universal constant $c > 0$ such that the following holds. Let C be a depth-d QAC circuit acting on any number of qubits, and let $|\nu\rangle$ be an n-nekomata such that at most $cn/(d+1)$ multi-qubit gates in C act on the targets of $|\nu\rangle$. Then the fidelity of $C|0\dots0\rangle$ and $|\nu\rangle$ is at most* $1/2 + \exp(-\Omega(n/(d+1)))$.

▶ **Corollary 1.6.** *Let c be the constant from Theorem 1.5. Let C be a depth-d QAC circuit acting on any number of qubits, and assume that, collectively, the first n of these qubits are acted on by at most $cn/(d+1)$ multi-qubit gates in C. Then for all states $|\psi\rangle$,*

- *for $|\phi_\oplus\rangle = |0, +^{n-1}\rangle$, the fidelity of $C|\phi_\oplus, 0\ldots 0\rangle$ and $U_\oplus|\phi_\oplus\rangle \otimes |\psi\rangle$ is at most $1/2 + \exp(-\Omega(n/(d+1)))$;*
- *for $|\phi_F\rangle = |+, 0^{n-1}\rangle$, the fidelity of $C|\phi_F, 0\ldots 0\rangle$ and $U_F|\phi_F\rangle \otimes |\psi\rangle$ is at most $1/2 + \exp(-\Omega(n/(d+1)))$;*
- *the fidelity of $C|0\ldots 0\rangle$ and $|🐱_n, \psi\rangle$ is at most $1/2 + \exp(-\Omega(n/(d+1)))$.*

(Perhaps surprisingly, a sharp "phase change" near the $cn/(d+1)$ threshold is in fact inherent to our proof. The $+1$ in $\exp(-\Omega(n/(d+1)))$ is necessary when $C = H$ and $|\nu\rangle = |+\rangle := \frac{|0\rangle+|1\rangle}{\sqrt{2}}$.) For example, Theorem 1.5 implies that a depth-2 QAC circuit constructing an approximate n-nekomata must have at least $\Omega(n)$ multi-qubit gates acting on the targets of that nekomata. This $\Omega(n)$ lower bound is tight, because Theorem 1.1 says that depth-2 QAC circuits can construct approximate n-nekomata, and a depth-d QAC circuit can have at most nd multi-qubit gates acting on any given set of n qubits. Similarly, Corollary 1.6 implies that depth-7 QAC circuits approximating n-qubit parity, fanout, or restricted fanout require at least $\Omega(n)$ multi-qubit gates acting on the n "input" qubits, and this $\Omega(n)$ lower bound is tight as well by Corollary 1.2.

Theorem 1.5 also implies that the *total* number of multi-qubit gates, a.k.a. the size, of a depth-d QAC circuit constructing an approximate n-nekomata must be at least $\Omega(n/(d+1))$. When d is $o(\log n)$, this lower bound is disappointingly far from the upper bounds of Theorem 1.1 and Corollary 1.2. However, Green et al. [8] observed that for some $d = \Theta(\log n)$, a depth-d QAC circuit of size $O(n)$ can construct an n-nekomata (specifically, the n-qubit cat state), so for this value of d our $\Omega(n/d)$ size lower bound is tight to within a logarithmic factor. Similarly, for some $d = \Theta(\log n)$, the minimum size of a depth-d QAC circuit that approximates n-qubit parity, fanout, or restricted fanout is between $\Omega(n/\log n)$ and $O(n)$, by Corollary 1.6 and upper bounds of Green et al.

If a QAC circuit has size $s \leq o(\sqrt{n})$ then its depth d satisfies $d \leq s \leq o(\sqrt{n})$, so $s \leq o(\sqrt{n}) \leq o(n/(d+1))$. It follows from Theorem 1.5 and Corollary 1.6 that QAC circuits of *arbitrary* depth require size at least $\Omega(\sqrt{n})$ to construct approximate n-nekomata, or to approximately compute n-qubit parity, fanout, or restricted fanout.[2]

Finally, we remark that Theorem 1.5 is actually a special case of a more general result, Theorem 5.2, about states $|\psi\rangle$ such that for some orthogonal projections[3] $Q_1, \ldots, Q_n$ on arbitrary numbers of qubits, $\langle\psi|\left(\bigotimes_{j=1}^n Q_j \otimes I\right)|\psi\rangle = \langle\psi|\left(\bigotimes_{j=1}^n (I - Q_j) \otimes I\right)|\psi\rangle = 1/2$. (For example, n-nekomata satisfy this criterion with $Q_j = |0\rangle\langle 0|$ for all j.) We will comment on this generalization of Theorem 1.5 again in Section 1.2.7.

1.2.6 A Normal Form for Quantum Circuits

Integral to our proof of Theorem 1.5 is a certain normal form for QAC circuits, which may be of independent interest since the standard quantum circuit model is that of QAC circuits whose gates have maximum arity 2. Here we give the underlying intuition, by way of analogy with well-known facts from classical circuit complexity. If we define AC circuits as consisting only of AND and NOT gates, then it cannot in general be assumed that the NOT gates are

[2] In the full paper we generalize this argument to hold for the number of multi-qubit gates acting on the target/input qubits.

[3] I.e. $Q_j = Q_j^2 = Q_j^\dagger$ for all j.

all adjacent to the inputs. However, by DeMorgan's laws we may equivalently allow OR gates in AC circuits as well, and then it *can* be assumed that the NOT gates are all adjacent to the inputs.[4] Similarly, we introduce a certain further generalization of generalized Toffoli gates which allows us to assume that the one-qubit gates in a QAC circuit are all adjacent to the input.

1.2.7 Depth-2 Lower Bounds

▶ **Theorem 1.7.** *Let C be a depth-2 QAC circuit of arbitrary size, acting on any number of qubits. Then for all states $|\psi\rangle$,*

(i) *for $|\phi_\oplus\rangle = |0, +^{n-1}\rangle$, the fidelity of $C|\phi_\oplus, 0 \dots 0\rangle$ and $U_\oplus|\phi_\oplus\rangle \otimes |\psi\rangle$ is at most $1/2 + \exp(-\Omega(n))$;*

(ii) *for $|\phi_F\rangle = |+, 0^{n-1}\rangle$, the fidelity of $C|\phi_F, 0 \dots 0\rangle$ and $U_F|\phi_F\rangle \otimes |\psi\rangle$ is at most $1/2 + \exp(-\Omega(n))$;*

(iii) *the fidelity of $C|0 \dots 0\rangle$ and $|\text{😸}_n, \psi\rangle$ is at most $1/2 + \exp(-\Omega(n))$.*

Our proof of Theorem 1.7 gives a multiplicative constant of roughly $1/10^{60000}$ implicit in the $\Omega(\cdot)$ notation in the above inequalities, which makes them trivial for small values of n. If n is sufficiently large however, then Theorem 1.7 implies that depth-2 QAC circuits cannot approximate n-qubit parity, fanout, or restricted fanout, or approximately construct the n-qubit cat state, even if these approximations are not required to be clean. Still taking n to be sufficiently large, this improves on the previously mentioned result of Padé et al. [13] that depth-2 QAC circuits cannot cleanly compute parity exactly on four or more qubits.

Theorem 1.7 and Corollary 1.2 imply that for all sufficiently large n, the minimum depth of a QAC circuit approximating n-qubit parity is between 3 and 7 inclusive, and likewise for fanout, restricted fanout, and constructing the cat state. By Theorem 1.1 there *is* a depth-2 QAC circuit that constructs an approximate n-nekomata for all n, so any proof of Theorem 1.7 must use some property of $|\text{😸}_n, \psi\rangle$ that does not hold for an arbitrary n-nekomata. Ours uses a property similar to the fact that if we measure some of the qubits in the "B" register of $|\text{😸}_n\rangle_A \otimes |\psi\rangle_B$ in an arbitrary basis, then the resulting state in registers A and B is still an n-nekomata.

Our proof of Theorem 1.7 mostly uses different techniques than those of Padé et al. An exception is the observation, of which we use a generalization, that if we define a "generalized Z gate" on any number of qubits by $Z = I - 2|1 \dots 1\rangle\langle 1 \dots 1|$ then $Z|0, \phi\rangle = |0, \phi\rangle$ and $Z|1, \phi\rangle = |1\rangle \otimes Z|\phi\rangle$ for all states $|\phi\rangle$. We also incorporate a variant of the proof given by Bene Watts, Kothari, Schaeffer and Tal [2, Theorem 16] that there is no QNC circuit (QAC circuit whose gates have maximum arity 2) of depth $o(\log n)$ that maps $|0 \dots 0\rangle$ to $|\text{😸}_n, 0 \dots 0\rangle$: Using a "light cone" argument they prove that out of any n output qubits, there are at least two whose standard-basis measurements would be independent, but the standard-basis measurements of any two qubits in $|\text{😸}_n\rangle$ are dependent.

Our proof of Theorem 1.7 goes roughly as follows. If there are only $o(n)$ multi-qubit gates acting on the n targets of $|\text{😸}_n, \psi\rangle$ then the result follows from Theorem 1.5. Otherwise, out of the multi-qubit gates acting on the targets, the average gate acts on $O(1)$ targets, as would be the case in a QNC circuit. Using a variant of a light cone argument, we choose $\Theta(n)$ pairwise disjoint sets of qubits on which to define orthogonal projections, and apply the generalization of Theorem 1.5 that was mentioned at the end of Section 1.2.5.

---

[4] Invoking this assumption results in a constant-factor blowup in size and no blowup in depth, where (as is customary) we do not count NOT gates toward the size or depth of AC circuits.

1.3 Organization

In Section 1.4 we introduce some miscellaneous notation and definitions. In Section 2 we give multiple equivalent characterizations of QAC circuits, including the previously mentioned normal form, and introduce some related definitions which we will use in more general contexts as well. In Section 3 we give reductions between parity, fanout, restricted fanout, and constructing nekomata; we also use these reductions to prove that the $d \geq 11$ case of Corollary 1.2 follows from Theorem 1.1 (the $d < 11$ case is proved in the full paper), that Corollary 1.6 follows from Theorem 1.5, and that Theorems 1.7(i) and 1.7(ii) follow from Theorem 1.7(iii). In Section 4 we prove our upper and lower bounds for constructing approximate nekomata in mostly classical circuits, Theorems 1.1 and 1.4. In Section 5 we prove our other main results, Theorem 1.5 and Theorem 1.7(iii). Sections 3 to 5 may be read in any order.

1.4 Preliminaries

We write log and ln to denote the logarithms base 2 and e respectively, and $(x_j)_j$ to denote the tuple of all x_j for j in some implicit index set. Also let $[n] = \{1, \ldots, n\}$ and $\|\psi\| = \sqrt{\psi^*\psi}$, i.e. $\|\cdot\|$ denotes the 2-norm. Anything written as $\langle\cdot|$ or $|\cdot\rangle$ is implicitly unit-length. "Proof sketch" environments are replaced by complete proofs in the full paper.

Orthogonal projections are linear transformations Q such that $Q = Q^2 = Q^\dagger$. For an orthogonal projection Q and a state $|\varphi\rangle$, we call $\langle\varphi|Q|\varphi\rangle$ "the probability that $|\varphi\rangle$ measures to Q". If $Q = |\psi\rangle\langle\psi|$ then we also call this "the probability that $|\varphi\rangle$ measures to $|\psi\rangle$", and it equals $|\langle\psi|\varphi\rangle|^2$, a.k.a. the *fidelity* of $|\varphi\rangle$ and $|\psi\rangle$. More generally, if Q is an orthogonal projection on some Hilbert space $\mathcal{H}$ then we call $\langle\varphi|(Q \otimes I)|\varphi\rangle$ "the probability that the $\mathcal{H}$ qubits of $|\varphi\rangle$ measure to Q".

We use standard notation for the Hadamard basis states $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, Hadamard gate $H = |+\rangle\langle 0| + |-\rangle\langle 1|$, and NOT gate $X = |0\rangle\langle 1| + |1\rangle\langle 0|$. We write I to denote the identity transformation, $I_{\mathcal{H}}$ for the identity on the Hilbert space $\mathcal{H}$, and I_n for the identity on some n-qubit Hilbert space.

To be thorough, we remind the reader that an n-*nekomata* is a state with n qubits (called *targets*) that measure to 0^n and to 1^n each with probability $1/2$, or equivalently a state of the form $\frac{1}{\sqrt{2}}\sum_{b=0}^{1} |b^n, \psi_b\rangle$ for some states $|\psi_0\rangle, |\psi_1\rangle$ on any number of qubits. For example, the n-qubit *cat state* is the state $|🐱_n\rangle = (|0^n\rangle + |1^n\rangle)/\sqrt{2}$.

2 QAC Circuits

Consider a quantum circuit C, written as $C = L_d M_d \cdots L_1 M_1 L_0$ such that each L_k consists only of one-qubit gates and each M_k is a layer (tensor product) of multi-qubit gates. We may assume that each L_k is a single layer as well, because the product of one-qubit gates is also a one-qubit gate. Define the *size* of C to be the number of multi-qubit gates in C, the *depth* of C to be the number of layers of multi-qubit gates in C (in this case, d), and the *topology* of C to be the set of pairs (S, k) such that S equals the support of some gate in M_k, where the *support* of a gate is the set of qubits acted on by that gate. Note that the topology of C encodes its depth, size, and more generally the number of multi-qubit gates acting on any given set of qubits.

Recall that QAC circuits are quantum circuits with arbitrary one-qubit gates and generalized Toffoli gates of arbitrary arity, where $(n+1)$-ary generalized Toffoli gates are defined by $|x, b\rangle \mapsto |x, b \oplus \bigwedge_{j=1}^{n} x_j\rangle$ for $x = (x_1, \ldots, x_n) \in \{0,1\}^n, b \in \{0,1\}$. Define an $(n+1)$-ary OR

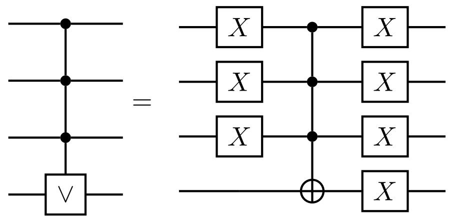

Figure 1 The multi-qubit gates on the left and right are OR and generalized Toffoli gates respectively, whose target qubits are on the bottom wire.

gate by $|x, b\rangle \mapsto |x, b \oplus \bigvee_j x_j\rangle$ for $x \in \{0,1\}^n, b \in \{0,1\}$, and call the qubit corresponding to b in these definitions the *target* qubit of the gate. By the construction of an OR gate from a generalized Toffoli gate and NOT gates in Figure 1, we may add OR gates to the set of allowed gates when defining QAC circuits, without changing the set of topologies of QAC circuits computing any given unitary transformation.

For a state $|\theta\rangle$ let $R_{|\theta\rangle} = R_\theta = I - 2|\theta\rangle\langle\theta|$ (the R stands for "reflection"). Let a *mono-product state* be a tensor product of any number of one-qubit states. When $|\theta\rangle$ is a mono-product state we call R_θ an $R_\otimes$ *gate*. For example, an $(n+1)$-qubit generalized Toffoli gate equals $R_{|1^n,-\rangle}$, because it acts on the basis $\{|0\rangle, |1\rangle\}^{\otimes n} \otimes \{|+\rangle, |-\rangle\}$ by multiplying $|1^n, -\rangle$ by -1 and leaving all other states in this basis unchanged.

Consider an $(n+1)$-qubit mono-product state $|\theta\rangle$, and let L be a layer of one-qubit gates such that $|\theta\rangle = L|1^n, -\rangle$. Then,

$$R_\theta = I - 2|\theta\rangle\langle\theta| = I - 2L|1^n, -\rangle\langle 1^n, -|L^\dagger = L(I - 2|1^n, -\rangle\langle 1^n, -|)L^\dagger = LR_{|1^n,-\rangle}L^\dagger, \quad (1)$$

i.e. R_θ equals the conjugation of a generalized Toffoli gate by a layer of one-qubit gates. (Fang et al. [5] observed Equation (1) in the case where $|\theta\rangle = |1^{n+1}\rangle$ and $L = I_n \otimes H$.) Therefore, similarly to the above, we may add arbitrary $R_\otimes$ gates to the set of allowed gates when defining QAC circuits.

In fact, a stronger statement holds. Let a QAC circuit be in $R_\otimes$ *normal form* if it can be written as CL such that C consists only of multi-qubit $R_\otimes$ gates and L is a layer of single-qubit gates. We will use the following in Section 5:

▶ **Proposition 2.1.** *Every QAC circuit computes the same unitary transformation as a circuit in $R_\otimes$ normal form with the same topology.*

Proof sketch. The proof is by induction on the depth of the circuit. Let $C = LMD$ be the circuit, where L (resp. M) is the top layer of one-qubit (resp. multi-qubit) gates in C. Then write $C = (LML^\dagger)(LD)$, and apply Equation (1) and the inductive hypothesis. ◀

3 Reductions to and from Constructing Nekomata

In Section 3.1 we define the problems mentioned in the following theorem, and in Sections 3.1 and 3.2 we prove the second and first paragraphs of this theorem respectively:

▶ **Theorem 3.1.** *For all $\varepsilon \geq 0$, if there is a QAC circuit of size s, depth d, and number of qubits acted on a that solves the $(1-\varepsilon)$-approximate n-nekomata problem, then there is a QAC circuit of size $O(s+n)$, depth $4d+3$, and number of ancillae a that solves the $(1 - O(\varepsilon))$-approximate $(n+1)$-qubit clean parity problem.*

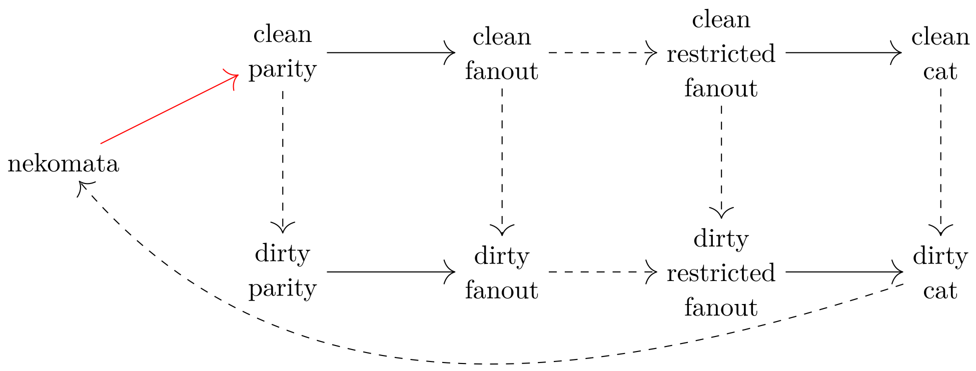

Figure 2 A visualization of Theorem 3.1; see the theorem statement for the meaning of the arrows.

For all $0 \le p \le 1$ and every non-red[5] *arrow from a problem P to a problem Q in Figure 2, if a QAC circuit C solves p-approximate, n-qubit P then there is a QAC circuit with the same topology as C that solves p-approximate, n-qubit Q. (If Q is the nekomata problem then substitute "n-nekomata" for "n-qubit nekomata" here.) Furthermore, if this arrow is dashed then C itself solves p-approximate, n-qubit Q.*

Then, using Theorem 3.1, in Section 3.3 we prove that the $d \ge 11$ case of Corollary 1.2 follows from Theorem 1.1. It is easy to prove Corollary 1.6 assuming Theorem 1.5, and to prove Theorems 1.7(i) and 1.7(ii) assuming Theorem 1.7(iii), using reasoning similar to that in the proof of Theorem 3.1.

3.1 Problem Definitions and Most Reductions

We omit certain elementary parts of the proof, which may be found in the full paper. Recall that we define the *phase-dependent fidelity* of states $|\varphi\rangle$ and $|\psi\rangle$ to be $1 - \||\varphi\rangle - |\psi\rangle\|^2$. This quantity is at most the fidelity of $|\varphi\rangle$ and $|\psi\rangle$, because

$$|\langle\psi|\varphi\rangle|^2 \ge \left(\frac{\langle\psi|\varphi\rangle + \langle\varphi|\psi\rangle}{2}\right)^2 = \left(1 - \frac{\||\varphi\rangle - |\psi\rangle\|^2}{2}\right)^2 \ge 1 - \||\varphi\rangle - |\psi\rangle\|^2. \tag{2}$$

▶ Remark. If $\langle\varphi|\psi\rangle$ is a real number close to 1, say $\langle\varphi|\psi\rangle = 1 - \varepsilon$, then $|\varphi\rangle$ and $|\psi\rangle$ have fidelity $1 - 2\varepsilon + \varepsilon^2$ and a nearly identical phase-dependent fidelity of $1 - 2\varepsilon$. On the other hand, if the phases of $|\varphi\rangle$ and $|\psi\rangle$ differ, then these states may have low phase-dependent fidelity even if their fidelity is close to 1.

The following two definitions are with respect to an arbitrary unitary transformation U on n qubits:

▶ **Problem 3.2** (p-approximate Clean U)**.** Construct a circuit C on at least n qubits such that for all n-qubit states $|\phi\rangle$, the phase-dependent fidelity of $C|\phi, 0 \ldots 0\rangle$ and $U|\phi\rangle \otimes |0 \ldots 0\rangle$ is at least p.

[5] Only the arrow from "nekomata" to "clean parity" is red.

▶ **Problem 3.3** (p-approximate Dirty U)**.** Construct a circuit C on at least n qubits such that for all n-qubit states $|\phi\rangle$, the first n qubits of $C|\phi, 0\dots0\rangle$ measure to $U|\phi\rangle$ with probability at least p.

Given n, recall that the unitary transformations for n-qubit parity and fanout are defined respectively by $U_\oplus|b,x\rangle = |b \oplus \bigoplus_{j=1}^{n-1} x_j, x\rangle$ and $U_F|b,x\rangle = |b, x_1 \oplus b, \dots, x_{n-1} \oplus b\rangle$ for $b \in \{0,1\}, x = (x_1,\dots,x_{n-1}) \in \{0,1\}^{n-1}$. Define the clean and dirty versions of approximating n-qubit parity and fanout as instances of Problems 3.2 and 3.3 with respect to $U_\oplus$ and U_F.

Green et al. [8] proved that $H^{\otimes n} U_\oplus H^{\otimes n} = U_F$. In the full paper we prove that if a circuit C computes p-approximate, n-qubit clean (resp. dirty) parity, then the circuit $(H^{\otimes n} \otimes I)C(H^{\otimes n} \otimes I)$ computes p-approximate, n-qubit clean (resp. dirty) fanout.

▶ **Problem 3.4** (p-approximate Clean Restricted Fanout)**.** Construct a circuit C on at least n qubits such that for all one-qubit states $|\phi\rangle$, the phase-dependent fidelity of $C|\phi, 0^{n-1}, 0\dots0\rangle$ and $U_F|\phi, 0^{n-1}\rangle \otimes |0\dots0\rangle$ is at least p.

▶ **Problem 3.5** (p-approximate Dirty Restricted Fanout)**.** Construct a circuit C on at least n qubits such that for all one-qubit states $|\phi\rangle$, the first n qubits of $C|\phi, 0^{n-1}, 0\dots0\rangle$ measure to $U_F|\phi, 0^{n-1}\rangle$ with probability at least p.

▶ **Problem 3.6** (p-approximate Clean $|\text{🐱}_n\rangle$)**.** Construct a circuit C on at least n qubits such that the phase-dependent fidelity of $C|0\dots0\rangle$ and $|\text{🐱}_n, 0\dots0\rangle$ is at least p.

▶ **Problem 3.7** (p-approximate Dirty $|\text{🐱}_n\rangle$)**.** Construct a circuit C on at least n qubits such that the first n qubits of $C|0\dots0\rangle$ measure to $|\text{🐱}_n\rangle$ with probability at least p.

▶ **Problem 3.8** (p-approximate n-nekomata)**.** Construct a circuit C such that for some n-nekomata $|\nu\rangle$, the fidelity of $C|0\dots0\rangle$ and $|\nu\rangle$ is at least p.

3.2 Reducing Clean Parity to Constructing Nekomata

Let C be a circuit on a qubits such that $|\nu\rangle \coloneqq C|0^a\rangle$ approximates an n-nekomata. A circuit for approximate $(n+1)$-qubit clean parity is shown in Figure 3, where the top n wires acted on by each of the C and $C^\dagger$ subcircuits correspond to the targets of $|\nu\rangle$. Within each dotted rectangle is a layer of n copies of $R_{|11\rangle}$, the i'th of which acts on the wires corresponding to the i'th input qubit and the i'th target of $|\nu\rangle$ for $i \in [n]$. The gate $R_{|11\rangle}$ is better known as a controlled Z gate, and acts as $|xy\rangle \mapsto (-1)^{xy}|xy\rangle$ for $x, y \in \{0,1\}$. In the middle of the circuit is an OR gate (recall Figure 1). The correctness of this circuit is proved in the full paper.

3.3 Proof of Corollary 1.2 ($d \geq 11$) Assuming Theorem 1.1

▶ **Lemma 3.9** (essentially Green et al. [8])**.** *For all $m \geq 2$ there is a quantum circuit of depth $\lceil \log_m n \rceil$ and size at most $n-1$, consisting only of restricted fanout gates of arity at most m, that computes n-qubit restricted fanout exactly using no ancillae.*

Proof. By linearity it suffices to consider input states of the form $|b, 0^{n-1}\rangle$ for $b \in \{0,1\}$. The proof is by induction on $d = \lceil \log_m n \rceil$, for a fixed value of m. Note that $d - 1 < \log_m n \leq d$, so $m^{d-1} < n \leq m^d$. If $d = 0$ then $n = 1$ and the identity circuit suffices. If $d > 0$ then by induction we can map $|b, 0^{n-1}\rangle$ to $|b^{m^{d-1}}, 0^{n-m^{d-1}}\rangle$ in depth $d-1$ and size at most $m^{d-1} - 1$. Let $n_1, \dots, n_{m^{d-1}} \in [m]$ be such that $\sum_i n_i = n$, and compute $\bigotimes_i U_F|b, 0^{n_i - 1}\rangle = |b^n\rangle$. Since $\bigotimes_i U_F$ has size at most $n - m^{d-1}$ (omitting one-qubit gates), the total size of the circuit is at most $(n - m^{d-1}) + (m^{d-1} - 1) = n - 1$. ◀

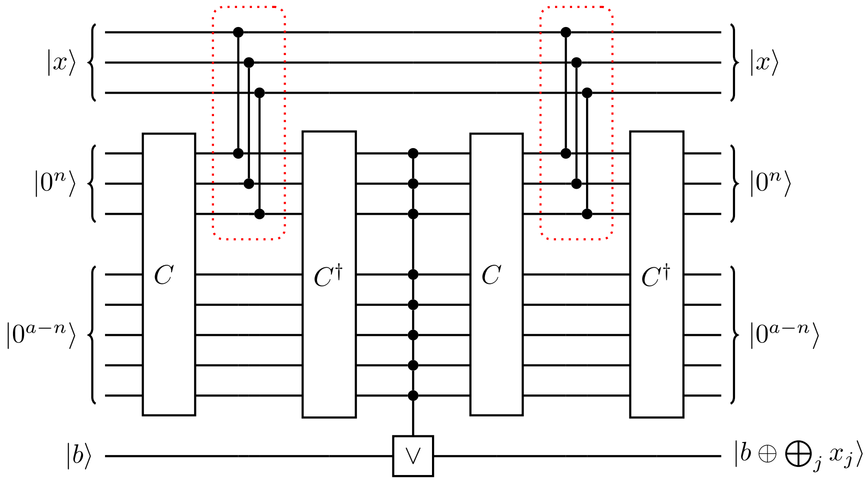

Figure 3 A circuit for parity, assuming C constructs an n-nekomata. See the surrounding text for further explanation.

Given Theorems 1.1 and 3.1 and Lemma 3.9, the proof of the $d \geq 11$ case of Corollary 1.2 is elementary (if slightly tedious) and may be found in the full paper.

4 Tight Bounds for Constructing Approximate Nekomata in "Mostly Classical" Circuits

Call a QAC circuit *purely classical* if it consists only of generalized Toffoli gates (including NOT gates, which are generalized Toffoli gates on one qubit). Call a QAC circuit *mostly classical* if it can be written as CL such that C is purely classical and L is a layer of $R_\otimes$ gates; by Equation (1) this is equivalent to the definition from Section 1.2.4. Call a mostly classical QAC circuit *nice* if it can be written as CL in this way such that every multi-qubit gate R_θ in L satisfies $|\langle 0 \dots 0|\theta\rangle|^2 \leq 1/4$. (The niceness condition will allow us to express certain quantities as convex combinations in a convenient way, by ensuring that the coefficients in these convex combinations are between 0 and 1.) We prove the following generalizations of Theorems 1.1 and 1.4 respectively:

▶ **Theorem 4.1.** *For all $2 \leq d \leq \log n$ and $\varepsilon > 0$ there exists a nice, mostly classical, depth-d QAC circuit C of size and number of qubits acted on $\exp(O(n2^{-d}\log(n2^{-d}/\varepsilon))) + O(n)$ such that $C|0\dots0\rangle$ has fidelity at least $1-\varepsilon$ with some n-nekomata.*

▶ **Theorem 4.2.** *Let C be a mostly classical circuit of size s and depth d.*

(i) *The fidelity of $C|0\dots0\rangle$ and any n-nekomata is at most*

$$\frac{1}{2} + \exp\left(-\Omega\left(\frac{n/(4^d \log n)}{\max\left(\log s, \sqrt{n/(4^d \log n)}\right)}\right)\right).$$

(ii) *If C is nice, then the fidelity of $C|0\dots0\rangle$ and any n-nekomata is at most*

$$\frac{1}{2} + \exp\left(-\Omega\left(\frac{n/2^d}{\max\left(\log s, \sqrt{n/2^d}\right)}\right)\right).$$

Theorems 4.1 and 4.2(ii) imply that for $d \geq 2$, the minimum size of a nice, mostly classical, depth-d QAC circuit that "constructs an approximate n-nekomata" (i.e. maps $|0\dots0\rangle$ to a state that has fidelity at least $3/4$ with some n-nekomata) is between $\exp(\Omega(n/2^d))$ and $\exp(\tilde{O}(n/2^d))+O(n)$. We prove the $d > 2$ case of Theorem 4.1 solely for the sake of comparison with Theorem 4.2(ii), as Theorem 4.1 gives a weaker upper bound than Corollary 1.2 when $\omega(1) \leq d \leq o(\log n)$. Theorem 4.2 makes a stronger statement about nice circuits than about non-nice circuits, since $a/\max(\log s, \sqrt{a}) = \min(a/\log s, \sqrt{a})$ for all $a > 0$.

In Section 4.1 we make some general observations about mostly classical circuits and "approximate nekomata", including observations common to the proofs of Theorems 4.1 and 4.2. In Section 4.2 we prove Theorem 4.1, and in Section 4.3 we prove Theorem 4.2(ii). We prove Theorem 4.2(i) in the full paper; its proof has a similar high-level idea to that of Theorem 4.2(ii), and is much more complicated.

4.1 Reduction to a Classical Sampling Problem

Collectively, the following observations reduce proving Theorems 4.1 and 4.2 to proving upper and lower bounds respectively for a certain type of sampling problem. This sampling problem can be succinctly characterized in purely classical and probabilistic terms, with only a transient reference to quantum circuits. Claims made in this subsection are proved in the full paper.

Recall that nekomata can be defined as states for which a standard-basis measurement of the targets is distributed in a certain way. The following two lemmas make similar statements about "approximate nekomata", and are used to prove Theorems 4.1 and 4.2 respectively:

▶ **Lemma 4.3.** *Let $|\varphi\rangle$ be a state with n "target" qubits that measure to all-zeros with probability exactly $1/2$ and all-ones with probability at least $1/2 - (2/3)\varepsilon$. Then there exists an n-nekomata $|\nu\rangle$ such that $|\langle\nu|\varphi\rangle|^2 \geq 1 - \varepsilon$.*

▶ **Lemma 4.4.** *Let $|\varphi\rangle$ be a state with n "target" qubits that measure to all-zeros with probability p and all-ones with probability q. Then $|\langle\nu|\varphi\rangle|^2 \leq 1/2 + \sqrt{\min(p,q)}$ for all n-nekomata $|\nu\rangle$ with the same targets as $|\varphi\rangle$.*

Consider a mostly classical circuit, written as CL such that C is purely classical and L is a layer of $R_\otimes$ gates. A standard-basis measurement of designated "target" qubits of $CL|0\dots0\rangle$ is distributed identically to an appropriate marginal distribution of a standard-basis measurement of *all* qubits of $CL|0\dots0\rangle$. It is easy to see that standard-basis measurements commute with generalized Toffoli gates, so we may first measure $L|0\dots0\rangle$ in the standard basis and then apply C to the result.

Finally, the following is straightforward to verify:

▶ **Lemma 4.5.** *Let $(|\theta_j\rangle)_j$ be one-qubit states, and let $p_j = |\langle 1|\theta_j\rangle|^2$ for all j. A standard-basis measurement of $R_{\bigotimes_j |\theta_j\rangle}|0\dots0\rangle$ outputs all-zeros with probability $\left(1 - 2\prod_j(1-p_j)\right)^2$, and any other boolean string $(y_j)_j$ with probability $4\prod_j(1-p_j)P(\mathrm{Bernoulli}(p_j) = y_j)$.*

For mostly classical circuits that are nice, the following is a more convenient characterization of this distribution:

▶ **Corollary 4.6.** *If $\prod_j(1-p_j) \leq 1/4$ then the distribution from Lemma 4.5 is a convex combination of all-zeros with probability $1 - 4\prod_j(1-p_j)$ and $(\mathrm{Bernoulli}(p_j))_j$ with probability $4\prod_j(1-p_j)$, where the $\mathrm{Bernoulli}(p_j)$ random variables are all independent.*

4.2 Proof of Theorem 4.1

Here we prove the depth-2 case of Theorem 4.1 (which is sufficient for all of our applications of Theorem 4.1); the generalization to depths greater than 2 is handled in the full paper.

▶ **Reminder** (depth-2 case of Theorem 4.1). *For all $\varepsilon > 0$ there exists a nice, mostly classical, depth-2 QAC circuit C of size and number of qubits acted on $\exp(O(n \log(n/\varepsilon)))$ such that $C|0\dots0\rangle$ has fidelity at least $1-\varepsilon$ with some n-nekomata.*

Proof. Let $M \in \mathbb{N}$ and $\delta \in (0,1)$ be parameters to be chosen later.[6] The circuit acts on $n(M+1)$ qubits, all initialized to $|0\rangle$, and arranged in a grid of dimensions $n \times (M+1)$ (Figure 4). Designate one column as the "target" column, and call the qubits in the M other columns "ancillae". First, to each ancilla column, apply $R_{\left(\sqrt{\delta}|0\rangle+\sqrt{1-\delta}|1\rangle\right)^{\otimes n}}$. Second, to each row, apply an $(M+1)$-qubit OR gate whose target qubit is in the target column. (A layer of OR gates is a depth-1 purely classical circuit, by the construction in Figure 1.)

All measurements described below are with respect to the state on the ancillae between the first and second layers of the above circuit. By Lemma 4.3 it suffices to choose M and δ such that if we measure the ancillae in the standard basis, then with probability exactly $1/2$ all of the ancillae measure to 0, and with probability at least $1/2 - (2/3)\varepsilon$ at least one ancilla in each row measures to 1. We now choose δ in terms of M such that the ancillae measure to all-zeros with probability $1/2$. By Lemma 4.5 and the independence of measurements of different columns, it suffices to ensure that $(1-2\delta^n)^{2M} = 1/2$. Choose $\delta \in (0, (1/2)^{1/n})$ that satisfies this equation.

Let $\varepsilon' = (2/3)\varepsilon$. Below we will choose M such that the probability that there exists an ancilla column measuring to neither all-zeros nor all-ones is at most ε'. Equivalently, with probability at least $1-\varepsilon'$, every ancilla column measures to either all-zeros or all-ones. Since the ancillae measure to all-zeros with probability $1/2$, it follows that with probability at least $1/2 - \varepsilon'$, every ancilla column measures to all-zeros or all-ones *and* at least one ancilla column measures to all-ones. Therefore the probability is at least $1/2-\varepsilon'$ that at least one ancilla in every row measures to 1, as desired.

By Lemma 4.5 and a union bound, the probability that there exists an ancilla column measuring to neither all-zeros nor all-ones is at most

$$M(1-(1-2\delta^n)^2 - 4\delta^n(1-\delta)^n) = 4M\delta^n(1-\delta^n-(1-\delta)^n) \le 4Mn\delta^{n+1}.$$

Since $1/2 = (1-2\delta^n)^{2M} \le \exp(-4\delta^n M)$, it follows that $\delta^n \le \ln(2)/4M$, so

$$4Mn\delta^{n+1} \le 4Mn(\ln(2)/4M)^{1+1/n} = \ln(2)n(\ln(2)/4M)^{1/n}.$$

To make this bound at most ε', let $M = \lceil(\ln(2)/4) \cdot (\ln(2)n/\varepsilon')^n\rceil \le \exp(O(n\log(n/\varepsilon)))$. Finally, the circuit is nice because $\delta^n \le \ln(2)/4M \le \ln(2)/4 < 1/4$. ◀

[6] Ultimately we will let $M = \exp(\Theta(n\log(n/\varepsilon)))$ and $\delta = \Theta(\varepsilon/n)$.

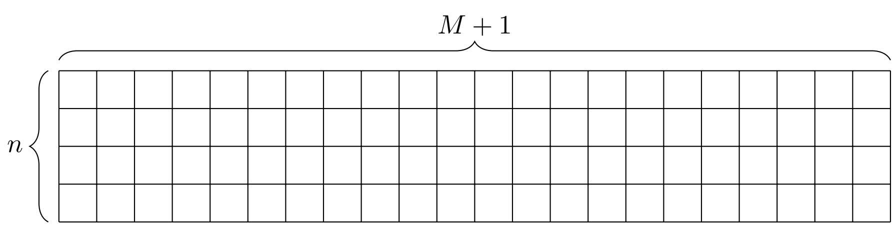

Figure 4 The layout of the circuit.

4.3 Proof of Theorem 4.2(ii)

We use the following concentration inequality of Gavinsky, Lovett, Saks and Srinivasan [6]:

▶ **Definition 4.7** ([6]). *Call a random string $(Y_1, \ldots, Y_n) \in \{0,1\}^n$ a* read-r family *if there exist $m \in \mathbb{N}$, independent random variables $X_1, \ldots, X_m$, sets $S_1, \ldots, S_n \subseteq [m]$ such that $|\{j \mid i \in S_j\}| \leq r$ for all $i \in [m]$, and functions $f_1, \ldots, f_n$ such that $Y_j = f_j((X_i)_{i \in S_j})$ for all $j \in [n]$.*

▶ **Theorem 4.8** ([6]). *Let $(Y_1, \ldots, Y_n)$ be a read-r family, and let $\mu = \mathbb{E}\left[\sum_{j=1}^n Y_j\right]$. Then for all $\varepsilon \geq 0$,*

$$P(Y_1 + \cdots + Y_n \geq \mu + \varepsilon n) \leq \exp(-2\varepsilon^2 n/r),$$
$$P(Y_1 + \cdots + Y_n \leq \mu - \varepsilon n) \leq \exp(-2\varepsilon^2 n/r).$$

▶ Remark. For example, if $r = 1$ then $Y_1, \ldots, Y_n$ are all independent and so Theorem 4.8 recovers a well-known Chernoff bound for sums of independent Bernoulli random variables. Theorem 4.8 also recovers this Chernoff bound when $n = rm$ and $Y_j = X_{\lceil j/r \rceil}$ for all j [6].

Consider a string x of independent Bernoulli random variables. If G is a generalized Toffoli gate then $G|x\rangle$ is a read-2 family, because for all i the i'th bit of x can only influence the i'th and target bits of $G|x\rangle$. More generally, if G is a generalized Toffoli gate and L_1, L_2 are layers of NOT gates acting on subsets of the support of G, then $L_1 G L_2 |x\rangle$ is a read-2 family. Even more generally, it follows by induction that if C is a depth-d purely classical circuit then $C|x\rangle$ is a read-2^d family.

Before proving Theorem 4.2(ii), as a warmup we briefly prove the following:

▶ **Proposition 4.9.** *If C is a depth-d purely classical circuit and $|\phi\rangle$ is a mono-product state, then $|\langle \nu | C | \phi \rangle|^2 \leq 1/2 + \exp(-\Omega(n/2^d))$ for all n-nekomata $|\nu\rangle$.*

Proof. Since standard-basis measurements of qubits in a mono-product state are independent, it follows from the above discussion that a standard-basis measurement of any n designated target qubits of $C|\phi\rangle$ is a read-2^d family. If the expected Hamming weight of a standard-basis measurement of the targets of $C|\phi\rangle$ is less (resp. greater) than or equal to $n/2$, then Theorem 4.8 implies that the targets of $C|\phi\rangle$ measure to all-ones (resp. all-zeros) with probability at most $\exp(-\Omega(n/2^d))$, and the result follows from Lemma 4.4. ◀

▶ **Reminder** (Theorem 4.2(ii)). *If C is a nice, mostly classical circuit of size s and depth d, then the fidelity of $C|0\ldots0\rangle$ and any n-nekomata is at most*

$$\frac{1}{2} + \exp\left(-\Omega\left(\frac{n/2^d}{\max\left(\log s, \sqrt{n/2^d}\right)}\right)\right).$$

Abridged proof. Designate n qubits of $C|0\ldots0\rangle$ as targets, and assume without loss of generality that $s \geq \exp\left(\sqrt{n/2^d}\right)$. We will prove that for some $a \in \{0,1\}$, the targets of $C|0\ldots0\rangle$ measure to a^n with probability at most $\exp(-\Omega(n2^{-d}/\log s))$. The result then follows from Lemma 4.4.

Write $C = D(L \otimes \bigotimes_{G \in \mathcal{G}} G)$ such that D is purely classical, L is a layer of single-qubit gates, and $\mathcal{G}$ is a set of multi-qubit $R_\otimes$ gates that each satisfy the precondition of Corollary 4.6. For all $G \in \mathcal{G}$, a standard-basis measurement of $G|0\ldots0\rangle$ is distributed identically to $(b_G \wedge x_{G,i})_i$ for some independent Bernoulli random variables $b_G, (x_{G,i})_i$, where $\mathbb{E}[b_G] = 4\prod_i(1 - \mathbb{E}[x_{G,i}])$. Let $\mu_G = \sum_i \mathbb{E}[x_{G,i}]$; then $\mathbb{E}[b_G] \leq 4\exp(-\mu_G)$.

By a union bound, the probability that there exists $G \in \mathcal{G}$ such that $\mu_G > 2\ln s$ and $b_G = 1$ is at most

$$\sum_{G:\mu_G > 2\ln s} 4\exp(-\mu_G) < 4s\exp(-2\ln s) = \exp(-\Omega(\log s)) \le \exp(-\Omega(n2^{-d}/\log s)).$$

Therefore it suffices to prove that for some $a \in \{0,1\}$, the targets of $|\varphi\rangle := D(L \otimes \bigotimes_{G:\mu_G\le 2\ln s} G \otimes I)|0\dots0\rangle$ measure to a^n with probability at most $\exp(-\Omega(n2^{-d}/\log s))$. Henceforth we will never refer to any gate G for which $\mu_G > 2\ln s$; phrases such as "for all G" and "$(\cdot_G)_G$" will implicitly quantify over only those gates G for which $\mu_G \le 2\ln s$.

Let $b = (b_G)_G$ and $x = (x_{G,i})_{G,i}$. Call x "good" if $\sum_i x_{G,i} \le c\ln s$ for all G, where $c > 2$ is a universal constant large enough so that $e(2e/c)^c < 1$. A well-known Chernoff bound states that if S is a sum of independent Bernoulli random variables, and $\mu = \mathbb{E}[S]$, then $P(S > t) < (e\mu/t)^t e^{-\mu}$ for all $t > \mu$. Therefore, by a union bound and the fact that $\max_G \mu_G \le 2\ln s$, the probability that x fails to be good is at most

$$\sum_G (e\mu_G/c\ln s)^{c\ln s} \le s(2e/c)^{c\ln s} = (e(2e/c)^c)^{\ln s} = e^{-\Omega(\log s)} \le \exp(-\Omega(n2^{-d}/\log s)).$$

Let y be a string of independent Bernoulli random variables distributed identically to a standard-basis measurement of $L|0\dots0\rangle$. Call the targets of $D|y,(b_G \wedge x_{G,i})_{G,i},0\dots0\rangle$ the "output bits", and note that they are distributed identically to a standard-basis measurement of the targets of $|\varphi\rangle$. If b is fixed then the output bits are a read-2^d family (as functions of the independent Bernoulli random variables in x and y). Alternatively, if x and y are fixed and x is good then the output bits are a read-$O(2^d\log s)$ family (as functions of the independent Bernoulli random variables in b).

The rest of the proof is given in the full paper, and involves the triangle inequality. ◀

5 Lower Bounds for General QAC Circuits

In Section 5.1 we prove a generalization of Theorem 1.5. The proof uses the following claim, which is proved in Section 5.2 (and is obtained as a corollary of a stronger result):

▶ **Corollary 5.1.** *For all $d \ge 1$, orthogonal projections $Q_1,\dots,Q_d$, and states $|\phi\rangle$,*

$$\|Q_d\cdots Q_1|\phi\rangle\| \le \exp\left(-\frac{\langle\phi|(I-Q_d)|\phi\rangle}{2d}\right).$$

Then, using this generalization of Theorem 1.5, in Sections 5.3 and 5.4 we prove Theorem 1.7(iii).

5.1 Proof of Theorem 1.5

Theorem 1.5 is the case of the following in which $\mathcal{H}_1,\dots,\mathcal{H}_n$ are single-qubit Hilbert spaces, $|\phi\rangle$ is the all-zeros state, $Q_j = |0\rangle\langle0|$ for all j, and $|\psi\rangle$ is an n-nekomata.

▶ **Theorem 5.2.** *There is a universal constant $c > 0$ such that the following holds. Let $\mathcal{H}_1,\dots,\mathcal{H}_n$ be Hilbert spaces, let $\mathcal{H}_T = \bigotimes_{j=1}^n \mathcal{H}_j$ (for "targets"), and let $\mathcal{H}_A$ be a Hilbert space (for "ancillae"). Let $|\phi\rangle = |\phi_1,\dots,\phi_n,\phi_A\rangle$ for some states $|\phi_j\rangle \in \mathcal{H}_j, j \in [n]\cup\{A\}$. Let Q_j be an orthogonal projection on $\mathcal{H}_j$ for $j \in [n]$, and let $|\psi\rangle$ be a state in $\mathcal{H}_T \otimes \mathcal{H}_A$ that measures to $\bigotimes_{j=1}^n Q_j \otimes I_{\mathcal{H}_A}$ and to $\bigotimes_{j=1}^n (I-Q_j)\otimes I_{\mathcal{H}_A}$ each with probability 1/2. Let C be a depth-d QAC circuit on $\mathcal{H}_T \otimes \mathcal{H}_A$ with at most $cn/(d+1)$ multi-qubit gates acting on $\mathcal{H}_T$. Then, $|\langle\psi|C|\phi\rangle|^2 \le 1/2 + \exp(-\Omega(n/(d+1)))$.*

Abridged proof. By Proposition 2.1 we may write $C = DL$ for some layer of single-qubit gates L and QAC circuit D, where D has the same topology as C and consists only of multi-qubit $R_\otimes$ gates. Since $L|\phi\rangle$ factors as a product state in the same way that $|\phi\rangle$ does, we may assume without loss of generality that C consists only of multi-qubit $R_\otimes$ gates, by replacing C and $|\phi\rangle$ with D and $L|\phi\rangle$ respectively.

We now generalize Lemma 4.4 from nekomata to states such as $|\psi\rangle$. Let $Q = \bigotimes_{j=1}^n Q_j \otimes I_{\mathcal{H}_A}$ and $Q' = \bigotimes_{j=1}^n (I - Q_j) \otimes I_{\mathcal{H}_A}$, and let $|\varphi\rangle = C|\phi\rangle$. Since $|\psi\rangle$ measures to $Q + Q'$ with probability 1, it follows from the triangle inequality and Cauchy-Schwarz that

$$\begin{aligned}|\langle\psi|\varphi\rangle|^2 &= |\langle\psi|(Q+Q')|\varphi\rangle|^2 \le (|\langle\psi|Q|\varphi\rangle| + |\langle\psi|Q'|\varphi\rangle|)^2\\ &\le (\|Q|\varphi\rangle\| \cdot \|Q|\psi\rangle\| + \|Q'|\varphi\rangle\| \cdot \|Q'|\psi\rangle\|)^2 = (\|Q|\varphi\rangle\|/\sqrt{2} + \|Q'|\varphi\rangle\|/\sqrt{2})^2\\ &= \langle\varphi|(Q+Q')|\varphi\rangle/2 + \|Q|\varphi\rangle\| \cdot \|Q'|\varphi\rangle\| \le 1/2 + \min(\|Q|\varphi\rangle\|, \|Q'|\varphi\rangle\|),\end{aligned}$$

so it suffices to prove that $\min(\|Q|\varphi\rangle\|, \|Q'|\varphi\rangle\|) \le \exp(-\Omega(n/(d+1)))$.

Since $\sum_{j=1}^n \langle\phi_j|Q_j|\phi_j\rangle + \sum_{j=1}^n \langle\phi_j|(I-Q_j)|\phi_j\rangle = n$, either $\sum_{j=1}^n \langle\phi_j|Q_j|\phi_j\rangle \ge n/2$ or $\sum_{j=1}^n \langle\phi_j|(I-Q_j)|\phi_j\rangle \ge n/2$. Assume without loss of generality that $\sum_{j=1}^n \langle\phi_j|(I-Q_j)|\phi_j\rangle \ge n/2$. We will prove that $\|Q|\varphi\rangle\| \le \exp(-\Omega(n/(d+1)))$.

Let $\mathcal{G}$ be the set of gates in C, ordered such that $C = \prod_{G\in\mathcal{G}} G$ (where each gate G is implicitly tensored with the identity). Also let $\mathcal{G}_T \subseteq \mathcal{G}$ be the set of gates in C that act on $\mathcal{H}_T$. For $G \in \mathcal{G}_T$ let $|\theta_G\rangle$ be the mono-product state, specified up to a phase factor, such that $G = R_{\theta_G} = I - 2|\theta_G\rangle\langle\theta_G|$. Let F be the set of functions with domain $\mathcal{G}$ that map each gate G in $\mathcal{G}_T$ to either I or $|\theta_G\rangle\langle\theta_G|$, and map each gate G in $\mathcal{G}\backslash\mathcal{G}_T$ to G itself. Then $C = \sum_{f\in F} (-2)^{|\{G: f(G)=|\theta_G\rangle\langle\theta_G|\}|} \prod_{G\in\mathcal{G}} f(G)$, so by the triangle inequality,

$$\|Q|\varphi\rangle\| = \|QC|\phi\rangle\| \le \sum_{f\in F} 2^{|\{G: f(G)=|\theta_G\rangle\langle\theta_G|\}|} \cdot \max_{f\in F} \left\| Q \prod_{G\in\mathcal{G}} f(G) \cdot |\phi\rangle \right\|.$$

By assumption, $|\mathcal{G}_T| \le cn/(d+1)$ (for a constant c to be specified later), so

$$\sum_{f\in F} 2^{|\{G: f(G)=|\theta_G\rangle\langle\theta_G|\}|} = \sum_{S\subseteq\mathcal{G}_T} 2^{|S|} = \prod_{G\in\mathcal{G}_T} (2^0 + 2^1) = 3^{|\mathcal{G}_T|} \le 3^{cn/(d+1)}.$$

Consider an arbitrary function $f \in F$. In the full paper we write $\left\|Q \prod_{G\in\mathcal{G}} f(G) \cdot |\phi\rangle\right\|$ as a product of $n+1$ terms, one of which we bound by 1, and the other n of which we bound individually using Corollary 5.1. The result is that

$$\left\| Q \prod_{G\in\mathcal{G}} f(G) \cdot |\phi\rangle \right\| \le \prod_{j=1}^n \exp\left(-\frac{\langle\phi_j|(I-Q_j)|\phi_j\rangle}{2(d+1)}\right) \le \exp\left(-\frac{n/2}{2(d+1)}\right).$$

Altogether this implies that $\|Q|\varphi\rangle\| \le \exp((c\ln 3 - 1/4) \cdot n/(d+1))$, and the result follows by taking $c < 1/(4\ln 3)$. ◀

5.2 Proof of Corollary 5.1

Let $\Delta(|\alpha\rangle, |\beta\rangle) = \arccos|\langle\alpha|\beta\rangle|$; we will abbreviate this as $\Delta(\alpha, \beta)$. In the full paper we prove the following:

▶ **Lemma 5.3.** *The function Δ satisfies the triangle inequality, i.e. $\Delta(\alpha,\gamma) \le \Delta(\alpha,\beta) + \Delta(\beta,\gamma)$ for all states $|\alpha\rangle, |\beta\rangle, |\gamma\rangle$.*

▶ Remark. For intuition as to why Lemma 5.3 is true, consider the similarly defined function $\Delta'(u,v) = \arccos\langle u,v\rangle$ for unit vectors $u, v \in \mathbb{R}^3$, where $\langle\cdot,\cdot\rangle$ denotes the usual inner product on $\mathbb{R}^3$. It is well known that $\Delta'(u,v)$ equals the angle between u and v, which equals the length of the arc (Figure 5) formed by traversing a great circle on the unit sphere from u to v in the shorter of the two directions. This arc is known to be the shortest path on the unit sphere between u and v, so Δ' represents distance on the unit sphere. See the full paper for more related discussion.

▶ **Proposition 5.4.** *For all $d \geq 1$, nonzero orthogonal projections Q_d, and states $|\phi\rangle$,*

$$\max_{Q_1,\ldots,Q_{d-1}} \|Q_d Q_{d-1} \cdots Q_1|\phi\rangle\| = \cos\left(\frac{\arccos\|Q_d|\phi\rangle\|}{d}\right)^d,$$

where the maximum is taken over all orthogonal projections $Q_1, \ldots, Q_{d-1}$.

Abridged proof. We first prove an analogous statement about rank-1 orthogonal projections, specifically that for all states $|\theta_0\rangle$ and $|\theta_d\rangle$,

$$\max_{|\theta_1\rangle,\ldots,|\theta_{d-1}\rangle} \left|\prod_{j=1}^{d}\langle\theta_{j-1}|\theta_j\rangle\right| = \cos\left(\frac{\arccos|\langle\theta_0|\theta_d\rangle|}{d}\right)^d. \tag{3}$$

Then, in the full paper, we prove that the original proposition follows from this rank-1 analogue.

On the image of Δ, i.e. on the interval $[0,\pi/2]$, the cosine function is decreasing and concave. Therefore for all states $|\theta_1\rangle, \ldots, |\theta_{d-1}\rangle$, by the AM-GM inequality, Jensen's inequality, and Lemma 5.3,

$$\left|\prod_{j=1}^{d}\langle\theta_{j-1}|\theta_j\rangle\right|^{1/d} \leq \frac{1}{d}\sum_{j=1}^{d}|\langle\theta_{j-1}|\theta_j\rangle| = \frac{1}{d}\sum_{j=1}^{d}\cos\Delta(\theta_{j-1},\theta_j) \leq \cos\left(\frac{1}{d}\sum_{j=1}^{d}\Delta(\theta_{j-1},\theta_j)\right)$$
$$\leq \cos\left(\frac{\Delta(\theta_0,\theta_d)}{d}\right) = \cos\left(\frac{\arccos|\langle\theta_0|\theta_d\rangle|}{d}\right).$$

In the full paper we give an example (Figure 6) which shows that this bound is tight. ◀

▶ **Reminder** (Corollary 5.1). *For all $d \geq 1$, orthogonal projections $Q_1, \ldots, Q_d$, and states $|\phi\rangle$,*

$$\|Q_d \cdots Q_1|\phi\rangle\| \leq \exp\left(-\frac{\langle\phi|(I-Q_d)|\phi\rangle}{2d}\right).$$

Proof sketch. The case $Q_d = 0$ is trivial. The case $Q_d \neq 0$ is handled using Proposition 5.4 and the Lagrange remainder theorem. ◀

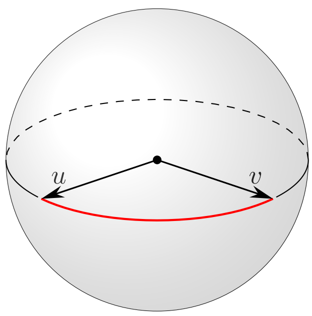

Figure 5 A geodesic on the sphere.

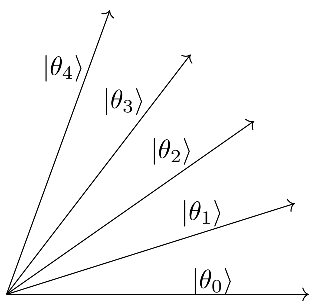

Figure 6 An optimal choice of $|\theta_1\rangle, \ldots, |\theta_{d-1}\rangle$ in the $d = 4$ case of Equation (3).

5.3 Simplifying Depth-2 QAC Circuits by Measuring Ancillae

For a one-qubit state $|\psi\rangle$, let *the* $|\psi\rangle$ *basis* be an orthonormal basis of $\mathbb{C}^2$ that includes $|\psi\rangle$. (We refer to "the" $|\psi\rangle$ basis because, up to a phase factor, there is a unique state orthogonal to $|\psi\rangle$.)

▶ **Lemma 5.5.** *Let $\mathcal{H}_1$ be a one-qubit Hilbert space, and let $\mathcal{H}_2$ and $\mathcal{H}_3$ be Hilbert spaces on arbitrary numbers of qubits. Then for all $|\psi\rangle \in \mathcal{H}_1, |\theta\rangle \in \mathcal{H}_2, |\phi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, the following two procedures generate identically distributed random states in $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$:*

- *measure the $\mathcal{H}_1$ qubit of $(R_{|\psi,\theta\rangle} \otimes I_{\mathcal{H}_3})|\phi\rangle$ in the $|\psi\rangle$ basis;*
- *measure the $\mathcal{H}_1$ qubit of $|\phi\rangle$ in the $|\psi\rangle$ basis, and then, conditioned on the outcome being $|\psi\rangle$, apply R_θ on $\mathcal{H}_2$.*

Proof. This follows easily from the fact that $R_{|\psi,\theta\rangle} = (I - |\psi\rangle\langle\psi|) \otimes I + |\psi\rangle\langle\psi| \otimes R_\theta$. ◀

Theorem 1.7(iii) is clearly equivalent to the statement that if C is a depth-2 QAC circuit, then any n designated "target" qubits of $C|0\ldots0\rangle$ measure to $|\text{😸}_n\rangle$ with probability at most $1/2 + \exp(-\Omega(n))$. The following is the starting point for our proof:

▶ **Proposition 5.6.** *Let p and $|\psi\rangle$ be such that for some depth-2 QAC circuit C, designated "target" qubits of $C|0\ldots0\rangle$ measure to $|\psi\rangle$ with probability p. Then there exist layers of $R_\otimes$ gates L_2, L_1 and a mono-product state $|\phi\rangle$ such that for some partition of the qubits of $L_2L_1|\phi\rangle$ into "targets" and "ancillae",*

(i) *the targets of $L_2L_1|\phi\rangle$ measure to $|\psi\rangle$ with probability at least p;*
(ii) *for all $k \in \{1, 2\}$, every ancilla is acted on by a gate in L_k, and every gate in L_k acts on at least one target.*

▶ Remark. Although not necessary for our purposes, using Proposition 2.1 it is easy to generalize the following argument to show that the gates in L_2 and L_1 may be assumed to be multi-qubit gates.

Proof sketch. By Proposition 2.1 there exist $L_2, L_1, |\phi\rangle$ satisfying (i) but not necessarily (ii). We measure selected ancillae in an appropriate product basis; doing so simplifies the circuit due to Lemma 5.5, and in expectation does not change the probability that the targets measure to $|\psi\rangle$. ◀

5.4 Proof of Theorem 1.7(iii)

The $\delta = 1$ case of the following is Markov's inequality:

▶ **Lemma 5.7.** *Let $0 < \delta \leq 1$, let $a > 0$, and let X be a nonnegative random variable. Then there exists $t \in [a, ae^{\delta^{-1}-1}]$ such that $P(X \geq t) \leq \delta\mathbb{E}[X]/t$.*

▶ Remark. The intuition behind our use of Lemma 5.7 is as follows. Theorem 5.2 implies that depth-2 QAC circuits require size at least $\Omega(n)$ to approximately construct $|\text{😸}_n, \psi\rangle$, and Proposition 5.6 implies that depth-2 QAC circuits that approximately construct $|\text{😸}_n, \psi\rangle$ have size at most $2n$ without loss of generality, so these bounds are "just a constant factor" away from implying that depth-2 QAC circuits of arbitrary size cannot approximately construct $|\text{😸}_n, \psi\rangle$. This is analogous to how Markov's inequality is "just a factor of δ" away from the conclusion of Lemma 5.7.

Proof sketch. Assume the contrary, write $\mathbb{E}[X] = \int_0^\infty P(X \geq t)dt$, and obtain the contradiction $\mathbb{E}[X] > \mathbb{E}[X]$ using the assumed lower bound on $P(X \geq t)$. ◀

▶ **Theorem 5.8** (Turán's theorem[7]). *Let $\mathcal{G}$ be a simple undirected graph on n vertices, and let d be the average degree of the vertices in $\mathcal{G}$. Then $\mathcal{G}$ contains an independent set of size at least $n/(d+1)$.*

In the full paper, we repeat the proof of Theorem 5.8 exposited by Alon and Spencer [1].

▶ Remark. For the intuition behind our use of Theorem 5.8, recall the discussion of disjoint light cones from Section 1.2.7.

Recall that $|\phi\rangle, C, |\psi\rangle, (Q_j)_j$ are variables from the statement of Theorem 5.2. In upcoming applications of Theorem 5.2 we will refer to $|\phi\rangle$ as the "input state", C as the "circuit", $|\psi\rangle$ as the "desired output state", and $(Q_j)_j$ as "projections".

▶ Remark. We will not actually use the full strength of Theorem 5.2, in the sense that we will always upper-bound the number of multi-qubit gates acting on the targets by upper-bounding the *total* number of gates. One could instead use the full strength of Theorem 5.2 in this regard, and forgo the use of Proposition 5.6 entirely by measuring selected ancillae all at once later in the proof, but we consider the current presentation to be simpler.

▶ **Reminder** (Theorem 1.7(iii), paraphrased). *If C is a depth-2 QAC circuit, then any n designated "target" qubits of $C|0\dots0\rangle$ measure to $|\text{🐱}_n\rangle$ with probability at most $1/2 + \exp(-\Omega(n))$.*

Abridged proof. Let L_2, L_1 be layers of $R_\otimes$ gates and let $|\phi\rangle$ be a mono-product state, with n qubits designated as targets and all other qubits designated as ancillae. Assume that for all $k \in \{1,2\}$, every ancilla is acted on by a gate in L_k, and every gate in L_k acts on at least one target. By Proposition 5.6 it suffices to prove that the targets of $L_2L_1|\phi\rangle$ measure to $|\text{🐱}_n\rangle$ with probability at most $1/2 + \exp(-\Omega(n))$.

Let c be the constant from Theorem 5.2, and let $\gamma = (c/2)(c/3)/(1+c/2)$ and $\delta = (c/2)\gamma^2$. Since Theorem 5.2 remains true if c is replaced by any constant between 0 and c, we may take c to be small enough so that $\gamma, \delta \leq 1$.

For a circuit C let $|C|$ denote the number of gates in C, and write "$G \in C$" to denote that G is a gate in C. First consider the case where $|L_2| \leq \gamma n$. It suffices to prove that $L_2L_1|\phi\rangle$ and $|\text{🐱}_n, \psi\rangle$ have fidelity at most $1/2+\exp(-\Omega(n))$ for all states $|\psi\rangle$. If $|L_1| \leq n(c/3)/(1+c/2)$ then $|L_1| + |L_2| \leq (c/3)n$, and the result follows from applying Theorem 5.2 with input state $|\phi\rangle$, circuit L_2L_1, desired output state $|\text{🐱}_n, \psi\rangle$, and n one-qubit projections $|0\rangle\langle 0|$ acting on the targets. Alternatively, if $|L_1| \geq n(c/3)/(1+c/2)$ then $|L_2| \leq (c/2)|L_1|$, and the result follows from applying Theorem 5.2 with input state $L_1|\phi\rangle$, circuit L_2, desired output state $|\text{🐱}_n, \psi\rangle$, and for every gate $G \in L_1$ the projection $|0\rangle\langle 0| \otimes I$ on the support of G, where $|0\rangle\langle 0|$ acts on one of the targets acted on by G. (Here we used the fact that $1/2 + \exp(-\Omega(|L_1|)) \leq 1/2 + \exp(-\Omega(n))$ by our assumption about $|L_1|$.)

The rest of the proof, i.e. the analysis of the case where $|L_2| \geq \gamma n$, is given in the full paper, and is a (relatively complicated) application of previously discussed ideas. ◀

[7] Usually Turán's theorem is phrased as saying that dense graphs have large cliques, whereas Theorem 5.8 says that sparse graphs have large independent sets. These statements are equivalent, because taking the complement of a graph turns cliques into independent sets and vice versa.

References

1 Noga Alon and Joel Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, 4 edition, 2016.

2 Adam Bene Watts, Robin Kothari, Luke Schaeffer, and Avishay Tal. Exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits. In *Symposium on the Theory of Computing*, pages 515–526, 2019. doi:10.1145/3313276.3316404.

3 Debajyoti Bera. A lower bound method for quantum circuits. *Information Processing Letters*, 111(15):723–726, 2011. doi:10.1016/j.ipl.2011.05.002.

4 Debajyoti Bera, Frederic Green, and Steven Homer. Small depth quantum circuits. *ACM SIGACT News*, 38(2):35–50, 2007. doi:10.1145/1272729.1272739.

5 Maosen Fang, Stephen Fenner, Frederic Green, Steven Homer, and Yong Zhang. Quantum lower bounds for fanout. *Quantum Information & Computation*, 6(1):46–57, 2006. arXiv:quant-ph/0312208.

6 Dmitry Gavinsky, Shachar Lovett, Michael Saks, and Srikanth Srinivasan. A tail bound for read-k families of functions. *Random Structures & Algorithms*, 47(1):99–108, 2015. doi:10.1002/rsa.20532.

7 Pranav Gokhale, Samantha Koretsky, Shilin Huang, Swarnadeep Majumder, Andrew Drucker, Kenneth R. Brown, and Frederic T. Chong. Quantum fan-out: circuit optimizations and technology modeling. arXiv, 2020. arXiv:2007.04246.

8 Frederic Green, Steven Homer, Cristopher Moore, and Christopher Pollett. Counting, fanout, and the complexity of quantum ACC. *Quantum Information & Computation*, 2(1):35–65, 2002. arXiv:quant-ph/0106017.

9 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. arXiv, 2020. arXiv:2007.00662.

10 Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Symposium on the Theory of Computing*, pages 6–20, 1986. doi:10.1145/12130.12132.

11 Peter Høyer and Robert Špalek. Quantum fan-out is powerful. *Theory of Computing*, 1(5):81–103, 2005. doi:10.4086/toc.2005.v001a005.

12 Alastair Kay. *Tutorial on the Quantikz package*, 2020. doi:10.17637/rh.7000520.

13 Daniel Padé, Stephen Fenner, Daniel Grier, and Thomas Thierauf. Depth-2 QAC circuits cannot simulate quantum parity. arXiv, 2020. arXiv:2005.12169.

14 Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. *Computational Complexity*, 25(4):849–881, 2016. doi:10.1007/s00037-016-0140-0.