Quantum State and Unitary Complexity

by

Gregory Rosenthal

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Department of Computer Science
University of Toronto

Quantum State and Unitary Complexity

Gregory Rosenthal
Doctor of Philosophy

Department of Computer Science
University of Toronto
2023

Abstract

Many natural problems in quantum computing involve constructing a quantum state or implementing a unitary transformation. However, relatively little is known about the computational complexity of these problems compared to that of computing boolean functions. In this thesis we do the following:

- We prove upper bounds on the complexity of constructing arbitrary states and implementing arbitrary unitaries with the help of a classical oracle.
- We prove bounds on the complexity of computing parity in QAC^0, a quantum analogue of AC^0, by way of a reduction to the task of constructing a certain type of state.
- We prove upper bounds on the circuit size and depth required to construct arbitrary states and implement arbitrary unitaries, in various quantum circuit models. Many of these bounds are tight.
- We define models of interactive proofs for constructing states and implementing unitaries with the help of an untrusted prover. We prove a quantum state analogue of the inclusion $\mathsf{PSPACE} \subseteq \mathsf{QIP}$, and obtain somewhat analogous results for unitaries and with multiple entangled provers.
- We present barriers to improving several of the above results.

Acknowledgments

> Come on down to South Park and meet some friends of mine.
>
> *South Park* Theme

First I'd like to thank my co-advisors Ben Rossman and Henry Yuen for all their support. They suggested interesting open problems and research areas to me, while also giving me space to pursue my own ideas. Additionally they funded lots of the academic travel discussed below, and were always available to talk to and were friendly and fun to be around.

Next I'd like to thank my other committee members François Le Gall, Toni Pitassi, Shubhangi Saraf, and Nathan Wiebe for their insightful questions and feedback regarding this thesis. And although most of the papers comprising this thesis are single-authored (excepting one paper with Henry), the technical content and/or writing quality of these papers benefited from specific conversations with Scott Aaronson, Srinivasan Arunachalam, Daniel Grier, Fermi Ma, Ian Mertz, Karen J. Morenz Korol, Eric Rosenthal, Benjamin Rossman, Rahul Santhanam, Adrian She, Nathan Wiebe, and Henry Yuen.

I'm grateful for being hosted by Tom Gur at Warwick (Summer 2022), by Henry Yuen at Columbia (Fall 2022), and by the Simons Institute (portions of the Fall 2018, Fall 2019, Spring 2023, and Summer 2023 semesters). These were great research experiences and the hosts went out of their way to make me feel welcome. I'm also grateful to my undergraduate advisor David Bindel for supervising my first research project.

Being part of the Theory Group at Toronto has been an outstanding experience. Credit belongs to my fellow students Alex, Coby, Deeksha, Deepanshu, Greg M., Hamoon, Harry, Ian, Jimmy, Kevan, Lawrence, Lily, Morgan, Noah, Robert, Sajjad, Suhail, and Yasaman, for all the time spent playing board games, going on walks and hikes, eating lunch in the Theory Lab, solving puzzle hunts (thanks Clément for organizing), playing board games, ice skating, drinking cocktails, playing squash, going to machine learning talks for the free food, and playing board games. (That the research environment is also great should go without saying.) Credit also belongs to the professors for helping to build such a great culture, and to Ingrid, Jankie, and the other staff for helping to make things run so smoothly.

Similar comments apply to my friends from the aforementioned visits to Warwick (A.P., Hugo, Marcel, Ninad, Sathya, Thejaswini), Columbia (Clay, Jason, John B., Maryam, Miranda, Natalie, Oliver, Rashida, Sayak, Shivam, Shyamal, Tim, William, Yuval E.), and Simons (Chinmay, Eli, Elizabeth, Fermi, Fran, Matthew, Noel, Peter, Rahul, Stefan, Yuval F.), as well as friends from various conferences (Daniel, Ewin, Joe, John K., Luke, Sujit, Xinyu), Hart House Orchestra friends (fellow cellists Chris A., Lynn, Nat, Rich, Tom; non-cellists Chris K., Hans, Jack, Jacob, Mihira, Rod; and the conductor Henry), miscellaneous friends and family who hosted me for some portion of my travels (Colin, Eric and Mia, Shao Min), and my intramural basketball teammates.[1] Thanks to Antonina for organizing music nights at Simons, and to Avishay for providing his keyboard at said music nights.

Finally, I'm incredibly grateful to my parents and to my brother Eric for all their love and support over the years. In particular, I'd like to thank my parents for hosting me for a year and a half during Covid in the middle of my PhD.

I was supported by an NSERC (PGS D) grant for the majority of my PhD. Circuit diagrams in this thesis were made using the Quantikz package [65].

[1] Friends falling under multiple of the above categories are listed just once.

Contents

List of Figures

Chapter 1

Introduction

> Why a four-year-old child could understand this report! Run out and find me a four-year-old child, I can't make head or tail out of it.
>
> Groucho Marx in *Duck Soup*

1.1 Quantum states and unitaries

One of the most basic tasks in computer science is to compute a given boolean function $f : \{0,1\}^* \to \{0,1\}^*$, where $\{0,1\}^*$ denotes the set of finite strings of zeros and ones. This task was formalized by Turing [97] in the 1930s, and captures many problems that arise in practice since any countable set can be identified with a subset of $\{0,1\}^*$. However not all computational problems can be expressed as simply computing a boolean function. For example, a *search problem* is described by a relation $R \subseteq \{0,1\}^* \times \{0,1\}^*$, where on input $x \in \{0,1\}^*$ the goal is to output *some* string y such that $(x, y) \in R$. Applications such as Monte Carlo methods [68] and randomized algorithms more generally [78] motivate the study of *sampling problems*, where the goal is to output a sample from a particular distribution over $\{0,1\}^*$ given independent uniform random bits.

Other examples of computational problems besides computing boolean functions arise in quantum computing. Although first suggested by Feynman [44] and Deutsch [40] in the 1980s, quantum computing did not receive widespread attention among the theoretical computer science community until the 1990s, when Shor [92] proved that quantum computers can efficiently factor integers. Since then, the nascent field of quantum complexity theory has mostly focused on bounding the resources required for quantum computers to compute boolean functions, or to perform other tasks that (perhaps with a loss in efficiency) classical computers can perform as well.

However there are also tasks that can *only* be performed on a quantum computer. Before describing these tasks we first give a high-level overview of quantum computing. An n-qubit *quantum state* is analogous to a probability distribution over $\{0,1\}^n$, except that each n-bit string is associated with a complex number instead of a nonnegative real number. The squared magnitudes of these complex numbers are required to form a probability distribution, i.e. they must sum to 1. The laws of quantum mechanics require that any physically realizable mapping from input states to output states be linear (excepting measurements), i.e. it must be a *unitary transformation*.[1] By identifying bit-strings with quantum states and identifying boolean functions with unitary transformations in a standard way, a hypothetical working quantum computer can simulate any classical computation with no loss in efficiency. The power of quantum computers comes from their ability to manipulate quantum states that do *not* simply encode classical strings, and to apply unitary transformations that do *not* simply encode boolean functions.

When the overall goal is to compute a boolean function, classical and quantum computers are equivalent in terms of computability, because a classical computer can simulate any quantum computation with an exponential blowup in time (e.g. $\mathsf{BQP} \subseteq \mathsf{EXP}$). However quantum computers can also perform tasks where the input and/or output is *itself* a quantum state that is not necessarily the encoding of a string. Even though a quantum state collapses to a string when measured (informally, when observed by a human), a computational task with a quantum output can still be interesting as a subroutine for computing a boolean function, or as part of a multi-round protocol where the output state of one round is not immediately measured. Such protocols arise in quantum cryptography for example [69], or whenever a quantum state is exchanged between two parties.

Some motivating examples are as follows. Decoders for quantum error-correcting codes transform noise-corrupted states into noise-free states [71]. Hamiltonian simulation algorithms implement unitaries that describe the evolution of physical systems [44]. Hawking radiation from a black hole is a quantum state, and decoding it requires implementing a unitary transformation [54]. The second step in the Linear Combinations of Unitaries (LCU) [23, 32] framework in quantum algorithms is to implement the associated unitaries in superposition. Algorithms for quantum state tomography [37] take as input copies of a quantum state, and output an approximation of some portion of the classical description of that state.

Sometimes the input is promised to be the all-zeros state, and the goal is to output a particular quantum state $|\psi\rangle$. We refer to this as the task of *constructing* $|\psi\rangle$, and it is a generalization of sampling problems. For example, states such as quantum money [2] and quantum pseudorandom states [61] are used in quantum cryptographic protocols. Variational quantum eigensolvers are algorithms that prepare ground states of physical systems [30]. Applying a Hamiltonian simulation algorithm on the all-zeros input can be used to study the long-term dynamics of that Hamiltonian [95]. The first step in the Linear Combinations of Unitaries (LCU) [23, 32] framework in quantum algorithms is to construct a state encoding the coefficients of that linear combination. If a family of states called QSampling states can be prepared in quantum polynomial time, then $\mathsf{SZK} \subseteq \mathsf{BQP}$ [3, 7].

These examples are representative of the progress that has been made in recent years toward understanding the complexity of certain types of states and unitaries, motivated by specific applications. However there is little in the way of a *general* theory of quantum state and unitary complexity like there is for boolean functions. By this we mean that boolean functions are not studied in isolation, but are instead grouped into well-motivated classes such as $\mathsf{P}, \mathsf{NP}, \mathsf{BQP}$ and so on; relationships between these classes are proved or conjectured, complete problems are often found, and this structure provides a framework in which to describe the complexity of any particular function of interest. A 2016 survey by Aaronson [3] arguably marks the beginning of quantum state and unitary complexity as a unified field, and this thesis is a continuation of that line of work.

[1] Strictly speaking, unitaries are less general than quantum channels for a given input size, but by Stinespring's dilation theorem [35] any channel can be simulated by a unitary transformation on at most double the number of qubits.

Constructing a quantum state or implementing a unitary transformation requires overcoming a number of obstacles that do not arise when computing a boolean function. For example, an n-qubit state $\sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$ is comprised of 2^n complex amplitudes α_x, making it an exponentially more complicated object than an n-bit string. Similarly an n-qubit unitary is comprised of $2^n \times 2^n = 4^n$ complex numbers, whereas the truth table of a function $f : \{0,1\}^n \to \{0,1\}$ consists of "only" 2^n bits. When implementing a unitary transformation, by the no-cloning theorem [36] it is impossible to copy or obtain a classical description of the input state, and even learning *anything* about the input state disturbs it.[2]

1.1.1 Ancillae

Another challenge is to reset any ancilla qubits to the all-zeros state at the end of the computation. *Ancillae* are extra workspace qubits that may be used by a quantum circuit, in addition to the input/output qubits. More formally we adopt the following definitions:

Definition 1.1.1 (Constructing a state and implementing a unitary transformation, cleanly or non-cleanly)**.** An $(n+a)$-qubit circuit C *cleanly* constructs an n-qubit state $|\psi\rangle$ if $C|0^{n+a}\rangle = |\psi\rangle|0^a\rangle$, and C *non-cleanly* constructs $|\psi\rangle$ if $C|0^{n+a}\rangle = |\psi\rangle|\phi\rangle$ for some a-qubit state $|\phi\rangle$. Similarly C *cleanly* implements an n-qubit unitary U if $C|\psi\rangle|0^a\rangle = U|\psi\rangle \otimes |0^a\rangle$ for all n-qubit states $|\psi\rangle$, and C *non-cleanly* implements U if there exists an a-qubit state $|\phi\rangle$ such that for all n-qubit states $|\psi\rangle$ it holds that $C|\psi\rangle|0^a\rangle = U|\psi\rangle \otimes |\phi\rangle$.[3] In all of these tasks the last a qubits are referred to as *ancillae*, and $|\phi\rangle$ is referred to as a *garbage state.*

One motivation for resetting ancillae to the all-zeros state is so that they can subsequently be used as ancillae for a different task. Another motivation is so that multiple similar tasks can be performed in superposition. For example, suppose that U_0 and U_1 are n-qubit unitaries and we wish to implement the unitary $U = |0\rangle\langle 0| \otimes U_0 + |1\rangle\langle 1| \otimes U_1$. If C_0 and C_1 are $(n+a)$-qubit circuits that cleanly implement U_0 and U_1 respectively, then by linearity the circuit $|0\rangle\langle 0| \otimes C_0 + |1\rangle\langle 1| \otimes C_1$ cleanly implements U, i.e.

$$(|0\rangle\langle 0| \otimes C_0 + |1\rangle\langle 1| \otimes C_1)|\psi\rangle|0^a\rangle = U|\psi\rangle \otimes |0^a\rangle$$

for all $(n+1)$-qubit states $|\psi\rangle$. But if C_0' and C_1' are $(n+a)$-qubit circuits that *non-cleanly* implement U_0 and U_1 respectively, with distinct garbage states $|\phi_0\rangle$ and $|\phi_1\rangle$, then

$$\big(|0\rangle\langle 0| \otimes C_0' + |1\rangle\langle 1| \otimes C_1'\big)|\psi\rangle|0^a\rangle = (|0\rangle\langle 0| \otimes U_0)|\psi\rangle \otimes |\phi_0\rangle + (|1\rangle\langle 1| \otimes U_1)|\psi\rangle \otimes |\phi_1\rangle$$

for all $(n+1)$-qubit states $|\psi\rangle$; thus the circuit $|0\rangle\langle 0| \otimes C_0' + |1\rangle\langle 1| \otimes C_1'$ does not even non-cleanly implement U, because the ancillae are entangled with the input/output register at the end of the computation.

Therefore ideally all of our upper bounds should hold for clean computations, and all of our lower bounds should hold for non-clean computations (which generalize clean computations), although we will not always achieve this.

It is well known that clean- and non-clean computations are equivalent for *boolean functions*, as shown in Fig. 1, but this approach does not generalize to operations with

[2] Analogously, given one sample from a probability distribution, there is no general procedure to generate independent samples from that same distribution.

[3] The requirement that $|\phi\rangle$ be independent of $|\psi\rangle$ is without loss of generality, because if $C|\psi_1\rangle|0^a\rangle = U|\psi_1\rangle \otimes |\phi_1\rangle$ and $C|\psi_2\rangle|0^a\rangle = U|\psi_2\rangle \otimes |\phi_2\rangle$, then by linearity the state $C(|\psi_1\rangle + |\psi_2\rangle)|0^a\rangle$ can only factor as the tensor product of $U(|\psi_1\rangle + |\psi_2\rangle)$ and an a-qubit state if $|\phi_1\rangle = |\phi_2\rangle$.

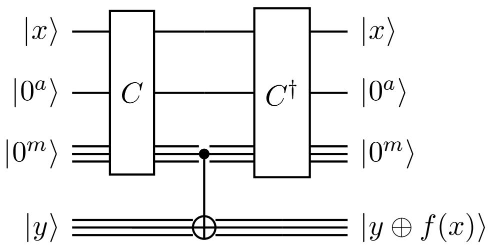

Figure 1: Let $f : \{0,1\}^n \to \{0,1\}^m$. If there exists an a-qubit state $|\phi\rangle$ such that $C|x, 0^{a+m}\rangle = |x\rangle|\phi\rangle|f(x)\rangle$ for all $x \in \{0,1\}^n$, then the circuit pictured here cleanly computes f.

a *quantum* output because by the no-cloning theorem such an output cannot be copied. Aaronson [3, Question 3.3.2] posed the question of finding a state that is easier to construct non-cleanly than to construct cleanly, and we pose the same question for unitaries:

Open Problem 1. Find a unitary that is "easier" to implement non-cleanly than to implement cleanly, according to some reasonable choice of complexity measure, or prove that no such unitary exists.

Throughout this Introduction, results will be stated to varying degrees of formality, and results of ours that are stated here informally will be restated formally in subsequent chapters. In particular we will mostly avoid specifying metrics for approximation error here. For upper bounds for constructing states and implementing unitaries, we usually consider error in the 2-norm (for pure states), operator 2-norm (for unitaries), or trace norm (for mixed states), and for lower bounds we usually consider related inner products (e.g. fidelity).

1.2 Reductions to boolean function complexity

Before proceeding further it is natural to ask: can questions about the complexity of quantum states and unitaries be reduced to questions about the complexity of boolean functions? After all, boolean function complexity has been studied for decades whereas quantum state and unitary complexity is a very young field. One facet of this question was posed by Aaronson and Kuperberg [6] in 2007 and named the "unitary synthesis problem" by Aaronson [3] in 2016:

Open Problem 2 (The unitary synthesis problem [3, 6])**.** Is there a polynomial-time quantum algorithm A such that for every unitary U, there exists a classical oracle f such that A^f approximately implements U?

We can also ask the analogous question for quantum states, although as we will explain shortly this question has been resolved in the affirmative:

Question 1.2.1 (The state synthesis problem [3])**.** Is there a polynomial-time quantum algorithm A such that for every state $|\psi\rangle$, there exists a classical oracle f such that A^f approximately constructs $|\psi\rangle$?

There are actually two versions of each of these questions, depending on whether or not the construction of $|\psi\rangle$ or implementation of U is required to be clean. By A^f we mean A with query access to the boolean function f, where queries can be made in superposition and may be adaptive. We refer to (not necessarily polynomial-time) upper bounds for Question 1.2.1 and Open Problem 2 as *state synthesis algorithms* and *unitary synthesis algorithms* respectively.

We can require f to have just one output bit without loss of generality, for reasons that will be explained in Section 2.1 when we define the quantum query model. The requirement that all queries be to the *same* function f is also without loss of generality, because if the j'th query is to a function f_j then the function $(j, x) \mapsto f_j(x)$ can simulate all queries. The requirement that a *single* algorithm A work for *all* states $|\psi\rangle$ or unitaries U is also without loss of generality, up to an additive constant blowup in the number of queries. This is because if C is a quantum circuit and f is a boolean function such that C^f approximately constructs $|\psi\rangle$ or implements U, then A can first query the description of C, then simulate C^f controlled on the description of C (assuming query access to f), and finally perform one more query to uncompute the description of C if a clean computation is desired. In the case where C makes no queries, this implies one-query non-clean and two-query clean state and unitary synthesis algorithms, but these algorithms are not *time-efficient* because most states and unitaries require exponentially large circuits as we will explain in Section 1.3.4.

Applications of state synthesis algorithms will appear throughout this thesis. As explained in a survey by Aaronson [1], the unitary synthesis problem has applications in areas such as the nonabelian hidden subgroup problem [41], decoding Hawking radiation [3, 54], and quantum copy-protection and quantum money [2]; in all of these cases the goal is to implement some unitary transformation U, and a positive solution to the unitary synthesis problem (with, say, an oracle $f \in \mathsf{PSPACE}$) would imply a small quantum circuit for U assuming a dramatic collapse of complexity classes (such as $\mathsf{BQP} = \mathsf{PSPACE}$).

Question 1.2.1 asks whether upper bounds for boolean functions imply upper bounds for constructing quantum states. We pose (but do not investigate) the converse question:

Open Problem 3 (Reverse state synthesis problem)**.** Is there a polynomial-time quantum algorithm A such that for every function $f : \{0,1\}^n \to \{0,1\}$, there exists a state $|\psi\rangle$ such for every unitary U that constructs $|\psi\rangle$ it holds that $A^{\text{ctrl-}U}$ approximately computes f?

By $A^{\text{ctrl-}U}$ we mean A with query access to controlled-U and controlled-$U^\dagger$. The special case where A makes just one query to U on the all-zeros state at the start of the algorithm is captured by the class $\mathsf{BQP/qpoly}$, consisting of functions computable by a bounded-error polynomial-time quantum algorithm with quantum advice. Aaronson and Drucker [4] proved that $\mathsf{BQP/qpoly} = \mathsf{YQP/poly}$, or in other words trusted quantum advice can be simulated by untrusted quantum advice along with trusted classical advice. Therefore Open Problem 3 can be interpreted as asking whether the *ability* to construct a state $|\psi\rangle$ yields more power than simply being *given* a copy of $|\psi\rangle$. The analogous "reverse unitary synthesis problem" is uninteresting, because any boolean function f can trivially be computed by making one query to a unitary that encodes f.

In Section 1.2.1 we present a polynomial-time state synthesis algorithm using just *one query*, and in Section 1.2.2 we present an $\tilde{O}(2^{n/2})$-time unitary synthesis algorithm. Finally in Section 1.2.3 we present a matching $\Omega(2^{n/2})$ query lower bound for a certain class of unitary synthesis algorithms, including the algorithm used to achieve our upper bound; thus new ideas are needed in order to make further progress on the unitary synthesis problem. Throughout we present comparisons to previous work, proof sketches, and open problems.

1.2.1 Upper bound for state synthesis

Recall that there is a trivial exponential-time state synthesis algorithm, making one query for a non-clean construction or two queries for a clean construction. The following algorithm

Algorithm	Queries	Size	Space	Error	Uniform	Clean
Trivial	1	exp	exp	1/exp	yes	no
	2					yes
Theorem 1.2.2	poly	poly	poly	1/exp	yes	yes
Theorem 1.2.3	1	exp	poly	1/poly	no	no
	2			1/exp		yes
Theorem 1.2.4 (this thesis)	1	poly	poly	1/exp	yes	no
	4					yes

Figure 2: Comparison of state synthesis algorithms.

improves on the trivial algorithm by running in polynomial time, but at the expense of requiring a super-constant number of queries:

Theorem 1.2.2 (Aaronson [3, Proposition 3.3.5])**.** *There is a uniform sequence $(C_n)_n$ of* $\mathrm{poly}(n)$*-size quantum circuits, each making $O(n)$ queries to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle f such that C_n^f cleanly constructs $|\psi\rangle$ to within exponentially small error.*

Theorem 1.2.2 generalizes a similar result of Grover and Rudolph [50]. The following algorithm also improves on the trivial algorithm, by running in polynomial rather than exponential *space*, *without* an increase in the number of queries:

Theorem 1.2.3 (Irani, Natarajan, Nirkhe, Rao and Yuen [59, Theorems 1.3 and 1.4])**.** *There is a nonuniform sequence $(C_n)_n$ of* $\mathrm{poly}(n)$*-qubit quantum circuits, each making one (resp. two) queries to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle f such that C_n^f non-cleanly (resp. cleanly) constructs $|\psi\rangle$ to within polynomially (resp. exponentially) small error.*

However Theorem 1.2.3 does not give an upper bound on the circuit size required to implement the non-query operations, besides the trivial exponential upper bound, and these circuits are nonuniform. Furthermore in the one-query version of Theorem 1.2.3, the approximation error is inverse polynomial rather than inverse exponential.

We prove the following upper bound, stated here informally:

Theorem 1.2.4. *There is a uniform sequence $(C_n)_n$ of* $\mathrm{poly}(n)$*-size quantum circuits, each making one (resp. four) queries to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle f such that C_n^f non-cleanly (resp. cleanly) constructs $|\psi\rangle$ to within exponentially small error.*

Fig. 2 compares all of these algorithms, of which ours is the only one that runs in polynomial time using a constant number of queries. Our result answers questions posed by Aaronson [3, Question 3.3.6] and Irani et al. [59, Section 7], who collectively asked whether there exists a polynomial-time one-query state synthesis algorithm with exponentially small error. The non-clean algorithm from Theorem 1.2.4 is a common improvement on all of the other non-clean algorithms mentioned in Fig. 2, but the clean algorithm from Theorem 1.2.4 makes two more queries than are made by some of the other clean algorithms mentioned in Fig. 2. This raises the following question:

Open Problem 4. For the clean version of the state synthesis problem, what is the optimal tradeoff between the complexity measures referred to in Fig. 2?

The proof of Theorem 1.2.4 goes roughly as follows. For simplicity, in this proof sketch we allow the circuit (as opposed to just the oracle) to depend on the state $|\psi\rangle$ being constructed. Call a state of the form $C \cdot 2^{-n/2} \sum_{x \in \{0,1\}^n} \pm|x\rangle$ where C is a Clifford unitary a "Clifford times phase state". Irani, Natarajan, Nirkhe, Rao and Yuen [59] proved that every state has fidelity $\Omega(1)$ with some Clifford times phase state, and observed that Clifford times phase states can be efficiently constructed with one query. (More generally this holds for C from any 2-design.) Thus all that remains is to decrease the approximation error.

We recursively define $|\phi_k\rangle$ for $k \geq 0$ as a Clifford times phase state that has fidelity $\Omega(1)$ with $\left(|\psi\rangle - \sum_{j=0}^{k-1} c_j|\phi_j\rangle\right) / \left\| |\psi\rangle - \sum_{j=0}^{k-1} c_j|\phi_j\rangle \right\|$, for appropriately chosen coefficients $c_0, c_1, \ldots$ tending to zero. We show that $\sum_{j=0}^{k-1} c_j|\phi_j\rangle$ is a good approximation of $|\psi\rangle$ for sufficiently large k. Furthermore, using Linear Combinations of Unitaries (LCU) [23, 32] we can construct this approximation of $|\psi\rangle$ with constant success probability. Finally we increase the success probability either by parallel repetition (in the one-query version of the theorem), with parallel queries merged into a single query, or by a hybrid of parallel repetition and amplitude amplification (in the four-query version).

We pose the question of whether a similar error reduction procedure exists for the *unitary* synthesis problem (Open Problem 2):

Open Problem 5. Assume there exists a polynomial-time unitary synthesis algorithm, but for some relatively weak notion of approximation (e.g. 0.1 error in the operator 2-norm or in the normalized Frobenius norm). Does this imply a polynomial-time unitary synthesis algorithm with *exponentially* small error?

A challenge to generalizing the above approach from state synthesis to unitary synthesis is that if we define V_j analogously to $|\phi_j\rangle$, then the remainder $U - \sum_{j=0}^{k-1} c_j V_j$ might be far from unitary (even up to rescaling) in which case we might not be able to define V_k appropriately.

Finally, we remark that Theorem 1.2.2 (or the clean version of Theorem 1.2.4) implies that the clean and non-clean versions of the unitary synthesis problem are equivalent. This is because a solution to the clean version of the state synthesis problem can be used to uncompute the garbage state following a non-clean solution to the unitary synthesis problem.

1.2.2 Upper bound for unitary synthesis

We prove the following:

Theorem 1.2.5. *There is a uniform sequence $(C_n)_n$ of $\tilde{O}(2^{n/2})$-size quantum circuits, each making $O(2^{n/2})$ queries to a classical oracle, such that for every n-qubit unitary U there exists a classical oracle f such that C_n^f cleanly implements U to within exponentially small error.*

As we will explain in Section 1.3, every n-qubit unitary can be implemented by a circuit of size $\tilde{O}(2^{2n})$ over a finite gate set, so Theorem 1.2.5 improves on the trivial algorithm by a quartic factor. Irani et al. [59, Section 7.2] also observed that *with postselection* there is a polynomial-time solution to the unitary synthesis problem, using the Choi–Jamiołkowski isomorphism and quantum teleportation.

Our proof of Theorem 1.2.5 involves a reduction from the task of implementing a unitary U to that of implementing what we call a "U-qRAM":

Definition 1.2.6 (U-qRAM)**.** Given an n-qubit unitary U, call a unitary A acting on $m \geq 2n$ qubits a *U-qRAM* if $A|x, 0^{m-n}\rangle = |x\rangle \otimes U|x\rangle \otimes \left|0^{m-2n}\right\rangle$ for all $x \in \{0,1\}^n$.

More generally, qRAMs are unitaries that map $|i\rangle|0\ldots0\rangle$ to $|i\rangle|\psi_i\rangle$ for all $i \in \mathcal{I}$, given an index set $\mathcal{I}$ and states $(|\psi_i\rangle)_{i\in\mathcal{I}}$ [46]. Informally, controlled on an input string $x \in \{0,1\}^n$, a U-qRAM cleanly constructs the corresponding output state $U|x\rangle$ of U in a separate register; if this separate register is not initialized to the all-zeros state then a U-qRAM's behavior is unspecified (subject to unitarity). Using a zero-error variant of Grover search we prove the following, where by C^A we mean C with A and $A^\dagger$ oracles:

Theorem 1.2.7. *There is a uniform family $(C_{n,m})_{n,m}$ for $m \geq 2n$ of quantum circuits, each making $O\left(2^{n/2}\right)$ queries to an m-qubit quantum oracle, such that for all n-qubit unitaries U and all m-qubit U-qRAMs A it holds that $C^A_{n,m}$ cleanly, exactly implements U.*

To see why Theorem 1.2.7 is nontrivial, suppose that we wish to apply a unitary U on the input state $\sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$. A natural first step is to query a U-qRAM to obtain the state $\sum_{x\in\{0,1\}^n} \alpha_x|x\rangle \otimes U|x\rangle$. But now to obtain $U\sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$, it is necessary to uncompute $|x\rangle$ in superposition controlled on $U|x\rangle$.

For simplicity we have omitted from Theorem 1.2.7 a bound on the circuit complexity of $C_{n,m}$. Such a bound will be given in Section 2.4.1, and in particular when $m = \mathrm{poly}(n)$ the size of $C_{n,m}$ is $\tilde{O}\left(2^{n/2}\right)$. Theorem 1.2.5 follows because the clean version of Theorem 1.2.4 trivially generalizes from constructing states to implementing $\mathrm{poly}(n)$-qubit U-qRAMs. We remark that Theorem 1.2.2 would also be suitable for this purpose, up to a $\mathrm{poly}(n)$ factor blowup in the number of queries, but the non-clean version of Theorem 1.2.4 would *not* be suitable for reasons discussed in Section 1.1.1.

1.2.3 Lower bound for unitary synthesis

We prove a matching $\Omega\left(2^{n/2}\right)$ query lower bound for Theorem 1.2.7 when U is Haar random and the U-qRAM A is defined appropriately:

Theorem 1.2.8. *For all sequences of quantum circuits $(C_n)_n$ making $o\left(2^{n/2}\right)$ queries to a $2n$-qubit quantum oracle, with probability $1 - o(1)$ over a Haar random n-qubit unitary U, there exists a $2n$-qubit U-qRAM A such that C^A_n is (in some sense) almost maximally far from implementing U, even non-cleanly.*

Since any U-qRAM tensored with the identity is also a U-qRAM, we may replace "$2n$-qubit" with "m-qubit" in Theorem 1.2.8 for any $m \geq 2n$. However, some restrictions on A are still necessary for a lower bound such as Theorem 1.2.8 to hold, at least if we allow A to act on more than $2n$ qubits. For example, the unitary A defined by

$$\forall x, y \in \{0,1\}^n, b \in \{0,1\} : \; A|x, y, b\rangle = \begin{cases} |x\rangle \otimes U|x \oplus y\rangle \otimes |0\rangle & \text{if } b = 0 \\ U|x\rangle \otimes |y\rangle \otimes |1\rangle & \text{if } b = 1 \end{cases}$$

is a U-qRAM, and can trivially be used to implement U when applied with $b = 1$.

It is well known that unstructured search on a list of length N requires $\Omega(\sqrt{N})$ quantum queries [80], but this does not immediately imply that Theorem 1.2.7 is tight, since there also exist algorithms that do not simulate unstructured search. As we will explain more precisely when we prove Theorem 1.2.8, it also takes $\Omega(\sqrt{N})$ quantum queries to compute $\sigma^{-1}(1)$ given query access to a permutation σ of $\{1,\ldots,N\}$ [10, 79], and this almost immediately implies an $\Omega\big(2^{n/2}\big)$ lower bound for Theorem 1.2.7 when U is a permutation matrix and A is defined appropriately. However this example is unsatisfying if our ultimate goal is to prove lower bounds for the unitary synthesis problem, since n-qubit permutation matrices can be efficiently implemented with two queries (on input x, first query $\sigma(x)$, and then use one more query to uncompute x controlled on $\sigma(x)$). In contrast, if *any* family of unitaries is hard to implement in the sense of the unitary synthesis problem, then Haar random unitaries are also hard to implement for the following reason:

Observation 1.2.9. *Any fixed unitary U can be written as $U = UR \cdot R^\dagger$ pointwise where R is Haar random, so since UR and $R^\dagger$ are also Haar random, the task of implementing U reduces to that of successively implementing two (dependent) Haar random unitaries.*

This reduction, along with the U-qRAM from our proof of Theorem 1.2.5 as discussed in Section 1.2.2, shows that an $o\big(2^{n/2}\big)$ upper bound for Theorem 1.2.7 in the case where U is Haar random would imply an $o\big(2^{n/2}\big)$ upper bound for the unitary synthesis problem in the general case. (Provided that in this hypothetical improvement to Theorem 1.2.7, the complexity of the non-query operations is not too large.) However, Theorem 1.2.8 rules out this approach.

On the other hand, Theorem 1.2.8 does not rule out the possibility of obtaining a tighter upper bound for the unitary synthesis problem by reducing to some *other* qRAM:

Open Problem 6. Is there a sequence $(C_n)_n$ of quantum circuits, each making $o\big(2^{n/2}\big)$ queries to a $(p(n) + q(n))$-qubit quantum oracle where $p(n), q(n) = \mathrm{poly}(n)$, such that for all n-qubit unitaries U there exists a family of $q(n)$-qubit states $\Psi = (|\psi_x\rangle)_{x\in\{0,1\}^{p(n)}}$ such that for all Ψ-qRAMs A it holds that C_n^A implements U?

If $\Psi = (|f(x)\rangle)_{x\in\{0,1\}^{p(n)}}$ for some function $f : \{0,1\}^{p(n)} \to \{0,1\}$, then plugging any Ψ-qRAM into Fig. 1 yields a clean computation of f, so Open Problem 6 is essentially a rephrasing of the unitary synthesis problem with a more modest runtime requirement.[4] However, this rephrasing suggests an approach to proving lower bounds for the unitary synthesis problem by considering increasingly general classes of state families Ψ.

We prove Theorem 1.2.8 by using Observation 1.2.9 to reduce to the previously mentioned lower bound for the case where U is a permutation matrix.

1.3 Quantum circuit complexity of states and unitaries

A central goal of computational complexity theory is to prove lower bounds on the time required for a Turing machine to compute a given boolean function. One approach lies in the fact that if a function $f : \{0,1\}^* \to \{0,1\}$ can be computed in time $T(n)$, then it can also be computed by a sequence of DeMorgan circuits $(C_n)_n$ of size $O(T(n)\log T(n))$ [12]. In other words the circuit $C_n : \{0,1\}^n \to \{0,1\}$ consists of $O(T(n)\log T(n))$ AND, OR and NOT gates and computes f on inputs of length n. This motivates the field of *circuit complexity* [64], which studies the resources required for circuits to compute explicit boolean functions. Although in 1949 Shannon [90] proved by a counting argument that most functions $f : \{0,1\}^n \to \{0,1\}$ require DeMorgan circuit size at least $2^n/n$, it remains an open problem to find an explicit function f that requires even superlinear-size DeMorgan circuits,

[4] Invoking Fig. 1 is necessary, because on input $|x,1\rangle$ a Ψ-qRAM might output $\phi_x|x, 1 \oplus f(x)\rangle$ for some unit-magnitude complex number $\phi_x \neq 1$.

so research in circuit complexity has instead focused on proving lower bounds in restricted circuit classes.

The above motivation applies to *quantum* circuit complexity as well; in particular, sequences of quantum circuits of size $O(T(n)^2)$ can simulate time-$T(n)$ quantum Turing machines [77, 104]. Another motivation for studying quantum circuit complexity is that, even more so than in classical computing, quantum circuits are a much cleaner model of computation to analyze than quantum Turing machines are. We will focus in particular on *low-depth* quantum circuits. Low-depth circuits are a model of fast parallel computation, and this is especially important in the quantum case, because quantum computations need to be fast relative to the decoherence time of the qubits in order to avoid error.

In Sections 1.3.1 and 1.3.2 we define and motivate various quantum circuit classes and complexity measures. In Section 1.3.3 we present upper and lower bounds on the complexity of computing the parity function in a class called QAC^0, by reducing to the task of constructing a certain type of quantum state. In Section 1.3.4 we present bounds on the circuit complexity of "worst-case" states and unitaries, according to various complexity measures. Finally in Section 1.3.5 we present a barrier to proving low-depth quantum circuit lower bounds for constructing explicit states.

1.3.1 The basics

The following quantum circuits are analogous to DeMorgan circuits:

Definition 1.3.1 (QNC circuits)**.** A QNC circuit is a quantum circuit consisting of one-qubit and CNOT gates.

The name "QNC circuit" is nonstandard but is in keeping with the names of other classical and quantum circuit classes that we will discuss. Analogously to DeMorgan circuits, QNC circuits with roughly 4^n gates can exactly implement any n-qubit unitary transformation [80], and this upper bound is tight by a dimension-counting argument. However there are uncountably many one-qubit gates, so it is impossible to provide a finite classical description of a QNC circuit. Furthermore it is not physically realistic to directly implement an arbitrary one-qubit gate in a lab. This motivates the following definition, where $\mathrm{SU}(2^k)$ denotes the group of k-qubit unitary transformations with determinant 1:

Definition 1.3.2 (Universal gate sets)**.** A set $\mathcal{G} \subseteq \mathrm{SU}(2^k)$ is a *universal gate set* if $|\mathcal{G}|$ is finite and the group generated by $\mathcal{G}$ is dense in $\mathrm{SU}(2^k)$.

The restriction to unitaries with determinant 1 is without loss of generality, since a global phase does not have physical significance. Universal gate sets exist [80], and circuits over any universal gate set can *efficiently* approximately simulate arbitrary QNC circuits:

Theorem 1.3.3 (The Solovay-Kitaev theorem[5] [26, 38])**.** *Let $\mathcal{G} \subseteq \mathrm{SU}(2^k)$ be a universal gate set. Then for all $\varepsilon > 0$, and all QNC circuits C with determinant 1 that consist of s gates and act on at least k qubits, there exists a circuit C' consisting of $s \cdot \operatorname{poly}\log(s/\varepsilon)$ gates from $\mathcal{G}$ (this can be improved to $O(s \log(s/\varepsilon))$ for certain universal gate sets $\mathcal{G}$ [55]) such that $\|C' - C\| \leq \varepsilon$.*

[5] The original statement of the Solovay-Kitaev theorem [38] required $\mathcal{G}$ to be closed under inverses, but Bouland and Giurgică-Tiron [26] showed that this is not necessary. The references that we cite here only state the $s = 1$ case of the theorem, but the general case follows immediately by Eq. (2.1.3).

Here $\|\cdot\|$ denotes the operator 2-norm. In particular, when $s \leq \exp(\mathrm{poly}(n))$ and $\varepsilon \geq \exp(-\mathrm{poly}(n))$ the multiplicative $\mathrm{poly}\log(s/\varepsilon)$ blowup in size is $\mathrm{poly}(n)$. This justifies QNC circuits as a model of computation. When we refer to a "quantum circuit" without specifying the type of circuit—as we have done, for example, in Section 1.2—we implicitly mean QNC circuits or circuits over a universal gate set.

One measure of the complexity of a quantum circuit is the number of ancillae, which we have discussed in Section 1.1.1. Two other measures are size and depth:

Definition 1.3.4 (Quantum circuit size and depth)**.** The *size* of a quantum circuit is the number of gates that it contains, not counting one-qubit gates. The *depth* of a quantum circuit is the number of layers of gates that it contains, not counting layers with only one-qubit gates.

Our decision not to count one-qubit gates is somewhat nonstandard, but allows for a cleaner statement of the lower bounds that we will discuss in Section 1.3.3 and does not significantly affect the statement of any of our upper bounds. A similar justification underlies the convention of not counting NOT gates toward size and depth in boolean circuit complexity. In fact, similarly to how we can push all of the NOT gates to the bottom of a DeMorgan circuit without loss of generality [64], in Section 3.1 we will show how to push all of the one-qubit gates to the bottom of certain types of quantum circuits without loss of generality (including QNC circuits). Furthermore size and depth can be interpreted as measures of the physical 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.

The size of a circuit is trivially at most its depth times number of qubits acted on, so we will often omit a size parameter from our circuit upper bound statements when it can be inferred in this way. In particular, in constant-depth circuit classes with unbounded-fanin gates a circuit can have polynomial "size" yet act on a superpolynomial number of ancillae, so for these circuit classes we will refer to polynomial-*space* circuits (which are necessarily polynomial-size) to denote efficient computation.

The multiplicative $\mathrm{poly}\log(1/\varepsilon)$ blowup in *depth* from the Solovay-Kitaev theorem is significant in certain contexts, so we pose the following question:

Open Problem 7. Prove a low-depth version of the Solovay-Kitaev theorem (with ancillae allowed, and for universal sets of *multi*-qubit gates so that the ancillae can be accessed). Relatedly, most circuit upper bounds discussed in the rest of this section concern circuits with arbitrary one-qubit gates; prove similar results using a universal gate set in place of these gates.

We will be more interested in quantum circuit lower bounds for *approximate* computation than for exact computation, for at least two reasons that are not applicable in boolean circuit complexity. One reason, which we have already discussed, is that there is some error inherent in applications of the Solovay-Kitaev theorem anyway. A second reason is that Jia and Wolf [63] found explicit states and unitaries that require exponential QNC circuit size to *exactly* construct and implement. Their proof involves the *transcendence degree* of a unitary U, defined as the maximum number of algebraically independent standard-basis elements of U. Jia and Wolf proved that unitaries implemented by small QNC circuits have low transcendence degree, while on the other hand there exist explicit unitaries with high transcendence degree. However their approach does not generalize to lower bounds for *approximately* implementing a unitary or constructing a state.

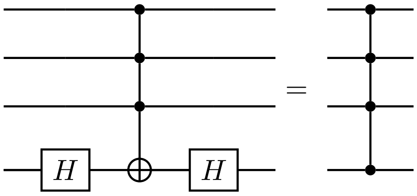

Figure 3: On the left, a generalized Toffoli gate with the target qubit conjugated by Hadamard gates. On the right, a generalized Z gate.

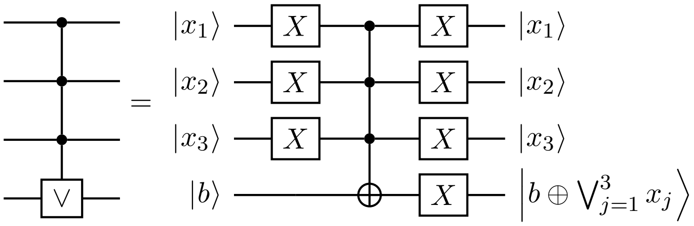

Figure 4: Implementation of an OR gate in QAC^0.

1.3.2 Quantum analogues of AC^0

One of the best understood boolean circuit classes is that of AC^0 circuits, which are constant-depth circuits consisting of NOT gates and unbounded-fanin AND and OR gates. A quantum analogue of AC^0 was defined by Green, Homer, Moore and Pollett [49], and is one of the weakest quantum circuit classes that is natural to define:[6]

Definition 1.3.5 (QAC^0 [49])**.** A QAC circuit is a quantum circuit consisting of arbitrary one-qubit gates, as well as *generalized Toffoli gates* of arbitrary arity defined by

$$|x, b\rangle \mapsto \left|x, b \oplus \prod_{j=1}^{n} x_j\right\rangle \quad \text{for} \quad x = (x_1, \ldots, x_n) \in \{0,1\}^n, b \in \{0,1\},$$

or equivalently *generalized Z gates* of arbitrary arity defined by

$$Z = I - 2|1 \ldots 1\rangle\langle 1 \ldots 1|.$$

(The equivalence between generalized Toffoli and Z gates is illustrated in Fig. 3, and was observed by Fang, Fenner, Green, Homer and Zhang [42].) A QAC^0 circuit is a constant-depth QAC circuit.

It may seem as if QAC^0 circuits can simulate AC^0 circuits. After all, generalized Toffoli gates can simulate unbounded-fanin AND, one-qubit gates can simulate NOT, unbounded-fanin OR can be obtained from these gates using DeMorgan's laws (Fig. 4), and we have already explained how to uncompute the garbage at the end (Fig. 1). The problem is that AC^0 circuits are granted the ability to feed an input bit or the output of a gate into arbitrarily many subsequent gates at no cost. To emulate this ability in QAC^0 it is necessary to make copies of a classical bit, i.e. to implement the *fanout transformation*:

Definition 1.3.6 (Fanout, restricted fanout, and $\mathsf{QAC}^0_\mathsf{f}$ [49])**.** An $(n+1)$-qubit FANOUT gate maps $|b, x\rangle$ to $|b, x \oplus b^n\rangle$ for $b \in \{0,1\}, x \in \{0,1\}^n$. We refer to the case where x is

[6] The even weaker class QNC^0 of constant-depth QNC circuits, for example, is easy to prove lower bounds against by light cone arguments. However for search and sampling problems there are still some interesting results surrounding this class (see Parham [82] for a survey).

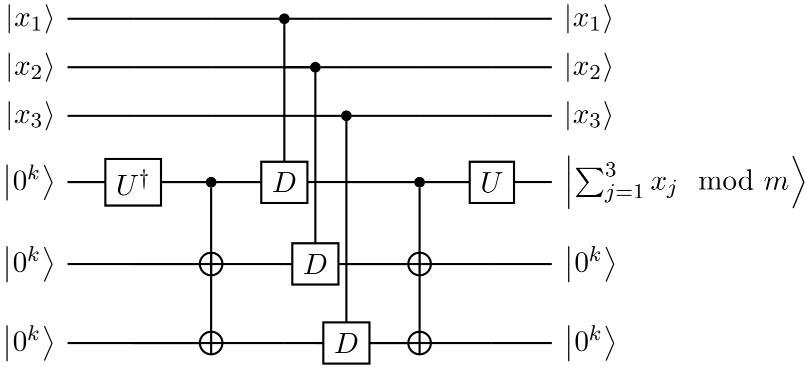

Figure 5: Implementation of a MOD_m gate in $\mathsf{QAC}^0_\mathsf{f}$. Let $k = \lceil \log(m-1) \rceil$, and let V be the k-qubit unitary transformation such that $V|r\rangle = |r+1 \mod m\rangle$ for $0 \le r < m$ and $V|r\rangle = |r\rangle$ for $m \le r < 2^k$. Let $V = UDU^\dagger$ be an eigendecomposition of V, i.e. U, D are unitary and D is diagonal. A MOD_m gate should compute $\sum_j x_j \mod m$, or equivalently $V^{\sum_j x_j}|0^k\rangle$, which equals $UD^{\sum_j x_j}U^\dagger|0^k\rangle$. Since m is constant, the unitaries U and D are fixed, so the pictured circuit can be implemented in $\mathsf{QAC}^0_\mathsf{f}$ by applying the controlled-D unitaries in parallel.

promised to be 0^n as *restricted fanout.* A QAC_f circuit is a QAC circuit with fanout gates of arbitrary arity, and a $\mathsf{QAC}^0_\mathsf{f}$ circuit is a constant-depth QAC_f circuit.

$\mathsf{QAC}^0_\mathsf{f}$ circuits may be physically realistic in certain quantum computing architectures such as ion traps [47, 52]. Restricted fanout makes copies of a bit $b \in \{0,1\}$, which does not violate the no-cloning theorem because b is classical. It follows that $\mathsf{QAC}^0_\mathsf{f}$ circuits can simulate AC^0 circuits gate by gate. In fact, $\mathsf{QAC}^0_\mathsf{f}$ circuits are *strictly* more powerful than AC^0 circuits, because polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuits can also compute threshold functions [57, 96] whereas AC^0 circuits require exponential size to do so [56]. In Fig. 5 we illustrate the weaker claim that polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuits can compute the MOD_m function for any constant m [49].

We will say more about QAC^0 and $\mathsf{QAC}^0_\mathsf{f}$ circuits shortly, and a survey by Bera, Green and Homer [22] discusses them in greater detail as well. For now we observe that QAC^0 circuits can be efficiently simulated by QNC circuits:

Lemma 1.3.7. *Every n-qubit, depth-d QAC_f circuit can be cleanly simulated by an $O(n)$-qubit, depth-$O(d \log n)$, size-$O(dn)$ QNC circuit.*

In particular, QNC^1 circuits (i.e. log-depth QNC circuits) can simulate $\mathsf{QAC}^0_\mathsf{f}$ circuits acting on polynomially many qubits, analogously to how $\mathsf{AC}^0 \subseteq \mathsf{NC}^1$. We perceive Lemma 1.3.7 to be folklore but could not find a reference, so we include a proof in Appendix A.1. It is an open problem whether the converse statement holds:

Open Problem 8. Can $\mathsf{QAC}^0_\mathsf{f}$ circuits simulate NC^1 circuits, or even QNC^1 or unbounded-depth QNC circuits, efficiently or otherwise?

For comparison, Valiant [98] proved that any linear-size NC^1 circuit with n input bits can be simulated by a $2^{O(n/\log\log n)}$-size depth-3 AC^0 circuit [64].

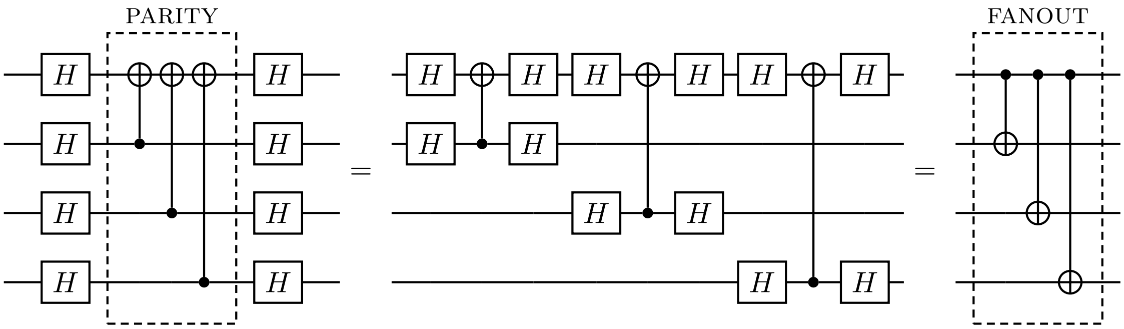

Figure 6: Equivalence of parity and fanout in QAC^0.

1.3.3 Are fanout and parity in QAC^0?

Green et al. [49] observed that fanout is equivalent to parity up to conjugation by Hadamard gates (Fig. 6), and is equivalent to restricted fanout in the sense that if C computes restricted fanout then the circuit in Fig. 1 computes fanout. This raises the following question:

Open Problem 9 (Green et al. [49])**.** Can parity (equivalently, fanout and restricted fanout) be computed by polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuits? Or at least by $\mathsf{QAC}^0_\mathsf{f}$ circuits of *arbitrary* size?

For comparison, Håstad famously proved that AC^0 circuits require exponential size to compute parity [19, 56]. But even an exponential-size AC^0 circuit for parity such as a CNF or DNF formula cannot be directly simulated in QAC^0, because this simulation requires implementing fanout to make copies of the input.

We *relate Open Problem 9 to quantum state complexity* by showing that in QAC^0, the tasks of computing n-qubit parity and fanout are equivalent to that of constructing (a purification of) the state $\frac{1}{2}|0^n\rangle\langle 0^n| + \frac{1}{2}|1^n\rangle\langle 1^n|$. The cat state $\frac{1}{\sqrt{2}}|0^n\rangle + \frac{1}{\sqrt{2}}|1^n\rangle$ can be constructed by applying restricted fanout to the state $\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right)|0^{n-1}\rangle$, but the converse direction of this equivalence (which is similar in spirit to Open Problem 3) is less obvious. Using this equivalence we prove the following result, as well as a generalization to superconstant-depth QAC circuits which we omit here for simplicity:

Theorem 1.3.8. QAC^0 *circuits of exponential size and depth* 7 *can cleanly, approximately compute parity to within exponentially small error. Furthermore,* QAC^0 *circuits that* either

(i) are of a certain specific form that also applies to the circuit from the above upper bound, but are of subexponential size;

(ii) are of sublinear size;

(iii) or are of depth at most 2*;*

cannot achieve more than an exponentially small improvement on the trivial 1/2 *approximation of parity, even non-cleanly. Similar results hold for fanout and* $\frac{1}{2}|0^n\rangle\langle 0^n|+\frac{1}{2}|1^n\rangle\langle 1^n|$.

The upper bound from Theorem 1.3.8 is the first known upper bound of *any* size for approximating parity in a sublogarithmic-depth QAC circuit. (The same holds for other equivalent tasks that we have mentioned, but for brevity we will leave this equivalence implicit in the following discussion and just refer to parity.) The following question remains open however:

Open Problem 10. Can QAC circuits of depth $o(\log n)$ *exactly* compute n-qubit parity?

Fang et al. [42] proved that QAC circuits with a ancillae require depth at least $\Omega(\log(n/(a+1)))$ to compute n-qubit parity, which is nontrivial when a is $o(n)$. Bera [21] used a different approach to prove something slightly weaker than the $a = 0$ case of this result. In contrast the lower bounds in Theorem 1.3.8 hold regardless of the number of ancillae, which may be much larger than the size of the circuit due to unbounded-arity gates.

Theorem 1.3.8(i) implies that new ideas are needed in order to obtain a subexponential-size upper bound for approximating parity in QAC^0. The "certain specific form" referred to can be explained as follows. Call a QAC^0 circuit "mostly classical" if it equals $CLML^\dagger$ where L is a layer of one-qubit gates, M is a layer of generalized Toffoli gates, and C consists of a constant number of layers of generalized Toffoli gates. The exponential-size construction of $\frac{1}{2}|0^n\rangle\langle 0^n| + \frac{1}{2}|1^n\rangle\langle 1^n|$ in our proof of Theorem 1.3.8 is achieved by a mostly classical circuit. On the other hand, we prove that mostly classical QAC^0 circuits *require* exponential size to approximately construct $\frac{1}{2}|0^n\rangle\langle 0^n| + \frac{1}{2}|1^n\rangle\langle 1^n|$, or more generally to sample from any probability distribution over $\{0,1\}^n$ for which the Hamming weight of a sample is not concentrated around its mean.

Although linear-size lower bounds are generally unimpressive in circuit complexity, Theorem 1.3.8(ii) is actually nontrivial, because circuits with unbounded-arity gates do not require linear size in order to "read" the entire input. We believe that the following question is interesting:

Open Problem 11. Can the proof techniques from Theorems 1.3.8(i) and 1.3.8(ii) be fused to obtain a superlinear-size lower bound for parity in arbitrary QAC^0 circuits?

Theorem 1.3.8(iii) is similar but incomparable to a result of Padé, Fenner, Grier and Thierauf [81], who also proved that depth-2 QAC circuits cannot compute parity. On the one hand, their result holds only for *exact, clean* computation whereas ours also holds for approximate, non-clean computation. On the other hand, their result holds when the number of input bits is as small as 4 whereas ours is only nontrivial for much larger inputs.

We now pose some additional open problems:

Open Problem 12. Is AC^0 in QAC^0, even if fanout isn't in QAC^0? More generally what boolean functions can polynomial-size QAC^0 circuits compute, besides read-k AC^0 formulas for constant k?

(A read-k formula is a boolean circuit where all gates have fanout 1, and k copies of each input bit are required.)

Open Problem 13. What is the QAC^0 complexity of computing the parity of a string x, given as input arbitrarily many copies of x? What about for other functions besides parity?

Open Problem 13 is motivated by the fact that if parity is not being used as a subroutine of a larger computation, and if we are not trying to compute the parities of multiple strings in superposition, then there is nothing to stop us from feeding multiple copies of x as input to a QAC^0 circuit.

Open Problem 14. Are there QAC^0 analogues of well-known techniques for obtaining AC^0 lower bounds, such as the switching lemma [56] or the polynomial method [83, 93]?

In contrast, all of the QAC^0 lower bounds that we have discussed are proved using techniques mostly unlike those used in AC^0 lower bounds.

1.3.4 Quantum analogues of Shannon and Lupanov's bounds

Recall that Shannon [90] proved that most functions $f : \{0,1\}^n \to \{0,1\}$ require DeMorgan circuit size at least $2^n/n$. A decade later Lupanov [74] proved that Shannon's lower bound is tight:

Theorem 1.3.9 (Lupanov [74])**.** *Every function $f : \{0,1\}^n \to \{0,1\}$ can be computed by a DeMorgan circuit of size $(1+o(1))2^n/n$.*

One can ask many analogous questions about the quantum circuit complexity of "worst-case" computational problems for a given input size. These include questions about upper or lower bounds on the complexity of cleanly or non-cleanly, exactly or approximately performing various types of tasks (computing a boolean function, constructing a state, or implementing a unitary transformation) in various circuit models (QAC, $\mathsf{QAC_f}$ or QNC circuits, or variants with a universal gate set instead of arbitrary one-qubit gates) according to various complexity measures (size, depth, number of ancillae, tradeoffs among these, or even the number of T gates in a Clifford + T circuit [33]). Some of these questions are uninteresting, such as the complexity of exactly implementing unitary transformations over a finite gate set, but even still there are too many questions to comment on individually. Therefore we will only discuss those questions that are related to results of this thesis and/or for which the author is aware of nontrivial results.

We will focus in particular on circuit depth. In *boolean* circuit complexity, Shannon's lower bound implies that most functions require DeMorgan circuits of depth at least $(1-o(1))n$, because a depth-d DeMorgan circuit has size less than 2^d. It is also easy to see that any function can be computed by a DeMorgan circuit of depth $(1+o(1))n$, such as a CNF or DNF formula.

Sun, Tian, Yang, Yuan and Zhang [94] used similar reasoning to prove that QNC circuits require depth at least $(1-o(1))n$ to non-cleanly,[7] exactly construct most n-qubit states, i.e. all states except for those in a set of Haar measure 0. This is because by a dimension-counting argument, QNC circuits require size $\Omega(2^n)$ to non-cleanly construct most n-qubit states (as was also observed by Knill [67, Theorem 3.4]), and by a light cone argument such circuits have depth at least $n(1-o(1))$ without loss of generality.

We prove the following:

Theorem 1.3.10. *Every n-qubit state can be cleanly, exactly constructed by a $\mathsf{QAC^0_f}$ circuit with $\tilde{O}(2^n)$ ancillae.*

The non-query operations from Theorem 1.2.4 can be efficiently implemented in $\mathsf{QAC^0_f}$, as we will discuss further in Section 1.3.5, so an analogue of Theorem 1.3.10 for *approximate* constructions follows by computing the queries from Theorem 1.2.4 with a $\mathsf{QAC^0_f}$ simulation of a CNF or DNF formula. Our proof of Theorem 1.3.10 instead uses a different approach inspired by the proof of Theorem 1.2.2. By Lemma 1.3.7 a statement similar to Theorem 1.3.10 holds for QNC circuits as well:

Corollary 1.3.11. *Every n-qubit state can be cleanly, exactly constructed by a QNC circuit of depth $O(n)$ with $\tilde{O}(2^n)$ ancillae.*

Sun et al. [94] and Zhang, Li and Yuan [106] independently proved Corollary 1.3.11, respectively shortly before and shortly after we did, and with just $O(2^n)$ ancillae and $O(2^n)$

[7] Their result is stated only for clean constructions, but their proof generalizes easily to non-clean constructions.

size. This matches the previously mentioned QNC size and depth lower bounds for exactly constructing states. Yuan and Zhang [105] proved in followup work that every n-qubit state can be cleanly, exactly constructed by a QNC circuit of size $O(2^n)$ and depth $O\left(n+\frac{2^n}{n+m}\right)$ using $m \geq 0$ ancillae, and that these size and depth upper bounds are tight for all n, m. This improves on a slightly weaker tradeoff of Sun et al. [94], and the proof of Yuan and Zhang's [105] upper bound cites ideas from our proof of Theorem 1.3.10.

Another $O(n)$-depth, $O(2^n)$-size upper bound for constructing arbitrary n-qubit states is due to Gui, Dalzell, Achille, Suchara and Chong [51]. Gui et al. refer to the *spacetime allocation* of a circuit as the number of pairs (j,k) such that the j'th qubit is not in the $|0\rangle$ state at the k'th time step, assuming the all-zeros input. Gui et al.'s construction also achieves the optimal $O(2^n)$ spacetime allocation, and as a corollary they obtain improved upper bounds for constructing tensor products of many unentangled states with limited ancillae.

Perhaps surprisingly, we show that the $\Omega(2^n)$ QNC size lower bound for *exactly* constructing most n-qubit states does not hold for *approximate* constructions:

Theorem 1.3.12. *There exists a finite gate set $\mathcal{G}$ such that for all $n \in \mathbb{N}, \varepsilon \geq \exp(-\mathrm{poly}(n))$ and n-qubit states $|\psi\rangle$, there exists a circuit consisting of $O(2^n \log(1/\varepsilon)/n)$ gates from $\mathcal{G}$ that cleanly constructs $|\psi\rangle$ to within error ε.*

To prove Theorem 1.3.12, we start with a more precise statement of the clean version of Theorem 1.2.4, and then apply Lupanov's upper bound (Theorem 1.3.9) to the oracle and apply the Solovay-Kitaev theorem (Theorem 1.3.3) to the non-query operations. We require $\varepsilon \geq \exp(-\mathrm{poly}(n))$ for convenience, but our proof technique implies a similar statement for smaller ε as well. The circuit from Theorem 1.3.12 uses exponentially many ancillae. Some ancillae are necessary, because Nielsen and Chuang [80, Section 4.5.4] proved by a counting argument that without ancillae, for every finite gate set $\mathcal{G}$ there exist states that require $\Omega(2^n \log(1/\varepsilon)/\log n)$ gates from $\mathcal{G}$ to construct to within error ε. We also prove an analogue of Nielsen and Chuang's lower bound *with* ancillae, by a similar counting argument, which matches the upper bound from Theorem 1.3.12 even for non-clean constructions:

Theorem 1.3.13. *Let $\mathcal{G}$ be a finite gate set. Then for all $n \in \mathbb{N}$ and $1/4 \geq \varepsilon \geq \exp(-\mathrm{poly}(n))$, there exists an n-qubit state $|\psi\rangle$ such that circuits over $\mathcal{G}$ require $\Omega(2^n \log(1/\varepsilon)/n)$ gates in order to non-cleanly construct $|\psi\rangle$ to within error ε.*

A slightly weaker lower bound of $2^n/\mathrm{poly}(n)$ holds if arbitrary one- and two-qubit gates are allowed, because by the Solovay-Kitaev theorem (Theorem 1.3.3) circuits consisting of $2^n n^{-\omega(1)}$ one- and two-qubit gates can be simulated to within exponentially small error by circuits consisting of $2^n n^{-\omega(1)}$ gates from a universal gate set, and therefore by Theorem 1.3.13 cannot construct arbitrary n-qubit states to within error ε. Closing this gap is an open problem:

Open Problem 15. Prove an analogue of Theorem 1.3.12 with an $o(2^n \log(1/\varepsilon)/n)$ upper bound for QNC circuits, or generalize Theorem 1.3.13 to QNC circuits.

We now turn our attention from states to unitaries. The non-query operations in Theorem 1.2.7 can be efficiently implemented in $\mathsf{QAC}^0_\mathsf{f}$, so by generalizing Theorem 1.3.10 from states to qRAMs we obtain the following:

Theorem 1.3.14. *Every n-qubit unitary transformation can be cleanly, exactly implemented by a QAC_f circuit of depth $O\big(2^{n/2}\big)$ with $\tilde{O}\big(2^{2n}\big)$ ancillae.*

By Lemma 1.3.7 a statement similar to Theorem 1.3.14 holds for QNC circuits as well:

Corollary 1.3.15. *Every n-qubit unitary transformation can be cleanly, exactly implemented by a QNC circuit of depth $\tilde{O}\big(2^{n/2}\big)$ with $\tilde{O}\big(2^{2n}\big)$ ancillae.*

Again, an analogous depth upper bound for *approximately* implementing unitaries also follows immediately from our upper bound for the unitary synthesis problem (Theorem 1.2.5), by using CNFs or DNFs to simulate the queries. And because Theorem 1.3.14 and by extension Corollary 1.3.15 are proved by reducing to Theorem 1.2.7, the following holds:

Observation 1.3.16. *The discussion in Section 1.2.3 is as relevant to the circuit depth of unitaries as it is to the unitary synthesis problem.*

Sun et al. [94] proved that every n-qubit unitary can be implemented by a QNC circuit of depth $\tilde{O}(2^n)$ with $O(2^n)$ ancillae, compared to which the circuit from Corollary 1.3.15 has lower depth but more ancillae. More generally, Sun et al. [94] proved that for $m \leq 2^n$, any n-qubit unitary can be implemented by a QNC circuit of size $O(4^n)$ and depth $\tilde{O}(4^n/m)$ with m ancillae. In followup work, Yuan and Zhang [105] generalized our proof of Corollary 1.3.15 to show that for $2^n \leq m \leq 4^n$, any n-qubit unitary can be implemented by a QNC circuit of depth $\tilde{O}\big(2^{3n/2}m^{-1/2}\big)$ with m ancillae; when $m = 4^n$ this matches Corollary 1.3.15 up to $\mathrm{poly}(n)$ factors.

The circuit from Corollary 1.3.15 has size $\tilde{O}\big(2^{2.5n}\big)$, whereas the optimal size is $O(4^n)$ [17, 29, 67, 80, 94]. This raises the following question:

Open Problem 16. Can every n-qubit unitary be implemented by a *single* QNC circuit that is *both* of size $\tilde{O}(4^n)$ and depth $\tilde{O}\big(2^{n/2}\big)$?

We also note following question:

Open Problem 17. Do there exist unitaries that require superlinear QNC circuit depth to implement, with no constraint on the size or number of ancillae? What about superconstant $\mathsf{QAC_f}$ circuit depth?

Note that Open Problem 17 is not asking to find an *explicit* unitary satisfying this requirement, although of course that would be preferable. Since there exist states that require $\Omega(n)$ depth to construct, there exist unitaries that require $\Omega(n)$ depth to implement.

Finally, the upper bound from Theorem 1.3.8 implies that exponentially large QAC^0 circuits can approximately simulate $\mathsf{QAC}^0_\mathsf{f}$ circuits, so we obtain the following as consequences of Theorem 1.3.10 and the fact that $\mathsf{QAC}^0_\mathsf{f}$ circuits can simulate AC^0 circuits:

Corollary 1.3.17. QAC^0 *circuits with double-exponentially many ancillae can cleanly, approximately construct any state and compute any boolean function, to within double-exponentially small error.*

Previously it was unknown whether QAC^0 circuits of arbitrary size could perform these tasks.

1.3.5 Barrier to proving $\mathsf{QAC}^0_\mathsf{f}$ lower bounds for constructing explicit states

In classical circuit complexity it is notoriously difficult to prove that an explicit boolean function is hard for a given circuit class. The same holds for quantum circuit complexity, since quantum circuits can simulate boolean circuits (with the possible exception of QAC^0 as we have discussed). However, this does not immediately imply that it should be difficult to prove quantum circuit lower bounds for *quantum* tasks with no classical analogue, such as constructing a quantum state.

Nevertheless, Aaronson [3] observed that Theorem 1.2.2 implies a barrier to finding an explicit sequence of quantum states that cannot be constructed by $\mathsf{BQP/poly}$ circuits (i.e. nonuniform polynomial-size QNC circuits) to within exponentially small error. Specifically, let $|\psi_n\rangle$ be an n-qubit state for all n and let f_n be the oracle associated with constructing $|\psi_n\rangle$ in Theorem 1.2.2. If $(f_n)_n$ can be computed in $\mathsf{BQP/poly}$, then plugging these circuits for $(f_n)_n$ into the algorithm from Theorem 1.2.2 yields a sequence of $\mathsf{BQP/poly}$ circuits for constructing $(|\psi_n\rangle)_n$. Conversely, if there are no $\mathsf{BQP/poly}$ circuits for constructing $(|\psi_n\rangle)_n$ then there are no $\mathsf{BQP/poly}$ circuits for computing $(f_n)_n$. This would be a breakthrough in circuit complexity, since finding an explicit function that is not in $\mathsf{BQP/poly}$ (or even $\mathsf{P/poly}$) is a longstanding open problem.

However this still leaves open the possibility of finding an explicit sequence of quantum states that cannot be constructed by $\mathcal{C}$ circuits, for some nonuniform quantum circuit class $\mathcal{C}$ such as $\mathsf{QAC}^0_\mathsf{f}$ that (as far as we know) is weaker than $\mathsf{BQP/poly}$. The non-query operations from Theorem 1.2.4 can be efficiently implemented in $\mathsf{QAC}^0_\mathsf{f}$, as previously mentioned after Theorem 1.3.10, so we can rule out this possibility by reasoning similar to the above:

Observation 1.3.18. $\mathsf{QAC}^0_\mathsf{f}$ *circuit lower bounds for cleanly constructing explicit states (to within exponentially small error) would imply* $\mathsf{QAC}^0_\mathsf{f}$ *circuit lower bounds for computing explicit boolean functions.*

Recall from Section 1.3.2 that $\mathsf{TC}^0 \subseteq \mathsf{QAC}^0_\mathsf{f}$ [57, 96], where TC^0 denotes the class of functions computable by non-uniform polynomial-size boolean circuits with NOT gates and unbounded-fanin AND, OR, and MAJORITY gates. It is an open problem to prove superpolynomial-size TC^0 lower bounds for an explicit function, so Observation 1.3.18 implies a barrier to proving similar $\mathsf{QAC}^0_\mathsf{f}$ lower bounds for constructing explicit states.

By embedding parity into the oracle from Theorem 1.2.4, it follows that Observation 1.3.18 holds with QAC^0 in place of $\mathsf{QAC}^0_\mathsf{f}$ as well. The analogous statement for QAC^0 is less of a barrier, since we have less reason to expect QAC^0 lower bounds for explicit boolean functions to be difficult to prove. However, Observation 1.3.18 does imply that a QAC^0 lower bound for some explicit state would strongly suggest that QAC^0 does not equal $\mathsf{QAC}^0_\mathsf{f}$.

Despite this, we can still hope to prove circuit lower bounds for constructing explicit states when the number of ancillae is limited, since the circuit from Theorem 1.2.4 requires ancillae:

Open Problem 18. Find an explicit state that cannot be constructed (to within exponentially small error) by $\mathsf{QAC}^0_\mathsf{f}$ or even QNC circuits of polynomial size without ancillae.

The following question asks whether there are any *other* barriers to proving lower bounds for constructing explicit states, and is related to Open Problem 3:

Open Problem 19. Suppose we had an explicit boolean function $f \notin \mathsf{BQP/poly}$. Can we define a sequence of quantum states in terms of f that cannot be constructed (to within exponentially small error) by polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ or even QNC circuits?

We also do not know of any barriers to proving circuit lower bounds for implementing explicit *unitaries*:

Open Problem 20. Find an explicit unitary that cannot be implemented (to within exponentially small error) by QAC^0 or even QNC circuits of polynomial size. Alternatively, find a barrier to such a lower bound.

For example, an upper bound for the unitary synthesis problem (Open Problem 2) would be such a barrier to QNC circuit lower bounds for explicit unitaries. Conversely, an upper bound for the following question would imply that circuit lower bounds for the unitaries $(V_n)_n$ referred to in the question are *necessary* in order to prove lower bounds for the unitary synthesis problem:

Open Problem 21 (Unitary synthesis problem with instance-independent quantum oracle)**.** Is there a sequence of unitaries $(V_n)_n$ and a polynomial-time quantum algorithm A such that for all $n \in \mathbb{N}$ and n-qubit unitaries U, there exists a classical oracle f such that A^{f,V_n} approximately implements U?

Similarly, an upper bound for the following question would be a barrier to $\mathsf{QAC}^0_\mathsf{f}$ circuit lower bounds for explicit unitaries:

Open Problem 22 (Constant-depth unitary synthesis problem)**.** Is there a solution to the unitary synthesis problem with a constant number of queries, where the non-query operations can be efficiently implemented in $\mathsf{QAC}^0_\mathsf{f}$?

Finally, the following is to Open Problem 22 as Open Problem 21 is to the unitary synthesis problem:

Open Problem 23 (Constant-depth unitary synthesis problem with instance-independent quantum oracle)**.** Is there a solution to Open Problem 21 with a constant number of queries, where the non-query operations can be efficiently implemented in $\mathsf{QAC}^0_\mathsf{f}$?

1.4 Quantum interactive proofs for states and unitaries

If a proof is a piece of text explaining why something is true, then an *interactive proof* is a *conversation* with the author of that text. More formally, the complexity class IP is the unbounded-round, randomized analogue of NP. Even more precisely, a language L is in IP if there exists a BPP verifier V such that the following conditions hold:

- *Completeness:* For all strings $x \in L$, there exists a prover P such that $V \leftrightarrows P$ accepts with probability 1.
- *Soundness:* For all strings $x \notin L$ and all provers P, it holds that $V \leftrightarrows P$ accepts with probability at most $1/2$.

Here the probabilities are over the randomness of V, and by $V \leftrightarrows P$ we mean the interaction between V and P. The soundness parameter can be made exponentially small by repeating the protocol polynomially many times, and accepting if and only if all instances accept. Interactive proofs have had applications in areas such as zero-knowledge proofs [48] and PCPs and hardness of approximation [13].

It is easy to see that $\mathsf{IP} \subseteq \mathsf{PSPACE}$, and less obviously the converse inclusion $\mathsf{PSPACE} \subseteq \mathsf{IP}$ [73, 89] holds as well, i.e. $\mathsf{IP} = \mathsf{PSPACE}$. One can also consider variants of IP. For example, the class MIP is defined similarly to IP except with multiple non-communicating provers, and is equal to NEXP [15]. The class MIP^* is defined similarly except that the provers share entangled qubits, and Ji, Natarajan, Vidick, Wright and Yuen [62] proved that $\mathsf{MIP}^* = \mathsf{RE}$.

What about interactive proofs with *quantum* verifiers? This is discussed in more detail in a survey by Vidick and Watrous [99], but a brief overview is as follows. The class QIP is defined similarly to IP, except with a BQP verifier. The inclusion $\mathsf{PSPACE} \subseteq \mathsf{QIP}$ holds trivially given that $\mathsf{PSPACE} \subseteq \mathsf{IP}$, and in fact Watrous [100] established the stronger inclusion $\mathsf{PSPACE} \subseteq \mathsf{QIP}(3)$ where $\mathsf{QIP}(3)$ denotes the three-message analogue of QIP. The converse inclusion $\mathsf{QIP} \subseteq \mathsf{PSPACE}$ is nontrivial and was proved by Jain, Ji, Upadhyay, Sarvagya and Watrous [60]:

Theorem 1.4.1 ([60, 100])**.** $\mathsf{QIP}(3) = \mathsf{QIP} = \mathsf{PSPACE}$.

The class QMIP with a quantum verifier and multiple entangled provers is (nontrivially) equal to the class MIP^* defined above [84], and is therefore equal to RE:

Theorem 1.4.2 ([62, 84])**.** $\mathsf{QMIP} = \mathsf{MIP}^* = \mathsf{RE}$.

What about quantum interactive proofs where the goal is not to solve a decision problem, but to (verifiably) perform an operation with a quantum input and/or output? This question was first asked in some of the work comprising this thesis. In Sections 1.4.1 and 1.4.2 we discuss interactive proofs for constructing states and implementing unitaries respectively.

1.4.1 Interactive state synthesis

We introduce a notion of interactive proofs for constructing a quantum state ρ, where a BQP verifier interacts with an unbounded-complexity but untrusted prover. At the end of the interaction the verifier accepts or rejects, and (unlike in previous models of interactive proofs) when accepting the verifier also outputs a quantum state. The *completeness* condition is that there should exist a prover such that the verifier accepts with probability 1. The *soundness* condition is that for every prover such that the verifier accepts with probability at least ε, the verifier's output state conditioned on accepting should be an approximation of ρ to within error δ.

We define $\mathsf{stateQIP}_{\varepsilon,\delta}$ as the class of state sequences $(\rho_n)_n$ that can be constructed in this way, and stateQIP as the intersection of $\mathsf{stateQIP}_{\varepsilon,\delta}$ over all functions $\varepsilon(n), \delta(n) \geq 1/\mathrm{poly}(n)$. We also define $\mathsf{stateQIP}(m)$ as the m-message version of stateQIP, and statePSPACE as the set of state sequences that can be constructed by a polynomial-space quantum algorithm (which may require exponential time) to within inverse polynomial error. These definitions are actually slight variants of our original ones, introduced in followup work by Metger and Yuen [76].

We prove the following quantum state analogue of one of the inclusions in Theorem 1.4.1:

Theorem 1.4.3. $\mathsf{statePSPACE} \subseteq \mathsf{stateQIP}(6)$.

We also proved a partial converse inclusion $\mathsf{stateQIP} \subseteq \mathsf{stateEXP}$. Metger and Yuen [76] then proved the full converse inclusion $\mathsf{stateQIP} \subseteq \mathsf{statePSPACE}$, establishing the equality $\mathsf{stateQIP} = \mathsf{stateQIP}(6) = \mathsf{statePSPACE}$.

The proof of Theorem 1.4.3 goes roughly as follows. Metger and Yuen [76] proved that statePSPACE is closed under purification, so it suffices to consider a pure state $|\psi\rangle$ that the verifier would like to construct. Let f be the oracle associated with constructing $|\psi\rangle$ in the one-query version of Theorem 1.2.4. Tomography of states in statePSPACE can be done in

PSPACE, and inspection of the proof of Theorem 1.2.4 reveals that f can be computed in PSPACE given query access to the description of $|\psi\rangle$, so it follows that f can be computed in PSPACE.[8] This suggests the following candidate protocol for constructing $|\psi\rangle$: simulate the algorithm from the one-query version of Theorem 1.2.4, with queries to f answered by running the $\mathsf{QIP}(3) = \mathsf{PSPACE}$ protocol from Theorem 1.4.1 in superposition.

However, controlled on an input string x to the $\mathsf{QIP}(3) = \mathsf{PSPACE}$ protocol, there is a garbage state (jointly held by the verifier and prover) associated with x at the end of the $\mathsf{QIP}(3) = \mathsf{PSPACE}$ protocol. The prover is required to help the verifier uncompute this garbage state, so that the verifier's output register is not entangled with the verifier's or prover's private workspace. The main challenge is to ensure that the prover uncomputes this garbage state honestly. Finally the soundness of the protocol is improved by repeating the above procedure polynomially many times in parallel, accepting if and only if every instance accepts, and then outputting the output state of a random instance.

We also prove the following, where stateR denotes the class of state sequences $(\rho_n)_n$ whose descriptions can be computed as a function of n, and stateQMIP is to stateQIP as QMIP is to QIP:

Theorem 1.4.4. $\mathsf{stateR} = \mathsf{stateQMIP}(6)$.

The proof is similar to that of Theorem 1.4.3, except using Theorem 1.4.2 instead of Theorem 1.4.1. The reason that $\mathsf{stateQMIP} \subseteq \mathsf{stateR}$, whereas $\mathsf{QMIP} \not\subseteq \mathsf{R}$, relates to the distinction between search and decision problems.

Followup work of Delavenne, Le Gall, Liu and Miyamoto [39] defined the classes stateQMA (i.e. $\mathsf{stateQIP}(1)$) and stateQCMA, which are to QMA and QCMA respectively as stateQIP is to QIP. More generally we can ask the following question:

Open Problem 24. How powerful are the classes $\mathsf{stateQIP}(k)$ for $k < 6$?

Recall that we defined stateQIP as the intersection of $\mathsf{stateQIP}_{\varepsilon,\delta}$ over all functions $\varepsilon(n), \delta(n) \geq 1/\mathrm{poly}(n)$. This raises the question of whether the soundness can be amplified to something stronger than inverse polynomial:

Open Problem 25. Is the class $\mathsf{statePSPACE} = \mathsf{stateQIP}$ contained in $\mathsf{stateQIP}_{\varepsilon,\delta}$ for some functions $\varepsilon(n), \delta(n) \leq n^{-\omega(1)}$?

Our protocol from Theorem 1.4.3 requires the honest prover to solve PSPACE-complete problems. In cryptography, however, interactive protocols require that the honest prover run in polynomial time (perhaps with advice), which raises the following question:

Open Problem 26. How powerful is the analogue of stateQIP with *efficient* provers?

Followup work of Bartusek, Khurana and Srinivasan [18] and Colisson, Muguruza and Speelman [34] introduced definitions of *zero-knowledge* proofs for constructing quantum states, motivated by cryptographic applications.

1.4.2 Interactive unitary synthesis

We define unitaryQIP and unitaryPSPACE analogously to stateQIP and statePSPACE respectively, except for unitaries rather than states. The following remains an open problem:

[8] More precisely, the sequence $(f_n)_n$ can be computed in PSPACE, where f_n is the oracle associated with constructing the n'th state in a given state sequence in statePSPACE.

Open Problem 27. Is $\mathsf{unitaryPSPACE} \subseteq \mathsf{unitaryQIP}$?

Part of the challenge is that the answers to the following questions are not obvious:

Open Problem 28. Assuming that $\mathsf{unitaryPSPACE} \subseteq \mathsf{unitaryQIP}_{1/2,1/2}$, would it follow that $\mathsf{unitaryPSPACE} \subseteq \mathsf{unitaryQIP}$? In other words, does soundness amplification hold for $\mathsf{unitaryQIP}$?

Open Problem 29. Assume that the unitary synthesis problem (Open Problem 2) has a polynomial-time solution, where for any sequence of unitaries in $\mathsf{unitaryPSPACE}$ the associated sequence of classical oracles is in PSPACE. Would it follow that $\mathsf{unitaryPSPACE} \subseteq \mathsf{unitaryQIP}_{1/2,1/2}$?

A reason these questions are nontrivial is that in the protocol from our proof of Theorem 1.4.3, we rely on parallel repetition for soundness amplification and to detect dishonest uncomputation of the garbage state (after the $\mathsf{QIP}(3) = \mathsf{PSPACE}$ protocol), and this is impossible with only one copy of the input state.[9] Unlike with interactive state synthesis, here the verifier must protect the unique copy of the input state from undetectable corruption by a dishonest prover.

Therefore we are only able to prove that certain subclasses of $\mathsf{unitaryPSPACE}$ are contained in $\mathsf{unitaryQIP}$. For example, we say that a unitary sequence $(U_n)_n$ has *polynomial action* if U_n acts nontrivially on only a $\mathrm{poly}(n)$-dimensional subspace. Examples include reflections $I - 2|\psi\rangle\langle\psi|$ acting on a one-dimensional subspace, which can be nontrivial to implement if $|\psi\rangle$ is difficult to construct. We prove the following, as well as an analogue with multiple entangled provers:

Theorem 1.4.5. *Every sequence of unitaries in* $\mathsf{unitaryPSPACE}$ *with polynomial action is in* $\mathsf{unitaryQIP}(6)$.

The proof of Theorem 1.4.5 goes roughly as follows. If U is in $\mathsf{unitaryPSPACE}$ then we can write $U = \exp(it\rho)$, where the Hamiltonian ρ is in $\mathsf{statePSPACE}$ and the evolution time t is in PSPACE. To implement U we first use Theorem 1.4.3 to construct copies of ρ and compute t, and then apply a Hamiltonian simulation algorithm [66, 72]. The assumption of polynomial action ensures that the evolution time t is at most $\mathrm{poly}(n)$, so the Hamiltonian simulation algorithm runs in $\mathrm{poly}(n)$ time. Generalizing this argument to Hamiltonian simulations with $t > n^{\omega(1)}$ seems potentially related to questions about "fast-forwarding of Hamiltonians" that were raised by Atia and Aharonov [14].

We also prove the following by reducing to Theorem 1.4.5:

Corollary 1.4.6. *An analogue of* $\mathsf{unitaryPSPACE} \subseteq \mathsf{unitaryQIP}$ *holds when the verifier's input state is promised to come from the* -1 *eigenspace of a reflection in* $\mathsf{unitaryPSPACE}$.

The following questions are related to Open Problem 27:

Open Problem 30. What other interesting subclasses of $\mathsf{unitaryPSPACE}$ are in $\mathsf{unitaryQIP}$? Examples might include more general Hamiltonian simulations, or qRAMs.

Open Problem 31. Would $\mathsf{unitaryPSPACE} \subseteq \mathsf{unitaryQIP}$ follow from a collapse of complexity classes of *decision* problems, such as $\mathsf{BQP} = \mathsf{PSPACE}$?

[9] If we relax the model such that the verifier is given *many* copies of an input state ρ and is only required to output a *single* copy of $U\rho U^\dagger$, then soundness amplification can be achieved but it is still nontrivial to detect dishonest uncomputation of the garbage state.

If $\mathsf{BQP} = \mathsf{PSPACE}$ then states in $\mathsf{statePSPACE}$ can be approximately constructed in quantum polynomial time, even without a prover, since for states in $\mathsf{statePSPACE}$ the oracle from Theorem 1.2.4 (or Theorem 1.2.2) is in PSPACE. However, absent a solution to the unitary synthesis problem, it is not clear that this helps with synthesizing unitaries in $\mathsf{unitaryPSPACE}$.

The converse to Open Problem 27 is also open:

Open Problem 32. Is $\mathsf{unitaryQIP} \subseteq \mathsf{unitaryPSPACE}$?

Finally, we remark that Bostanci, Efron, Metger, Poremba, Qian and Yuen [24] proved an anlogue of $\mathsf{unitaryQIP} = \mathsf{unitaryPSPACE}$ in the *average-case* setting, where the input state comes from an efficiently-sampleable distribution.

1.5 Conclusion

A common theme throughout this thesis is that upper bounds for constructing states can imply upper bounds for implementing unitaries. To repeat some previously discussed examples, we give a unitary synthesis algorithm by reducing to a state synthesis algorithm, we prove a $\mathsf{QAC^0}$ upper bound for parity and fanout by reducing to a $\mathsf{QAC^0}$ upper bound for the state $\frac{1}{2}|0^n\rangle\langle 0^n| + \frac{1}{2}|1^n\rangle\langle 1^n|$, we prove circuit depth upper bounds for arbitrary unitaries by reducing to circuit depth upper bounds for arbitrary states, and we give interactive proofs for implementing certain unitaries by reducing to interactive proofs for constructing certain states. This raises the following question:

Open Problem 33 (Very informal)**.** Are there *other* useful “generic” ways to obtain upper bounds for implementing unitaries, *besides* reducing to upper bounds for constructing states?

A solution to the following problem would imply solutions to both Open Problems 2 and 20:

Open Problem 34. Find an explicit sequence of unitaries $(U_n)_n$ such that for all boolean functions f, the unitaries $(U_n)_n$ cannot be efficiently implemented (approximately) even given query access to f.

An intermediate challenge between constructing states and implementing unitaries is that of implementing *isometries*, which are length-preserving linear transformations between two possibly distinct vector spaces. In other words, given an isometry $V : S \to \left(\mathbb{C}^2\right)^{\otimes n}$ where $S \subseteq \left(\mathbb{C}^2\right)^{\otimes n}$ is a subspace, implement *some* unitary U such that $U|\psi\rangle = V|\psi\rangle$ for all $|\psi\rangle \in S$. We have already seen some examples of isometries, such as qRAMs, Corollary 1.4.6, and any circuit with ancillae that are promised to start in the all-zeros state. However these cases are far from exhaustive:

Open Problem 35 (Very informal)**.** What is the complexity of implementing a “generic” isometry from S to $\left(\mathbb{C}^2\right)^{\otimes n}$, for various subspaces S and notions of complexity?

Delavenne et al. [39] posed the following question:

Open Problem 36 ([39])**.** Do there exist natural notions of reductions and completeness for state complexity classes?

The analogous classical problem would be to define reductions and completeness for sampling problems, which has also not been done to our knowledge. The analogous problem for unitaries is more straightforward, e.g. a trivial complete problem for $\mathsf{unitaryPSPACE}$ is "given an input length n, and a polynomial-space Turing machine M such that $M(1^n)$ is the description of a $\mathrm{poly}(n)$-qubit circuit, apply that circuit". Bostanci et al. [24] also gave *nontrivial* complete problems for many unitary complexity classes, relating to the Uhlmann transformation.

Open Problem 37. Besides the Uhlmann transformation, what are some other interesting, nontrivial complete problems for unitary complexity classes?

Finally, we provide some references to areas in quantum state and unitary complexity that are not otherwise discussed in this thesis. As of this writing, there has been a great deal of recent work on property testing of quantum channels [16, 31, 43, 53, 91]. A survey by Anshu and Arunachalam [11] discusses the complexity of learning quantum states. Le Gall, Miyamoto and Nishimura [70] consider the distributed complexity of constructing quantum states. To our knowledge no work has been done on the communication complexity of quantum states and unitaries, but in our opinion this is a natural area to explore:

Open Problem 38. Define an analogue of communication complexity for problems with a quantum input and/or output.

In catalytic computing in classical complexity theory [28], an otherwise space-bounded Turing machine is given a larger amount of "catalytic space" with the caveat that this space is initialized to an arbitrary string s (not all-zeros) and must end in that same string s. Quantum catalytic computing for states and unitaries (and even boolean functions) has barely if at all been explored:

Open Problem 39. Can ancillae be useful even if their starting state is adversarially chosen rather than all-zeros?

This thesis is based on the manuscripts "Bounds on the QAC^0 complexity of approximating parity" [85], "Interactive proofs for synthesizing quantum states and unitaries" [88], "Query and depth upper bounds for quantum unitaries via Grover search" [87], and "Efficient quantum state synthesis with one query" [86]. The paper "Interactive proofs..." [88] is joint work with Henry Yuen, and the rest are single-authored.

Section 1.6 is the preliminaries. In Chapter 2 we prove our results about state and unitary synthesis (from Section 1.2), and give the $\mathsf{QAC}^0_\mathsf{f}$ implementation of our state synthesis algorithms. In Chapter 3 we prove our results about the QAC^0 complexity of parity and related problems (from Section 1.3.3), in Chapter 4 we prove our other results about quantum circuit complexity (from Sections 1.3.4 and 1.3.5), and in Chapter 5 we prove our results about interactive state and unitary synthesis (from Section 1.4).

1.6 Preliminaries

Let log and ln denote the logarithms base 2 and e respectively, and let $[n] = \{1, \dots, n\}$. We write $(x_n)_n$ to denote the tuple of x_n for all n in some implicit index set, where the index set is usually the natural numbers. We denote probability by $\Pr(\cdot)$, expectation by $\mathbb{E}[\cdot]$, and the indicator variable of an event A by $\mathbb{1}_A$. We write $x \sim S$ to denote that x is sampled uniformly at random from a set S.

We assume basic familiarity with computational complexity theory [12]. Turing machines in this thesis have a read-only input tape, read-write work tapes, and a write-only output tape. Let $\{0,1\}^*$ denote the set of finite strings over $\{0,1\}$, and for $x \in \{0,1\}^*$ let $|x|$ denote the length of x. For $s : \mathbb{N} \to \mathbb{N}$ a Turing machine M uses space s if for all $x \in \{0,1\}^*$, at most $s(|x|)$ cells are used on the work tapes in the computation of $M(x)$. If M uses space s and halts then M uses time $O(2^s)$, so $|M(x)| \leq O\big(2^{s(|x|)}\big)$ for all x.

We also assume basic familiarity with quantum computing [80]. A *register* R is a named finite-dimensional complex Hilbert space. If $\mathsf{A}, \mathsf{B}, \mathsf{C}$ are registers, for example, then the concatenation ABC denotes the tensor product of the associated Hilbert spaces. For a linear transformation L and register R, we write L_R to indicate that L acts on R, and similarly we write ρ_R to indicate that a state ρ is in the register R. We write $\mathrm{tr}(\cdot)$ to denote trace, $\mathrm{tr}_\mathsf{R}(\cdot)$ to denote the partial trace over a register R, and $\mathrm{tr}_{>n}(\cdot)$ to denote the partial trace over all but the first n qubits.

We write I_n to denote the n-qubit identity transformation (or just I when n is implicit), H and X to denote the Hadamard and NOT gates respectively, ctrl-$U = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$ to denote controlled-U, and $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$ to denote the Hadamard basis states. For a (not necessarily normalized) vector $|\psi\rangle$ we write $\psi = |\psi\rangle\langle\psi|$.

We write $\|\cdot\|$ to denote the 2-norm of a vector, i.e. $\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$, and also to denote the operator 2-norm of a matrix, i.e. $\|M\| = \max_{|\psi\rangle} \|M|\psi\rangle\|$ where $|\psi\rangle$ ranges over all unit vectors. The value $\|M\|$ also equals the largest singular value of M. Let $\|M\|_1 = \mathrm{tr}(|M|)$ denote the trace norm of a matrix M, and let $\mathrm{td}(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1$ denote the trace distance between mixed states ρ and σ. We use the fact that

$$\mathrm{td}(\Phi(\rho), \Phi(\sigma)) \leq \mathrm{td}(\rho,\sigma) \tag{1.6.1}$$

for all channels Φ and states ρ, σ. We also use the following special case of the Fuchs-van de Graaf inequality [80]: if ρ is a mixed state and $|\psi\rangle$ is a pure state then

$$\mathrm{td}(\rho,\psi) \leq \sqrt{1-\mathrm{tr}(\rho\psi)} = \sqrt{\mathrm{tr}(\rho(I-\psi))}. \tag{1.6.2}$$

In particular, if $|\psi\rangle$ and $|\phi\rangle$ are pure states then

$$\mathrm{td}(\psi,\phi) \leq \sqrt{1-|\langle\psi|\phi\rangle|^2} = \sqrt{(1+|\langle\psi|\phi\rangle|)(1-|\langle\psi|\phi\rangle|)} \leq \sqrt{2(1-\mathrm{Re}(\langle\psi|\phi\rangle))} = \||\psi\rangle - |\phi\rangle\|. \tag{1.6.3}$$

For $|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$ and $\varepsilon > 0$, we define an *ε-precision description of* $|\psi\rangle$ to be a tuple $(\tilde{\alpha}_x)_{x\in\{0,1\}^n}$ of complex numbers specified exactly in binary such that $|\tilde{\alpha}_x - \alpha_x| \leq \varepsilon$ for all x. We will often leave ε implicit and simply refer to "the description of $|\psi\rangle$", by which we mean an $\exp(-p(n))$-precision description of $|\psi\rangle$ where p is a polynomial that may be chosen to be as large as desired; in this case $\mathrm{poly}(n)$ bits of precision are needed to specify $\tilde{\alpha}_x$.

Chapter 2

Quantum query complexity and state and unitary synthesis

> Just pass the work I assign you along to somebody else and trust to luck.
>
> Joseph Heller, *Catch-22*

In Section 2.1 we define the quantum query model and make some basic observations about it. In Section 2.2 we present a one-query state synthesis algorithm, given the ability to postselect on a measurement outcome that occurs with approximately constant probability. By reducing to this algorithm in different ways, in Section 2.3 we remove the postselection and present our clean and non-clean state synthesis algorithms. In Section 2.4 we prove our results regarding unitary synthesis and U-qRAMs.

Although upper bounds in this chapter are phrased in terms of $\mathsf{QAC_f}$ circuits, similar upper bounds for QNC circuits can be obtained using Lemma 1.3.7.

2.1 The quantum query model

By a *quantum circuit making k queries to an n-qubit quantum oracle*, we mean a circuit of the form $C = C_k Q_k C_{k-1} Q_{k-1} \cdots C_0$ where each C_j is a unitary and each Q_j is a placeholder for either a "forward" or "backward" query. For an n-qubit unitary A, by C^A we mean the unitary defined by substituting A and $A^\dagger$ respectively for the forward and backward queries in C. Claims about the quantum circuit complexity of C are in reference to the circuit $C_k C_{k-1} \cdots C_0$ defined by removing the queries from C. Let $C^\dagger = C_0^\dagger Q_1^\dagger C_1^\dagger Q_2^\dagger \cdots C_k^\dagger$, where the "conjugate transpose" of the forward query symbol is the backward query symbol and vice versa, and note that $\left(C^\dagger\right)^A = \left(C^A\right)^\dagger$.

Queries to a classical oracle (i.e. a boolean function) can be modeled in either of two standard ways. In the first, a function $f : \{0,1\}^n \mapsto \{0,1\}^m$ is encoded as the oracle U_f defined by $U_f|x, y\rangle = |x, y \oplus f(x)\rangle$. In the second, which is only applicable when $m = 1$, the function f is instead encoded as the oracle V_f defined by $V_f|x\rangle = (-1)^{f(x)}|x\rangle$. These models are equivalent, because $V_f = (I_n \otimes \langle -|)U_f(I_n \otimes |-\rangle)$, and if $g(x, y) = \bigoplus_{j=1}^m f(x)_j y_j$ (where the subscript j indicates the j'th bit of an m-bit string) then $U_f = (I_n \otimes H^{\otimes m})V_g(I_n \otimes H^{\otimes m})$. We write C^f to abbreviate C^{U_f} or C^{V_f}; since U_f and V_f are Hermitian we do not need to distinguish between forward and backward queries to a classical oracle.

We use the fact that parallel queries to classical oracles can be merged into a single query to a classical oracle, i.e.

$$V_{f_1} \otimes \cdots \otimes V_{f_k} = V_F \qquad \text{for} \qquad F\Big(x^{(1)}, \ldots, x^{(k)}\Big) = \bigoplus_{j=1}^{k} f_j\Big(x^{(j)}\Big) \tag{2.1.1}$$

for all functions $f_1, \ldots, f_k$. More generally, a collection of parallel queries of the form $\bigotimes_j U_{f_j} \otimes \bigotimes_k V_{g_k}$ can be merged into a single query to a classical oracle, using the above equivalence between the query models.

Sometimes we will want to implement $C_j^{f_j}$ controlled on an index $j \in [m]$, where C_j is a quantum circuit making k_j queries to a classical oracle f_j. This can be achieved by making $k = \max_j k_j$ queries to the classical oracle

$$|0\rangle\langle 0| \otimes I + \sum_{j=1}^{m} |j\rangle\langle j| \otimes U_j, \tag{2.1.2}$$

where the first register is set to $|j\rangle$ in the first k_j queries and $|0\rangle$ in the remaining $k - k_j$ queries. Interspersed with the queries are the non-query components of C_j controlled on j, e.g. using Lemma A.2.2 in the $\mathsf{QAC_f}$ circuit model, and with $k - k_j$ dummy components equal to the identity at the end.

2.1.1 Error propagation across queries

Often we will have a circuit that performs some task when given query access to *any* quantum oracle that implements a certain isometry (i.e. a rectangular matrix U such that $U^\dagger U = I$). If we instead use an oracle that *approximately* implements that isometry, then the error in the output can be bounded as follows:

Lemma 2.1.1. *Let C be an $(m+a)$-qubit quantum circuit making k queries to an n-qubit quantum oracle, and let J be an isometry from m qubits to $m+a$ qubits. Assume there exists a subspace $S \subseteq \big(\mathbb{C}^2\big)^{\otimes n}$ and an isometry $A : S \to \big(\mathbb{C}^2\big)^{\otimes n}$ such that for all n-qubit unitaries U consistent with A it holds that $C^U(I_m \otimes |0^a\rangle) = J$. Then for all isometries $B : S \to \big(\mathbb{C}^2\big)^{\otimes n}$ and all n-qubit unitaries V consistent with B, it holds that $\big\|C^V(I_m \otimes |0^a\rangle) - J\big\| \leq \sqrt{2} \cdot k\|A - B\|$.*

Some common use cases of Lemma 2.1.1 will be as follows. When the oracle is supposed to construct a state $|\psi\rangle$ when some control qubit is 1 and act as the identity when the control qubit is 0, the associated isometry is $A = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes |\psi\rangle\langle 0^n|$. If the isometry $B = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes |\phi\rangle\langle 0^n|$ instead constructs a state $|\phi\rangle$, then $\|A - B\| = \||\psi\rangle - |\phi\rangle\|$. When C is supposed to construct a certain state, the parameter m should be set to 0 and J should equal that state (up to an isomorphism between $\mathbb{C}^N$ and $\mathbb{C}^{N\times 1}$).

Our proof of Lemma 2.1.1 uses the fact that

$$\|U_m U_{m-1} \cdots U_1 - V_m V_{m-1} \cdots V_1\| \leq \sum_{j=1}^{m} \|U_j - V_j\| \tag{2.1.3}$$

for all unitaries U_j, V_j by the triangle inequality [80, Eq. 4.69]. (The reason that Lemma 2.1.1 does not trivially follow from Eq. (2.1.3), and without the $\sqrt{2}$ factor, is that some of the states acted on by applications of V in $C^V(I_m \otimes |0^a\rangle)$ might be far from S.)

For a unitary $U \in \mathbb{C}^{n\times n}$ and a subspace $S \subseteq \mathbb{C}^n$, we write $U_{|S}$ to denote the isometry from S to $\mathbb{C}^n$ defined by restricting the domain of U to S. The following inequality is trivially tight to within a $\sqrt{2}$ factor:

Claim 2.1.2. *For all unitaries $V \in \mathbb{C}^{n\times n}$, subspaces $S \subseteq \mathbb{C}^n$, and isometries $W : S \to \mathbb{C}^n$, there exists a unitary $U \in \mathbb{C}^{n\times n}$ such that $U_{|S} = W$ and $\|U - V\| \le \sqrt{2}\left\|W - V_{|S}\right\|$.*

Lemma 2.1.1 follows immediately from Eq. (2.1.3) and Claim 2.1.2, along with the fact that $\left\|U^\dagger - V^\dagger\right\| = \|U - V\|$ (to handle backward queries).

Proof of Claim 2.1.2. Let $P\Sigma Q^\dagger$ be a singular value decomposition of $W^\dagger V_{|S}$, and write $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_k)$ where $k = \dim(S)$ and $\sigma_1 \ge \cdots \ge \sigma_k \ge 0$. By definition $P, Q : \mathbb{C}^k \to S$ are (bijective) isometries such that $P\Sigma Q^\dagger = W^\dagger V_{|S}$. Then

$$\begin{aligned}\left\|W - V_{|S}\right\| \ge \left\|\left(W - V_{|S}\right)Q|k\rangle\right\| \ge \sqrt{2 - 2\left|\langle k|Q^\dagger W^\dagger V_{|S} Q|k\rangle\right|} = \sqrt{2 - 2\sigma_k\left|\langle k|Q^\dagger P|k\rangle\right|}\\ \ge \sqrt{2 - 2\sigma_k},\end{aligned}$$

where the last inequality is by Cauchy-Schwarz.

Define isometries $A, B \in \mathbb{C}^{n\times k}$ by $A = WP$ and $B = V_{|S}Q$, and note that $A^\dagger B = P^\dagger W^\dagger V_{|S} Q = \Sigma$. Let

$$\overline{A} = \operatorname{trnc}\Big[(B - A\Sigma)\big(I - \Sigma^2\big)^{-1/2}\Big], \qquad \overline{B} = \operatorname{trnc}\Big[(B\Sigma - A)\big(I - \Sigma^2\big)^{-1/2}\Big],$$

where the "truncation" operator $\operatorname{trnc}[\cdot]$ removes the j'th column of a matrix for all j such that $\sigma_j = 1$ (thus avoiding division by zero). It is straightforward to verify that $\overline{A}, \overline{B}$ are isometries satisfying $A^\dagger\overline{A} = B^\dagger\overline{B} = 0$ and $\overline{A}^\dagger\overline{B} = \operatorname{trnc}[\Sigma]$, so

$$\left\|\overline{A} - \overline{B}\right\| = \max_{\||\psi\rangle\|=1}\left\|\overline{A}|\psi\rangle - \overline{B}|\psi\rangle\right\| = \max_{\||\psi\rangle\|=1}\sqrt{2 - 2\langle\psi|\operatorname{trnc}[\Sigma]|\psi\rangle} = \sqrt{2 - 2\sigma_k} \le \left\|W - V_{|S}\right\|.$$

Let $S' \subseteq \mathbb{C}^n$ be the subspace that V maps to the image of $\overline{B}$. The subspaces S and S' are orthogonal, because $B^\dagger\overline{B} = 0$ and V maps S to the image of B, so we can write $\mathbb{C}^n = S \oplus S' \oplus S''$ for some subspace S''. Let $R : S' \to \mathbb{C}^{|\{j:\sigma_j \ne 1\}|}$ be the (bijective) isometry such that $V_{|S'} = \overline{B}R$, and define a unitary $U \in \mathbb{C}^{n\times n}$ by

$$U_{|S} = W, \qquad U_{|S'} = \overline{A}R, \qquad U_{|S''} = V_{|S''}.$$

To see that U is in fact unitary, note that $W^\dagger \cdot \overline{A}R = PA^\dagger\overline{A}R = 0$, and that the image of $U_{|S\oplus S'}$ (i.e. the span of the columns of A and $\overline{A}$) equals the image of $V_{|S\oplus S'}$ (i.e. the span of the columns of B and $\overline{B}$).

Let $|\psi\rangle \in \mathbb{C}^n$ be a unit vector such that $\|U - V\| = \|(U - V)|\psi\rangle\|$, and for a subspace $T \subseteq \mathbb{C}^n$ let $|\psi_T\rangle$ be the projection of $|\psi\rangle$ onto T. Then by the triangle inequality and Cauchy-Schwarz,

$$\begin{aligned}\|U - V\| = \|(U - V)|\psi\rangle\| \le \sum_{T\in\{S,S',S''\}}\left\|(U - V)_{|T}|\psi_T\rangle\right\| \le \left\|W - V_{|S}\right\| \cdot \||\psi_S\rangle\| + \left\|\overline{A} - \overline{B}\right\| \cdot \||\psi_{S'}\rangle\|\\ \le \left\|W - V_{|S}\right\|(\||\psi_S\rangle\| + \||\psi_{S'}\rangle\|) \le \sqrt{2}\left\|W - V_{|S}\right\| \cdot \||\psi_{S\oplus S'}\rangle\| \le \sqrt{2}\left\|W - V_{|S}\right\|. \qquad \square\end{aligned}$$

2.2 One-query state synthesis with postselection

In this section we prove the following lemma, which implies an efficient one-query state synthesis algorithm given the ability to postselect on a measurement outcome that occurs with approximately constant probability:

Lemma 2.2.1. *There is a real number $\gamma \approx 0.18$ such that the following holds. Let $\varepsilon : \mathbb{N} \to (0, 1/2)$ be a function such that $\varepsilon(n) \geq \exp(-\mathrm{poly}(n))$ and $\varepsilon(n)$ is computable in $\mathrm{poly}(n)$ time for all n, and let $t(n) = \lceil \log\log(1/\varepsilon(n)) \rceil + 7$. Then there is a uniform sequence of $\mathrm{poly}(n)$-qubit $\mathsf{QAC}^0_{\mathsf{f}}$ circuits $(A_n)_n$, each making one query to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle $f = f_{|\psi\rangle}$, a $(t(n)+n)$-qubit state $|\tau\rangle$ such that $\left(\langle 0^{t(n)}| \otimes I_n\right)|\tau\rangle = 0$, and a string $z \in \{0,1\}^{\mathrm{poly}(n)}$ such that*

$$\left\| A_n^f |0\dots 0\rangle - \left(\gamma \left|0^{t(n)}\right\rangle|\psi\rangle + \sqrt{1-\gamma^2}|\tau\rangle\right)|z\rangle \right\| \leq \varepsilon(n). \tag{2.2.1}$$

Furthermore there is an algorithm that takes as input the description of an n-qubit state $|\psi\rangle$ and a string x, runs in $\mathrm{poly}(n)$ space, and outputs $f_{|\psi\rangle}(x)$.

The proof is organized as follows. In Section 2.2.1 we describe A_n and f, and in Section 2.2.2 we prove that Eq. (2.2.1) holds. In Section 2.2.3 we prove that f can be computed in $\mathrm{poly}(n)$ space; actually we prove this for a slightly different oracle f' due to a subtlety involving floating-point arithmetic, but we show that Eq. (2.2.1) also holds with f' in place of f (and with slightly different values of $|\tau\rangle$ and z). In Section 2.2.4 we sketch an alternate proof of a statement similar to Lemma 2.2.1.

2.2.1 The algorithm

Our algorithm uses *Clifford unitaries*, which are products of Hadamard, phase (i.e. $S = |0\rangle\langle 0| + i|1\rangle\langle 1|$), and CNOT gates. Let

$$\alpha = 0.35, \qquad \beta = \sqrt{1-\alpha^2} \approx 0.94, \qquad \gamma = (1-\beta)/\alpha \approx 0.18.$$

For a complex number c, let $\mathrm{sgnRe}(c) = 1$ if the real part of c is nonnegative, and let $\mathrm{sgnRe}(c) = -1$ otherwise. For a vector $|\eta\rangle \in \left(\mathbb{C}^2\right)^{\otimes n}$ and a Clifford unitary C, let

$$|p_{\eta,C}\rangle = C \cdot 2^{-n/2} \sum_{x \in \{0,1\}^n} \mathrm{sgnRe}(\langle \eta | C | x\rangle)|x\rangle.$$

The following is implicit in Irani et al. [59], as explained in Appendix B:

Lemma 2.2.2 ([59])**.** *For all states $|\eta\rangle$ there exists a Clifford unitary C such that $\mathrm{Re}(\langle \eta | p_{\eta,C}\rangle) \geq \alpha$.*

Remark. In Section 2.2.4 we prove an analogue of Lemma 2.2.2 for a class of states other than $|p_{\eta,C}\rangle$, which can be used to give an alternate proof of a statement similar to Lemma 2.2.1. The idea to use $|p_{\eta,C}\rangle$ was suggested to us by Fermi Ma [75] after we sketched the argument in Section 2.2.4 to him.

Fix $n, |\psi\rangle, t = t(n)$ as in Lemma 2.2.1. For $k \geq 0$, given states $|\phi_0\rangle, \dots, |\phi_{k-1}\rangle$, let $|\eta_k\rangle = |\psi\rangle - \alpha \sum_{j=0}^{k-1} \beta^j |\phi_j\rangle$, and (using Lemma 2.2.2) let C_k be a Clifford unitary such that the state $|\phi_k\rangle = |p_{\eta_k, C_k}\rangle$ satisfies $\mathrm{Re}(\langle \eta_k | \phi_k\rangle) \geq \alpha \| |\eta_k\rangle \|$.

Let $T = 2^t$ and $|\sigma\rangle = \sqrt{\frac{1-\beta}{1-\beta^T}} \cdot \sum_{j=0}^{T-1} \sqrt{\beta^j} |j\rangle$. Observe that

$$|\sigma\rangle = \sqrt{\frac{1-\beta}{1-\beta^T}} \Big(|0\rangle + \beta^{2^{t-2}}|1\rangle\Big) \otimes \Big(|0\rangle + \beta^{2^{t-1}}|1\rangle\Big) \otimes \cdots (|0\rangle + \beta|1\rangle) \otimes \Big(|0\rangle + \beta^{1/2}|1\rangle\Big),$$

so there is a tensor product L of t one-qubit gates such that $L\big|0^t\big\rangle = |\sigma\rangle$.

The circuit A_n is described in Procedure 1, where A is a t-qubit register and B is an n-qubit register. Although the algorithm is phrased in terms of multiple queries, these can be merged into a single query using Eq. (2.1.1) and the surrounding discussion. Aaronson and Gottesman [5, Theorem 8] proved that every Clifford unitary can be written as a round of Hadamard gates, then a round of CNOT gates, then a round of phase gates, and so on in the sequence H-C-P-C-P-C-H-P-C-P-C with no ancillae (a "round" may consist of any number of layers of the given gate type). On Lines 5 and 7, by the description of a Clifford unitary we mean the concatenation of the descriptions of the rounds comprising that unitary, defined as follows:

- Since $H^2 = I$ a round of Hadamard gates equals $\bigotimes_{j=1}^n H^{x_j}$ for some string $x = (x_1, \ldots, x_n) \in \{0,1\}^n$. We call x the description of this round.
- Similarly since $S^4 = I$, a round of phase gates can be described by a string in $\{0,1,2,3\}^n$.
- A round of CNOT gates acts on the standard basis as $|x\rangle \mapsto |Mx\rangle$ for some $M \in \mathrm{GL}_n(\mathbb{F}_2)$, because this holds for a single CNOT gate and $\mathrm{GL}_n(\mathbb{F}_2)$ is closed under multiplication. We call the pair (M, M^{-1}) the description of this round.

Procedure 1 Circuit and oracle for Lemma 2.2.1

```
1: Construct $|\sigma\rangle_\mathsf{A}|+^n\rangle_\mathsf{B}$.                              ▷ Using $L$.
2: controlled on the classical state $|j\rangle_\mathsf{A}|x\rangle_\mathsf{B}$,
3:     apply a phase of sgnRe($\langle\eta_j|C_j|x\rangle$) by querying the oracle.
4: end control
5: Query descriptions of $C_0, \ldots, C_{T-1}$.  ▷ Merge with the Line 3 query using Eq. (2.1.1).
6: controlled on the classical state $|j\rangle_\mathsf{A}$,
7:     Apply $(C_j)_\mathsf{B}$ using the queried description of $C_j$.
8: end control
9: Apply $L_\mathsf{A}^\dagger$.
```

We now describe the $\mathsf{QAC}^0_\mathsf{f}$ implementation of Line 7 in greater detail. Given an index j and descriptions of Clifford unitaries $C_0, \ldots, C_{T-1}$, the description of C_j can be computed using Lemma A.2.1. A polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuit can then implement C_j by successively implementing the rounds comprising C_j. Rounds of Hadamard and phase gates can be implemented trivially. To implement a round of CNOT gates acting as $|x\rangle \mapsto |Mx\rangle$, first compute $y = Mx$, and then uncompute $x = M^{-1}y$ controlled on y, using that parity is in $\mathsf{QAC}^0_\mathsf{f}$ [49]. Since $\varepsilon(n) \geq \exp(-\mathrm{poly}(n))$ it holds that $t \leq O(\log n)$ and $T \leq \mathrm{poly}(n)$, so A_n requires $\mathrm{poly}(n)$ qubits.

Remark. Aaronson and Gottesman [5, Section 6] used similar reasoning to prove that Clifford unitaries can be implemented in QNC^1. The purpose of querying the description of C_j in Line 5, rather than applying it nonuniformly in Line 7, is to make the circuit (unlike the oracle) independent of $|\psi\rangle$. The purpose of querying *all* of $C_0, \ldots, C_{T-1}$, rather than just

C_j, is so that the register holding the description of C_j is unentangled with the rest of the system.

2.2.2 Proof of correctness

The string z referred to in the lemma is the concatenation of the descriptions of $C_0, \ldots, C_{T-1}$ along with some number of zeros. Let $|\varphi\rangle$ denote the final state in AB, let $|\theta\rangle = \langle 0^t|_{\mathsf{A}}|\varphi\rangle$, and let

$$|\tau\rangle = \frac{(I - |0^t\rangle\langle 0^t|)_{\mathsf{A}}|\varphi\rangle}{\|(I - |0^t\rangle\langle 0^t|)_{\mathsf{A}}|\varphi\rangle\|} = \frac{(I - |0^t\rangle\langle 0^t|)_{\mathsf{A}}|\varphi\rangle}{\sqrt{1 - \||\theta\rangle\|^2}}.$$

Then

$$\begin{aligned}
&\left\|A_n^f|0\ldots 0\rangle - \left(\gamma|0^t\rangle|\psi\rangle + \sqrt{1-\gamma^2}|\tau\rangle\right)|z\rangle\right\|^2 \\
&= \left\||\varphi\rangle - \left(\gamma|0^t\rangle|\psi\rangle + \sqrt{1-\gamma^2}|\tau\rangle\right)\right\|^2 \\
&= \left\||0^t\rangle(|\theta\rangle - \gamma|\psi\rangle) + \left(\sqrt{1-\||\theta\rangle\|^2} - \sqrt{1-\gamma^2}\right)|\tau\rangle\right\|^2 \\
&= \||\theta\rangle - \gamma|\psi\rangle\|^2 + \left(\sqrt{1-\||\theta\rangle\|^2} - \sqrt{1-\gamma^2}\right)^2 \\
&= \||\theta\rangle - \gamma|\psi\rangle\|^2 + \left(\frac{(\||\theta\rangle\| + \gamma)(\||\theta\rangle\| - \gamma)}{\sqrt{1-\||\theta\rangle\|^2} + \sqrt{1-\gamma^2}}\right)^2 \\
&\leq \||\theta\rangle - \gamma|\psi\rangle\|^2 + \frac{(1+\gamma)^2}{1-\gamma^2}(\||\theta\rangle\| - \gamma)^2 && \text{because } \||\theta\rangle\| \leq 1 \\
&\leq \left(1 + \frac{(1+\gamma)^2}{1-\gamma^2}\right)\||\theta\rangle - \gamma|\psi\rangle\|^2 && \text{by the triangle inequality} \\
&\leq 2.45\||\theta\rangle - \gamma|\psi\rangle\|^2,
\end{aligned}$$

so

$$\left\|A_n^f|0\ldots 0\rangle - \left(\gamma|0^t\rangle|\psi\rangle + \sqrt{1-\gamma^2}|\tau\rangle\right)|z\rangle\right\| \leq 1.58\||\theta\rangle - \gamma|\psi\rangle\|.$$

Inspection of Procedure 1 reveals that

$$|\varphi\rangle = L_{\mathsf{A}}^{\dagger}\left(\sum_{j<T}\left(j_{\mathsf{A}} \otimes C_j \sum_{x\in\{0,1\}^n} \mathrm{sgnRe}(\langle\eta_j|C_j|x\rangle)x_{\mathsf{B}}\right)\right)|\sigma\rangle_{\mathsf{A}}|+^n\rangle_{\mathsf{B}},$$

so

$$\begin{aligned}
|\theta\rangle &= \langle\sigma|_{\mathsf{A}}\left(\sum_{j<T}\left(j_{\mathsf{A}} \otimes C_j \cdot 2^{-n/2}\sum_{x\in\{0,1\}^n} \mathrm{sgnRe}(\langle\eta_j|C_j|x\rangle)|x\rangle_{\mathsf{B}}\right)\right)|\sigma\rangle_{\mathsf{A}} \\
&= \sum_{j<T} |\langle j|\sigma\rangle|^2 |p_{\eta_j,C_j}\rangle_{\mathsf{B}} = \frac{1-\beta}{1-\beta^T}\sum_{j<T}\beta^j|\phi_j\rangle = \frac{1-\beta}{1-\beta^T}\cdot\frac{|\psi\rangle - |\eta_T\rangle}{\alpha} = \frac{\gamma(|\psi\rangle - |\eta_T\rangle)}{1-\beta^T},
\end{aligned}$$

and therefore by the triangle inequality

$$\||\theta\rangle - \gamma|\psi\rangle\| = \frac{\gamma}{1-\beta^T}\left\|(|\psi\rangle - |\eta_T\rangle) - \left(1-\beta^T\right)|\psi\rangle\right\| = \frac{\gamma}{1-\beta^T}\left\|\beta^T|\psi\rangle - |\eta_T\rangle\right\|$$

$$\leq \frac{\gamma}{1-\beta}\big(\beta^T + \||\eta_T\rangle\|\big) \leq 2.86\big(\beta^T + \||\eta_T\rangle\|\big).$$

We prove by induction on k that $\||\eta_k\rangle\| \leq \beta^k$ for all k. The base case $k = 0$ holds because $|\eta_0\rangle = |\psi\rangle$. If the claim holds for k, then

$$\begin{aligned}\||\eta_{k+1}\rangle\|^2 = \Big\||\eta_k\rangle - \alpha\beta^k|\phi_k\rangle\Big\|^2 &= \||\eta_k\rangle\|^2 - 2\alpha\beta^k \mathrm{Re}(\langle\eta_k|\phi_k\rangle) + \alpha^2\beta^{2k} \\ &\leq \||\eta_k\rangle\|^2 - 2\alpha^2\beta^k\||\eta_k\rangle\| + \alpha^2\beta^{2k},\end{aligned}$$

where the inequality is by the definition of $|\phi_k\rangle$. This bound is convex as a function of $\||\eta_k\rangle\|$, so it achieves its maximum over $0 \leq \||\eta_k\rangle\| \leq \beta^k$ at either $\||\eta_k\rangle\| = 0$ or $\||\eta_k\rangle\| = \beta^k$. In both cases it follows straightforwardly that $\||\eta_{k+1}\rangle\| \leq \beta^{k+1}$, using in the $\||\eta_k\rangle\| = 0$ case the fact that $\alpha < \beta$, and using in the $\||\eta_k\rangle\| = \beta^k$ case the fact that $1 - \alpha^2 = \beta^2$.

Finally, writing $\varepsilon = \varepsilon(n)$ it holds that

$$\beta^T = \beta^{2^t} = \beta^{2^{\lceil \log\log(1/\varepsilon)\rceil + 7}} \leq \beta^{128\log(1/\varepsilon)} = \varepsilon^{128\log(1/\beta)} \leq \varepsilon^{8.36} \leq \varepsilon \cdot (1/2)^{7.36} \leq 0.01\varepsilon,$$

so

$$\Big\|A_n^f|0\dots0\rangle - \Big(\gamma|0^t\rangle|\psi\rangle + \sqrt{1-\gamma^2}|\tau\rangle\Big)|z\rangle\Big\| \leq 1.58 \cdot 2.86 \cdot 2\beta^T \leq \varepsilon.$$

2.2.3 Computing f in $\mathrm{poly}(n)$ space

Recall from Procedure 1 that the oracle f encodes, for each $0 \leq j < T$, a description of the Clifford unitary C_j and the values $\mathrm{sgnRe}(\langle\eta_j|C_j|x\rangle)$ for $x \in \{0,1\}^n$. The problem is that $\mathrm{sgnRe}(\langle\eta_j|C_j|x\rangle)$ depends discontinuously on $\langle\eta_j|C_j|x\rangle$, and $\langle\eta_j|C_j|x\rangle$ can only be computed approximately due to (exponentially small) error in the description of $|\eta_j\rangle$ and in floating-point arithmetic. Therefore we will use a slightly different oracle f'. Let $\delta = 0.01 \cdot \beta^{2T} \geq \exp(-\mathrm{poly}(n))$; we will use δ to define bounds on the floating-point error in certain calculations.

The new oracle f' For a vector $|\eta\rangle \in \big(\mathbb{C}^2\big)^{\otimes n}$ and a Clifford unitary C, let

$$\big|p'_{\eta,C}\big\rangle = C \cdot 2^{-n/2}\sum_{x\in\{0,1\}^n} \mathrm{sgnRe}\Big(\widetilde{\langle\eta|C|x\rangle}\Big)|x\rangle,$$

where $\widetilde{\langle\eta|C|x\rangle}$ is a value computable in $\mathrm{poly}(n)$ space (given descriptions of $|\eta\rangle$ and C) such that $\Big|\widetilde{\langle\eta|C|x\rangle} - \langle\eta|C|x\rangle\Big| \leq 2^{-n/2}\delta$. For example we may compute $\widetilde{\langle\eta|C|x\rangle}$ by a "sum over histories" argument, i.e. write $C = R_1 \cdots R_{13}$ as the product of the rounds R_i comprising the description of C, and use that

$$\langle\eta|C|x\rangle = \sum_{y_0,\dots,y_{12}\in\{0,1\}^n} \langle\eta|y_0\rangle \cdot \prod_{i=1}^{12}\langle y_{i-1}|R_i|y_i\rangle \cdot \langle y_{12}|R_{13}|x\rangle.$$

For a state $|\eta\rangle$ and Clifford unitary C, if $|\mathrm{Re}(\langle\eta|C|x\rangle)| > 2^{-n/2}\delta$ then $\mathrm{sgnRe}\Big(\widetilde{\langle\eta|C|x\rangle}\Big) =$

$\mathrm{sgnRe}(\langle\eta|C|x\rangle)$, so by the triangle inequality

$$\begin{aligned}\left|\mathrm{Re}\big(\langle\eta|p'_{\eta,C}\rangle\big) - \mathrm{Re}(\langle\eta|p_{\eta,C}\rangle)\right| &= \left|2^{-n/2}\sum_{x\in\{0,1\}^n}\Big(\mathrm{sgnRe}\Big(\widetilde{\langle\eta|C|x\rangle}\Big) - \mathrm{sgnRe}(\langle\eta|C|x\rangle)\Big)\mathrm{Re}(\langle\eta|C|x\rangle)\right| \\ &\le 2^{-n/2}\sum_{x\in\{0,1\}^n} 2\cdot 2^{-n/2}\delta = 2\delta.\end{aligned}$$

Therefore by Lemma 2.2.2, for all states $|\eta\rangle$ there exists a Clifford unitary C such that $\mathrm{Re}\Big(\Big\langle\eta\Big|p'_{\eta,C}\Big\rangle\Big) \ge \alpha - 2\delta$.

For $k \ge 0$, given states $|\phi'_0\rangle, \ldots, |\phi'_{k-1}\rangle$, let $|\eta'_k\rangle = |\psi\rangle - \alpha\sum_{j=0}^{k-1}\beta^j\Big|\phi'_j\Big\rangle$, and let C'_k be a Clifford unitary such that the state $|\phi'_k\rangle = \Big|p'_{\eta'_k,C'_k}\Big\rangle$ satisfies $\mathrm{Re}(\langle\eta'_k|\phi'_k\rangle) \ge (\alpha-2\delta)\||\eta'_k\rangle\| - \delta$. Let f' be the oracle that encodes, for each $0 \le j < T$, the description of C'_j and $\mathrm{sgnRe}\left(\widetilde{\Big\langle\eta'_j\Big|C'_j|x\Big\rangle}\right)$ for $x \in \{0,1\}^n$.

Computing f' in $\mathrm{poly}(n)$ space Given the description of $|\eta'_k\rangle$, a valid Clifford unitary C'_k can be found in $\mathrm{poly}(n)$ space by performing a brute-force search for a Clifford unitary C such that $\mathrm{Re}\Big(\Big\langle\eta'_k\Big|p'_{\eta'_k,C}\Big\rangle\Big) \ge (\alpha-2\delta)\||\eta'_k\rangle\|$. (The "extra" δ term in the definition of C'_k allows for floating-point error in the calculation of $\mathrm{Re}\Big(\Big\langle\eta'_k\Big|p'_{\eta'_k,C}\Big\rangle\Big) - (\alpha-2\delta)\||\eta'_k\rangle\|$ during this search.) The description of the vector $\big|\eta'_{k+1}\big\rangle = |\eta'_k\rangle - \alpha\beta^k|\phi'_k\rangle$ can be subsequently computed in $\mathrm{poly}(n)$ space. Since $|\eta'_0\rangle = |\psi\rangle$, it follows by induction that descriptions of C'_k and $\big|\eta'_{k+1}\big\rangle$ for $k \ge 0$ can be computed in $(k+1)\mathrm{poly}(n)$ space, by answering queries to individual bits of the description of $|\eta'_k\rangle$ recursively. Since $T \le \mathrm{poly}(n)$, descriptions of C'_k and $|\eta'_k\rangle$ for $0 \le k < T$ can be computed in $\mathrm{poly}(n)$ space. Finally, the value $\mathrm{sgnRe}\Big(\widetilde{\langle\eta'_k|C'_k|x\rangle}\Big)$ can by definition be computed in $\mathrm{poly}(n)$ space given descriptions of C'_k and $|\eta'_k\rangle$.

Constructing $|\psi\rangle$ using f' We now show that Eq. (2.2.1) still holds with f' substituted for f. By reasoning similar to that in Section 2.2.2, there exist a state $|\tau'\rangle$ and a string z' such that $\big(\langle 0^t\big|\otimes I\big)|\tau'\rangle = 0$ and

$$\left\|A_n^{f'}|0\ldots0\rangle - \Big(\gamma\big|0^t\big\rangle|\psi\rangle + \sqrt{1-\gamma^2}\big|\tau'\big\rangle\Big)\big|z'\big\rangle\right\| \le 1.58\cdot 2.86\big(\beta^T + \||\eta'_T\rangle\|\big).$$

We prove by induction on k that $\||\eta'_k\rangle\|^2 \le \beta^{2k} + 0.1\beta^{2T}\sum_{j=0}^{k-1}\beta^j$ for all k. The case $k=0$ holds trivially. If the claim holds for k, then (similarly to in Section 2.2.2)

$$\begin{aligned}\big\||\eta'_{k+1}\rangle\big\|^2 &\le \big\||\eta'_k\rangle\big\|^2 - 2\alpha\beta^k\big((\alpha-2\delta)\big\||\eta'_k\rangle\big\| - \delta\big) + \alpha^2\beta^{2k} \\ &= \big\||\eta'_k\rangle\big\|^2 - 2\alpha^2\beta^k\big\||\eta'_k\rangle\big\| + \alpha^2\beta^{2k} + \beta^k\cdot 2\alpha\big(2\big\||\eta'_k\rangle\big\| + 1\big)\delta.\end{aligned}$$

By the triangle inequality

$$\||\eta'_k\rangle\| = \left\||\psi\rangle - \alpha\sum_{j=0}^{k-1}\beta^j|\phi'_j\rangle\right\| \leq 1+\alpha\sum_{j=0}^{k-1}\beta^j \leq 1+\frac{\alpha}{1-\beta} = 1+\frac{1}{\gamma},$$

so recalling that $\delta = 0.01\beta^{2T}$ it holds that

$$2\alpha(2\||\eta'_k\rangle\|+1)\delta \leq 2\alpha\left(2\left(1+\frac{1}{\gamma}\right)+1\right)\cdot 0.01\beta^{2T} < 0.1\beta^{2T},$$

and therefore

$$\||\eta'_{k+1}\rangle\|^2 \leq \||\eta'_k\rangle\|^2 - 2\alpha^2\beta^k\||\eta'_k\rangle\| + \alpha^2\beta^{2k} + \beta^k\cdot 0.1\beta^{2T}.$$

The rest of the inductive argument follows by reasoning similar to that in Section 2.2.2.

Therefore $\||\eta'_T\rangle\| \leq \sqrt{\beta^{2T}+0.1\beta^{2T}/(1-\beta)} < 1.7\beta^T$, and the rest of the proof is similar to that in Section 2.2.2.

2.2.4 Alternate proof

Below we argue that in the proof of (a statement similar to) Lemma 2.2.1, instead of using states of the form $C\cdot 2^{-n/2}\sum_{x\in\{0,1\}^n}\pm|x\rangle$ where C is a Clifford unitary, we could alternatively use what we call “hash states”:

Definition 2.2.3 (Hash states)**.** A hash state is an n-qubit state $|\phi\rangle$ such that there exists a set $S\subseteq\{0,1\}^n$, with $|S| = 2^k$ a power of 2, such that $|\phi\rangle = |S|^{-1/2}\sum_{x\in S}\sigma_x|x\rangle$ where $\sigma_x\in\{1,-1\}$, and furthermore there exists a linear transformation $A:\mathbb{F}_2^n\to\mathbb{F}_2^k$ that is one-to-one on S. In particular if $k=0$ then A exists vacuously.

Remark. The resulting variant of Lemma 2.2.1 would have a lower $\mathsf{QAC}^0_\mathsf{f}$ circuit depth, but a measurement of the first t qubits would output 0^t with probability $\Theta(1/n)$ instead of $\Theta(1)$. This is not a problem for our proofs of Theorems 1.2.4 and 1.4.3 and Observation 1.3.18, but would be in our proof of Theorem 1.3.12.

A hash state $|\phi\rangle$ can be constructed with one query as follows. First prepare $\left|+^k\right\rangle$ in a register R. Then controlled on the state $|y\rangle_\mathsf{R}$ where $y\in\{0,1\}^k$, query the unique string $x\in S$ such that $Ax=y$, while simultaneously making a query to apply a phase of σ_x. Finally use A to uncompute y controlled on x, using that parity is in $\mathsf{QAC}^0_\mathsf{f}$ [49].

More generally, for $0\leq j<T$ let $|\phi_j\rangle$ be a hash state and let $A_j\in\mathbb{F}_2^{k_j\times n}$ be the linear transformation associated with $|\phi_j\rangle$. To construct $|\phi_j\rangle$ controlled on j, first construct $|+^n\rangle_\mathsf{R}$, and then proceed as above controlled on j. Here the oracle ignores the last $n-k_j$ qubits of R, and also outputs descriptions of $A_0,\ldots,A_{T-1}$. Finally uncompute $\left|+^{n-k_j}\right\rangle$ in the last $n-k_j$ qubits of R, controlled on k_j (which is implicit in the description of A_j).

All that remains is to write an arbitrary n-qubit state $|\psi\rangle$ as a linear combination of hash states, in a manner suitable to an LCU application like that in Section 2.3.1. (It will be apparent from our proof that the queries can be computed in poly(n) space given the description of $|\psi\rangle$, by reasoning similar to that in Section 2.2.3.) By writing $|\psi\rangle = |\psi_R\rangle + i|\psi_I\rangle$ where $|\psi_R\rangle$ and $|\psi_I\rangle$ are real-valued vectors, it suffices to write a *real-valued* vector with norm at most 1 as such a linear combination of hash states. To do this we will need the following lemma, which is proved using the probabilistic method:

Lemma 2.2.4. *Let $n > 0$. For all $S \subseteq \mathbb{F}_2^n$ with $|S| = 2^k$ a power of 2, there exists a matrix $A \in \mathbb{F}_2^{k\times n}$ satisfying $|\{Ax : x \in S\}| > \frac{1}{2} \cdot 2^k$.*

We remark that Alon, Dietzfelbinger, Miltersen, Petrank and Tardos [8] also investigated the properties of random linear hash functions from $S \subseteq \mathbb{F}_2^n$ to $\mathbb{F}_2^k$. However, they did not bound the number of nonempty buckets when $|S| = 2^k$.

Proof of Lemma 2.2.4. Let $A \in \mathbb{F}_2^{k\times n}$ be uniform random conditioned on having rank k. The kernel of A has dimension $n-k$ and therefore contains 2^{n-k} elements, one of which is the all-zeros vector. Therefore any fixed nonzero vector is in $\ker(A)$ with probability $p := \frac{2^{n-k}-1}{2^n-1}$.[1] We say that distinct strings $x, y \in S$ *collide* if $Ax = Ay$. Since any distinct $x, y \in S$ collide with probability $\Pr(A(x+y) = 0) = p$, the expected number of collisions is

$$\binom{|S|}{2} \cdot p = \frac{2^k(2^k-1)}{2} \cdot \frac{2^{n-k}-1}{2^n-1} = \frac{2^k-1}{2} \cdot \frac{2^n-2^k}{2^n-1} < \frac{2^k}{2}.$$

Therefore there exists a fixed matrix A with less than $2^k/2$ collisions.

Let $T = \{Ax : x \in S\}, t = |T|$ and for $y \in T$ let $S_y = \{x \in S : Ax = y\}$. The sets S_y form a partition of S, so by Jensen's inequality the number of collisions is[2]

$$\sum_{y\in T} \binom{|S_y|}{2} \geq t \cdot \binom{\sum_{y\in T}|S_y|/t}{2} = t \cdot \binom{2^k/t}{2} = \frac{2^k}{2} \cdot \left(\frac{2^k}{t} - 1\right).$$

Since $2^k/2$ is greater than the number of collisions which is at least $2^k/2 \cdot (2^k/t - 1)$, it follows that $t > 2^k/2$. □

Using Lemma 2.2.4 we prove the following:

Lemma 2.2.5. *For all n-qubit states $|\psi\rangle$ with real (standard-basis) amplitudes, there exists a hash state $|\phi\rangle$ such that $\langle\phi|\psi\rangle \geq \Omega(1/\sqrt{n})$.*

Proof of Lemma 2.2.5. Write $|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$. By a limiting argument we can assume without loss of generality that the $|\alpha_x|$ are all distinct. Let $0 \leq k \leq n$ be a parameter to be chosen later, and let S be the set of the 2^k largest elements of $\{0,1\}^n$ according to the total order defined by $x > y$ when $|\alpha_x| > |\alpha_y|$. By Lemma 2.2.4 there exists a matrix $A \in \mathbb{F}_2^{k\times n}$ such that $|\{Ax : x \in S\}| > \frac{1}{2} \cdot 2^k$.

Define a function $f : \{0,1\}^k \to \{0,1\}^n$ as follows: for all $y \in \{0,1\}^k$, if there exists $x \in S$ such that $Ax = y$ then let $f(y)$ be the lexicographically first $x \in S$ such that $Ax = y$, and otherwise let $f(y)$ be the lexicographically first $x \in \{0,1\}^n$ such that $Ax = y$. (To see that such an x exists in the latter case, note that A has rank k because the image of A has cardinality greater than 2^{k-1}.) Let $|\phi\rangle = 2^{-k/2}\sum_{x\in \operatorname{im} f} \operatorname{sgn}(\alpha_x)|x\rangle$ where $\operatorname{im} f$ denotes the image of f. Clearly $|\phi\rangle$ is a hash state, and

$$\langle\phi|\psi\rangle = 2^{-k/2}\sum_{x\in \operatorname{im} f} |\alpha_x| \geq 2^{-k/2}\sum_{x\in \operatorname{im} f\cap S} |\alpha_x| \geq 2^{-k/2} \cdot |\operatorname{im} f \cap S| \cdot \min_{x\in S} |\alpha_x|$$

[1] One way to see this is as follows. Let $x, y \in \mathbb{F}_2^n$ be nonzero vectors, and let $B \in \mathbb{F}_2^{n\times n}$ be an invertible matrix such that $Bx = y$. Then AB is distributed identically to A, so $\Pr(Ax = 0) = \Pr(ABx = 0) = \Pr(Ay = 0)$.

[2] Define $\binom{r}{2} = r(r-1)/2$ even for non-integer values of r.

$$= 2^{-k/2} \cdot |\{Ax : x \in S\}| \cdot \min_{x \in S} |\alpha_x| \geq \frac{1}{2} \cdot 2^{k/2} \cdot \min_{x \in S} |\alpha_x|.$$

For $j \in [2^n]$ let β_j be the j'th largest element of the set $\{|\alpha_x| : x \in \{0,1\}^n\}$, and let $\mu = \max_{j \in [2^n]} \beta_j \sqrt{j}$. Then

$$1 = \sum_{x \in \{0,1\}^n} \alpha_x^2 = \sum_{j=1}^{2^n} \beta_j^2 \leq \sum_{j=1}^{2^n} (\mu/\sqrt{j})^2 = \mu^2 \sum_{j=1}^{2^n} 1/j \leq O(\mu^2 n),$$

so $\mu \geq \Omega(1/\sqrt{n})$. Let $j \in [2^n]$ be such that $\mu = \beta_j \sqrt{j}$, and choose k such that $2^k \leq j < 2^{k+1}$. Then,

$$\langle \phi | \psi \rangle \geq \frac{1}{2} \cdot \sqrt{2^k} \cdot \beta_{2^k} \geq \frac{1}{2} \cdot \sqrt{j/2} \cdot \beta_j = \frac{1}{2\sqrt{2}} \cdot \mu \geq \Omega(1/\sqrt{n}). \qquad \square$$

2.3 State synthesis algorithms: removing the postselection

In this section we formally state and prove both the non-clean, one-query version and the clean, four-query version of Theorem 1.2.4. We also give a clean, ten-query state synthesis algorithm that has two advantages compared to the four-query algorithm. First, the ten-query algorithm is simpler. Second, the oracle in the ten-query algorithm requires fewer input bits, which will be relevant when we prove circuit upper bounds for approximately constructing arbitrary states (i.e. Theorem 1.3.12).

All of these algorithms invoke the algorithm from Lemma 2.2.1. Let γ be the constant from Lemma 2.2.1. Given an n-qubit state $|\psi\rangle$ and parameter ε, when we say "define $A, f, |\tau\rangle, z, t$ as in Lemma 2.2.1 with respect to $|\psi\rangle$ and error tolerance ε", we mean that A is the circuit A_n from Lemma 2.2.1 and all other variables have the same meaning as in Lemma 2.2.1. We will write $\varepsilon = \varepsilon(n)$ and $t = t(n)$ when n is implicit. In the four- and ten-query algorithms not all of the queries will be to precisely the same function, but this can easily be addressed as discussed shortly after Question 1.2.1.

2.3.1 The one-query algorithm

We prove the following by using parallel repetition to boost the success probability from Lemma 2.2.1:

Theorem 2.3.1. *Let ε be a function such that $\varepsilon(n) \geq \exp(-\mathrm{poly}(n))$ and $\varepsilon(n)$ is computable in $\mathrm{poly}(n)$ time for all n. Then there is a uniform sequence of $\mathrm{poly}(n)$-qubit $\mathsf{QAC}^0_\mathsf{f}$ circuits $(C_n)_n$, each making one query to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle $f = f_{|\psi\rangle}$ such that the reduced state ρ on the first n qubits of $C_n^f|0\ldots0\rangle$ satisfies $\mathrm{td}(\rho, \psi) \leq \varepsilon(n)$. Furthermore there is an algorithm that takes as input the description of an n-qubit state $|\psi\rangle$ and a string x, runs in $\mathrm{poly}(n)$ space, and outputs $f_{|\psi\rangle}(x)$.*

Proof. Let $s = \lceil 2\ln(2/\varepsilon)/\gamma^2 \rceil \leq \mathrm{poly}(n)$, and define $A, f, |\tau\rangle, z, t$ as in Lemma 2.2.1 with respect to $|\psi\rangle$ and error tolerance $\varepsilon/(2s) \geq \exp(-\mathrm{poly}(n))$. The algorithm is presented in Procedure 2, where A_k is a t-qubit register, B_k is an n-qubit register, and C_k is a $|z|$-qubit register for all $k \in [s]$. Procedure 2 is phrased in terms of multiple parallel queries, but these can be merged into a single query using Eq. (2.1.1). The $\mathsf{QAC}^0_\mathsf{f}$ implementation of Line 5 uses Lemma A.2.1.

Procedure 2 One-query state synthesis

```
1: for $k \in [s]$ in parallel do
2:     Apply $A^f$ in $\mathsf{A}_k\mathsf{B}_k\mathsf{C}_k$.        ▷ Merge queries using Eq. (2.1.1).
3: end for
4: controlled on the classical state $|x_1\rangle_{\mathsf{A}_1} \cdots |x_s\rangle_{\mathsf{A}_s}$
5:     if there exists $k$ such that $x_k = 0^t$ then return $\mathsf{B}_k$ for the smallest such $k$.
6:     else return an arbitrary $n$-qubit state.
7:     end if
8: end control
```

Let

$$|\tilde{\varphi}\rangle = \bigotimes_{k=1}^{s} \Big(\gamma |0^t\rangle_{\mathsf{A}_k} |\psi\rangle_{\mathsf{B}_k} + \sqrt{1-\gamma^2} |\tau\rangle_{\mathsf{A}_k\mathsf{B}_k}\Big) |z\rangle_{\mathsf{C}_k},$$

and let $\tilde{\rho}$ denote the n-qubit output state produced by running Lines 4 to 8 on $|\tilde{\varphi}\rangle$. If the A_k registers of $|\tilde{\varphi}\rangle$ are measured in the standard basis, then the probability that none of the measurement outcomes are 0^t is $\left(1-\gamma^2\right)^s \leq \exp(-\gamma^2 s)$, so by Eq. (1.6.2)

$$\mathrm{td}(\psi, \tilde{\rho}) \leq \exp(-\gamma^2 s/2) \leq \varepsilon/2.$$

Let $|\varphi\rangle$ denote the state of the system after Line 3. Then by Eqs. (1.6.1) and (1.6.3) and the triangle inequality

$$\mathrm{td}(\tilde{\rho}, \rho) \leq \mathrm{td}(\tilde{\varphi}, \varphi) \leq \| |\tilde{\varphi}\rangle - |\varphi\rangle \| \leq s \cdot \varepsilon/(2s) = \varepsilon/2,$$

so by the triangle inequality

$$\mathrm{td}(\rho, \psi) \leq \mathrm{td}(\rho, \tilde{\rho}) + \mathrm{td}(\tilde{\rho}, \psi) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \qquad \square$$

2.3.2 The ten-query algorithm

We prove the following by using amplitude amplification to boost the success probability from Lemma 2.2.1:

Theorem 2.3.2. *Let ε be a function such that $\varepsilon(n) \geq \exp(-\mathrm{poly}(n))$ and $\varepsilon(n)$ is computable in $\mathrm{poly}(n)$ time for all n. Then there is a uniform sequence of $\mathrm{poly}(n)$-qubit $\mathsf{QAC}^0_\mathsf{f}$ circuits $(C_n)_n$, each making ten queries to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle $f = f_{|\psi\rangle}$ such that $\left\| C^f_n |0\ldots0\rangle - |\psi\rangle|0\ldots0\rangle \right\| \leq \varepsilon(n)$. Furthermore there is an algorithm that takes as input the description of an n-qubit state $|\psi\rangle$ and a string x, runs in $\mathrm{poly}(n)$ space, and outputs $f_{|\psi\rangle}(x)$.*

Proof. Define $A, f, |\tau\rangle, z, t$ as in Lemma 2.2.1 with respect to $|\psi\rangle$ and error tolerance $\varepsilon/(9\sqrt{2})$. Since $\sin(\pi/18) < 0.174 < 0.18 < \gamma$, there exists a one-qubit gate G such that

$$G|0\rangle = \frac{\sin(\pi/18)}{\gamma}|0\rangle + \sqrt{1 - \left(\frac{\sin(\pi/18)}{\gamma}\right)^2}|1\rangle.$$

Let $|\theta\rangle = \left(G \otimes A^f\right)|0\ldots0\rangle$. The algorithm is described in Procedure 3.[3]

[3] Inspection of the proof of Lemma 2.2.1 reveals that the last query (i.e. uncomputing z) can be computed

Procedure 3 Ten-query state synthesis

```
Construct $\big((2\theta - I)\cdot\big((I - 2\big|0^{1+t}\big\rangle\big\langle 0^{1+t}\big|)\otimes I\big)\big)^4|\theta\rangle$.
Use one more query to uncompute $z$.
```

By Lemma 2.1.1 it suffices to prove that if we substitute the state

$$\left|\tilde{\theta}\right\rangle = G|0\rangle \otimes \Big(\gamma\big|0^t\big\rangle|\psi\rangle + \sqrt{1-\gamma^2}|\tau\rangle\Big)|z\rangle$$

for each occurrence of $|\theta\rangle$ in Procedure 3, then the output state is exactly $|\psi\rangle|0\dots0\rangle$. Since $\big(\big\langle 0^{1+t}\big|\otimes I\big)\left|\tilde{\theta}\right\rangle = \sin(\pi/18)|\psi\rangle|z\rangle$, we may write

$$\left|\tilde{\theta}\right\rangle = \big(\sin(\pi/18)\big|0^{1+t}\big\rangle|\psi\rangle + \cos(\pi/18)|\varphi\rangle\big)|z\rangle$$

for some state $|\varphi\rangle$ such that $\big(\big\langle 0^{1+t}\big|\otimes I\big)|\varphi\rangle = 0$. By well-known arguments (cf. the proof of correctness of Grover's algorithm [80]) it follows that if $\left|\tilde{\theta}\right\rangle$ is substituted for $|\theta\rangle$, then the output state is

$$\big(\sin(9\cdot\pi/18)\big|0^{1+t}\big\rangle|\psi\rangle + \cos(9\cdot\pi/18)|\varphi\rangle\big)|0\dots0\rangle = |\psi\rangle|0\dots0\rangle. \qquad \square$$

2.3.3 The four-query algorithm

The following statement is identical to Theorem 2.3.2 except with "four" instead of "ten", and is proved by a combination of the ideas from Sections 2.3.1 and 2.3.2:

Theorem 2.3.3. *Let ε be a function such that $\varepsilon(n) \geq \exp(-\mathrm{poly}(n))$ and $\varepsilon(n)$ is computable in $\mathrm{poly}(n)$ time for all n. Then there is a uniform sequence of $\mathrm{poly}(n)$-qubit $\mathsf{QAC}^0_\mathsf{f}$ circuits $(C_n)_n$, each making four queries to a classical oracle, such that for every n-qubit state $|\psi\rangle$ there exists a classical oracle $f = f_{|\psi\rangle}$ such that $\left\|C_n^f|0\dots0\rangle - |\psi\rangle|0\dots0\rangle\right\| \leq \varepsilon(n)$. Furthermore there is an algorithm that takes as input the description of an n-qubit state $|\psi\rangle$ and a string x, runs in $\mathrm{poly}(n)$ space, and outputs $f_{|\psi\rangle}(x)$.*

Proof. Let $\delta = \sqrt{1-\gamma^2}$. Let s be the smallest power of 2 that is at least $\log(4/\varepsilon)/\log(1/\delta)$, and observe that $s \leq 2\log(4/\varepsilon)/\log(1/\delta) \leq \mathrm{poly}(n)$. Define $A, f, |\tau\rangle, z, t$ as in Lemma 2.2.1 with respect to $|\psi\rangle$ and error tolerance $\varepsilon/(\sqrt{2}\cdot 8s)$.

First we show how to approximately construct $|\tau\rangle|z\rangle$ with three queries using amplitude amplification. Since $\sin(\pi/6) = 1/2 < 0.98 \approx \delta$, there exists a one-qubit gate G such that

$$G|0\rangle = \frac{\sin(\pi/6)}{\delta}|0\rangle + \sqrt{1 - \left(\frac{\sin(\pi/6)}{\delta}\right)^2}|1\rangle.$$

Let $|\theta\rangle = \big(G\otimes A^f\big)|0\dots0\rangle$ and

$$U^f = (2\theta - I)\big((I - 2|0\rangle\langle 0| \otimes \big(I - \big|0^t\big\rangle\big\langle 0^t\big|\big)\big)\otimes I\big)\Big(G\otimes A^f\Big).$$

in $\mathrm{poly}(n)$ space, as required by the theorem. The idea to use G to artificially decrease the initial "success" amplitude was suggested to us by Wiebe [102].

Then U^f can be efficiently implemented with three queries, and if A^f *exactly* constructs $\big(\gamma\big|0^t\big\rangle|\psi\rangle + \delta|\tau\rangle\big)|z\rangle$ then U^f *exactly* constructs $|0\rangle|\tau\rangle|z\rangle$ by reasoning similar to that in Section 2.3.2.

Since

$$\sum_{k=0}^{s-1} \delta^k |k\rangle = \Big(|0\rangle + \delta^{s/2}|1\rangle\Big) \otimes \Big(|0\rangle + \delta^{s/4}|1\rangle\Big) \otimes \cdots \otimes (|0\rangle + \delta|1\rangle),$$

there exists a tensor product L of one-qubit gates such that

$$L\Big|0^{\log s}\Big\rangle = \frac{\gamma}{\sqrt{1-\delta^{2s}}} \sum_{k=0}^{s-1} \delta^k |k\rangle.$$

The algorithm is presented in Procedure 4, where A_k is a t-qubit register, B_k is an n-qubit register, and C_k is a $|z|$-qubit register for all $0 \le k < s$; additionally K is a $(\log s)$-qubit register and O is an n-qubit register. The extra ancilla qubit in Line 12 accounts for the fact that U^f acts on one more qubit than A^f does. The $\mathsf{QAC}^0_\mathsf{f}$ implementation of Line 9 uses Lemma A.2.1, and the $\mathsf{QAC}^0_\mathsf{f}$ implementations of Lines 12 and 13 use Lemma A.2.2.[4]

Procedure 4 Four-query state synthesis

```
 1: for $0 \le k < s$ in parallel do
 2:     Apply $A^f$ in $\mathsf{A}_k\mathsf{B}_k\mathsf{C}_k$.                ▷ Merge queries using Eq. (2.1.1).
 3: end for
 4: controlled on the classical state $|x_0\rangle_{\mathsf{A}_0} \cdots |x_{s-1}\rangle_{\mathsf{A}_{s-1}}$
 5:     if there exists $k$ such that $x_k = 0^t$ then $\mathsf{K} \leftarrow$ the smallest such $k$.
 6:     end if
 7: end control
 8: controlled on the classical state $|k\rangle_\mathsf{K}$
 9:     Swap $\mathsf{B}_k$ and $\mathsf{O}$.
10:     Uncompute $|z\rangle_{\mathsf{C}_k}$, controlled on $|z\rangle_{\mathsf{C}_j}$ for some $j \ne k$.
11:     for $0 \le j < s$ in parallel do                         ▷ Merge queries using Eq. (2.1.1).
12:         if $j < k$ then then apply $\big(U^f\big)^\dagger$ in $\mathsf{A}_j\mathsf{B}_j\mathsf{C}_j$ (with one extra ancilla qubit).
13:         else if $j > k$ then then apply $\big(A^f\big)^\dagger$ in $\mathsf{A}_j\mathsf{B}_j\mathsf{C}_j$.
14:         end if
15:     end for
16: end control
17: Apply $L^\dagger$ in $\mathsf{K}$.
18: return $\mathsf{O}$.
```

Assume for now that A^f *exactly* constructs $\big(\gamma\big|0^t\big\rangle|\psi\rangle + \delta|\tau\rangle\big)|z\rangle$. Let $|\Psi_\ell\rangle$ denote the state of the system after line ℓ, up to omitting registers in the all-zeros state for brevity. Then

$$|\Psi_3\rangle = \bigotimes_{k=0}^{s-1} A^f|0\ldots0\rangle_{\mathsf{A}_k\mathsf{B}_k\mathsf{C}_k} = \bigotimes_{k=0}^{s-1}\Big(\gamma\big|0^t\big\rangle_{\mathsf{A}_k}|\psi\rangle_{\mathsf{B}_k} + \delta|\tau\rangle_{\mathsf{A}_k\mathsf{B}_k}\Big)|z\rangle_{\mathsf{C}_k},$$

[4] Although not directly implied by Lemma A.2.2, inspection of the proof of Lemma A.2.2 reveals that if $(B_n)_n$ is a *uniform* sequence of polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuits then $(\text{ctrl-}B_n)_n$ can be implemented by a *uniform* sequence of polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuits.

so

$$\left\| |\Psi_7\rangle - \sum_{k=0}^{s-1} \delta^k \gamma |k\rangle_{\mathsf{K}} \otimes \bigotimes_{j=0}^{k-1} |\tau\rangle_{\mathsf{A}_j\mathsf{B}_j} |z\rangle_{\mathsf{C}_j} \otimes \left|0^t\right\rangle_{\mathsf{A}_k} |\psi\rangle_{\mathsf{B}_k} |z\rangle_{\mathsf{C}_k} \otimes \bigotimes_{j=k+1}^{s-1} A^f |0\dots 0\rangle_{\mathsf{A}_j\mathsf{B}_j\mathsf{C}_j} \right\| \le \delta^s,$$

so

$$\left\| |\Psi_{16}\rangle - |\psi\rangle_{\mathsf{O}} \otimes \sum_{k=0}^{s-1} \delta^k \gamma |k\rangle_{\mathsf{K}} \otimes |0\dots 0\rangle_{\mathsf{A}_0\mathsf{B}_0\mathsf{C}_0\cdots\mathsf{A}_{s-1}\mathsf{B}_{s-1}\mathsf{C}_{s-1}} \right\| \le \delta^s,$$

so

$$\left\| |\Psi_{18}\rangle - \sqrt{1-\delta^{2s}} |\psi\rangle |0\dots 0\rangle \right\| \le \delta^s.$$

By the triangle inequality it follows that

$$\| |\Psi_{18}\rangle - |\psi\rangle|0\dots 0\rangle \| \le 1 - \sqrt{1-\delta^{2s}} + \delta^s \le \delta^{2s} + \delta^s \le 2\delta^s.$$

Now remove the assumption that A^f constructs $\left(\gamma\left|0^t\right\rangle|\psi\rangle + \delta|\tau\rangle\right)|z\rangle$ exactly. Procedure 4 makes $4s$ queries to ctrl-A^f and its inverse, so by Lemma 2.1.1 it follows that the *actual* output state $|\Psi_{18}\rangle$ satisfies

$$\| |\Psi_{18}\rangle - |\psi\rangle|0\dots 0\rangle \| \le 2\delta^s + \sqrt{2} \cdot 4s \cdot \varepsilon / (\sqrt{2} \cdot 8s) \le \varepsilon/2 + \varepsilon/2 = \varepsilon. \qquad \square$$

2.4 U-qRAMs and unitary synthesis

We repeat the following definition for convenience:

Definition 1.2.6 (U-qRAM)**.** Given an n-qubit unitary U, call a unitary A acting on $m \ge 2n$ qubits a *U-qRAM* if $A|x, 0^{m-n}\rangle = |x\rangle \otimes U|x\rangle \otimes \left|0^{m-2n}\right\rangle$ for all $x \in \{0,1\}^n$.

In Section 2.4.1 we reduce the task of implementing U to that of implementing a U-qRAM. Then in Section 2.4.2 we obtain an upper bound for the unitary synthesis problem as a corollary, and also present a weaker upper bound for the unitary synthesis problem using quantum teleportation. Finally in Section 2.4.3 we prove a matching query lower bound for implementing U using a U-qRAM.

2.4.1 Upper bound for implementing U given a U-qRAM

We will use a variant of Grover search that finds the marked string with certainty rather than just with high probability:

Lemma 2.4.1. *There is a uniform sequence of* $\mathsf{QAC_f}$ *circuits* $(G_n)_n$*—each of depth* $O\left(2^{n/2}\right)$*, making* $O\left(2^{n/2}\right)$ *queries, and with one ancilla—such that for all* $x \in \{0,1\}^n$ *it holds that* $G_n^{I-2|x,1\rangle\langle x,1|}$ *cleanly, exactly constructs* $|x\rangle$*.*

Imre and Balázs [58] survey several proofs of the query upper bound from Lemma 2.4.1 in detail, and below we include an alternate proof of Lemma 2.4.1 due to Wiebe [102]:

Proof. Let $x \in \{0,1\}^n$ denote the marked string that we wish to find, and let

$$t = \left\lceil \frac{\pi}{4} 2^{n/2} \right\rceil, \qquad \theta = \frac{\pi/2}{2t+1}, \qquad p = 2^n \sin^2 \theta.$$

Since $p \leq 2^n\theta^2 \leq 2^n\left(\frac{\pi/2}{2t}\right)^2 \leq 1$ we can define states

$$|\psi_0\rangle = |+^n\rangle \otimes \left(\sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle\right), \qquad |\psi_t\rangle = ((2|\psi_0\rangle\langle\psi_0| - I)(I - 2|x,1\rangle\langle x,1|))^t|\psi_0\rangle,$$

and since $\langle x,1|\psi_0\rangle = 2^{-n/2}\sqrt{p} = \sin\theta$ we may write $|\psi_0\rangle = \cos\theta|\alpha\rangle + \sin\theta|\beta\rangle$ where $|\beta\rangle = |x,1\rangle$ and $|\alpha\rangle$ is a superposition of standard basis states besides $|x,1\rangle$. By reasoning similar to that in the proof of correctness of Grover's algorithm [27, 80] it follows that

$$|\psi_t\rangle = \cos((2t+1)\theta)|\alpha\rangle + \sin((2t+1)\theta)|\beta\rangle = \cos(\pi/2)|\alpha\rangle + \sin(\pi/2)|\beta\rangle = |x,1\rangle. \qquad \square$$

Now we prove the following:

Theorem 2.4.2 (formal version of Theorem 1.2.7)**.** *There is a uniform family of* $\mathsf{QAC_f}$ *circuits* $(C_{n,m})_{n,m}$ *for* $m \geq 2n$*—each of depth* $O(2^{n/2})$*, making* $O(2^{n/2})$ *queries to an* m*-qubit quantum oracle, and acting on* $O(m)$ *qubits—such that for all* n*-qubit unitaries* U *and all* m*-qubit* U*-qRAMs* A *it holds that* $C_{n,m}^A$ *cleanly, exactly implements* U*.*

Proof. By linearity we may assume that the input is a string $x \in \{0,1\}^n$; our goal is to output $U|x\rangle \otimes |0\dots0\rangle$. First apply A, yielding $|x\rangle \otimes U|x\rangle \otimes |0\dots0\rangle$. The challenge now is to uncompute $|x\rangle$.

Let

$$C = (A \otimes I_1)\big(I_n \otimes (I_{m-n+1} - 2\big|0^{m-n},1\big\rangle\big\langle 0^{m-n},1\big|)\big)\left(A^\dagger \otimes I_1\right),$$

and observe that C can be implemented by a $\mathsf{QAC_f^0}$ circuit making two queries. By the definition of A we have that

$$\begin{aligned} C &= (A \otimes I_1)\left(I_{m+1} - 2\sum_{y\in\{0,1\}^n}|y\rangle\langle y| \otimes \big|0^{m-n}\big\rangle\big\langle 0^{m-n}\big| \otimes |1\rangle\langle 1|\right)\left(A^\dagger \otimes I_1\right) \\ &= I_{m+1} - 2\sum_{y\in\{0,1\}^n}|y\rangle\langle y| \otimes U|y\rangle\langle y|U^\dagger \otimes \big|0^{m-2n}\big\rangle\big\langle 0^{m-2n}\big| \otimes |1\rangle\langle 1| \\ &= I_{m+1} - 2\sum_{y\in\{0,1\}^n}|y,1\rangle\langle y,1| \otimes U|y\rangle\langle y|U^\dagger \otimes \big|0^{m-2n}\big\rangle\big\langle 0^{m-2n}\big|, \end{aligned}$$

where in the last line we reorder the qubits for future convenience. Therefore

$$C\big(I_{n+1} \otimes U|x\rangle \otimes \big|0^{m-2n}\big\rangle\big) = (I_{n+1} - 2|x,1\rangle\langle x,1|) \otimes U|x\rangle \otimes \big|0^{m-2n}\big\rangle,$$

so using our copy of $U|x\rangle$, the circuit C can implement the reflection $I_{n+1} - 2|x,1\rangle\langle x,1|$ in a disjoint register without disturbing the copy of $U|x\rangle$.

We could therefore simulate the circuit G_n from Lemma 2.4.1, with queries to $I - 2|x,1\rangle\langle x,1|$ answered in this manner, to construct a copy of $|x\rangle$. Instead perform this simulation *in reverse*, to *uncompute* the existing copy of $|x\rangle$ while preserving the copy of $U|x\rangle$. Finally swap $U|x\rangle$ into the appropriate register. $\square$

2.4.2 Upper bound for unitary synthesis

Theorem 2.4.3 (formal version of Theorem 1.2.5)**.** *Let* $\varepsilon(n) = \exp(-\mathrm{poly}(n))$*. Then there is a uniform sequence of* $\mathsf{QAC_f}$ *circuits* $(C_n)_n$*—each of depth* $O(2^{n/2})$*, making* $O(2^{n/2})$

queries, and with $\mathrm{poly}(n)$ *ancillae—such that for all n-qubit unitaries U there exists a classical oracle f such that* $\left\| C_n^f(I_n \otimes |0\dots0\rangle) - U \otimes |0\dots0\rangle \right\| \le \varepsilon(n)$.

Proof. By generalizing Theorem 2.3.3 from states to qRAMs (using Eq. (2.1.2)), there is a uniform sequence of $\mathrm{poly}(n)$-qubit $\mathsf{QAC}^0_{\mathsf{f}}$ circuits $(A_n)_n$, each making four queries to a classical oracle, such that for all n-qubit unitaries U there exists a classical oracle f such that

$$\max_{x\in\{0,1\}^n} \left\| A_n^f |x, 0^{n+m}\rangle - |x\rangle \otimes U|x\rangle \otimes |0^m\rangle \right\| \le \varepsilon / \left(c2^n \cdot \sqrt{2} \right)$$

for $m = \mathrm{poly}(n)$. Here c is a constant such that the circuit in Theorem 2.4.2 makes at most $c2^{n/2}$ queries. Since the operator norm of a matrix is at most the Frobenius norm, it follows that

$$\begin{aligned}\left\| A_n^f \big(I_n \otimes |0^{n+m}\rangle\big) - \sum_{x\in\{0,1\}^n} |x\rangle\langle x| \otimes U|x\rangle \otimes |0^m\rangle \right\| &\le \sqrt{\sum_{x\in\{0,1\}^n} \left\| A_n^f |x, 0^{n+m}\rangle - |x\rangle \otimes U|x\rangle \otimes |0^m\rangle \right\|^2} \\ &\le \varepsilon / \left(c2^{n/2} \cdot \sqrt{2} \right),\end{aligned}$$

so the result follows by Theorem 2.4.2 and Lemma 2.1.1. □

2.4.3 Lower bound for implementing U given a U-qRAM

For linear transformations L, M from n qubits to m qubits where $n \le m$ let $\langle L, M\rangle = 2^{-n}\,\mathrm{tr}\big(L^\dagger M\big)$, i.e. $\langle \cdot, \cdot \rangle$ is the Frobenius inner product normalized such that $\langle A, A\rangle = 1$ for all isometries A.

Theorem 2.4.4 (formal version of Theorem 1.2.8)**.** *For all sequences of quantum circuits $(C_n)_n$ making $o\big(2^{n/2}\big)$ queries to a $2n$-qubit quantum oracle, with probability $1 - o(1)$ over a Haar random n-qubit unitary U, there exists a $2n$-qubit U-qRAM A such that for all states $|\psi\rangle$,*

$$\left| \big\langle C_n^A (I_n \otimes |0\dots0\rangle), U \otimes |\psi\rangle \big\rangle \right| \le o(1).$$

Proof. For a permutation σ of $\{0,1\}^n$ let A_σ be the unitary defined by $A_\sigma|x, y\rangle = |x, y \oplus \sigma(x)\rangle$ for all $x, y \in \{0,1\}^n$. Nayak [79, Corollary 1.2] proved that any quantum circuit making $o\big(2^{n/2}\big)$ queries to A_σ outputs $\sigma^{-1}(0^n)$ with probability less than $1/2$, where the probability is over a uniform random permutation σ of $\{0,1\}^n$ as well as the randomness of the output measurement. (We remark that Ambainis [10] previously proved a similar result using different techniques.)

Let $\varepsilon, \delta > 0$ be universal constants, and assume for the sake of contradiction that there exists a quantum circuit C making $o\big(2^{n/2}\big)$ queries to a $2n$-qubit quantum oracle, such that with probability at least ε over a Haar random n-qubit unitary U, for all $2n$-qubit U-qRAMs A, there exists a state $|\psi\rangle$ such that $\left| \big\langle C^A (I_n \otimes |0\dots0\rangle), U \otimes |\psi\rangle \big\rangle \right| \ge \delta$. We prove that there exists a quantum oracle circuit making $o\big(2^{n/2}\big)$ queries to A_σ that outputs $\sigma^{-1}(0^n)$ with probability $\Omega(1)$, where the probability is over a uniform random permutation σ of $\{0,1\}^n$ as well as the randomness of the output measurement. By executing this circuit constantly many times until it outputs $\sigma^{-1}(0^n)$, we can boost the success probability to be greater than $1/2$ which contradicts Nayak's result. Therefore no such circuit C exists.

Write $C = C_s Q_s C_{s-1} Q_{s-1} \cdots C_0$ for $s = o\big(2^{n/2}\big)$, where each C_i is a unitary and each Q_i is a placeholder for either a forward or backward query. For an n-qubit unitary R, define a quantum circuit C_R by replacing each forward query Q_i in C with $(I_n \otimes R)Q_i$, and replacing each backward query Q_i in C with $Q_i\big(I_n \otimes R^\dagger\big)$. For a permutation σ of $\{0,1\}^n$ let P_σ denote the corresponding permutation matrix on n qubits, i.e. $P_\sigma|x\rangle = |\sigma(x)\rangle$ for all $x \in \{0,1\}^n$. Clearly for all R, σ it holds that $C_R^{A_\sigma} = C^{(I_n \otimes R)A_\sigma}$, and that $(I_n \otimes R)A_\sigma$ is a $2n$-qubit RP_σ-qRAM. If σ is fixed and R is Haar random, then RP_σ is also Haar random and so

$$\mathrm{Pr}_R\Big(\Big|\Big\langle C_R^{A_\sigma}(I_n \otimes |0\dots0\rangle), RP_\sigma \otimes |\psi\rangle\Big\rangle\Big| \geq \delta\Big) \geq \varepsilon.$$

Call a fixed unitary R "good with respect to σ" if $\Big|\Big\langle C_R^{A_\sigma}(I_n \otimes |0\dots0\rangle), RP_\sigma \otimes |\psi\rangle\Big\rangle\Big| \geq \delta$. Also let $D_R = C_R^\dagger(R \otimes |\psi\rangle)$. (For intuition, if R is good with respect to σ then $C_R^{A_\sigma}$ approximately implements RP_σ, and so D_R approximately implements $P_{\sigma^{-1}}$.) If R is good with respect to σ then

$$\begin{aligned}
\delta &\leq \left|2^{-n}\,\mathrm{tr}\Big((I_n \otimes \langle 0\dots0|)\Big(C_R^{A_\sigma}\Big)^\dagger (RP_\sigma \otimes |\psi\rangle)\Big)\right| && \text{(definition of } \langle\cdot,\cdot\rangle)\\
&= \left|2^{-n}\sum_{x\in\{0,1\}^n}\langle x, 0\dots0|\Big(C_R^{A_\sigma}\Big)^\dagger (RP_\sigma|x\rangle \otimes |\psi\rangle)\right| && \text{(definition of trace)}\\
&\leq 2^{-n}\sum_{x\in\{0,1\}^n}\Big|\langle x, 0\dots0|D_R^{A_\sigma}|\sigma(x)\rangle\Big| && \text{(triangle ineq., definitions of } D_R, P_\sigma)\\
&\leq 2^{-n}\sum_{x\in\{0,1\}^n}\Big\|\big(\langle\sigma^{-1}(x)| \otimes I\big)D_R^{A_\sigma}|x\rangle\Big\| && \text{(Cauchy-Schwarz, } x \leftarrow \sigma^{-1}(x)).
\end{aligned}$$

For $x \in \{0,1\}^n$ let $p_{\sigma,R,x}$ be the probability that if we run $D_R^{A_\sigma}$ on input x and measure the first n qubits of the output state, then the result is $\sigma^{-1}(x)$. Then we can phrase the above inequality as $\delta \leq 2^{-n}\sum_{x\in\{0,1\}^n}\sqrt{p_{\sigma,R,x}}$, and by Cauchy-Schwarz it follows that $\delta^2 \leq 2^{-n}\sum_{x\in\{0,1\}^n} p_{\sigma,R,x}$.

Therefore for every fixed permutation σ of $\{0,1\}^n$, for Haar random R and uniform random $x \in \{0,1\}^n$, it holds that

$$\varepsilon \leq \Pr_R(R \text{ is good w.r.t. } \sigma) \leq \Pr_R\Big(\delta^2 \leq \mathop{\mathbb{E}}_x[p_{\sigma,R,x}]\Big) \leq \delta^{-2}\mathop{\mathbb{E}}_{R,x}[p_{\sigma,R,x}]$$

where the last step is by Markov's inequality. If we also take σ to be uniform random then $\mathbb{E}_{\sigma,R,x}[p_{\sigma,R,x}] \geq \varepsilon\delta^2$, so there exist *fixed* values of R and x such that $\mathbb{E}_\sigma[p_{\sigma,R,x}] \geq \varepsilon\delta^2$. Thus there exists a quantum circuit (specifically $D_R^{A_\sigma}|x\rangle$ for these fixed values of R and x) making $o\big(2^{n/2}\big)$ queries to A_σ that outputs $\sigma^{-1}(x)$ with probability at least $\varepsilon\delta^2$, where the probability is over a uniform random permutation σ of $\{0,1\}^n$ as well as the randomness of the output measurement. By symmetry such a circuit exists with $x = 0^n$ as desired. □

Chapter 3

Bounds on the QAC^0 complexity of parity and fanout

> Cats and tomatoes don't mix!
>
> *Code Geass* S2E05

We repeat the definition of QAC^0 for convenience:

Definition 1.3.5 (QAC^0 [49])**.** A QAC circuit is a quantum circuit consisting of arbitrary one-qubit gates, as well as *generalized Toffoli gates* of arbitrary arity defined by

$$|x, b\rangle \mapsto \left| x, b \oplus \prod_{j=1}^{n} x_j \right\rangle \quad \text{for} \quad x = (x_1, \ldots, x_n) \in \{0,1\}^n, b \in \{0,1\},$$

or equivalently *generalized Z gates* of arbitrary arity defined by

$$Z = I - 2|1 \ldots 1\rangle\langle 1 \ldots 1|.$$

(The equivalence between generalized Toffoli and Z gates is illustrated in Fig. 3, and was observed by Fang, Fenner, Green, Homer and Zhang [42].) A QAC^0 circuit is a constant-depth QAC circuit.

We use the following notation:

Definition 3.0.1 (Parity, fanout, and the cat state)**.** We denote the n-qubit parity and fanout transformations respectively by

$$\begin{aligned} P_n|b, x_1, \ldots, x_{n-1}\rangle &= \left| b \oplus \bigoplus_{j=1}^{n-1} x_j, x_1, \ldots, x_{n-1} \right\rangle, \\ F_n|b, x_1, \ldots, x_{n-1}\rangle &= |b, x_1 \oplus b, \ldots, x_{n-1} \oplus b\rangle, \end{aligned}$$

and the n-qubit cat state by $\left|\text{🐱}_n\right\rangle = \frac{1}{\sqrt{2}}|0^n\rangle + \frac{1}{\sqrt{2}}|1^n\rangle$.

Observe that $\left|\text{🐱}_n\right\rangle = F_n(H \otimes I_{n-1})|0^n\rangle$, so lower bounds for constructing the cat state imply lower bounds for computing fanout (and therefore for computing parity, by Fig. 6). We also introduce the following generalizations of the cat state:

Definition 3.0.2 (Nekomata)**.** An n-*nekomata* is a purification of $\frac{1}{2}|0^n\rangle\langle 0^n| + \frac{1}{2}|1^n\rangle\langle 1^n|$, or equivalently a state of the form $\frac{1}{\sqrt{2}}|0^n\rangle|\psi_0\rangle + \frac{1}{\sqrt{2}}|1^n\rangle|\psi_1\rangle$ for some states $|\psi_0\rangle, |\psi_1\rangle$ on any number of qubits. We refer to the n qubits on which the reduced state is $\frac{1}{2}|0^n\rangle\langle 0^n|+\frac{1}{2}|1^n\rangle\langle 1^n|$ as the *targets* of an n-nekomata.

We choose the name "nekomata" because nekomata are two-tailed cats from Chinese and Japanese folklore, and the states $|\psi_0\rangle$ and $|\psi_1\rangle$ can be thought of as "tails" of the cat state.

In Section 3.1 we introduce a normal form for QAC circuits which will be used in some of our proofs. In Section 3.2 we prove that exponentially large depth-2 QAC^0 circuits can approximately construct an n-nekomata, and we prove a matching lower bound for a certain class of QAC^0 circuits (which we call "mostly classical circuits") that generalizes the circuit from our upper bound. Then in Section 3.3 we reduce the task of computing parity to that of constructing a nekomata, and obtain an exponentially large QAC^0 circuit for parity (and fanout) as a corollary. In Section 3.4 we prove that QAC^0 circuits require $\Omega(n)$ size to approximately construct an n-nekomata, and therefore to non-cleanly compute n-qubit parity and fanout. Finally in Section 3.5 we prove that QAC^0 circuits of *arbitrary* size require depth at least three to non-cleanly, approximately construct the n-qubit cat state for sufficiently large n, and therefore to non-cleanly, approximately compute n-qubit parity and fanout.

We note two subtleties in these results. First, our QAC^0 reduction from computing parity to constructing a nekomata does not preserve the "mostly classical" property of QAC^0 circuits, so our lower bound for constructing nekomata in mostly classical circuits does not imply lower bounds for parity and fanout in mostly classical circuits. Second, our depth-2 lower bound for constructing the cat state does not contradict our depth-2 upper bound for constructing a nekomata, because the latter result is for a nekomata *other than* the cat state.

3.1 A normal form for QAC circuits

Consider a QAC 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 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 *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 (recall Definition 1.3.4), and more generally the number of multi-qubit gates acting on any given set of qubits.

Definition 3.1.1 ($R_\otimes$ gates and normal form)**.** For a state $|\theta\rangle$ that factors as a tensor product of one-qubit states, let $R_{|\theta\rangle} = R_\theta = I - 2|\theta\rangle\langle\theta|$ and call this transformation an "$R_\otimes$ gate" (the R stands for "reflection"). We say that a QAC circuit is 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 one-qubit gates.

Every k-qubit $R_\otimes$ gate R_θ can be implemented by a size-1, k-qubit QAC^0 circuit, because there exists a layer L of k one-qubit gates such that $L\left|1^k\right\rangle = |\theta\rangle$ and therefore $R_\theta = LZL^\dagger$ (where $Z = I - 2\left|1^k\right\rangle\!\left\langle 1^k\right|$).

Proposition 3.1.2. *Every* QAC *circuit computes the same unitary transformation as a circuit in* $R_\otimes$ *normal form with the same topology.*

Proof. The proof is by induction on the depth d of a QAC circuit C. If $d = 0$ then C is a layer of one-qubit gates, which is already in $R_\otimes$ normal form. Otherwise write $C = LMD$ such that L is a layer of one-qubit gates, M is a layer of multi-qubit generalized Z gates, and D is a depth-$(d-1)$ QAC circuit. Since $C = LML^\dagger LD$ it suffices to prove that LD is equivalent to a circuit in $R_\otimes$ normal form with the same topology as D, and that $LML^\dagger$ is equivalent to a layer of $R_\otimes$ gates that has the same topology as M. The first claim follows by the inductive hypothesis. To prove the second claim note that $LML^\dagger = \bigotimes_j L_j Z_j L_j^\dagger = \bigotimes_j R_{L_j|1^n\rangle}$, where Z_j ranges over all generalized Z gates in M, and L_j is the tensor product of the gates in L that act on the support of Z_j. □

3.2 Bounds for constructing nekomata in "mostly classical" circuits

Definition 3.2.1 (Mostly classical circuits)**.** Call a QAC circuit *purely classical* if it consists only of generalized Toffoli gates. 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. 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 $|\langle 0\dots 0|\theta\rangle|^2 \le 1/4$.

Purely classical circuits can include NOT gates, which are generalized Toffoli gates on one qubit. Purely classical circuits are roughly analogous to AC circuits with bounded fanout. The niceness condition will allow us to express certain random variables as convex combinations of certain other random variables, by ensuring that the coefficients in these convex combinations are between 0 and 1. We prove the following:

Theorem 3.2.2 (Upper bound)**.** *For all $2 \le d \le \log n$ and $\varepsilon > 0$ there exists a nice, mostly classical, depth-d QAC circuit C acting on $\exp(O(n2^{-d}\log(n2^{-d}/\varepsilon))) + O(n)$ qubits, and an n-nekomata $|\nu\rangle$ such that $\||C|0\dots 0\rangle - |\nu\rangle\| \le \varepsilon$.*

Theorem 3.2.3 (Lower bound)**.** *Let C be a mostly classical circuit of size s and depth d.*

(i) For every n-nekomata $|\nu\rangle$,

$$|\langle\nu|C|0\dots 0\rangle|^2 \le \frac{1}{2} + \exp\left(-\Omega\left(\min\left(\frac{n/(4^d\log n)}{\log s}, \sqrt{n/(4^d\log n)}\right)\right)\right). \qquad (3.2.1)$$

(ii) If C is nice then for every n-nekomata $|\nu\rangle$,

$$|\langle\nu|C|0\dots 0\rangle|^2 \le \frac{1}{2} + \exp\left(-\Omega\left(\min\left(\frac{n/2^d}{\log s}, \sqrt{n/2^d}\right)\right)\right). \qquad (3.2.2)$$

Theorems 3.2.3(ii) and 3.2.2 imply that for $d \ge 2$, the minimum size of a nice, mostly classical, depth-d QAC circuit that constructs an n-nekomata to within 0.1 error is between $\exp\big(\Omega\big(n/2^d\big)\big)$ and $\exp\Big(\tilde{O}\big(n/2^d\big)\Big) + O(n)$. We will use only the $d = 2$ case of Theorem 3.2.2, but the general case is not much more difficult to prove. The upper bounds in Eqs. (3.2.1) and (3.2.2) are tight (up to the value being exponentiated) because the identity transformation is a nice, mostly classical circuit. By Uhlmann's theorem these bounds could also be phrased in terms of the fidelity of $\frac{1}{2}|0^n\rangle\langle 0^n| + \frac{1}{2}|1^n\rangle\langle 1^n|$ with the reduced state on the first n qubits of $C|0\dots 0\rangle$. Finally we remark that 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 bound for parity which we will prove—then an easy generalization of our proof of Theorem 3.2.3(ii) implies an identical statement with r^d in in place of 2^d (and similarly for Theorem 3.2.3(i)).

In Section 3.2.1 we make some general observations about mostly classical circuits and approximations of nekomata, including observations common to the proofs of Theorems 3.2.2 and 3.2.3. Then in Sections 3.2.2 to 3.2.4 we prove Theorems 3.2.2, 3.2.3(ii) and 3.2.3(i) respectively. The proofs of Theorems 3.2.3(i) and 3.2.3(ii) have a similar high-level idea, but the proof of Theorem 3.2.3(ii) is the simpler of the two so we present it first.

3.2.1 Reduction to a classical sampling problem

Collectively, the following observations reduce proving Theorems 3.2.2 and 3.2.3 to proving upper and lower bounds respectively for a certain type of sampling problem, which can be succinctly characterized with only a transient reference to quantum circuits.

An n-nekomata can be defined as a state for which a standard-basis measurement of the targets outputs 0^n and 1^n both with probability $1/2$. The following two lemmas make similar statements about approximations of nekomata, and are used to prove Theorems 3.2.2 and 3.2.3 respectively:

Lemma 3.2.4. *Let $|\varphi\rangle$ be a state for which a standard-basis measurement of the first n qubits outputs 0^n with probability $1/2$ and 1^n with probability $1/2-\varepsilon$. Then there exists an n-nekomata $|\nu\rangle$ such that $\||\varphi\rangle - |\nu\rangle\| \leq 2\sqrt{\varepsilon}$.*

Proof. Write

$$|\varphi\rangle = \sqrt{\frac{1}{2}} \cdot |0^n\rangle|\psi_0\rangle + \sqrt{\frac{1}{2}-\varepsilon} \cdot |1^n\rangle|\psi_1\rangle + \sqrt{\varepsilon} \cdot |\tau\rangle,$$

where $|\tau\rangle$ is a state such that $(\langle 0^n| \otimes I)|\tau\rangle = (\langle 1^n| \otimes I)|\tau\rangle = 0$, and let

$$|\nu\rangle = \frac{1}{\sqrt{2}}|0^n\rangle|\psi_0\rangle + \frac{1}{\sqrt{2}}|1^n\rangle|\psi_1\rangle.$$

Then by the triangle inequality

$$\||\varphi\rangle - |\nu\rangle\| \leq \left|\sqrt{\frac{1}{2}-\varepsilon} - \sqrt{\frac{1}{2}}\right| + \sqrt{\varepsilon} \leq 2\sqrt{\varepsilon}. \qquad \square$$

Lemma 3.2.5. *Let $|\varphi\rangle$ be a state for which a standard-basis measurement of the first n qubits outputs 0^n with probability p and 1^n with probability q. Then $|\langle\nu|\varphi\rangle|^2 \leq 1/2 + \sqrt{\min(p,q)}$ for all n-nekomata $|\nu\rangle$.*

Proof. Let $Q_b = |b^n\rangle\langle b^n| \otimes I$ for $b \in \{0,1\}$. By the triangle inequality and Cauchy-Schwarz,

$$|\langle\nu|\varphi\rangle| = |\langle\nu|(Q_0+Q_1)|\varphi\rangle| \leq \sum_{b=0}^{1} |\langle\nu|Q_b|\varphi\rangle| \leq \sum_{b=0}^{1} \|Q_b|\nu\rangle\| \cdot \|Q_b|\varphi\rangle\| = \sqrt{p/2} + \sqrt{q/2},$$

so $|\langle\nu|\varphi\rangle|^2 \leq p/2 + q/2 + \sqrt{pq} \leq 1/2 + \sqrt{\min(p,q)}$. $\square$

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 the first n 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 3.2.6. *Let $|\theta_j\rangle$ be a one-qubit state and let $p_j = |\langle 1|\theta_j\rangle|^2$ for $j \in [k]$. A standard-basis measurement of $R_{\bigotimes_j|\theta_j\rangle}\left|0^k\right\rangle$ outputs 0^k with probability $\left(1-2\prod_j(1-p_j)\right)^2$, and any other string $(y_j)_{j\in[k]}$ with probability $4\prod_j(1-p_j)\Pr(\mathrm{Bernoulli}(p_j)=y_j)$.*

Proof. Let $|\theta\rangle = \bigotimes_j|\theta_j\rangle$ and recall that $R_\theta = I - 2\theta$. Clearly

$$\left|\left\langle 0^k\middle|R_\theta\middle|0^k\right\rangle\right|^2 = \left|1-2\left\langle 0^k\middle|\theta\middle|0^k\right\rangle\right|^2 = \left|1-2\prod_j|\langle 0|\theta_j\rangle|^2\right|^2 = \left(1-2\prod_j(1-p_j)\right)^2.$$

Similarly, if $y=(y_j)_j$ is any string besides the all-zeros string, then

$$\left|\langle y|R_\theta\middle|0^k\right\rangle\right|^2 = \left|2\langle y|\theta\rangle\left\langle\theta\middle|0^k\right\rangle\right|^2 = \left|2\prod_j\langle y_j|\theta_j\rangle\langle\theta_j|0\rangle\right|^2 = 4\prod_j|\langle y_j|\theta_j\rangle|^2|\langle 0|\theta_j\rangle|^2$$
$$= 4\prod_j(1-p_j)\Pr(\mathrm{Bernoulli}(p_j)=y_j). \qquad \square$$

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

Corollary 3.2.7. *If $\prod_j(1-p_j)\le 1/4$ then the distribution from Lemma 3.2.6 is a convex combination of 0^k with probability $1-4\prod_j(1-p_j)$ and $(\mathrm{Bernoulli}(p_j))_{j\in[k]}$ with probability $4\prod_j(1-p_j)$, where the $\mathrm{Bernoulli}(p_j)$ random variables are all independent.*

Proof. $\left(1-2\prod_j(1-p_j)\right)^2 = \left(1-4\prod_j(1-p_j)\right)+4\prod_j(1-p_j)^2$, and $4\prod_j(1-p_j)^2 = 4\prod_j(1-p_j)\Pr(\mathrm{Bernoulli}(p_j)=0)$. □

3.2.2 Upper bound

Theorem 3.2.2 (Upper bound)**.** *For all $2\le d\le\log n$ and $\varepsilon>0$ there exists a nice, mostly classical, depth-d QAC circuit C acting on $\exp(O(n2^{-d}\log(n2^{-d}/\varepsilon)))+O(n)$ qubits, and an n-nekomata $|\nu\rangle$ such that $\||C|0\dots0\rangle-|\nu\rangle\|\le\varepsilon$.*

Proof. First we prove the $d=2$ case of the theorem. Let

$$M = \left\lceil\frac{\ln 2}{4}\cdot\left(\frac{4\ln(2)n}{\varepsilon^2}\right)^n\right\rceil \le \exp(O(n\log(n/\varepsilon))),$$

and let $\delta\in(0,1)$ be such that $(1-2\delta^n)^{2M}=1/2$. The circuit acts on a grid of $n\times(M+1)$ qubits as illustrated in Fig. 7, with one designated "target" column and M "ancilla"

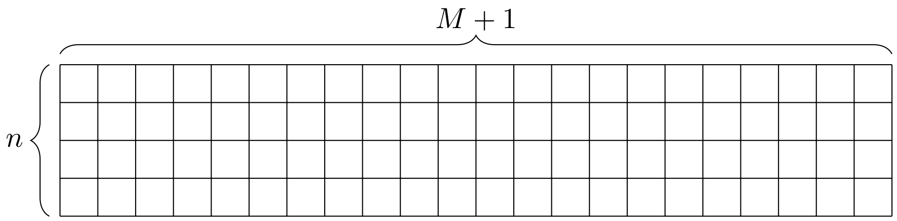

Figure 7: The qubits acted on in the depth-2 case of Theorem 3.2.2.

columns. First, in each ancilla column, apply $R_{(\sqrt{\delta}|0\rangle+\sqrt{1-\delta}|1\rangle)^{\otimes n}}$. Second, in each row, apply an $(M+1)$-qubit OR gate (Fig. 4) whose output qubit is in the target column.

All measurements described below are standard-basis measurements as discussed in Section 3.2.1. By Lemma 3.2.4 it suffices to prove that with probability exactly $1/2$ all of the ancillae measure to 0, and with probability at least $1/2-\varepsilon^2/4$ at least one ancilla in each row measures to 1. The probability that all ancillae measure to 0 equals $(1-2\delta^n)^{2M}$ by Lemma 3.2.6 and the independence of measurements of different columns, and this equals $1/2$ by the definition of δ.

Let $\varepsilon' = \varepsilon^2/4$. Below we will prove 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 3.2.6 and a union bound, the probability that there exists an ancilla column measuring to neither all-zeros nor all-ones is at most

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

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

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

Finally, the circuit is nice because $\delta^n \le \ln(2)/4M \le \ln(2)/4 < 1/4$.

Now we prove the $d>2$ case of the theorem. Use the circuit from the depth-2 case to construct a state $|\psi\rangle$ such that $\||\psi\rangle-|\nu\rangle\| \le \varepsilon$ for some m-nekomata $|\nu\rangle$ where $m=\lceil n/2^{d-2}\rceil$. Partition the m target qubits of $|\psi\rangle$ and $n-m$ new qubits into m sets, each of size at most $\lceil n/m\rceil \le 2^{d-2}$, and each including one qubit from $|\psi\rangle$. To each of these sets of qubits, apply the depth-($\le d-2$) circuit for restricted fanout from Lemma A.1.1, where the qubit being "fanned out" is the qubit from $|\psi\rangle$ in that set. □

3.2.3 Lower bound in the "nice" case

We use the following concentration inequality of Gavinsky, Lovett, Saks and Srinivasan [45]:

Definition 3.2.8 ([45])**.** Call a random string $(Y_1,\dots,Y_n)\in\{0,1\}^n$ a *read-r family* if there exist $m\in\mathbb{N}$, independent random variables $X_1,\dots,X_m$, sets $S_1,\dots,S_n\subseteq[m]$ such that $|\{j\in[n]: i\in S_j\}|\le r$ for all $i\in[m]$, and functions $f_1,\dots,f_n$ such that $Y_j=f_j((X_i)_{i\in S_j})$ for all $j\in[n]$.

Theorem 3.2.9 ([45])**.** *Let $(Y_1, \dots, Y_n)$ be a read-r family, and let $\mu = \mathbb{E}\Big[\sum_{j=1}^n Y_j\Big]$. Then for all $\varepsilon \geq 0$,*

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

For example, if $r = 1$ then $Y_1, \dots, Y_n$ are independent and so Theorem 3.2.9 recovers a well-known Chernoff bound for sums of independent Bernoulli random variables. More generally Theorem 3.2.9 recovers this Chernoff bound when $n = rm$ and $Y_j = X_{\lceil j/r \rceil}$ for all j [45].

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 3.2.3(ii), as a warmup we briefly prove the following:

Proposition 3.2.10. *If C is a depth-d purely classical circuit and $|\phi\rangle$ is a tensor product of one-qubit states, then $|\langle\nu|C|\phi\rangle|^2 \leq 1/2 + \exp\bigl(-\Omega\bigl(n/2^d\bigr)\bigr)$ for all n-nekomata $|\nu\rangle$.*

Proof. Since standard-basis measurements of unentangled qubits are independent, it follows from the above discussion that a standard-basis measurement of the first n qubits of $C|\phi\rangle$ is a read-2^d family. If the expected Hamming weight of a standard-basis measurement of these qubits is less (resp. greater) than or equal to $n/2$, then Theorem 3.2.9 implies that these qubits measure to 1^n (resp. 0^n) with probability at most $\exp(-\Omega(n/2^d))$, and the result follows by Lemma 3.2.5. □

Recall that Theorem 3.2.3(ii) says that if C is a nice, mostly classical circuit of size s and depth d then

$$|\langle\nu|C|0\dots0\rangle|^2 \leq \frac{1}{2} + \exp\left(-\Omega\left(\min\left(\frac{n/2^d}{\log s}, \sqrt{n/2^d}\right)\right)\right).$$

for all n-nekomata $|\nu\rangle$.

Proof of Theorem 3.2.3(ii). Assume without loss of generality that $s \geq \exp\Bigl(\sqrt{n/2^d}\Bigr)$. We will prove that for some $a \in \{0,1\}$, any n designated "target" qubits of $C|0\dots0\rangle$ measure to a^n with probability at most $\exp(-\Omega(n2^{-d}/\log s))$. The result then follows by Lemma 3.2.5.

Write $C = D\bigl(L \otimes \bigotimes_{G\in\mathcal{G}} G\bigr)$ such that D is purely classical, L is a layer of one-qubit gates, and $\mathcal{G}$ is a set of multi-qubit $R_\otimes$ gates that each satisfy the precondition of Corollary 3.2.7. For all $G \in \mathcal{G}$, a standard-basis measurement of $G|0\dots0\rangle$ 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

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

Therefore it suffices to prove that for some $a \in \{0,1\}$, the target qubits 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 $\Pr(S > t) < (e\mu/t)^t e^{-\mu}$ 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

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

Let y be a string of independent Bernoulli random variables distributed identically to a standard-basis measurement of $L|0\dots0\rangle$. Call the target qubits 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 target qubits 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).

For $r_1, r_2 \in \mathbb{R}$ let $r_1 \approx r_2$ if $|r_1 - r_2| \le 0.1n$. Let $W(b,z)$ be the Hamming weight of the output bits as a function of b and $z := (x,y)$, and let z' be an independent copy of z. We now argue that

$$\Pr_{b,z,z'}\big(W(b,z) \approx \mathbb{E}[W(b,\cdot) \mid b] \approx W(b,z') \approx \mathbb{E}[W(\cdot,z') \mid z']\big) \ge 1 - \exp(-\Omega(n2^{-d}/\log s)), \tag{3.2.3}$$

where the expectations are over independent copies of z and b respectively, substituted for "$\cdot$" as inputs to W. For any fixed value of b, Theorem 3.2.9 implies that $W(b,z) \approx \mathbb{E}[W(b,\cdot) \mid b]$ except with probability at most $\exp\big(-\Omega\big(n2^{-d}\big)\big)$ over z, and the same statement holds with z' in place of z. Similarly, for any fixed value of $z' = (x',y')$ such that x' is good, Theorem 3.2.9 implies that $W(b,z') \approx \mathbb{E}[W(\cdot,z') \mid z']$ except with probability at most $\exp(-\Omega(n2^{-d}/\log s))$ over b. Since x' is good except with probability at most $\exp\big(-\Omega\big(n2^{-d}/\log s\big)\big)$, Eq. (3.2.3) follows by a union bound.

Therefore by the triangle inequality,

$$\Pr_{b,z,z'}\big(\big|W(b,z) - \mathbb{E}[W(\cdot,z') \mid z']\big| \ge 0.3n\big) \le \exp(-\Omega(n2^{-d}/\log s)).$$

It follows that there exists a fixed value of z' such that the above inequality holds with the probability being over b,z. For this fixed value of z', if $\mathbb{E}[W(\cdot,z') \mid z']$ is at most (resp. at least) $n/2$, then the output bits are 1^n (resp. 0^n) with probability at most $\exp(-\Omega(n2^{-d}/\log s))$. □

3.2.4 Lower bound in the general case

Below we prove Theorem 3.2.3(i), i.e. that if CL is a mostly classical circuit of size s and depth d (where C is purely classical and L is a layer of $R_\otimes$ gates), then

$$|\langle\nu|CL|0\dots0\rangle|^2 \le \frac{1}{2} + \exp\left(-\Omega\left(\min\left(\frac{n/(4^d\log n)}{\log s}, \sqrt{n/(4^d\log n)}\right)\right)\right).$$

for all n-nekomata $|\nu\rangle$. The proof is similar to that of Theorem 3.2.3(ii), except that here our procedure for simulating a standard-basis measurement of the n designated target qubits is more complicated. For brevity we will omit some proof steps with clear analogues in the proof of Theorem 3.2.3(ii), i.e. in Section 3.2.3.

Proof. Consider a gate G in L. Write $G = R_{\bigotimes_j |\theta_j\rangle}$ for one-qubit states $(|\theta_j\rangle)_j$, and let $p_j = p_j^{(G)} = |\langle 1|\theta_j\rangle|^2$. We may assume that $p_j \neq 0$ for all j, because $G|0\dots0\rangle = \left(R_{\bigotimes_{j:p_j\neq 0}|\theta_j\rangle} \otimes I\right)|0\dots0\rangle$. Then by Lemma 3.2.6, a standard-basis measurement of $G|0\dots0\rangle$ is distributed identically to $(B \wedge X_j)_j$, where the X_j are independent Bernoulli(p_j) random variables conditioned on $(X_j)_j$ not being the all-zeros string, and

$$B \sim \text{Bernoulli}\left(4\prod_j (1-p_j) - 4\prod_j (1-p_j)^2\right)$$

is independent of $(X_j)_j$.

Let $R = (R_j)_j$ where each R_j is independently 1 with probability $1 - p_j$ and uniform random on $[0, 1)$ with probability p_j. Then $(X_j)_j$ is distributed identically to $(\mathbb{1}_{R_j<1})_j$ conditioned on $\min R < 1$. Let $\operatorname{argmin} R$ be a value of j such that $R_j = \min R$, and note that if we condition on $\min R < 1$ then $\operatorname{argmin} R$ is unique with probability 1. To sample R conditioned on $\min R < 1$, one may first sample $J = \operatorname{argmin} R$ conditioned on $\min R < 1$, next sample $\mu = \min R$ conditioned on $R_J = \min R < 1$, and finally, for all $j \neq J$, independently sample R_j conditioned on $R_j > \mu$.

Rather than sampling $(\operatorname{argmin} R \mid \min R < 1)$ (i.e. $\operatorname{argmin} R$ conditioned on $\min R < 1$) directly, we may do so as follows. Identifying C with the function from $\{0,1\}^*$ to $\{0,1\}^*$ that it computes, say that the j'th input bit "influences" the k'th output bit if there exist strings x, y differing only in position j such that $C|x\rangle$ and $C|y\rangle$ differ in position k. Recall that no input bit influences more than 2^d output bits. Let $\tau^{(G)}$ be the (non-random) tree constructed in the following two steps:

- Start with a rooted binary tree with $\binom{n}{2^d}$ leaves and depth $\lceil \log \binom{n}{2^d} \rceil$, and identify each leaf with a distinct set of 2^d targets of CL.
- Then, for each qubit v acted on by G, for some set u of 2^d targets including all of the targets influenced by v, add the node v and edge (u, v) to the tree.

For each non-leaf node u in $\tau^{(G)}$ such that $(\operatorname{argmin} R \mid \min R < 1)$ is descended from u with nonzero probability, independently "highlight" a random edge from u to one of its children, where the probability of highlighting an edge (u, v) equals the probability that $(\operatorname{argmin} R \mid \min R < 1)$ is descended from v divided by the probability that $(\operatorname{argmin} R \mid \min R < 1)$ is descended from u. Then there is a unique root-to-leaf path consisting only of highlighted edges, and the leaf at the end of this path is distributed identically to $(\operatorname{argmin} R \mid \min R < 1)$.

Altogether this implies the following procedure for simulating a standard-basis measurement of $G|0\dots0\rangle$. If G acts on a single qubit then simply output $Y^{(G)} \sim \text{Bernoulli}(|\langle 1|G|0\rangle|^2)$. Otherwise, first sample the following random variables, all independently:

- Sample $B^{(G)} \sim \text{Bernoulli}\left(4\prod_j \left(1-p_j^{(G)}\right) - 4\prod_j \left(1-p_j^{(G)}\right)^2\right)$.

- Highlight random edges in $\tau^{(G)}$, in the manner described above.
- For all j, sample $M_j^{(G)}$ from the distribution of $\min R$ conditioned on $R_j = \min R < 1$;
- For all j, sample $S_j^{(G)}$ from the uniform distribution on $[0, 1]$.

Then let $J^{(G)}$ be the leaf in the root-to-leaf path consisting of highlighted edges in $\tau^{(G)}$, and output

$$\left(\left(B^{(G)} = 1\right) \wedge \left(\left(J^{(G)} = j\right) \vee \left(S_j^{(G)} \leq \Pr\left(R_j < 1 \,\middle|\, R_j > M_{J^{(G)}}^{(G)}\right)\right)\right)\right)_j.$$

For $0 \leq k \leq \lceil \log \binom{n}{2^d} \rceil$ let $E_k^{(G)}$ be the set of highlighted edges between nodes at depths k and $k+1$ in $\tau^{(G)}$, where we define the root to have depth 0, children of the root to have depth 1, and so on. Note that $\left(E_k^{(G)}\right)_k$ is a partition of the set of highlighted edges in $\tau^{(G)}$. Let $Y = \left(Y^{(G)}\right)_G, B = \left(B^{(G)}\right)_G, E_k = \left(E_k^{(G)}\right)_G, M = \left(M_j^{(G)}\right)_{j,G}, S = \left(S_j^{(G)}\right)_{j,G}$.

Recall that s is defined as the size of CL, and assume (without loss of generality, given the theorem we are proving) that $s \geq \exp\left(\sqrt{n/(4^d \log n)}\right)$. Since $\mathbb{E}[B^{(G)}] \leq 4 \exp\left(-\sum_j p_j^{(G)}\right)$ for all G, we may assume that $\max_G \sum_j p_j^{(G)} \leq 2 \ln s$, by the same reasoning as in Section 3.2.3. Call S "good" if $\left|\left\{j : S_j^{(G)} \leq p_j^{(G)}\right\}\right| \leq c \log s$ for all G, where c is an appropriately large universal constant. As in Section 3.2.3, by a Chernoff bound, the probability that S fails to be good is at most $s^{-\Omega(1)}$. For all G,

$$\Pr\left(R_j < 1 \,\middle|\, R_j > M_{J^{(G)}}^{(G)}\right) \leq \Pr(R_j < 1) = p_j^{(G)}$$

(where the definition of R_j here implicitly depends on G), so if S is fixed and good then there are at most $O(\log s)$ indices j such that the boolean value $S_j^{(G)} \leq \Pr\left(R_j < 1 \,\middle|\, R_j > M_{J^{(G)}}^{(G)}\right)$ is not identically false.

Let $V = (Y, B, (E_k)_k, M, S)$, and note that the targets of a standard-basis measurement of $CL|0 \ldots 0\rangle$ are a read-$O(2^d \log s)$ family if $V \backslash Y$ is fixed, or if $V \backslash B$ is fixed and S is good, or if $V \backslash M$ is fixed and S is good, or if $V \backslash S$ is fixed. (We will consider the case where $V \backslash E_k$ is fixed and S is good shortly.) Let $W = W(V)$ be the Hamming weight of a standard-basis measurement of the targets of $CL|0 \ldots 0\rangle$. Then, by Lemma 3.2.5, Theorem 3.2.9, and an argument involving the triangle inequality[1] similar to that in Section 3.2.3, it suffices to prove the following:

Claim 3.2.11. *Fix $V \backslash (E_k)_k$ such that S is good, and let $\mu = \mathbb{E}[W \mid V \backslash (E_k)_k]$. Then for all $\varepsilon > 0$,*

$$\Pr(W \geq \mu + \varepsilon n) \leq \exp\left(-\Omega\left(\varepsilon^2 n / \left(4^d \log s \log n\right)\right)\right),$$
$$\Pr(W \leq \mu - \varepsilon n) \leq \exp\left(-\Omega\left(\varepsilon^2 n / \left(4^d \log s \log n\right)\right)\right),$$

where the probabilities are over $(E_k)_k$.

[1] In slightly greater detail: sample an independent copy V' of V, use a hybrid argument to show that $|W(V) - W(V')|$ is small with high probability, and then fix a value of V' such that $W(V)$ is concentrated around $W(V')$.

Observe that for all k, if $V \backslash E_k$ is fixed and S is good then the targets of a standard-basis measurement of $CL|0\dots 0\rangle$ are a read-$O(2^d \log s)$ family. Before proving Claim 3.2.11, we remark that a similar statement[2] with weaker parameters can be proved using another similar argument involving the triangle inequality.

Let $\mathcal{G}(v)$ be the set of real-valued random variables X such that $\mathbb{E}\big[e^{\lambda X}\big] \le \exp\big(\lambda^2 v/2\big)$ for all $\lambda \in \mathbb{R}$. (This definition is motivated by the fact that if X is Gaussian with mean 0 and variance v then $\mathbb{E}\big[e^{\lambda X}\big] = \exp\big(\lambda^2 v/2\big)$ for all $\lambda \in \mathbb{R}$.) Boucheron, Lugosi and Massart [25] noted that random variables obeying "sub-Gaussian" tail bounds also have sub-Gaussian moment generating functions, and vice versa:

Lemma 3.2.12 ([25, Chapter 2.3])**.** *Let X be a real-valued random variable such that $\mathbb{E}[X] = 0$.*

(i) If $\max(\Pr(X > t), \Pr(X < -t)) \le \exp\big(-t^2/(2v)\big)$ for all $t > 0$, then $X \in \mathcal{G}(16v)$.

(ii) If $X \in \mathcal{G}(v)$, then $\max(\Pr(X > t), \Pr(X < -t)) \le \exp\big(-t^2/(2v)\big)$ for all $t > 0$.

The following lemma is basically implicit in the martingale proof of McDiarmid's inequality [25], and is proved below for completeness:

Lemma 3.2.13. *Let $X_1, \dots, X_m$ be independent random variables, let $v_1, \dots, v_m > 0$, and let f be a function such that*

$$f(x_1, \dots, x_{i-1}, X_i, x_{i+1}, \dots, x_m) - \mathbb{E} f(x_1, \dots, x_{i-1}, X_i, x_{i+1}, \dots, x_m) \in \mathcal{G}(v_i)$$

for all $i \in [m]$ and fixed $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m$. Then,

$$f(X_1, \dots, X_m) - \mathbb{E} f(X_1, \dots, X_m) \in \mathcal{G}\left(\sum_{i=1}^m v_i\right).$$

Remark. Lemma 3.2.13 is tight when $(X_i)_i$ are independent Gaussians and f is the summation function.

Proof of Claim 3.2.11 assuming Lemma 3.2.13. It follows from Theorem 3.2.9 and Lemma 3.2.12(i) that $W - \mathbb{E}[W \mid V \backslash E_k] \in \mathcal{G}(O(n 2^d \log s))$ for all k. Since $\log \binom{n}{2^d} \le 2^d \log n$, it then follows from Lemma 3.2.13 that $W - \mathbb{E}[W \mid V \backslash (E_k)_k] \in \mathcal{G}\big(O\big(n 4^d \log n \log s\big)\big)$. Finally, apply Lemma 3.2.12(ii). □

Proof of Lemma 3.2.13. Let $Y = f(X_1, \dots, X_m)$, and for $i \in [m]$ let $X_{[i]} = (X_j)_{j \le i}$ and $E_i = \mathbb{E}[Y \mid X_{[i]}]$. Fix $\lambda \in \mathbb{R}$, and let $\varphi(x) = e^{\lambda x}$. We will prove that $\mathbb{E}\varphi(E_i - E_0) \le \exp(\lambda^2 v_i/2) \mathbb{E}\varphi(E_{i-1} - E_0)$ for all $i \in [m]$, from which it follows by induction that

$$\mathbb{E}\varphi(Y - \mathbb{E}Y) = \mathbb{E}\varphi(E_m - E_0) \le \exp\left(\lambda^2 \sum_i v_i/2\right),$$

as desired. Since

$$\mathbb{E}\varphi(E_i - E_0) = \mathbb{E}\mathbb{E}\big[\varphi(E_i - E_0) \mid X_{[i-1]}\big] = \mathbb{E}\big[\varphi(E_{i-1} - E_0) \mathbb{E}\big[\varphi(E_i - E_{i-1}) \mid X_{[i-1]}\big]\big],$$

[2] $\Pr(|W - W'| \ge \varepsilon n) \le O(2^d \log n) \cdot \exp(-\Omega(\varepsilon^2 n/(8^d \log s \log^2 n)))$, where W' is an independent copy of W. This may be proved by writing $W - W'$ as a sum of $\Theta(\log \binom{n}{2^d})$ terms that are each required to be of magnitude $O(\varepsilon n/ \log \binom{n}{2^d})$.

it suffices to prove that $\mathbb{E}\big[\varphi(E_i - E_{i-1}) \mid X_{[i-1]}\big] \le \exp(\lambda^2 v_i/2)$. Let $X_{\setminus i} = (X_j)_{j\neq i}$ and $E_{\setminus i} = \mathbb{E}\big[Y \mid X_{\setminus i}\big]$. By Jensen's inequality

$$\varphi(E_i - E_{i-1}) = \varphi\big(\mathbb{E}\big[Y - E_{\setminus i} \mid X_{[i]}\big]\big) \le \mathbb{E}\big[\varphi\big(Y - E_{\setminus i}\big) \mid X_{[i]}\big],$$

so

$$\begin{aligned}\mathbb{E}\big[\varphi(E_i - E_{i-1}) \mid X_{[i-1]}\big] &\le \mathbb{E}\big[\varphi(Y - E_{\setminus i}) \mid X_{[i-1]}\big] \le \sup_{X_{\setminus i}} \mathbb{E}\big[\varphi(Y - E_{\setminus i}) \mid X_{\setminus i}\big] \\ &\le \exp\big(\lambda^2 v_i/2\big). \qquad \square\end{aligned}$$

□

3.3 Upper bounds for parity and fanout

Theorem 3.3.1. *For all $\varepsilon > 0$ and $7 \le d \le O(\log n)$ there exists a depth-d QAC circuit C with $\exp\big(n^{O(1/d)}\log(n/\varepsilon)\big)$ ancillae such that*

$$\|C(I_n \otimes |0\dots0\rangle) - P_n \otimes |0\dots0\rangle\| \le \varepsilon. \tag{3.3.1}$$

The same statement holds with F_n in place of P_n.

Proof. First we show that it suffices to prove the theorem for parity. Suppose Eq. (3.3.1) holds for a circuit C, and let $C' = (H^{\otimes n} \otimes I)C(H^{\otimes n} \otimes I)$. Then by Fig. 6 it holds that

$$\big\|C'(I_n \otimes |0\dots0\rangle) - F_n \otimes |0\dots0\rangle\big\| = \|C(I_n \otimes |0\dots0\rangle) - P_n \otimes |0\dots0\rangle\| \le \varepsilon$$

as required.

We claim that if U is an a-qubit unitary that constructs an n-nekomata, then the circuit in Fig. 8 computes $(n+1)$-qubit parity. The gates after times 1 and 5 in Fig. 8 are two-qubit Z gates (i.e. $I - 2|11\rangle\langle 11|$), and the gate after time 3 is an OR gate (Fig. 4). Let

$$|\nu\rangle = U|0^a\rangle = \frac{1}{\sqrt{2}}|0^n\rangle|\psi_0\rangle + \frac{1}{\sqrt{2}}|1^n\rangle|\psi_1\rangle \qquad \text{and} \qquad \Big|\nu^\perp\Big\rangle = \frac{1}{\sqrt{2}}|0^n\rangle|\psi_0\rangle - \frac{1}{\sqrt{2}}|1^n\rangle|\psi_1\rangle.$$

If the input string x has parity 0, then the layer of Z gates after time 1 acts as the identity on $|x\rangle|\nu\rangle$, so the applications of U and $U^\dagger$ cancel out and the circuit acts as the identity on $|x, 0\dots0, b\rangle$. If instead the input string x has parity 1, then the state after time 2 is $|x\rangle\big|\nu^\perp\big\rangle|b\rangle$, and since $\langle 0^a|U^\dagger\big|\nu^\perp\big\rangle = \big\langle\nu\big|\nu^\perp\big\rangle = \frac{1}{2} - \frac{1}{2} = 0$ it follows that the state after time 4 is $|x\rangle \otimes U^\dagger\big|\nu^\perp\big\rangle \otimes |b \oplus 1\rangle$. The rest of the circuit uncomputes the garbage state $U^\dagger\big|\nu^\perp\big\rangle$.[3] By linearity it follows that the circuit computes parity on all entangled inputs.

By Theorem 3.2.2 there exists a depth-2 QAC circuit A acting on $\exp(O(n\log(n/\varepsilon)))$ qubits such that $A|0\dots0\rangle$ is within 2-norm distance $\varepsilon/(4\sqrt{2})$ of some n-nekomata. By Lemma 2.1.1, plugging A into the circuit from Fig. 8 yields a depth-11 QAC circuit C with $\exp(O(n\log(n/\varepsilon)))$ ancillae satisfying Eq. (3.3.1). We can decrease the depth from 11 to 7 as follows. First, if A approximately constructs an n-nekomata to within 2-norm error $\varepsilon/(4\sqrt{2})$ then trivially so does $(X^{\otimes n} \otimes I)A$. Plug $(X^{\otimes n} \otimes I)A$ into the circuit from Fig. 8,

[3] This is necessary even for a non-clean computation of parity, because if the input is a superposition of standard-basis states x with different parities then the garbage state is entangled with x.

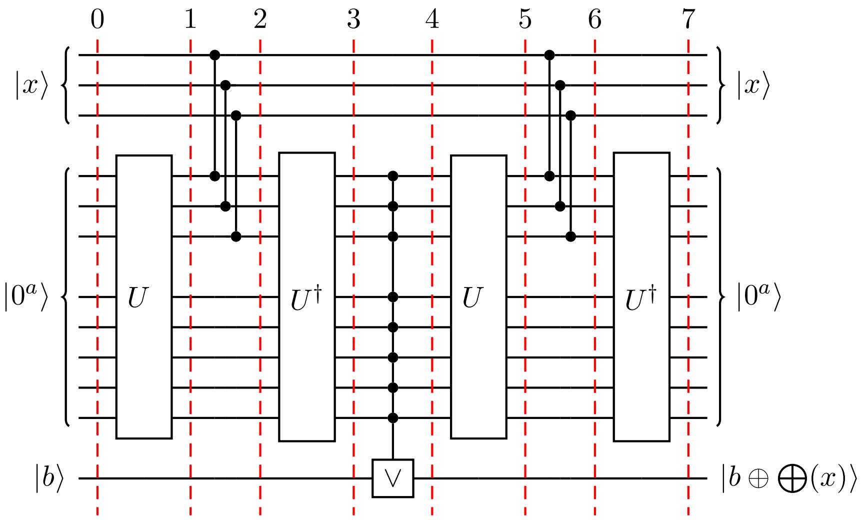

Figure 8: A circuit for parity, assuming U constructs an n-nekomata and $x \in \{0,1\}^n$.

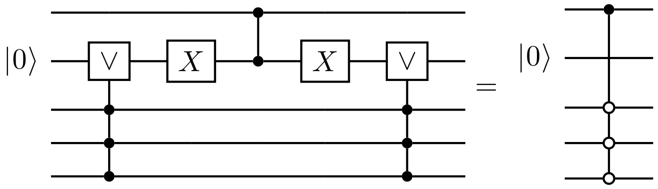

Figure 9: The gate on the right is $R_{|1,0\ldots0\rangle}$.

and then (using our specific circuit for A from Section 3.2.2) apply the equality pictured in Fig. 9 to decrease the depth from 11 to 7.

Finally we prove the depth-d case of the theorem by reducing to the depth-7 case. Let $D = \lfloor d/14 \rfloor$ and $N = \lceil n^{1/D} \rceil$. As illustrated in Fig. 10, there is a depth-D formula for n-bit parity consisting of $(\leq N)$-bit parity gates. This formula has at most n gates, so by the depth-7 case and Lemma 2.1.1 there exists a QAC circuit C of depth $2 \cdot 7D \leq d$ (the extra factor of 2 allows for uncomputing the garbage at the end) with

$$O(n) \cdot \exp(O(N \log(N/(\varepsilon/n)))) \leq \exp\Big(n^{O(1/d)} \log(n/\varepsilon)\Big)$$

ancillae that satisfies Eq. (3.3.1).

□

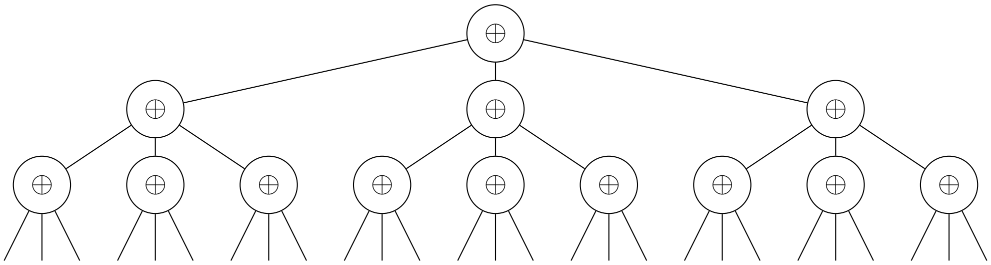

Figure 10: Downward self-reducibility of parity.

3.4 Linear-size lower bounds

In this section we prove the following, where the circuit C may act on any number of qubits:

Theorem 3.4.1. *There is a universal constant $c > 0$ such that the following holds. Let C be a depth-d QAC circuit where at most $cn/(d+1)$ multi-qubit gates act on the first n qubits. Then $|\langle\nu|C|0\ldots0\rangle|^2 \leq 1/2 + \exp(-\Omega(n/(d+1)))$ for all n-nekomata $|\nu\rangle$.*

It follows that for all circuits C satisfying the preconditions of Theorem 3.4.1,

$$\begin{aligned}\|C(I_n \otimes |0\ldots0\rangle) - F_n \otimes |0\ldots0\rangle\| &\geq \big\|C|+\rangle|0\ldots0\rangle - F_n\big(|+\rangle\big|0^{n-1}\big\rangle\big) \otimes |0\ldots0\rangle\big\| \\ &= \big\|C(H \otimes I)|0\ldots0\rangle - |🐱_n\rangle|0\ldots0\rangle\big\|,\end{aligned}$$

and since $\||\psi\rangle - |\phi\rangle\|^2 \geq 1 - |\langle\psi|\phi\rangle|^2$ for all states $|\psi\rangle$ and $|\phi\rangle$ (see Eq. (1.6.3)) the following is immediate:

Corollary 3.4.2. *There is a universal constant $c > 0$ such that the following holds. Let C be a depth-d QAC circuit where at most $cn/(d+1)$ multi-qubit gates act on the first n qubits. Then $\|C(I_n \otimes |0\ldots0\rangle) - P_n \otimes |0\ldots0\rangle\|^2 \geq 1/2 - \exp(-\Omega(n/(d+1)))$, and similarly for F_n.*

For example, Theorem 3.4.1 implies that depth-2 QAC circuits require $\Omega(n)$ multi-qubit acting on the first n qubits in order to approximately construct an n-nekomata. This lower bound is tight, because Theorem 3.2.2 says that depth-2 QAC circuits can approximately construct an n-nekomata, and at most $2n$ multi-qubit gates in a depth-2 circuit can act on the first n qubits. Similarly Corollary 3.4.2 implies that depth-7 QAC circuits approximating n-qubit parity or fanout require $\Omega(n)$ multi-qubit gates acting on the input/output register, and this is tight by Theorem 3.3.1.

Corollary 3.4.2 also implies that the *total* number of multi-qubit gates, a.k.a. the size, of a depth-d QAC circuit approximately computing n-qubit parity or fanout must be at least $\Omega(n/(d+1))$. When d is constant this is far from the exponential upper bound from Theorem 3.3.1, but when $d = \Theta(\log n)$ the trivial $O(n)$-size upper bound almost matches the $\Omega(n/\log n)$ lower bound. Similar observations apply to constructing an n-nekomata, using Theorem 3.4.1.

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 3.4.1 and Corollary 3.4.2 that QAC circuits of *arbitrary* depth require size at least $\Omega(\sqrt{n})$ to approximately construct an n-nekomata or approximately compute n-qubit parity or fanout.

More generally, inspection of its proof reveals that Theorem 3.4.1 also holds if "depth" is replaced by the maximum number of multi-qubit gates acting on any one of the first n qubits.

The bound $\|C(I_n \otimes |0\ldots0\rangle) - P_n \otimes |0\ldots0\rangle\|^2 \gtrapprox 1/2$ from Corollary 3.4.2 is tight, as witnessed by the circuit $C = (|+\rangle\langle+| + i|-\rangle\langle-|) \otimes I_{n-1}$. To see this, note that if Π_0 and Π_1 denote the projections onto $(n-1)$-bit strings with parities 0 and 1 respectively, then

$$\begin{aligned}\|P_n - C\| &= \|\Pi_0 \otimes I + \Pi_1 \otimes X - (\Pi_0 + \Pi_1) \otimes (|+\rangle\langle+| + i|-\rangle\langle-|)\| \\ &= \|((1-i)\Pi_0 + (-1-i)\Pi_1) \otimes |-\rangle\langle-|\| = \sqrt{2}.\end{aligned}$$

In Section 3.4.1 we prove a generalization of Theorem 3.4.1, which we will also use in Section 3.5. The proof uses the following claim, which is proved in Section 3.4.2 (and is obtained as a corollary of a stronger result):[4]

Corollary 3.4.3. *For all $d \geq 1$, orthogonal projections $Q_1, \dots, Q_d$, and states $|\phi\rangle$,*

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

3.4.1 Proof of Theorem 3.4.1

Theorem 3.4.1 is the case of the following in which $\mathsf{T}_1, \dots, \mathsf{T}_n$ are one-qubit registers, $|\phi\rangle$ is the all-zeros state, $Q_j = |0\rangle\langle 0|$ for all j, and $|\psi\rangle$ is an n-nekomata:

Theorem 3.4.4. *There is a universal constant $c > 0$ such that the following holds. Let $\mathsf{T}_1, \dots, \mathsf{T}_n$ be registers (for "targets") and $\mathsf{T} = \mathsf{T}_1 \cdots \mathsf{T}_n$, and let A be a register (for "ancillae"). Let $|\phi\rangle = |\phi_1\rangle \cdots |\phi_n\rangle|\phi_A\rangle$ for some states $|\phi_j\rangle_{\mathsf{T}_j}, j \in [n]$ and $|\phi_A\rangle_{\mathsf{A}}$. Let Q_j be an orthogonal projection on T_j for $j \in [n]$, and let $|\psi\rangle$ be a state in TA such that*

$$\langle\psi|\left(\bigotimes_{j=1}^{n} Q_j \otimes I_{\mathsf{A}}\right)|\psi\rangle = \langle\psi|\left(\bigotimes_{j=1}^{n} (I - Q_j) \otimes I_{\mathsf{A}}\right)|\psi\rangle = 1/2.$$

Let C be a depth-d QAC circuit on TA with at most $cn/(d+1)$ multi-qubit gates acting on T. Then $|\langle\psi|C|\phi\rangle|^2 \leq 1/2 + \exp(-\Omega(n/(d+1)))$.

Proof. By Proposition 3.1.2 we may write $C = DL$ for some layer of one-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 3.2.5 from nekomata to states such as $|\psi\rangle$. Let $Q = \bigotimes_{j=1}^n Q_j \otimes I_{\mathsf{A}}$ and $Q' = \bigotimes_{j=1}^n (I - Q_j) \otimes I_{\mathsf{A}}$, and let $|\varphi\rangle = C|\phi\rangle$. Since $(Q + Q')|\psi\rangle = |\psi\rangle$ it follows from the triangle inequality and Cauchy-Schwarz that

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

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

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

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 T. For $G \in \mathcal{G}_T$ let $|\theta_G\rangle$ be the state, specified up to a global phase, such that $G = R_{\theta_G} = I - 2\theta_G$. Let F be the set of functions with domain $\mathcal{G}$ that map each gate G in $\mathcal{G}_T$ to either I or θ_G, and map each gate G in $\mathcal{G} \backslash \mathcal{G}_T$ to G itself. Then $C = \sum_{f \in F} (-2)^{|\{G : f(G) = \theta_G\}|} \prod_{G \in \mathcal{G}} f(G)$,

[4] An orthogonal projection is a linear transformation Q such that $Q = Q^2 = Q^\dagger$.

so by the triangle inequality

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

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

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

Consider an arbitrary function $f \in F$. For all $G \in \mathcal{G}$ we may write $f(G) = f_T(G) \otimes f_A(G)$, where $f_T(G)$ is an orthogonal projection on T and $f_A(G)$ is either an orthogonal projection or a unitary transformation on A. (This can be seen by considering all three cases: $f(G) = I$, $f(G) = \theta_G$, or $G \notin \mathcal{G}_T$ and $f(G) = G$.) Furthermore if $G \notin \mathcal{G}_T$ then $f_T(G) = I$. Therefore letting $|\phi_T\rangle = |\phi_1\rangle \cdots |\phi_n\rangle$,

$$\left\|Q\prod_{G\in\mathcal{G}} f(G)\cdot|\phi\rangle\right\| = \left\|\bigotimes_j Q_j \cdot \prod_{G\in\mathcal{G}_T} f_T(G)\cdot|\phi_T\rangle\right\| \cdot \left\|\prod_{G\in\mathcal{G}} f_A(G)\cdot|\phi_A\rangle\right\|.$$

Clearly $\left\|\prod_{G\in\mathcal{G}} f_A(G)\cdot|\phi_A\rangle\right\| \leq 1$. For $k \in [d]$ let M_k be the tensor product of $f_T(G)$ over all "depth-k" gates $G \in \mathcal{G}_T$, i.e. $M_1, \ldots, M_d$ are layers of one-qubit orthogonal projections such that $\prod_{G\in\mathcal{G}_T} f_T(G) = M_d \cdots M_1$. Write $M_k = \bigotimes_{j=1}^n M_{jk}$, where M_{jk} is an orthogonal projection on T_j for all $j \in [n]$. Then by Corollary 3.4.3,

$$\begin{aligned}\left\|\bigotimes_j Q_j \cdot \prod_{G\in\mathcal{G}_T} f_T(G)\cdot|\phi_T\rangle\right\| &= \prod_{j=1}^n \|Q_j M_{jd}\cdots M_{j1}|\phi_j\rangle\| \leq \prod_{j=1}^n \exp\left(-\frac{\langle\phi_j|(I-Q_j)|\phi_j\rangle}{2(d+1)}\right) \\ &= \exp\left(-\frac{1}{2(d+1)}\sum_{j=1}^n\langle\phi_j|(I-Q_j)|\phi_j\rangle\right) \leq \exp\left(-\frac{n/2}{2(d+1)}\right).\end{aligned}$$

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

3.4.2 Proof of Corollary 3.4.3

For states $|\alpha\rangle, |\beta\rangle$ let $\Delta(|\alpha\rangle, |\beta\rangle) = \arccos|\langle\alpha|\beta\rangle|$; we will abbreviate this as $\Delta(\alpha, \beta)$.

Lemma 3.4.5. *The function Δ satisfies the triangle inequality, i.e. $\Delta(\alpha,\gamma) \leq \Delta(\alpha,\beta) + \Delta(\beta,\gamma)$ for all states $|\alpha\rangle, |\beta\rangle, |\gamma\rangle$.*

Remark. The intuition behind our ultimate use of Lemma 3.4.5 is that, up to normalization, the total amount of "progress" made by $Q_{d-1}\cdots Q_1$ in interpolating between $|\phi\rangle$ and Q_d is at most the sum of the amounts of progress made by the individual projections $Q_1, \ldots, Q_{d-1}$.

For intuition as to why Lemma 3.4.5 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 (Fig. 11) formed by traversing a great circle on the unit sphere

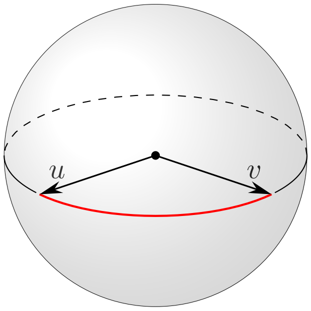

Figure 11: A geodesic on the sphere.

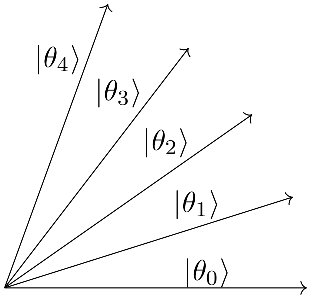

Figure 12: An optimal choice of $|\theta_1\rangle, \ldots, |\theta_{d-1}\rangle$ in the $d = 4$ case of Eq. (3.4.2).

from u to v in the shorter of the two directions. This arc is the shortest path on the unit sphere between u and v, so Δ' represents distance on the unit sphere.

We make two more unrelated remarks. First, if we count states differing only by a global phase as equivalent, then Lemma 3.4.5 implies that Δ is a metric on the set of all states on a given number of qubits. Second, the results in this subsection generalize easily from $\mathbb{C}^{2^n}$ to arbitrary Hilbert spaces.

Proof. Let $|\psi\rangle$ be a state orthogonal to $|\alpha\rangle$ such that $|\gamma\rangle$ is in the span of $|\alpha\rangle$ and $|\psi\rangle$, and let $\eta = \arccos \frac{|\langle\beta|\alpha\rangle|}{\sqrt{|\langle\beta|\alpha\rangle|^2+|\langle\beta|\psi\rangle|^2}}$. By the triangle inequality,

$$\begin{aligned}|\langle\beta|\gamma\rangle| &= |\langle\beta|(\alpha+\psi)|\gamma\rangle| \le |\langle\beta|\alpha\rangle| \cdot |\langle\gamma|\alpha\rangle| + |\langle\beta|\psi\rangle| \cdot |\langle\gamma|\psi\rangle| \\ &\le \frac{|\langle\beta|\alpha\rangle|}{\sqrt{|\langle\beta|\alpha\rangle|^2+|\langle\beta|\psi\rangle|^2}} \cdot |\langle\gamma|\alpha\rangle| + \frac{|\langle\beta|\psi\rangle|}{\sqrt{|\langle\beta|\alpha\rangle|^2+|\langle\beta|\psi\rangle|^2}} \cdot |\langle\gamma|\psi\rangle| \\ &= \cos\eta \cdot \cos\Delta(\alpha,\gamma) + \sin\eta \cdot \sin\Delta(\alpha,\gamma) = \cos(\Delta(\alpha,\gamma) - \eta),\end{aligned}$$

so $\Delta(\beta,\gamma) \ge |\Delta(\alpha,\gamma) - \eta| \ge \Delta(\alpha,\gamma) - \eta$. Similarly, $|\langle\beta|\alpha\rangle| \le \frac{|\langle\beta|\alpha\rangle|}{\sqrt{|\langle\beta|\alpha\rangle|^2+|\langle\beta|\psi\rangle|^2}} = \cos\eta$ so $\Delta(\alpha,\beta) \ge \eta$. Therefore $\Delta(\alpha,\beta) + \Delta(\beta,\gamma) \ge \eta + (\Delta(\alpha,\gamma) - \eta) = \Delta(\alpha,\gamma)$. □

Proposition 3.4.6. *For all $d \ge 1$, nonzero orthogonal projections Q_d, and states $|\phi\rangle$,*

$$\max_{Q_1,\ldots,Q_{d-1}} \|Q_d Q_{d-1} \cdots Q_1|\phi\rangle\| = \cos\left(\frac{\arccos \|Q_d|\phi\rangle\|}{d}\right)^d, \tag{3.4.1}$$

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

We will only use the fact that the left side of Eq. (3.4.1) is at most the right side, but we prove the converse inequality as well because it is easy to do so.

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

$$\max_{|\theta_1\rangle,\ldots,|\theta_{d-1}\rangle} \left|\prod_{j=1}^{d} \langle\theta_{j-1}|\theta_j\rangle\right| = \cos\left(\frac{\arccos|\langle\theta_0|\theta_d\rangle|}{d}\right)^d. \tag{3.4.2}$$

We then prove that the original proposition follows from this rank-1 analogue.

First we prove that the left side of Eq. (3.4.2) is at most the right side. On the image of Δ, i.e. on the interval $[0, \pi/2]$, the cosine function is decreasing and concave. Therefore for all states $|\theta_1\rangle, \dots, |\theta_{d-1}\rangle$, by the AM-GM inequality, Jensen's inequality, and Lemma 3.4.5,

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

Next we give an example (Fig. 12) which shows that the left side of Eq. (3.4.2) is at least the right side. For ease of notation let $|\sigma\rangle = |\theta_0\rangle$ and $|\tau\rangle = |\theta_d\rangle$. By multiplying $|\tau\rangle$ by a global phase we may assume that $\langle\sigma|\tau\rangle$ is a nonnegative real number. Let $\eta = \arccos(\langle\sigma|\tau\rangle)/d$, let $|\psi\rangle = \frac{(I-\sigma)|\tau\rangle}{\|(I-\sigma)|\tau\rangle\|} = \frac{|\tau\rangle - |\sigma\rangle\langle\sigma|\tau\rangle}{\sqrt{1-\langle\sigma|\tau\rangle^2}}$, and for $j \in [d-1]$ let $|\theta_j\rangle = \cos(j\eta)|\sigma\rangle + \sin(j\eta)|\psi\rangle$. The latter equation also holds for $j = 0$ and $j = d$, respectively because $|\sigma\rangle = |\theta_0\rangle$ and

$$\cos(d\eta)|\sigma\rangle + \sin(d\eta)|\psi\rangle = \langle\sigma|\tau\rangle \cdot |\sigma\rangle + \sqrt{1-\langle\sigma|\tau\rangle^2} \cdot |\psi\rangle = |\tau\rangle = |\theta_d\rangle.$$

Since $\langle\sigma|\psi\rangle = 0$, it follows that for all $j \in [d]$,

$$\langle\theta_{j-1}|\theta_j\rangle = \cos((j-1)\eta)\cos(j\eta) + \sin((j-1)\eta)\sin(j\eta) = \cos(j\eta - (j-1)\eta) = \cos(\eta),$$

so $\prod_{j=1}^{d}\langle\theta_{j-1}|\theta_j\rangle = \cos(\eta)^d$ as desired.

Finally we prove that the original proposition follows from Eq. (3.4.2). For $j \in [d-1]$ we may assume that Q_j is rank-1, because if $Q_j \cdots Q_1|\phi\rangle \neq 0$ then $Q_j \cdots Q_1|\phi\rangle = \theta_j Q_{j-1}\cdots Q_1|\phi\rangle$ for $|\theta_j\rangle = \frac{Q_j\cdots Q_1|\phi\rangle}{\|Q_j\cdots Q_1|\phi\rangle\|}$, and if $Q_j \cdots Q_1|\phi\rangle = 0$ then clearly we cannot decrease $\|Q_d \cdots Q_1|\phi\rangle\|$ by replacing Q_j with an arbitrary rank-1 orthogonal projection. For any state $|\varphi\rangle$, the norm $\|Q_d|\varphi\rangle\|$ equals the maximum of $|\langle\psi|\varphi\rangle|$ over all states $|\psi\rangle$ such that $Q_d|\psi\rangle = |\psi\rangle$.[5] (Here we used the fact that $Q_d \neq 0$ to ensure that there exists a *state* in the 1-eigenspace of Q_d, rather than just the zero vector.) Therefore

$$\begin{aligned}\max_{Q_1,\dots,Q_{d-1}}\|Q_d\cdots Q_1|\phi\rangle\| &= \max_{|\theta_1\rangle,\dots,|\theta_{d-1}\rangle}\|Q_d|\theta_{d-1}\rangle\cdots\langle\theta_1|\phi\rangle\| = \max_{\substack{|\theta_1\rangle,\dots,|\theta_{d-1}\rangle\\ |\psi\rangle = Q_d|\psi\rangle}}|\langle\psi|\theta_{d-1}\rangle\cdots\langle\theta_1|\phi\rangle|\\ &= \max_{|\psi\rangle=Q_d|\psi\rangle}\cos\left(\frac{\arccos|\langle\psi|\phi\rangle|}{d}\right)^d = \cos\left(\frac{\arccos\left(\max_{|\psi\rangle=Q_d|\psi\rangle}|\langle\psi|\phi\rangle|\right)}{d}\right)^d\\ &= \cos\left(\frac{\arccos\|Q_d|\phi\rangle\|}{d}\right)^d.\end{aligned}$$

□

Corollary 3.4.3. *For all $d \geq 1$, orthogonal projections $Q_1, \dots, Q_d$, and states $|\phi\rangle$,*

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

[5] By Cauchy-Schwarz, $|\langle\psi|\varphi\rangle| = |\langle\psi|Q_d|\varphi\rangle| \leq \|Q_d|\varphi\rangle\|$, with equality if $|\psi\rangle = Q_d|\varphi\rangle/\|Q_d|\varphi\rangle\|$ or if $Q_d|\varphi\rangle = 0$.

Proof. The claim is trivial if $Q_d = 0$, so assume otherwise. Since

$$\arccos \|Q_d|\phi\rangle\| \geq \sin\arccos \|Q_d|\phi\rangle\| = \sqrt{1 - \|Q_d|\phi\rangle\|^2} = \sqrt{\langle\phi|(I - Q_d)|\phi\rangle},$$

it follows from Proposition 3.4.6 that

$$\|Q_d \cdots Q_1|\phi\rangle\| \leq \cos\left(\frac{\arccos\|Q_d|\phi\rangle\|}{d}\right)^d \leq \cos\left(\frac{\sqrt{\langle\phi|(I - Q_d)|\phi\rangle}}{d}\right)^d,$$

so it suffices to prove that $\cos r \leq \exp(-r^2/2)$ for all $0 \leq r \leq 1$.

A special case of the Lagrange remainder theorem states that if $f : \mathbb{R} \to \mathbb{R}$ is n times differentiable on all of $\mathbb{R}$, then for all $x \in \mathbb{R}$ there exists h between 0 and x such that

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!}x^k + \frac{f^{(n)}(h)}{n!}x^n,$$

where $f^{(k)}$ denotes the k'th derivative of f. An application with $f = \cos(\cdot), x = r, n = 4$ reveals that

$$\cos r \leq 1 - \frac{r^2}{2} + \frac{\max\{\cos h : 0 \leq h \leq r\}}{24} \cdot r^4 = 1 - \frac{r^2}{2} + \frac{r^4}{24},$$

and an application with $f = \exp(\cdot), x = -r^2/2, n = 3$ reveals that

$$e^{-r^2/2} \geq 1 - r^2/2 + \frac{1}{2}(-r^2/2)^2 + \frac{\max\{e^h : -r^2/2 \leq h \leq 0\}}{6}(-r^2/2)^3 = 1 - \frac{r^2}{2} + \frac{r^4}{8} - \frac{r^6}{48}.$$

Finally, since $r^2 \leq 1$ it follows that $r^6 \leq r^4$, so

$$e^{-r^2/2} \geq 1 - \frac{r^2}{2} + \frac{r^4}{8} - \frac{r^4}{48} \geq 1 - \frac{r^2}{2} + \frac{r^4}{24} \geq \cos r.$$ □

3.5 Depth-2 lower bounds

We prove the following, where the circuit C may be of arbitrary size:

Theorem 3.5.1. *Let C be a depth-2 QAC circuit; then $\left\|\left(\langle \text{🐱}_n| \otimes I\right)C|0\ldots 0\rangle\right\|^2 \leq 1/2 + \exp(-\Omega(n))$.*

By reasoning similar to that in the beginning of Section 3.4.2, analogous lower bounds follow immediately for parity and fanout:

Corollary 3.5.2. *Let C be a depth-2 QAC circuit; then $\|C_n(I \otimes |0\ldots 0\rangle) - P_n \otimes |0\ldots 0\rangle\|^2 \geq 1/2 - \exp(-\Omega(n))$, and similarly for F_n.*

We remark that our proof of Theorem 3.5.1 gives a multiplicative constant of roughly $1/10^{60000}$ implicit in the $\Omega(\cdot)$ notation in the inequality, which makes it trivial for small values of n. Corollary 3.5.2 and Theorem 3.3.1 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 similarly for fanout.

Our proof of Theorem 3.5.1 goes roughly as follows. If there are only $o(n)$ multi-qubit gates acting on the n designated "target qubits", then the result follows from Theorem 3.4.1.

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^0 circuit. Using a variant of a light cone argument [20], we choose $\Theta(n)$ pairwise disjoint sets of qubits on which to define orthogonal projections, and apply Theorem 3.4.4.

3.5.1 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 global phase, there is a unique state orthogonal to $|\psi\rangle$.)

Lemma 3.5.3. *Let* A *be a one-qubit register, and let* B *and* C *be registers on arbitrary numbers of qubits. Then for all states* $|\psi\rangle_{\mathsf{A}}, |\theta\rangle_{\mathsf{B}}, |\phi\rangle_{\mathsf{ABC}}$, *the following two procedures generate identically distributed random states in* ABC*:*

- *measure the* A *qubit of* $\big(R_{|\psi\rangle|\theta\rangle} \otimes I_{\mathsf{C}}\big)|\phi\rangle$ *in the* $|\psi\rangle$ *basis;*
- *measure the* A *qubit of* $|\phi\rangle$ *in the* $|\psi\rangle$ *basis, and then, conditioned on the outcome being* $|\psi\rangle$*, apply* $R_{|\theta\rangle}$ *on* B*.*

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

Theorem 3.5.1 is clearly equivalent to the statement that if C is a depth-2 QAC circuit, and the projective measurement $\big(\text{😸}_n, I - \text{😸}_n\big)$ is performed on any n designated "target" qubits of $C|0\dots0\rangle$, then the outcome is 😸_n with probability at most $1/2 + \exp(-\Omega(n))$. The following is the starting point for our proof:

Lemma 3.5.4. *Let* $|\psi\rangle$ *be an* n*-qubit state, and let* C *be a depth-*2 QAC *circuit acting on at least* n *qubits. Partition these qubits into* n *"targets" followed by some number of ancillae. Then there exist layers of* $R_\otimes$ *gates* L_2, L_1 *and a tensor product* $|\phi\rangle$ *of one-qubit states such that the following conditions hold;*

(i) $\|(\langle\psi| \otimes I)L_2L_1|\phi\rangle\|^2 \geq \|(\langle\psi| \otimes I)C|0\dots0\rangle\|^2$*;*

(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 3.1.2 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. Let a "construction" be a tuple of the form $(L_2, L_1, |\phi\rangle)$ where L_2 and L_1 are layers of $R_\otimes$ gates and $|\phi\rangle$ is a tensor product of one-qubit states. By Proposition 3.1.2 there exists a construction satisfying (i). Below we describe a procedure that takes as input a construction satisfying (i) but not (ii), and outputs a construction satisfying (i) using fewer ancillae than the original construction. It then suffices to iterate this procedure on a construction satisfying (i) until the construction also satisfies (ii), because the number of ancillae can only decrease finitely many times.

Let $(L_2, L_1, |\phi\rangle)$ be a construction satisfying (i) but not (ii), and let $|\varphi\rangle = L_2L_1|\phi\rangle$. For all $k \in \{1, 2\}$ and gates G in L_k, write $G = R_{\bigotimes_{\mathsf{A}}|\theta^k_{\mathsf{A}}\rangle}$, where A ranges over all one-qubit registers acted on by G, and $\left|\theta^k_{\mathsf{A}}\right\rangle$ is a state in A. (Since A and k uniquely determine G, this does not assign conflicting definitions to any of the $\left|\theta^k_{\mathsf{A}}\right\rangle$.)

First consider the case where an ancilla A is not acted on by L_2 (that is, by any gate in L_2). If A is also not acted on by L_1 then we may simply remove A from the construction. Otherwise, measure the A qubit of $|\varphi\rangle$ in the $\left|\theta^1_{\mathsf{A}}\right\rangle$ basis. By Lemma 3.5.3, the resulting state on the qubits besides A equals $L_2'L_1'|\phi'\rangle$ for some random construction $(L_2', L_1', |\phi'\rangle)$. Furthermore, the expectation over $(L_2', L_1', |\phi'\rangle)$ of $\|(\langle\psi| \otimes I)L_2'L_1'|\phi'\rangle\|^2$ equals $\|(\langle\psi| \otimes I)L_2L_1|\phi\rangle\|^2$, which is at least $\|(\langle\psi| \otimes I)C|0\dots0\rangle\|^2$. Therefore there exists a fixed construction in the support of $(L_2', L_1', |\phi'\rangle)$ that satisfies (i), and the procedure may output this construction.

If an ancilla A is acted on by L_2 but not by L_1, then measure the A qubit of $|\varphi\rangle$ in the $\left|\theta^2_{\mathsf{A}}\right\rangle$ basis, and the rest of the argument is similar to the above. If every ancilla is acted on by L_2, and a gate G in L_1 does not act on any targets, then for all qubits A acted on by G, measure the A qubit of $|\varphi\rangle$ in the $\left|\theta^2_{\mathsf{A}}\right\rangle$ basis, and again the rest of the argument is similar to the above. Finally, if a gate G in L_2 does not act on any targets, then G acts on at least one ancilla, and also we may remove G from L_2 without changing $\|(\langle\psi| \otimes I)L_2L_1|\phi\rangle\|^2$, so this reduces to the previously considered case in which an ancilla is not acted on by L_2. □

3.5.2 Proof of Theorem 3.5.1

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

Lemma 3.5.5. *Let $0 < \delta \leq 1$, let $a > 0$, and let X be a nonnegative random variable. Then there exists $t \in \left[a, ae^{\delta^{-1}-1}\right]$ such that $\Pr(X \geq t) \leq \delta\mathbb{E}[X]/t$.*

Remark. The intuition behind our use of Lemma 3.5.5 is as follows. Theorem 3.4.4 implies that depth-2 QAC circuits require size at least $\Omega(n)$ to approximately construct $\left|\text{🐱}_n\right\rangle$, and Lemma 3.5.4 implies that depth-2 QAC circuits that approximately construct $\left|\text{🐱}_n\right\rangle$ 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 $\left|\text{🐱}_n\right\rangle$. This is analogous to how Markov's inequality is "just a factor of δ" away from the conclusion of Lemma 3.5.5.

Proof. Assume the contrary, and let $b = ae^{\delta^{-1}-1}$. Then,

$$\mathbb{E}[X] = \int_0^\infty \Pr(X \geq t)dt \geq \int_0^a \Pr(X \geq t)dt + \int_a^b \Pr(X \geq t)dt,$$

and

$$\int_0^a \Pr(X \geq t)dt \geq \int_0^a \Pr(X \geq a)dt = a\Pr(X \geq a) > \delta\mathbb{E}[X],$$

and

$$\int_a^b \Pr(X \geq t)dt > \int_a^b \delta\mathbb{E}[X]/t \cdot dt = \delta\mathbb{E}[X]\ln(b/a) = \delta\mathbb{E}[X](\delta^{-1} - 1),$$

so $\mathbb{E}[X] > \mathbb{E}[X]$, which is a contradiction. □

Theorem 3.5.6 (Turán's theorem[6])**.** *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 $n/(d+1)$.*

[6] Often Turán's theorem is phrased as saying that dense graphs have large cliques, whereas Theorem 3.5.6 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.

Remark. For the intuition behind our use of Theorem 3.5.6, recall the discussion of disjoint light cones from the proof overview earlier in this section.

Proof exposited by Alon and Spencer [9]. Identify the vertex set of $\mathcal{G}$ with $[n]$. Let σ be a uniform random permutation of $[n]$, and let $\mathcal{I}$ be the set of vertices u such that $\sigma(u) < \sigma(v)$ for all edges $\{u, v\}$. Then $\mathcal{I}$ is an independent set, because for every edge $\{u, v\}$, either $\sigma(u) < \sigma(v)$ or $\sigma(v) < \sigma(u)$. A vertex u with degree d_u is in $\mathcal{I}$ with probability $1/(d_u + 1)$, because any vertex out of u and its neighbors is equally likely to be assigned the lowest value by σ out of these vertices. By linearity of expectation it follows that $\mathbb{E}|\mathcal{I}| = \sum_{u\in[n]} 1/(d_u+1)$, and by Jensen's inequality this is at least $n/(d + 1)$. □

Recall that $|\phi\rangle, C, |\psi\rangle, (Q_j)_j$ are variables from the statement of Theorem 3.4.4. In upcoming applications of Theorem 3.4.4 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 3.4.4, 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 3.4.4 in this regard, and forgo the use of Lemma 3.5.4 entirely by measuring selected ancillae all at once later in the proof, but we consider the current presentation to be simpler.

Theorem 3.5.1. *Let C be a depth-2 QAC circuit; then $\left\|\left(\langle \text{😺}_n| \otimes I\right)C|0\ldots0\rangle\right\|^2 \leq 1/2 + \exp(-\Omega(n))$.*

Proof. Let L_2, L_1 be layers of $R_\otimes$ gates and let $|\phi\rangle$ be a tensor product of one-qubit states, with the first 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 Lemma 3.5.4 it suffices to prove that $\left\|\left(\langle \text{😺}_n| \otimes I\right)L_2L_1|\phi\rangle\right\|^2 \leq 1/2 + \exp(-\Omega(n))$.

Let c be the constant from Theorem 3.4.4, and let $\gamma = (c/2)(c/3)/(1 + c/2)$ and $\delta = (c/2)\gamma^2$. Since Theorem 3.4.4 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 $\left|\langle\phi|L_1^\dagger L_2^\dagger \cdot \left|\text{😺}_n\right\rangle|\psi\rangle\right|^2 \leq 1/2 + \exp(-\Omega(n))$ 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 by applying Theorem 3.4.4 with input state $|\phi\rangle$, circuit L_2L_1, desired output state $\left|\text{😺}_n\right\rangle|\psi\rangle$, 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 3.4.4 with input state $L_1|\phi\rangle$, circuit L_2, desired output state $\left|\text{😺}_n\right\rangle|\psi\rangle$, 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 $1/2 + \exp(-\Omega(|L_1|)) \leq 1/2 + \exp(-\Omega(n))$.)

Now consider the case where $|L_2| \geq \gamma n$. This time we will measure some carefully chosen ancillae before applying Theorem 3.4.4. Let X be the number of targets acted on by a uniform random gate in L_1. By Lemma 3.5.5 there exists a number $t \in [1, \exp(1/\delta)]$ such that $\Pr(X \geq t) \leq \delta\mathbb{E}[X]/t$. Fix such a t. Write $L_1 = L_1^B \otimes L_1^S$, for "big" and "small" respectively, where L_1^B (resp. L_1^S) consists of the gates in L_1 acting on at least (resp. fewer than) t targets. Then $\left|L_1^B\right| = |L_1|\Pr(X \geq t) \leq \delta|L_1|\mathbb{E}[X]/t \leq \delta n/t = (c/2)\gamma^2 n/t$.

Let $\mathcal{G}$ be the graph whose vertices are the gates in L_2, and whose edges are the pairs e of distinct vertices such that for some gate $G \in L_1^S$, for both vertices V in e, there exists a target that both G and V act on. Since $t \geq 1$, the degree of a vertex is at most $t-1$ times the number of targets acted on by that vertex. Therefore the average degree of the vertices in $\mathcal{G}$ is at most $(t-1)n/|L_2|$, so by Theorem 3.5.6 there exists an independent set $\mathcal{I}$ in $\mathcal{G}$ of size

$$|\mathcal{I}| \geq \frac{|L_2|}{(t-1)n/|L_2|+1} \geq \frac{\gamma n}{(t-1)n/(\gamma n)+1} = \frac{\gamma^2 n}{t-1+\gamma} \geq \gamma^2 n/t.$$

Fix such a set $\mathcal{I}$. It follows that $\left|L_1^B\right| \leq (c/2)|\mathcal{I}|$, and also that $|\mathcal{I}| \geq \gamma^2 n/\exp(1/\delta) \geq \Omega(n)$.

For $V \in \mathcal{I}$ let A_V be the register consisting of the following two types of qubits: targets acted on by a gate in L_1^S that acts on one of the same targets as V, and qubits acted on by V that are not acted on by L_1^S. The A_V are registers on pairwise disjoint sets of qubits, because $\mathcal{I}$ is an independent set in $\mathcal{G}$ and because a qubit cannot be acted on by multiple gates in any given layer.

For $G \in L_2$ write $G = R_{\bigotimes_{\mathsf{A}}|\theta_{\mathsf{A}}\rangle}$, where A ranges over all one-qubit Hilbert spaces acted on by G, and $|\theta_{\mathsf{A}}\rangle$ is a state in A. This defines $|\theta_{\mathsf{A}}\rangle$ for every ancilla A, because L_2 acts on every ancilla. For all ancillae A acted on by L_1^S, measure the A qubit of $L_2 L_1|\phi\rangle$ in the $|\theta_{\mathsf{A}}\rangle$ basis. By Lemma 3.5.3, the resulting state $|\varphi\rangle$ on the qubits that were not measured satisfies $|\varphi\rangle = L_2' L_1^B|\phi'\rangle$, where L_2' and L_1^B are implicitly tensored with the identity, and

- $|\phi'\rangle$ is the tensor product of (i) a tensor product of one-qubit states on the qubits that were not acted on by L_1^S, and (ii) the tensor product over $G \in L_1^S$ of a state on the targets that were acted on by G. In particular, $|\phi'\rangle$ factors as $|\phi'\rangle = \bigotimes_{V \in \mathcal{I}}|\phi_V'\rangle \otimes |\phi_A'\rangle$, for some states $|\phi_V'\rangle \in \mathsf{A}_V$ (none of the qubits in A_V were measured) and a state $|\phi_A'\rangle$ on all other qubits in $|\varphi\rangle$.
- $L_2' = \bigotimes_{G \in L_2} U_G$, where U_G is a Hermitian unitary transformation (specifically, the identity or an $R_{\otimes}$ gate) on the qubits in $|\varphi\rangle$ that were acted on by G.

It suffices to prove that $\left\|\left(\langle \text{🐱}_n\right| \otimes I\right)|\varphi\rangle\right\|^2 \leq 1/2 + \exp(-\Omega(n))$, or equivalently that $\left|\left(\langle \text{🐱}_n\right|\langle\psi|\right) L_2' L_1^B|\phi'\rangle\right|^2 \leq 1/2 + \exp(-\Omega(n))$ for all states $|\psi\rangle$. For $V \in \mathcal{I}$, the transformation U_V acts on a subset of the qubits in A_V, including at least one target because V acted on at least one target and none of the targets were measured. Therefore we may define an orthogonal projection on A_V by $Q_V = U_V(|0\rangle\langle 0| \otimes I)U_V \otimes I$, where $|0\rangle\langle 0|$ acts on a target. Observe that $I - Q_V = U_V(|1\rangle\langle 1| \otimes I)U_V \otimes I$, and that the state $|\tau\rangle = L_2'\left|\text{🐱}_n\right\rangle|\psi\rangle$ satisfies

$$\langle\tau|\left(\bigotimes_{V \in \mathcal{I}} Q_V \otimes I\right)|\tau\rangle = \langle\tau|\left(\bigotimes_{V \in \mathcal{I}}(I - Q_V) \otimes I\right)|\tau\rangle = 1/2.$$

Therefore the result follows by applying Theorem 3.4.4 with input state $|\phi'\rangle$, circuit L_1^B, desired output state $|\tau\rangle$, and projections $(Q_V)_V$, recalling that $|L_1^B| \leq (c/2)|\mathcal{I}|$ and that $|\mathcal{I}| \geq \Omega(n)$. □

Chapter 4

Quantum circuit upper bounds for states, unitaries, and functions

> And now for something completely different.
>
> *Monty Python's Flying Circus*

In Section 4.1 we present $\mathsf{QAC_f}$ size and depth upper bounds for arbitrary states and unitaries. In Section 4.2 we present size upper and lower bounds for constructing arbitrary states over a universal gate set. In Section 4.3 we present $\mathsf{QAC^0}$ size upper bounds for arbitrary functions and states. Finally in Section 4.4 we present a barrier to $\mathsf{QAC^0_f}$ size lower bounds for constructing explicit states.

4.1 $\mathsf{QAC_f}$ upper bounds for states and unitaries

Theorem 1.3.10. *Every n-qubit state can be cleanly, exactly constructed by a $\mathsf{QAC^0_f}$ circuit with $\tilde{O}(2^n)$ ancillae.*

A proof sketch is as follows. First consider the analogous problem of sampling a string s from a given distribution over $\{0,1\}^n$. One way to sample s is to first sample

$$b_x \sim \text{Bernoulli}(\Pr(s \text{ begins with } x1 \mid s \text{ begins with } x))$$

independently for all binary strings x of length less than n, and then output the string y defined by $y_i = b_{y_1 y_2 \cdots y_{i-1}}$ for i from 1 to n. Furthermore each bit of y can be computed by a DNF formula of size $\tilde{O}(2^n)$ as a function of $(b_x)_x$. Similarly we can construct a quantum state $\sum_{y \in \{0,1\}^n} \alpha_y |y\rangle$ using unentangled one-qubit states in place of $(b_x)_x$; this actually yields a state of the form $\sum_{y \in \{0,1\}^n} \alpha_y |y\rangle \big|\text{garbage}_y\big\rangle$, but it turns out that $\big|\text{garbage}_y\big\rangle$ can be efficiently uncomputed controlled on y.

Our proof will use the following notation. Let $\{0,1\}^{\leq n}$ (resp. $\{0,1\}^{<n}$) denote the set of strings of length at most (resp. less than) n over $\{0,1\}$, including the empty string ϵ. For $x, y \in \{0,1\}^*$ let $x_k, x_{<k}, x_{\leq k}$ respectively denote the k'th bit, first $k-1$ bits, and first k bits of x, and let xy denote the concatenation of x and y.

Proof. Let $|\psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$ denote the n-qubit state to be constructed, and define "conditional amplitudes" β_x for $x \in \{0,1\}^{\leq n} \backslash \{\epsilon\}$ as follows: Let $|\psi_\epsilon\rangle = |\psi\rangle$, and for

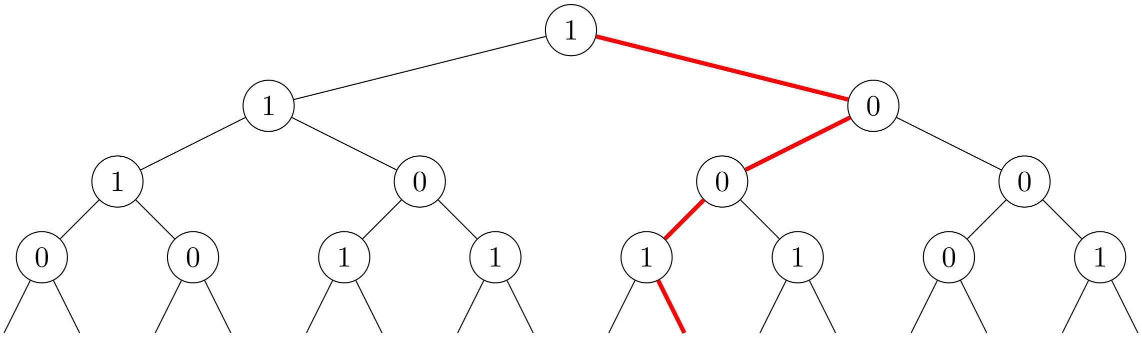

Figure 13: The nodes are labeled with the inputs to f. The highlighted path represents the output of f, and is defined by starting at the root and repeatedly walking to the left or right child depending on whether the current node is labeled 0 or 1.

$x \in \{0,1\}^{<n}$, given an $(n-|x|)$-qubit state $|\psi_x\rangle$, write

$$|\psi_x\rangle = \begin{cases} \beta_{x0}|0\rangle|\psi_{x0}\rangle + \beta_{x1}|1\rangle|\psi_{x1}\rangle & \text{if } |x| \leq n-2, \\ \beta_{x0}|0\rangle + \beta_{x1}|1\rangle & \text{if } |x| = n-1 \end{cases}$$

for $(n-|x|-1)$-qubit states $|\psi_{x0}\rangle, |\psi_{x1}\rangle$ (if $|x| \leq n-2$) and complex numbers β_{x0}, β_{x1} such that $|\beta_{x0}|^2 + |\beta_{x1}|^2 = 1$. Let

$$|\phi_x\rangle = \beta_{x0}|0\rangle + \beta_{x1}|1\rangle$$

for $x \in \{0,1\}^{<n}$, and observe that $\alpha_x = \prod_{i=1}^n \beta_{x_{\leq i}}$ for all $x \in \{0,1\}^n$.

Let $f : \{0,1\}^{\{0,1\}^{<n}} \to \{0,1\}^n$ be the function defined by $f(x)_i = x_{f(x)_{<i}}$ for i from 1 to n. The function f is illustrated in Fig. 13 and can be computed by the following $\mathsf{AC^0}$ formula of leafsize $\tilde{O}(2^n)$:

$$f(x)_j = \bigvee_{\substack{t \in \{0,1\}^j \\ t_j = 1}} \bigwedge_{1 \leq i \leq j} \mathbb{1}_{x_{t_{<i}} = t_i} \qquad \text{for } 1 \leq j \leq n.$$

(The conjunction indicates whether t equals the first j bits of $f(x)$, and the disjunction indicates whether the satisfying t is such that $t_j = 1$.) Therefore the unitary U_f defined by

$$U_f|x, a\rangle = |x, a \oplus f(x)\rangle \qquad \text{for } x \in \{0,1\}^{\{0,1\}^{<n}}, a \in \{0,1\}^n$$

can be computed by a $\mathsf{QAC^0_f}$ circuit on $\tilde{O}(2^n)$ qubits.

Let $(\mathsf{R}_x)_{x \in \{0,1\}^{<n}}$ be one-qubit registers and let S be an n-qubit register. The first step toward constructing $|\psi\rangle$ is to construct the state

$$U_f \left(\bigotimes_{x \in \{0,1\}^{<n}} |\phi_x\rangle_{\mathsf{R}_x} \otimes |0^n\rangle_{\mathsf{S}} \right),$$

using a layer of one-qubit gates followed by the aforementioned circuit for U_f. Here, when computing U_f, the x'th input bit to f is in R_x for all x, and the output register of f is S.

Observe that

$$U_f(I \otimes |0^n\rangle) = \sum_{x\in\{0,1\}^{\{0,1\}^{<n}}} |x\rangle\langle x| \otimes |f(x)\rangle = \sum_{t\in\{0,1\}^n}\left(\sum_{x\in f^{-1}(t)}|x\rangle\langle x|\right)\otimes|t\rangle = \sum_{t\in\{0,1\}^n}\left(\bigotimes_{i=1}^{n} |t_i\rangle\langle t_i|_{\mathsf{R}_{t_{<i}}}\right)\otimes|t\rangle_{\mathsf{S}},$$

where the t'th tensor product above implicitly acts as the identity on all R_x for which x does not equal $t_{<i}$ for any i. Therefore

$$U_f\left(\bigotimes_{x\in\{0,1\}^{<n}}|\phi_x\rangle_{\mathsf{R}_x}\otimes|0^n\rangle_{\mathsf{S}}\right) = \sum_{t\in\{0,1\}^n}\ \bigotimes_{x\in\{0,1\}^{<n}}\begin{cases}|t_i\rangle\langle t_i|\phi_{t_{<i}}\rangle_{\mathsf{R}_x} & \text{if } x = t_{<i} \text{ for some } i\\ |\phi_x\rangle_{\mathsf{R}_x} & \text{otherwise}\end{cases}\otimes|t\rangle_{\mathsf{S}}.$$

By the definition of $|\phi_{t_{<i}}\rangle$ it holds that $\langle t_i|\phi_{t_{<i}}\rangle = \beta_{t_{<i}t_i} = \beta_{t_{\le i}}$, so since $\alpha_t = \prod_{i=1}^n \beta_{t_{\le i}}$ for all $t \in \{0,1\}^n$ it follows that

$$U_f\left(\bigotimes_{x\in\{0,1\}^{<n}}|\phi_x\rangle_{\mathsf{R}_x}\otimes|0^n\rangle_{\mathsf{S}}\right) = \sum_{t\in\{0,1\}^n}\alpha_t\bigotimes_{x\in\{0,1\}^{<n}}\begin{cases}|t_i\rangle_{\mathsf{R}_x} & \text{if } x = t_{<i} \text{ for some } i\\ |\phi_x\rangle_{\mathsf{R}_x} & \text{otherwise}\end{cases}\otimes|t\rangle_{\mathsf{S}}.$$

All that remains to construct the state $|\psi\rangle = \sum_{t\in\{0,1\}^n}\alpha_t|t\rangle$ is to uncompute the above content of $(\mathsf{R}_x)_{x\in\{0,1\}^{<n}}$ controlled on the state $|t\rangle$ of S. To do so, first make $|\{0,1\}^{<n}|$ copies of t using fanout. Then for each $x \in \{0,1\}^{<n}$ in parallel, controlled on one of these copies of t, if $x = t_{<i}$ for some i then perform in R_x an operation that maps $|t_i\rangle$ to $|0\rangle$, and otherwise perform in R_x an operation that maps $|\phi_x\rangle$ to $|0\rangle$. Finally, uncompute the extra copies of t using fanout. □

Theorem 1.3.14. *Every n-qubit unitary transformation can be cleanly, exactly implemented by a $\mathsf{QAC_f}$ circuit of depth $O\big(2^{n/2}\big)$ with $\tilde{O}\big(2^{2n}\big)$ ancillae.*

Proof. Let U denote the n-qubit unitary to be implemented. By Theorem 2.4.2 it suffices to implement a U-qRAM with an $\tilde{O}\big(2^{2n}\big)$-qubit $\mathsf{QAC_f^0}$ circuit, and this can be achieved as follows. On input $x \in \{0,1\}^n$ to the U-qRAM, for all $y \in \{0,1\}^n$ in parallel, in a register R_y use Theorem 1.3.10 and Lemma A.2.2 to construct $U|y\rangle$ controlled on $x = y$. Then swap R_x into the output register using Lemma A.2.1. □

4.2 Size bounds for approximately constructing states over a universal gate set

In this section we define the size of a circuit to be the *total* number of gates (including one-qubit gates), in contrast to Definition 1.3.4 in which only multi-qubit gates count toward size. First we prove the upper bound:

Theorem 4.2.1 (formal version of Theorem 1.3.12)**.** *There exists a finite gate set $\mathcal{G}$ such that for all $n \in \mathbb{N}, \varepsilon \ge \exp(-\mathrm{poly}(n))$ and n-qubit states $|\psi\rangle$, there exists a size-$O(2^n\log(1/\varepsilon)/n)$ circuit C over $\mathcal{G}$ such that $\||C|0\ldots0\rangle - |\psi\rangle|0\ldots0\rangle\| \le \varepsilon$.*

Proof. Let $\mathcal{G}$ be any universal gate set that includes the Toffoli and NOT gates. By Theorem 2.3.2 and the Solovay-Kitaev theorem (Theorem 1.3.3) there exists a poly(n)-size circuit A over $\mathcal{G}$, making ten queries to a boolean function f, such that $\big\|A^f|0\ldots0\rangle - |\psi\rangle|0\ldots0\rangle\big\| \le$

ε. Inspection of the proof of Theorem 2.3.2 reveals that f has $n + \log\log(1/\varepsilon) + O(1)$ input bits, and that only the first output bit of f depends on the input to f. By Lupanov's upper bound (Theorem 1.3.9) it follows that f can be computed by an $O(2^n \log(1/\varepsilon)/n)$-size boolean circuit, where the output bits not depending on the input are hard-coded into the circuit. Since boolean circuits can be cleanly simulated by quantum circuits consisting only of Toffoli and NOT gates with a constant-factor blowup in size, it follows that f can be computed by an $O(2^n \log(1/\varepsilon)/n)$-size circuit over $\mathcal{G}$. Combining this circuit with A yields the desired result. □

Now we prove the lower bound:

Theorem 4.2.2 (formal version of Theorem 1.3.13)**.** *Let $\mathcal{G}$ be a finite gate set. Then for all $n \in \mathbb{N}$ and $1/4 \geq \varepsilon \geq \exp(-\mathrm{poly}(n))$, there exists an n-qubit state $|\psi\rangle$ such that circuits C over $\mathcal{G}$ require size $\Omega(2^n \log(1/\varepsilon)/n)$ in order for the reduced state ρ on the first n qubits of $C|0\dots0\rangle$ to satisfy $\mathrm{td}(\rho, |\psi\rangle\langle\psi|) \leq \varepsilon$.*

To properly compare Theorems 1.3.12 and 1.3.13 it is necessary to convert the error bound in Theorem 1.3.12 from 2-norm error to trace distance error. Identifying a pure state $|\phi\rangle$ with the density matrix $|\phi\rangle\langle\phi|$, the trace distance between two pure states is at most the 2-norm distance between those states (see Eq. (1.6.3)), so the conclusion of Theorem 1.3.12 implies that the trace distance between $|\psi\rangle|0\dots0\rangle$ and $C|0\dots0\rangle$ is at most ε. Therefore by Eq. (1.6.1) the trace distance between $|\psi\rangle$ and the reduced state on the first n qubits of $C|0\dots0\rangle$ is at most ε, so the lower bound from Theorem 1.3.13 matches the upper bound from Theorem 1.3.12.

Proof of Theorem 4.2.2. Let $S_n(r) = \{x \in \mathbb{R}^{n+1} : \|x\| = r\}$ and $S_n = S_n(1)$. The set of n-qubit pure states can be identified with $S_{2^{n+1}-1}$, because an n-qubit pure state is described by 2^n complex amplitudes, each of which has a real part and an imaginary part, and these 2^{n+1} real numbers form a unit vector. Let μ_n denote n-dimensional volume; then $\mu_n(S_n)$ obeys the recurrence

$$\mu_0(S_0) = 2, \qquad \mu_1(S_1) = 2\pi, \qquad \mu_{n+1}(S_{n+1}) = 2\pi\mu_{n-1}(S_{n-1})/n \quad \text{for } n \geq 1$$

and $\mu_n(S_n(r)) = r^n \mu(S_n)$ [103]. We will write $\mu = \mu_n$ when n is clear from the context.

For an n-qubit mixed state ρ and $\varepsilon \geq 0$, let $N_\varepsilon(\rho)$ denote the set of pure states $|\psi\rangle$ such that $\mathrm{td}(\rho, \psi) \leq \varepsilon$. If ρ itself is rank-1, say $\rho = |\rho\rangle\langle\rho|$, then for all pure states $|\psi\rangle$ it is well known that $\mathrm{td}(\rho, \psi) = \sqrt{1 - |\langle\rho|\psi\rangle|^2}$, and so $|\psi\rangle$ is in $N_\varepsilon(\rho)$ if and only if $|\langle\rho|\psi\rangle|^2 \geq 1 - \varepsilon^2$. Therefore

$$\mu(N_\varepsilon(\rho)) = \int_{\theta=0}^{\arcsin\varepsilon} \mu(S_1(\cos\theta))\mu(S_{2^{n+1}-3}(\sin\theta))d\theta,$$

because $\langle\rho|\psi\rangle$ is described by two real numbers whose squares sum to a value $\cos^2\theta$ between 1 and $1 - \varepsilon^2$, and the rest of $|\psi\rangle$ is described by $2^{n+1} - 2$ real numbers whose squares sum to $\sin^2\theta$. It follows that for $m = 2^{n+1}$,

$$\begin{aligned}\mu(N_\varepsilon(\rho)) &= \int_{\theta=0}^{\arcsin\varepsilon} \cos\theta \sin^{m-3}\theta d\theta \cdot \mu(S_1)\mu(S_{m-3}) = \int_{u=0}^{\varepsilon} u^{m-3}du \cdot \mu(S_1)\mu(S_{m-3}) \\ &= \varepsilon^{m-2}\mu(S_1)\mu(S_{m-3})/(m-2) = \varepsilon^{m-2}\mu(S_{m-1}).\end{aligned}$$

More generally, consider an n-qubit mixed state ρ of arbitrary rank. If $N_\varepsilon(\rho)$ is nonempty then there exists a state $|\psi\rangle \in N_\varepsilon(\rho)$, so for all $|\phi\rangle \in N_\varepsilon(\rho)$, by the triangle inequality

$\mathrm{td}(\psi,\phi) \le \mathrm{td}(\psi,\rho) + \mathrm{td}(\rho,\phi) \le 2\varepsilon$. In other words $N_\varepsilon(\rho) \subseteq N_{2\varepsilon}(\psi)$. It follows from the case proved above that

$$\mu(N_\varepsilon(\rho)) \le \mu(N_{2\varepsilon}(\psi)) \le (2\varepsilon)^{m-2}\mu(S_{m-1}) \le \varepsilon^{(m-2)/2}\mu(S_{m-1}),$$

where the last inequality holds because $\varepsilon \le 1/4$.

For $s \in \mathbb{N}$ let $\mathcal{C}_s$ denote the set of size-s circuits over $\mathcal{G}$. Circuits in $\mathcal{C}_s$ act on $O(s)$ qubits without loss of generality, and there are $\mathrm{poly}(s)$ ways to choose a gate from $\mathcal{G}$ and the qubits that it acts on out of $O(s)$ total qubits, so $|\mathcal{C}_s| \le \mathrm{poly}(s)^s \le 2^{O(s\log s)}$. In particular, if $s \le o(2^n \log(1/\varepsilon)/n)$ then $\log s \le O(n) + \log\log(1/\varepsilon) \le O(n)$ and so $2^{O(s\log s)} \le (1/\varepsilon)^{o(2^n)}$; therefore

$$\begin{aligned}\mu\left(\bigcup_{C\in\mathcal{C}_s} N_\varepsilon\left(\mathrm{tr}_{>n}\left(C|0\dots0\rangle\langle0\dots0|C^\dagger\right)\right)\right) &\le \sum_{C\in\mathcal{C}_s}\mu\left(N_\varepsilon\left(\mathrm{tr}_{>n}\left(C|0\dots0\rangle\langle0\dots0|C^\dagger\right)\right)\right)\\ &\le \sum_{C\in\mathcal{C}_s}\varepsilon^{(m-2)/2}\mu(S_{m-1}) \le \varepsilon^{(m-2)/2-o(m)}\mu(S_{m-1}) \le o(\mu(S_{m-1})).\end{aligned}$$

□

4.3 QAC^0 upper bounds for functions and states

Lemma 4.3.1. *For all n-qubit $\mathsf{QAC}^0_\mathsf{f}$ circuits A and all $\varepsilon \ge \exp(-\mathrm{poly}(n))$, there exists an $\exp(\mathrm{poly}(n))$-qubit QAC^0 circuit C such that $\|C(I_n \otimes |0\dots0\rangle) - A \otimes |0\dots0\rangle\| \le \varepsilon$.*

Proof. Let $d \le O(1)$ be the depth of A, and note that A has at most dn multi-qubit gates, each of which acts on at most n qubits. By Theorem 3.3.1 each gate can be simulated to within error ε/dn by a QAC^0 circuit with $\exp(\mathrm{poly}(n))$ ancillae, and by the triangle inequality (more specifically, Eq. (2.1.3)) it follows that the product of these simulations of individual gates of A simulates A to within error ε. □

Since every n-bit boolean function can be computed by a $\mathsf{QAC}^0_\mathsf{f}$ circuit with $\exp(O(n))$ ancillae (by fanning out copies of the input and simulating a CNF or DNF), and every n-qubit state can be constructed by a $\mathsf{QAC}^0_\mathsf{f}$ circuit with $\exp(O(n))$ ancillae (by Theorem 1.3.10), the following formalizations of Corollary 1.3.17 follow immediately from Lemma 4.3.1:

Corollary 4.3.2. *For all functions $f : \{0,1\}^n \to \{0,1\}$ and all $\varepsilon \ge \exp(-\exp(O(n)))$, there exists a QAC^0 circuit C with $\exp(\exp(O(n)))$ ancillae such that*

$$\|C(I_{n+1} \otimes |0\dots0\rangle) - U_f \otimes |0\dots0\rangle\| \le \varepsilon,$$

where $U_f|x,b\rangle = |x, b\oplus f(x)\rangle$ for all $x \in \{0,1\}^n, b \in \{0,1\}$.

Corollary 4.3.3. *For all n-qubit states $|\psi\rangle$ and all $\varepsilon \ge \exp(-\exp(O(n)))$, there exists a QAC^0 circuit C with $\exp(\exp(O(n)))$ ancillae such that $\|C|0\dots0\rangle - |\psi\rangle|0\dots0\rangle\| \le \varepsilon$.*

4.4 Barrier to $\mathsf{QAC}^0_\mathsf{f}$ lower bounds for constructing explicit states

Call a state sequence $(|\psi_n\rangle)_n$ *explicit* if $|\psi_n\rangle$ is an n-qubit state whose description can be computed in time $\exp(\mathrm{poly}(n))$ as a function of n. For example, every pure state sequence in the class $\mathsf{statePSPACE}_\mathsf{exp}$ (which we will define in Section 5.1.2) is explicit up to global phases, by Lemmas 5.1.2 and 5.1.6 and the fact that $\mathsf{PSPACE} \subseteq \mathsf{EXP}$. We say that a language is in $\mathsf{QAC}^0_\mathsf{f}$ if it can be decided with bounded error by a nonuniform sequence of polynomial-size $\mathsf{QAC}^0_\mathsf{f}$ circuits. The following is one way to more formally state Observation 1.3.18:

Theorem 4.4.1. *Assume there exists an explicit state sequence* $(|\psi_n\rangle)_n$ *and function* $\varepsilon(n) = \exp(-\mathrm{poly}(n))$ *such that for all sequences* $(C_n)_n$ *of polynomial-size* $\mathsf{QAC}^0_\mathsf{f}$ *circuits, it holds that* $\| C_n|0\ldots0\rangle - |\psi_n\rangle|0\ldots0\rangle\| \geq \varepsilon(n)$*. Then* $\mathsf{EXP} \not\subseteq \mathsf{QAC}^0_\mathsf{f}$*.*

Proof. We prove the contrapositive statement: if $\mathsf{EXP} \subseteq \mathsf{QAC}^0_\mathsf{f}$ then for all functions $\varepsilon(n) = \exp(-\mathrm{poly}(n))$, every explicit state sequence $(|\psi_n\rangle)_n$ can be constructed to within error ε in $\mathsf{QAC}^0_\mathsf{f}$. Let $C_n^{f_n}$ be the circuit-oracle combination for constructing $|\psi_n\rangle$ from Theorem 2.3.3. We argue that $(f_n)_n$ is in EXP: given n, first compute the description of $|\psi_n\rangle$ (which takes exponential time since $(|\psi_n\rangle)_n$ is explicit) and then run the assumed algorithm for f_n from Theorem 2.3.3 (which takes polynomial space and therefore exponential time). By the assumption that $\mathsf{EXP} \subseteq \mathsf{QAC}^0_\mathsf{f}$ it follows that $(f_n)_n \in \mathsf{QAC}^0_\mathsf{f}$, and therefore $\left(C_n^{f_n}\right)_n$ can be implemented in $\mathsf{QAC}^0_\mathsf{f}$. □

Chapter 5

Interactive state and unitary synthesis

> If I lied the first time, I'm not going to tell you the truth just because you ask twice.
>
> Eliezer Yudkowsky, *Harry Potter and the Methods of Rationality*

In Section 5.1 we define various state complexity classes, in Section 5.2 we prove that $\mathsf{statePSPACE} \subseteq \mathsf{stateQIP}(6)$, in Section 5.3 we prove our results about interactive unitary synthesis after defining the relevant classes, and in Section 5.4 we prove our results about multiple entangled provers after defining the relevant classes.

The history of these results is as follows. First the author and Yuen [88] proved similar results except with polynomially many rounds of interaction, using the state synthesis algorithm from Theorem 1.2.2 which makes polynomially many queries. Then in followup work, Metger and Yuen [76] simplified some of our proofs and adopted slightly different definitions of state complexity classes (in particular, they generalized our definitions from pure states to mixed states). After that, the author [86] used the one-query state synthesis algorithm from Theorem 1.2.4 to decrease the number of rounds of interaction from polynomial to constant, and incorporated the simplifications and revised definitions introduced by Metger and Yuen [76]. This latter work of the author [86] considered only *state* synthesis, but here we rephrase the author and Yuen's [88] results about *unitary* synthesis and their proofs to use the revised definitions and simplifications from the aforementioned followup work [76, 86]. Below we present only the simplest, most modern versions of the results and proofs, while pointing out many of the differences with the original versions.

5.1 State complexity classes

In this section we define various state complexity classes and establish some basic facts about them as preparation for the proof that $\mathsf{statePSPACE} \subseteq \mathsf{stateQIP}(6)$. Although for simplicity these classes are defined in terms of state sequences where the n'th state is on n qubits, the definitions (and related results) generalize easily to the case where the n'th state is on $\text{poly}(n)$ qubits. Some of the language in this section is modeled on passages from Metger and Yuen [76].

5.1.1 polyL-explicit state sequences

Recall from Section 1.6 that we define an ε-precision description of a pure state $\sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$ to be a tuple $(\tilde{\alpha}_x)_{x\in\{0,1\}^n}$ of complex numbers specified exactly in binary such that $|\tilde{\alpha}_x - \alpha_x| \leq \varepsilon$ for all x. We define a similar notion for mixed states: an *ε-precision description* of a mixed state $\sum_{x,y\in\{0,1\}^n} \rho_{x,y}|x\rangle\langle y|$ is a tuple $(\tilde{\rho}_{x,y})_{x,y\in\{0,1\}^n}$ of complex numbers specified exactly in binary such that $|\tilde{\rho}_{x,y} - \rho_{x,y}| \leq \varepsilon$ for all x, y.

Definition 5.1.1 (polyL-explicit state sequences)**.** Let $|\psi_n\rangle$ be an n-qubit pure state for all n. We call the sequence $(|\psi_n\rangle)_n$ polyL-*explicit* if for all functions of the form $\varepsilon(n) = \exp(-\mathrm{poly}(n))$, there is an algorithm that on input n outputs an $\varepsilon(n)$-precision description of $|\psi_n\rangle$ using space $\mathrm{poly}(n)$ (i.e. space polylogarithmic in the output length).

Similarly, let ρ_n be an n-qubit mixed state for all n. We call the sequence $(\rho_n)_n$ polyL-*explicit* if for all functions of the form $\varepsilon(n) = \exp(-\mathrm{poly}(n))$, there is an algorithm that on input n outputs an $\varepsilon(n)$-precision description of ρ_n using space $\mathrm{poly}(n)$.

Lemma 5.1.2. *Let $(\rho_n)_n$ be a* polyL-*explicit sequence of rank-1 mixed states. Then there is a* polyL-*explicit sequence of pure states $(|\psi_n\rangle)_n$ such that $\rho_n = |\psi_n\rangle\langle\psi_n|$ for all n.*

Proof. Fix n and write $\rho = \rho_n = \sum_{x,y\in\{0,1\}^n} \rho_{x,y}|x\rangle\langle y|$. Let $(\tilde{\rho}_{x,y})_{x,y\in\{0,1\}^n}$ be a $\left(\frac{1}{4}\cdot 2^{-n}\right)$-precision description of ρ computable in $\mathrm{poly}(n)$ space. Since $\mathrm{tr}(\rho) = 1$ there exists a string x such that $\rho_{x,x} \geq 2^{-n}$, implying that $\tilde{\rho}_{x,x} \geq \rho_{x,x} - \frac{1}{4}\cdot 2^{-n} \geq \frac{3}{4}\cdot 2^{-n}$. Let y be the lexicographically first string such that $\tilde{\rho}_{y,y} \geq \frac{3}{4}\cdot 2^{-n}$ (which we have just shown to exist) and observe that $\rho_{y,y} \geq \tilde{\rho}_{y,y} - \frac{1}{4}\cdot 2^{-n} \geq \frac{1}{2}\cdot 2^{-n}$. Let

$$|\psi\rangle = |\psi_n\rangle = \frac{\rho|y\rangle}{\sqrt{\rho_{y,y}}} = \sum_{x\in\{0,1\}^n} \frac{\rho_{x,y}}{\sqrt{\rho_{y,y}}}|x\rangle.$$

Since ρ is rank-1 it is easy to see that $\rho = \psi$.

For $\varepsilon = \exp(-\mathrm{poly}(n))$ an ε-precision description of $|\psi\rangle$ can be computed in $\mathrm{poly}(n)$ space as follows. Let $\delta = \frac{1}{64}\cdot 2^{-2n}\varepsilon^2 \geq \exp(-\mathrm{poly}(n))$ and let $(\sigma_{x,y'})_{x,y'\in\{0,1\}^n}$ be a δ-precision description of ρ computable in $\mathrm{poly}(n)$ space. First compute y (using that $\tilde{\rho}$ can be computed in $\mathrm{poly}(n)$ space), and then output $\left(\sigma_{x,y}/\sqrt{\sigma_{y,y}}\right)_{x\in\{0,1\}^n}$.

This algorithm is correct, because by the triangle inequality

$$\begin{aligned}\left|\frac{\sigma_{x,y}}{\sqrt{\sigma_{y,y}}} - \frac{\rho_{x,y}}{\sqrt{\rho_{y,y}}}\right| &= \left|\frac{\sigma_{x,y}\sqrt{\rho_{y,y}} - \sqrt{\sigma_{y,y}}\rho_{x,y}}{\sqrt{\sigma_{y,y}\rho_{y,y}}}\right| \leq \frac{\sqrt{\rho_{y,y}}\cdot|\sigma_{x,y} - \rho_{x,y}| + |\rho_{x,y}|\cdot\left|\sqrt{\rho_{y,y}} - \sqrt{\sigma_{y,y}}\right|}{\sqrt{(\rho_{y,y} - \delta)\rho_{y,y}}} \\ &\leq \frac{\delta + \sqrt{|\rho_{y,y} - \sigma_{y,y}|}}{\sqrt{\left(\frac{1}{2}\cdot 2^{-n} - \delta\right)\cdot\frac{1}{2}\cdot 2^{-n}}} \leq \frac{2\sqrt{\delta}}{\sqrt{\frac{1}{8}\cdot 2^{-2n}}} \leq \varepsilon,\end{aligned}$$

where the second-to-last inequality uses that $\delta \leq \frac{1}{4}\cdot 2^{-n}$. □

5.1.2 The class statePSPACE

For convenience we use the universal gate set $\{H, CNOT, T\}$ [80] in the following definition, although our results hold for any universal gate set consisting of gates with algebraic entries.

Definition 5.1.3 (General quantum circuits and space-uniformity)**.** A *general quantum circuit* is a circuit consisting of gates from the set $\{H, CNOT, T\}$ as well as non-unitary gates that (a) introduce new qubits initialized in the zero state, (b) trace them out, or (c) measure them in the standard basis. A general quantum circuit uses space s if at most s qubits are involved at any time step of the computation. The description of a general quantum circuit is the sequence of its gates (unitary or non-unitary) along with a specification of which qubits they act on.

We call a sequence $(C_n)_n$ of general quantum circuits *space-uniform* if C_n uses space $\mathrm{poly}(n)$, and there is an algorithm that on input n uses space $\mathrm{poly}(n)$ and outputs the (possibly exponentially long) description of C_n.

Definition 5.1.4 (statePSPACE and variants thereof)**.** For $\delta : \mathbb{N} \to [0, \infty)$, let $\mathsf{statePSPACE}_\delta$ be the class of all sequences of mixed states $(\rho_n)_n$ such that each ρ_n is a state on n qubits, and there exists a space-uniform sequence of general quantum circuits $(C_n)_n$ such that for all sufficiently large n, the circuit C_n takes no inputs and C_n outputs a mixed state σ_n such that $\mathrm{td}(\sigma_n, \rho_n) \leq \delta(n)$. Let $\mathsf{statePSPACE} = \bigcap_p \mathsf{statePSPACE}_{1/p}$ and $\mathsf{statePSPACE}_{\mathsf{exp}} = \bigcap_p \mathsf{statePSPACE}_{\exp(-p)}$ where p ranges over all polynomials.

We abuse notation and write $(|\psi_n\rangle)_n \in \mathsf{statePSPACE}_\delta$ if $(|\psi_n\rangle\langle\psi_n|)_n$ is in $\mathsf{statePSPACE}_\delta$. Also recall that the definitions of state complexity classes such as $\mathsf{statePSPACE}_\delta$ generalize easily to sequences where the n'th state is on $\mathrm{poly}(n)$ qubits. With this in mind we can state the following result, which in particular implies that $\mathsf{statePSPACE}_{\mathsf{exp}}$ is closed under purification:

Lemma 5.1.5 ([76, part of Theorem 6.1][1])**.** *Let $(\rho_n)_n \in \mathsf{statePSPACE}_\delta$ be a sequence of mixed states for some function δ. Then there exists a sequence of pure states $(|\psi_n\rangle)_n \in \bigcap_{\varepsilon(n)=\exp(-\mathrm{poly}(n))} \mathsf{statePSPACE}_{2\sqrt{\delta}+\varepsilon}$ such that $|\psi_n\rangle$ is a purification of ρ_n for all n.*

We also use the following:

Lemma 5.1.6. *Every sequence of mixed states in $\mathsf{statePSPACE}_{\mathsf{exp}}$ is polyL-explicit.*

Proof. Metger and Yuen [76, Lemma 6.2] proved that every sequence of mixed states in $\mathsf{statePSPACE}_0$ is polyL-explicit. The general case follows by the triangle inequality. □

We remark that [76, Lemma 6.2] generalizes and simplifies previous work of the author and Yuen [88, Section 5]. The high-level idea behind the proof is that tomography of states in $\mathsf{statePSPACE}_0$ can be done in $\mathsf{BQPSPACE}$, and $\mathsf{BQPSPACE} = \mathsf{PSPACE}$ [101]. The proof of $\mathsf{BQPSPACE} = \mathsf{PSPACE}$ relies on the assumption that the gates used in Definition 5.1.3 have algebraic entries, which is why we imposed this requirement.

5.1.3 Quantum interactive protocols

Since in quantum computing the standard model of computation is the quantum circuit model (rather than quantum Turing machines), we model the verifier in a quantum interactive protocol as a sequence of *verifier circuits*, one for each input length. A verifier circuit is itself a tuple of quantum circuits that correspond to the operations performed by the verifier in each round of the protocol. Below we describe this more formally.

The case where the verifier sends the first message is illustrated in Fig. 14. For a register A let $\mathrm{D}(\mathsf{A})$ denote the set of density matrices on A. A $2r$*-message quantum verifier circuit*

[1] As of this writing the ε term is omitted from [76, Theorem 6.1], but inspection of their proof reveals that this omission is an error.

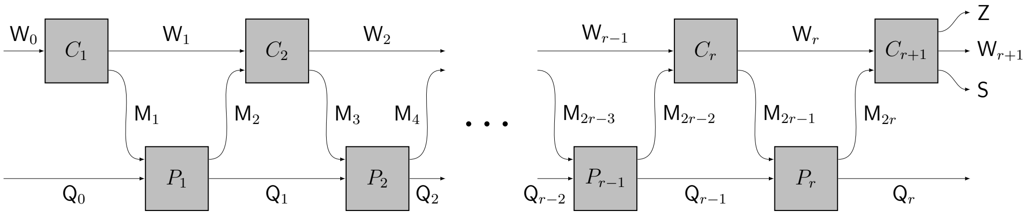

Figure 14: Generic quantum interactive protocol.

$C = (C_j)_{j\in[r+1]}$ is a tuple of general quantum circuits, where $C_1 : \mathrm{D}(\mathsf{W}_0) \to \mathrm{D}(\mathsf{W}_1\mathsf{M}_1)$, and $C_j : \mathrm{D}(\mathsf{W}_{j-1}\mathsf{M}_{2j-2}) \to \mathrm{D}(\mathsf{W}_j\mathsf{M}_{2j-1})$ for $2 \le j \le r$, and $C_{r+1} : \mathrm{D}(\mathsf{W}_r\mathsf{M}_{2r}) \to \mathrm{D}(\mathsf{Z}\mathsf{W}_{r+1}\mathsf{S})$. A *quantum prover* P for such a verifier circuit C is a tuple of quantum channels $(P_j)_{j\in[r]}$ where $P_j : \mathrm{D}(\mathsf{Q}_{j-1}\mathsf{M}_{2j-1}) \to \mathrm{D}(\mathsf{Q}_j\mathsf{M}_{2j})$. We think of W_j (resp. Q_j) as the verifier's (resp. prover's) private memory at a given time, and we think of M_j as the j'th message. At the end of the protocol, the verifier produces a one-qubit register Z indicating whether to accept or reject, and a register S containing an output state.

Let x denote a string whose length is at most the number of qubits in W_0. We write $C(x)\leftrightarrows P$ to denote the interaction between the verifier circuit C and the prover P on input x, which means applying the channels C_j and P_j as pictured in Fig. 14 to the initial state $|x,0\dots0\rangle_{\mathsf{W}_0}|0\dots0\rangle_{\mathsf{Q}_0}$. We say that $C(x)\leftrightarrows P$ accepts (resp. rejects) if measuring Z in the standard basis yields the outcome 1 (resp. 0). If $C(x)\leftrightarrows P$ accepts with nonzero probability, then by the *output of* $C(x)\leftrightarrows P$ *conditioned on accepting* we mean the reduced state in S conditioned on $C(x)\leftrightarrows P$ accepting. In other words if ρ denotes the output of C_{r+1}, then the output state conditioned on accepting is

$$\mathrm{tr}_{\mathsf{W}_{r+1}}\left(\frac{\langle 1|_{\mathsf{Z}}\rho|1\rangle_{\mathsf{Z}}}{\mathrm{tr}(\langle 1|_{\mathsf{Z}}\rho|1\rangle_{\mathsf{Z}})}\right).$$

By dilating we can assume without loss of generality that the prover's channels are all unitary, i.e. $P_j(A) = U_jAU_j^\dagger$ for some unitary U_j, and similarly for the verifier. (This is the purpose of the registers $\mathsf{Q}_0, \mathsf{Q}_r, \mathsf{W}_{r+1}$.) We always assume that the prover is unitary, but only sometimes assume that the verifier is unitary.

We can model interactions in which the prover sends the first (nontrivial) message by requiring M_1 to only convey the input string x that was in W_0. In this case there are only $2r-1$ (nontrivial) messages.

We say that a sequence of quantum verifier circuits $(V_n)_n$ is *uniform* if the total number gates in all circuits in V_n is $\mathrm{poly}(n)$, and the descriptions of the circuits in V_n can be computed in $\mathrm{poly}(n)$ time as a function of n. For $m : \mathbb{N} \to \mathbb{N}$, an *$m$-message quantum verifier* is a uniform sequence $V = (V_n)_n$ of quantum verifier circuits where V_n defines a protocol with $m(n)$ messages. These $m(n)$ messages include messages sent by both the verifier and prover, and do not include the trivial first message sent by the verifier if $m(n)$ is odd.

5.1.4 The class $\mathsf{QIP}(3)$

The class QIP is the standard quantum analogue of the complexity class IP. For our purposes we will only need to define the three-message version of QIP, known as $\mathsf{QIP}(3)$. Below we abbreviate $V_{|x|}(x)\leftrightarrows P$ by $V(x)\leftrightarrows P$.

Definition 5.1.7 ($\mathsf{QIP}(3)$)**.** For $\varepsilon : \mathbb{N} \to [0,1]$, the class $\mathsf{QIP}_\varepsilon(3)$ is the set of languages $L \subseteq \{0,1\}^*$ for which there exists a three-message quantum verifier $V = (V_n)_n$ (with no output state) satisfying the following conditions:

- *Completeness:* For all $x \in L$, there exists a quantum prover P (called an *honest prover*) such that $\Pr(V(x) \leftrightarrows P \text{ accepts}) = 1$.[2]
- *Soundness:* For all $x \notin L$ and all quantum provers P, it holds that $\Pr(V(x) \leftrightarrows P \text{ accepts}) \leq \varepsilon(|x|)$.

Here the probability is over the randomness of the interaction. Define $\mathsf{QIP}(3) = \bigcap_p \mathsf{QIP}_{2^{-p}}(3)$ where p ranges over all polynomials.

Theorem 1.4.1 ([60, 100])**.** $\mathsf{QIP}(3) = \mathsf{QIP} = \mathsf{PSPACE}$.

It is straightforward to generalize the inclusion $\mathsf{PSPACE} \subseteq \mathsf{QIP}(3)$ in Theorem 1.4.1 from decision problems to functions:

Corollary 5.1.8. *Let $f : \{0,1\}^* \to \{0,1\}^*$ be a PSPACE-computable function such that $|f(x)| \leq \mathrm{poly}(|x|)$ for all x, and let ε be a function of the form $\varepsilon(n) = \exp(-\mathrm{poly}(n))$. Then there exists a three-message quantum verifier $V = (V_n)_n$ satisfying the following conditions:*

- Completeness: *For all $x \in \{0,1\}^*$, there exists a quantum prover P (called an* honest prover*) such that $\Pr(V(x) \leftrightarrows P$ accepts and outputs $f(x)) = 1$.*
- Soundness: *For all $x \in \{0,1\}^*$ and all quantum provers P,*

$$\Pr(V(x) \leftrightarrows P \text{ accepts and outputs a string other than } f(x)) \leq \varepsilon(|x|).$$

Proof. The language $L = \{(x, f(x)) : x \in \{0,1\}^*\}$ is clearly in PSPACE, so by Theorem 1.4.1 there exists a $\mathsf{QIP}_\varepsilon(3)$ verifier V_L for L. A verifier V_f for f can be described as follows. First V_f sends the input string x to the prover. Then V_f receives a register M from the prover, measures M in the standard basis to obtain a string y, and simulates V_L on input (x, y). (Here the prover is expected to send both y and the first nontrivial message from the simulation of V_L in the same message, so that the total number of nontrivial messages is still three.) If V_L accepts then V_f accepts and outputs y, otherwise V_f rejects.

Completeness holds because an honest prover for V_f can send $y = f(x)$ and then simulate an honest prover for V_L. Soundness holds because conditioned on any string $y \neq f(x)$ that the verifier measures in M, the probability that V_L accepts is at most $\varepsilon(|x|)$ by the soundness of V_L. □

5.1.5 The classes $\mathsf{stateQIP}(m)$ and $\mathsf{stateQIP}$

Definition 5.1.9 ($\mathsf{stateQIP}(m)$ and $\mathsf{stateQIP}$)**.** Let $\varepsilon, \delta : \mathbb{N} \to [0, \infty)$ and $m : \mathbb{N} \to \mathbb{N}$ be functions. The class $\mathsf{stateQIP}_{\varepsilon,\delta}(m)$ is the set of mixed state sequences $(\rho_n)_n$ (where ρ_n is on n qubits) for which there exists an m-message quantum verifier $(V_n)_n$ satisfying the following for all sufficiently large n:

- *Completeness:* There exists a quantum prover P (called an *honest prover*) such that $\Pr(V_n \leftrightarrows P \text{ accepts}) = 1$.

[2] The reader may wonder whether the definition of $\mathsf{QIP}(3)$ here is sensitive to the assumption of perfect completeness; it is known that if the verifier uses the universal gate set $\{H, CNOT, T\}$, then we can assume perfect completeness without loss of generality [99, Section 4.2].

- *Soundness:* For all quantum provers P such that $\Pr(V_n \leftrightarrows P \text{ accepts}) \geq \varepsilon(n)$, it holds that $\mathrm{td}(\sigma, \rho_n) \leq \delta(n)$ where σ denotes the output of $V_n \leftrightarrows P$ conditioned on accepting.

Here the probabilities are over the randomness of the interaction.

Finally, define

$$\mathsf{stateQIP}(m) = \bigcap_{p,q} \mathsf{stateQIP}_{\frac{1}{p},\frac{1}{q}}(m), \qquad \mathsf{stateQIP} = \bigcup_{m'} \mathsf{stateQIP}(m')$$

where p, q, m' range over all polynomials.

Remark. Metger and Yuen [76] fixed p to 2 in their definition of $\mathsf{stateQIP}$, i.e. they considered the class $\mathsf{stateQIP}' = \bigcup_m \bigcap_q \mathsf{stateQIP}_{\frac{1}{2},\frac{1}{q}}(m)$. However our definitions are equivalent because

$$\mathsf{statePSPACE} \subseteq \mathsf{stateQIP}(6) \subseteq \mathsf{stateQIP} \subseteq \mathsf{stateQIP}' \subseteq \mathsf{statePSPACE},$$

where the first inclusion is Theorem 1.4.3, the second and third inclusions are trivial, and the fourth inclusion was proved by Metger and Yuen [76].

5.2 Proof that $\mathsf{statePSPACE} \subseteq \mathsf{stateQIP}(6)$

In this section we use the background from Section 5.1 to prove Theorem 1.4.3, i.e. that $\mathsf{statePSPACE} \subseteq \mathsf{stateQIP}(6)$. Let $(\rho_n)_n \in \mathsf{statePSPACE}$ and let $\varepsilon(n), \delta(n) = 1/\mathrm{poly}(n)$; below we prove that $(\rho_n)_n$ is in $\mathsf{stateQIP}_{\varepsilon,\delta}(6)$ which establishes the theorem.

5.2.1 The protocol

Since $(\rho_n)_n$ is in $\mathsf{statePSPACE}$ there exists a sequence $(\sigma_n)_n \in \mathsf{statePSPACE}_0$ such that $\mathrm{td}(\rho_n, \sigma_n) \leq \delta(n)/2$. By Lemma 5.1.5 there exists a sequence of pure states $(|\psi_n\rangle)_n \in \mathsf{statePSPACE}_{\mathsf{exp}}$ such that the reduced state on the first n qubits of $|\psi_n\rangle$ equals σ_n. By Lemma 5.1.6 the sequence $(\psi_n)_n$ is polyL-explicit, so by Lemma 5.1.2 the sequence $(|\psi_n\rangle)_n$ is polyL-explicit up to global phases. Therefore by Theorem 2.3.1 there exists a uniform sequence of polynomial-size quantum circuits $(A_n)_n$, making one query to a PSPACE-computable function f, such that the reduced state on the initial qubits of $A_n^f|0\ldots0\rangle$ is within 2^{-n} trace distance of ψ_n, and furthermore $(A_n)_n$ does not depend on $(\rho_n)_n$. Henceforth we will fix n and write $\rho = \rho_n, \varepsilon = \varepsilon(n)$ and so on for brevity.

Let $m = \mathrm{poly}(n)$ be the number of qubits on which A acts. By the discussion in Section 2.1, we can assume without loss of generality that f has a single output bit and that the query in A^f is of the form $D = \sum_{x\in\{0,1\}^m}(-1)^{f(x)}|x\rangle\langle x|$. Write $A^f|0^m\rangle = CD|\phi\rangle$ where C is the portion of A applied after the query, and $|\phi\rangle$ is the state constructed by the portion of A applied before the query.

Let $t = \mathrm{poly}(n)$ be a parameter to be chosen later, and for $x_1, \ldots, x_t \in \{0,1\}^m$ let $F(x_1, \ldots, x_t) = (f(x_1), \ldots, f(x_t))$. Since f is PSPACE-computable, so is F. Let V_F be the three-message quantum verifier circuit for F guaranteed to exist by Corollary 5.1.8, with soundness parameter 2^{-2n}. As mentioned in Section 5.1.3 we can assume without loss of generality that V_F is unitary. We can also assume without loss of generality that V_F preserves the classical state $|x\rangle$ of its input register, e.g. by defining a verifier circuit that makes a copy of x and simulates V_F on the copy.

We assign names to certain registers associated with V_F as follows. Let A be the input register, and write $\mathsf{A} = \mathsf{A}_1 \cdots \mathsf{A}_t$ where each A_j is an m-qubit register. Let S be the output register (which on input x, ideally holds $F(x)$), and write $\mathsf{S} = \mathsf{S}_1 \cdots \mathsf{S}_t$ where each S_j is a one-qubit register. Let Z be the one-qubit register indicating whether to accept or reject, and let W be the register disjoint from AZS that holds the rest of the output of V_F's final circuit.

Procedure 5 describes a verifier circuit for constructing ρ. There are six messages in total, because Line 4 requires four messages (including sending x to the prover) and Lines 10 and 11 each require one message.

Procedure 5 $\mathsf{stateQIP}_{\varepsilon,\delta}(6)$ verifier circuit for ρ

```
 1: for $j \in [t]$ do construct $|\phi\rangle_{\mathsf{A}_j}$.
 2: end for
 3: controlled on the classical state $|x\rangle_{\mathsf{A}}$,
 4:     Simulate $V_F$ on input $x$.
 5: end control
 6: if a standard-basis measurement of $\mathsf{Z}$ outputs 0 then reject and abort.
 7: end if
 8: Sample $k \in [t]$ uniformly at random.
 9: Apply the Pauli matrix $Z$ in $\mathsf{S}_k$.        ▷ $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$
10: Send $\mathsf{SW}$ to the prover.
11: Receive a $tm$-qubit register $\mathsf{M}$ from the prover.
12: controlled on the classical state $|x\rangle_{\mathsf{A}}$,
13:     XOR $x$ into $\mathsf{M}$.
14: end control
15: for $j \in [t] \backslash \{k\}$ do perform the projective measurement $(\phi, I - \phi)$ in $\mathsf{A}_j$.
16:     if the measurement outcome is $I - \phi$ then reject and abort.
17:     end if
18: end for
19: Apply $C_{\mathsf{A}_k}$.
20: accept and return the first $n$ qubits of $\mathsf{A}_k$.
```

Remark. The projective measurement in Line 15 does not generalize to protocols where multiple, *adaptive* invocations of the $\mathsf{QIP}(3) = \mathsf{PSPACE}$ protocol are made. Therefore our earlier version of the proof, which was based on a multiple-query state synthesis algorithm (Theorem 1.2.2), instead used a more complicated sequence of swap tests to achieve a similar effect [88, Section 6]. Our earlier version of the proof also used a more general soundness amplification procedure [88, Lemma 4.4] that applies to *any* $\mathsf{stateQIP}$ protocol, but that cannot trivially be parallelized.

5.2.2 Proof of completeness

We describe an honest prover P. On Line 4 P simulates an honest prover P_F for V_F. We can assume without loss of generality that if x denotes V_F's input string, then the final state of P_F's workspace includes a copy of x (e.g. by having P_F make an extra copy of x at the beginning of its computation). Write $|\phi\rangle^{\otimes t} = \sum_{x \in \{0,1\}^{tm}} \alpha_x |x\rangle$; then we can write the state of the system immediately after Line 4 as

$$\sum_{x\in\{0,1\}^{tm}} \alpha_x |x\rangle_{\mathsf{A}} |F(x)\rangle_{\mathsf{S}} |1\rangle_{\mathsf{Z}} |x\rangle_{\mathsf{M}} |\theta_x\rangle_{\mathsf{WQ}}.$$

Here M is a register held by P (which will later be sent to the verifier in Line 11), the register Q denotes the remainder of P's private workspace, and $|\theta_x\rangle$ is some state.

Let k be the value chosen by the verifier in Line 8. Given the above state, clearly applying Z_{S_k} has the same effect that applying D_{A_k} would have, so the state of the system after Line 10 is

$$D_{\mathsf{A}_k} \cdot \sum_{x\in\{0,1\}^{tm}} \alpha_x |x\rangle_{\mathsf{A}} |F(x)\rangle_{\mathsf{S}} |x\rangle_{\mathsf{M}} |\theta_x\rangle_{\mathsf{WQ}}$$

where A is held by the verifier and SMWQ is held by P.

Next P uncomputes the state $|F(x)\rangle_{\mathsf{S}} |\theta_x\rangle_{\mathsf{WQ}}$ controlled on $|x\rangle_{\mathsf{M}}$, and then sends M to the verifier in Line 11. After Line 14 the verifier holds the state

$$D_{\mathsf{A}_k} \cdot \sum_{x\in\{0,1\}^{tm}} \alpha_x |x\rangle_{\mathsf{A}} = D_{\mathsf{A}_k} \cdot \bigotimes_{j\in[t]} |\phi\rangle_{\mathsf{A}_j},$$

which clearly passes the subsequent measurements with probability 1.

5.2.3 Proof of soundness

It will be convenient to refer to the output register in a manner independent of the random variable k from Line 8. To this end, let O be an m-qubit register, and imagine that the verifier's final action is to apply the channel Φ_k that acts as the identity on the system except that Φ_k renames A_k as O. Fix a prover such that the verifier accepts with probability $\varepsilon' \geq \varepsilon$. Let τ denote the state of the system at the end of the protocol, conditioned on accepting, and let τ^O denote the reduced state of τ on O. Then $\mathrm{tr}_{>n}(\tau^O)$ is the output state conditioned on accepting.

Let n' be the number of qubits comprising $|\psi\rangle$. By the triangle inequality, Eqs. (1.6.1) and (1.6.2), and various definitions from Section 5.2.1, it holds that

$$\mathrm{td}\big(\mathrm{tr}_{>n}(\tau^O), \rho\big) \leq \mathrm{td}\big(\mathrm{tr}_{>n}(\tau^O), \sigma\big) + \mathrm{td}(\sigma, \rho) \leq \mathrm{td}\big(\mathrm{tr}_{>n}(\tau^O), \mathrm{tr}_{>n}(\psi)\big) + \delta/2 \qquad (5.2.1)$$

and that

$$\begin{aligned}\mathrm{td}\big(\mathrm{tr}_{>n}(\tau^O), \mathrm{tr}_{>n}(\psi)\big) &\leq \mathrm{td}\big(\mathrm{tr}_{>n'}(\tau^O), \psi\big)\\ &\leq \mathrm{td}\Big(\mathrm{tr}_{>n'}(\tau^O), \mathrm{tr}_{>n'}\Big(CD\phi DC^\dagger\Big)\Big) + \mathrm{td}\Big(\mathrm{tr}_{>n'}\Big(CD\phi DC^\dagger\Big), \psi\Big)\\ &\leq \mathrm{td}\Big(\tau^O, CD\phi DC^\dagger\Big) + 2^{-n} \leq \sqrt{\mathrm{tr}(\tau \cdot (I - CD\phi DC^\dagger)_{\mathsf{O}})} + 2^{-n}.\end{aligned} \qquad (5.2.2)$$

Let $|\varphi\rangle$ denote the state of the system after Line 5, and let U be the unitary jointly applied by the verifier and prover from Line 10 to Line 14. Then

$$\varepsilon'\tau = \frac{1}{t}\sum_{k=1}^{t} \Phi_k(\theta_k) \qquad \text{for} \qquad |\theta_k\rangle = \bigotimes_{j\neq k} \langle\phi|_{\mathsf{A}_j} \cdot C_{\mathsf{A}_k} U Z_{\mathsf{S}_k} \langle 1|_{\mathsf{Z}} |\varphi\rangle,$$

where $|\theta_k\rangle$ is (in general) subnormalized and $\theta_k = |\theta_k\rangle\langle\theta_k|$. Let

$$Q = \sum_{x\in\{0,1\}^{tm}} x_{\mathsf{A}} \otimes F(x)_{\mathsf{S}},$$

and similarly define a matrix $\tilde{\tau}$ as follows:

$$\varepsilon'\tilde{\tau} = \frac{1}{t}\sum_{k=1}^{t} \Phi_k\Big(\tilde{\theta}_k\Big) \qquad \text{for} \qquad \Big|\tilde{\theta}_k\Big\rangle = \bigotimes_{j\neq k}\langle\phi|_{\mathsf{A}_j} \cdot C_{\mathsf{A}_k} U Z_{\mathsf{S}_k} Q\langle 1|_{\mathsf{Z}}|\varphi\rangle.$$

We now argue that $\tilde{\tau}$ is a close approximation of τ, using the soundness property of V_F. For $k \in [t]$ it holds that $\Big\|\Big|\tilde{\theta}_k\Big\rangle - |\theta_k\rangle\Big\|^2 \leq \|(I-Q)\langle 1|_{\mathsf{Z}}|\varphi\rangle\|^2$. This bound equals the probability that if the register ASZ of $|\varphi\rangle$ is measured in the standard basis, then the measurement outcome is of the form $|x\rangle_{\mathsf{A}}|y\rangle_{\mathsf{S}}|1\rangle_{\mathsf{Z}}$ where $y \neq F(x)$. Conditioning on x and applying the soundness of V_F shows that this event has probability at most 2^{-2n}, so $\Big\|\Big|\tilde{\theta}_k\Big\rangle - |\theta_k\rangle\Big\| \leq 2^{-n}$. Therefore by the triangle inequality,

$$\begin{aligned}
\varepsilon'\|\tilde{\tau}-\tau\|_1 &\leq \frac{1}{t}\sum_{k=1}^{t}\Big\|\Big|\tilde{\theta}_k\Big\rangle\Big\langle\tilde{\theta}_k\Big| - |\theta_k\rangle\langle\theta_k|\Big\|_1 \\
&\leq \frac{1}{t}\sum_{k=1}^{t}\Big(\Big\|\Big(\Big|\tilde{\theta}_k\Big\rangle - |\theta_k\rangle\Big)\Big\langle\tilde{\theta}_k\Big|\Big\|_1 + \Big\||\theta_k\rangle\Big(\Big\langle\tilde{\theta}_k\Big| - \langle\theta_k|\Big)\Big\|_1\Big) \\
&\leq \frac{1}{t}\sum_{k=1}^{t}\Big(2^{-n}\Big\|\Big|\tilde{\theta}_k\Big\rangle\Big\| + 2^{-n}\Big) \leq 2^{-n}\big(1+2^{-n}\big) + 2^{-n} \leq \exp(-\Omega(n)).
\end{aligned}$$

Since $\varepsilon' \geq \varepsilon \geq 1/\mathrm{poly}(n)$ it follows that $\|\tilde{\tau}-\tau\|_1 \leq \exp(-\Omega(n))$.

Let $P = (I - CD\phi DC^\dagger)_{\mathsf{O}}$. Since P is an orthogonal projection,

$$\mathrm{tr}(\tau P) \leq \mathrm{tr}(\tilde{\tau}P) + \frac{\|\tilde{\tau}-\tau\|_1}{2} \leq \mathrm{tr}(\tilde{\tau}P) + \exp(-\Omega(n)). \tag{5.2.3}$$

By reasoning similar to that in Section 5.2.2, it holds that $UZ_{\mathsf{S}_k}Q = UD_{\mathsf{A}_k}Q = D_{\mathsf{A}_k}UQ$, so defining the subnormalized vector $|\varphi'\rangle = UQ\langle 1|_{\mathsf{Z}}|\varphi\rangle$ it holds that

$$\Big|\tilde{\theta}_k\Big\rangle = \bigotimes_{j\neq k}\langle\phi|_{\mathsf{A}_j} \cdot (CD)_{\mathsf{A}_k}|\varphi'\rangle.$$

Therefore since trace is linear,

$$\begin{aligned}
\varepsilon'\,\mathrm{tr}(\tilde{\tau}P) &= \frac{1}{t}\sum_{k=1}^{t}\mathrm{tr}\Big(\Phi_k\Big(\tilde{\theta}_k\Big)P\Big) = \frac{1}{t}\sum_{k=1}^{t}\mathrm{tr}\Big(\tilde{\theta}_k\cdot(I - CD\phi DC^\dagger)_{\mathsf{A}_k}\Big) \\
&= \frac{1}{t}\,\mathrm{tr}\left(\varphi'\cdot\sum_{k=1}^{t}\bigotimes_{j\neq k}\phi_j\otimes(I-\phi_k)\right) \leq \frac{1}{t}\,\mathrm{tr}\big(\varphi'\big) \leq \frac{1}{t},
\end{aligned}$$

where we used that $\sum_{k=1}^{t}\bigotimes_{j\neq k}\phi_j\otimes(I-\phi_k)$ is an orthogonal projection. Since $\varepsilon' \geq \varepsilon$ it follows that

$$\operatorname{tr}(\tilde{\tau}P) \leq 1/(\varepsilon t). \tag{5.2.4}$$

Choose $t = \lceil 16/(\varepsilon\delta^2) \rceil \leq \operatorname{poly}(n)$. Then for all sufficiently large n, it follows from Eqs. (5.2.1) to (5.2.4) that

$$\operatorname{td}\big(\operatorname{tr}_{>n}(\tau^O), \rho\big) \leq \sqrt{\frac{1}{\varepsilon t} + \exp(-\Omega(n)) + 2^{-n} + \delta/2} \leq \frac{2}{\sqrt{\varepsilon t}} + \frac{\delta}{2} \leq \delta$$

as desired.

5.3 Interactive unitary synthesis

We define the following classes, where by space-uniform sequences of unitary quantum circuits, we mean sequences of circuits like those in Definition 5.1.3 except with only unitary gates.

Definition 5.3.1 (unitaryPSPACE)**.** The class unitaryPSPACE consists of all space-uniform sequences $(U_n)_n$ of unitary quantum circuits such that each U_n acts on n qubits.

Our definition of unitaryPSPACE does not allow any error in the circuit, making it more analogous to statePSPACE$_0$ than it is to statePSPACE. This is just for convenience, and the definition can easily be generalized to allow some error as Bostanci et al. [24] do in followup work.[3]

Definition 5.3.2 (unitaryQIP(m) and unitaryQIP)**.** Let $\varepsilon, \delta : \mathbb{N} \to [0, \infty)$ and $m : \mathbb{N} \to \mathbb{N}$ be functions. The class unitaryQIP$_{\varepsilon,\delta}(m)$ is the set of unitary sequences $(U_n)_n$ (where U_n acts on n qubits) for which there exists an m-message quantum verifier $(V_n)_n$ satisfying the following for all sufficiently large n:

- *Completeness:* There exists a quantum prover P (called an *honest prover*) such that for all n-qubit states $|\psi\rangle$, it holds that $\Pr(V_n(|\psi\rangle) \leftrightarrows P \text{ accepts}) = 1$.
- *Soundness:* For all quantum provers P and n-qubit states $|\psi\rangle$ such that $\Pr(V_n(|\psi\rangle) \leftrightarrows P \text{ accepts}) \geq \varepsilon(n)$, it holds that $\operatorname{td}(\sigma, U_n \psi U_n^\dagger) \leq \delta(n)$ where σ denotes the output of $V_n \leftrightarrows P$ conditioned on accepting.

Here the probabilities are over the randomness of the interaction.

Finally, define

$$\mathsf{unitaryQIP}(m) = \bigcap_{p,q} \mathsf{stateQIP}_{\frac{1}{p},\frac{1}{q}}(m), \qquad \mathsf{stateQIP} = \bigcup_{m'} \mathsf{stateQIP}(m')$$

where p, q, m' range over all polynomials.

In this section we present our unitary synthesis protocol for unitaries in unitaryPSPACE with polynomial action, i.e. Theorem 1.4.5, and for unitaries in unitaryPSPACE when the verifier also receives a succinct description of a polynomial-dimensional subspace that contains the input state, i.e. Corollary 1.4.6.

When we define an algorithm whose output is a real number, the output is implicitly to poly(n) bits of precision (relative to the implicit parameter n). We define $\mathbb{D}_m$ to be

[3] Bostanci et al. [24]'s definition differs from ours in some other ways as well.

the set of integer multiples of 2^{-m} in the interval $[0,1)$. The *uniformly mixed state* is the state $2^{-n}I_n$, and can be constructed by sampling a uniform random string x, since $2^{-n}\sum_{x\in\{0,1\}^n}|x\rangle\langle x| = 2^{-n}I$.

5.3.1 Protocol for unitaries with polynomial action

Our unitary synthesis protocol is based on the following algorithm, which we denote `LMR` after Lloyd, Mohseni, and Rebentrost who formulated it [72].

Theorem 5.3.3 ([72, 66])**.** *There exists a quantum polynomial-time algorithm `LMR` that takes as input a state $\tau \otimes \rho^{\otimes k} \otimes |t\rangle\langle t|$, where τ and ρ are n-qubit mixed states and $t \geq 0$, and outputs an n-qubit mixed state σ such that*

$$\mathrm{td}(\sigma, W\tau W^\dagger) \leq O(t^2/k) \qquad \textit{for} \qquad W = \exp(2\pi i \cdot t \cdot \rho).$$

We introduce the following notation related to Theorem 5.3.3. If U is a unitary acting on n qubits, then for $t, \varepsilon \geq 0$ we call an n-qubit mixed state ρ a *program state for U with evolution time t and error ε* if for all n-qubit states $|\phi\rangle$,

$$\mathrm{td}\Big(U\phi U^\dagger, W\phi W^\dagger\Big) \leq \varepsilon \qquad \text{for} \qquad W = \exp(2\pi i \cdot t \cdot \rho).$$

For example, if

$$U = \sum_{j=1}^{2^n} e^{2\pi i \theta_j}|v_j\rangle\langle v_j| \tag{5.3.1}$$

is an eigendecomposition of U where $0 \leq \theta_j < 1$ for all j, then for

$$t = \sum_j \theta_j, \qquad\qquad \rho = \frac{1}{t}\sum_j \theta_j |v_j\rangle\langle v_j| \tag{5.3.2}$$

the state ρ is a program state for U with evolution time t and error 0 (if $t > 0$). We call ρ and t respectively the *canonical program state for U* and the *canonical evolution time for U*, and note that ρ and t do not depend on the chosen eigendecomposition of U.

Observe that if U has action a (i.e. U acts nontrivially on a subspace of dimension a) then the canonical evolution time for U is at most a, because

$$t = \sum_j \theta_j \leq \sum_j \lceil \theta_j \rceil = a.$$

Therefore the following theorem implies that unitary sequences in $\mathsf{unitaryPSPACE}$ with polynomial action are in $\mathsf{unitaryQIP}(6)$:

Theorem 5.3.4 (formal, more general version of Theorem 1.4.5)**.** *Let $(U_n)_n \in \mathsf{unitaryPSPACE}$ be a sequence of unitaries such that U_n has canonical evolution time at most $\mathrm{poly}(n)$ for all n. Then $(U_n)_n$ is in $\mathsf{unitaryQIP}(6)$.*

In the rest of this subsection we prove Theorem 5.3.4. Our proof uses the following lemma:

Lemma 5.3.5. *Let $(U_n)_n \in \mathsf{unitaryPSPACE}$. Then there exists*

- *a* PSPACE*-computable function f such that $0 \leq f(1^n) \leq 2^n$ and $f(1^n) \leq t_n + e^{-\Omega(n)}$, where t_n is the canonical evolution time for U_n,*
- *a sequence $(\rho_n)_n \in$* statePSPACE*,*

such that for all n the state ρ_n is a program state for U_n with evolution time $f(1^n)$ and error $e^{-\Omega(n)}$.

Our unitary synthesis protocol prepares copies of ρ_n and computes $f(1^n)$ (for ρ_n, f as defined in Lemma 5.3.5) using our state synthesis protocol, and then applies the LMR algorithm. First we prove Lemma 5.3.5 in a certain special case (explained below), then we prove Lemma 5.3.5 in the general case by reducing to the special case, and finally we prove Theorem 5.3.4 using Lemma 5.3.5.

A "problem" with the definitions of ρ and t in (5.3.2) is that they are sensitive to small perturbations in the unitary U from (5.3.1). For example, if $\theta_j = 0$ for some j, then an arbitrarily small perturbation to the j'th eigenvalue of U could change θ_j to near 1. Furthermore ρ is undefined when $t = 0$, and is sensitive to a slight increase in any of the θ_j when t is near zero. Motivated by these concerns, we define a notion of *stability* of a unitary, and first prove Lemma 5.3.5 in the case where U_n is stable for all n.

Definition 5.3.6 (Stability of unitaries)**.** An n-qubit unitary U is *stable* if all of its eigenvalues are of the form $e^{2\pi i\theta}$ where $2^{-3n} \leq \theta \leq 1 - 2^{-3n}$.

The following equivalent definition of stability will sometimes be more convenient to work with. Define a metric Δ on the real numbers as follows:

$$\Delta(r, s) = \min_{k\in\mathbb{Z}} |r - s + k| = \min(r - s - \lfloor r - s\rfloor, \lceil r - s\rceil - (r - s)) .$$

Intuitively, this corresponds to mapping the real line to the unit circle by identifying all integer points with each other, and measuring the distance between two points on the resulting 1-dimensional torus. Thus, an n-qubt unitary U is stable if all of its eigenvalues are of the form $e^{2\pi i\theta}$ where $\Delta(\theta, 0) \geq 2^{-3n}$.

Remark. Our original proof [88] of statePSPACE ⊆ stateQIP applied only to sequences of *pure* states in statePSPACE, whereas in Section 5.2 we used the closure of statePSPACE under purification [76] (Lemma 5.1.5) to show that this also holds for *mixed* states. The following proof is a simplified version of our original proof, based on this observation.

Proof of Lemma 5.3.5 when U_n is stable

The proof is organized as follows. First we review the well-known *phase estimation algorithm*. Then, using the phase estimation algorithm, we define and analyze a sequence of quantum circuits $(C_n)_n$ which are used to define both $(\rho_n)_n$ and f. More specifically, we define ρ_n as the output state of C_n when C_n accepts, and $f(1^n)$ as 2^n times a PSPACE-computable approximation of the acceptance probability of C_n. Then we prove that $(\rho_n)_n$ is in statePSPACE. Finally we prove that f satisfies the required properties.

The phase estimation algorithm. For an n-qubit unitary U and number $m \in \mathbb{N}$, we denote by $\mathtt{PE}^{(U,m)}$ the instance of the phase estimation algorithm that acts on an n-qubit register A (the "eigenvector register") and an m-qubit register B (the "eigenvalue register") and makes oracle calls to U_{A} controlled on the content of B. If $|v\rangle$ is an eigenvector of U with eigenvalue $e^{2\pi i\theta}$, then

$$\mathtt{PE}^{(U,m)}|v\rangle_{\mathsf{A}}|0^m\rangle_{\mathsf{B}} = |v\rangle_{\mathsf{A}}|\eta\rangle_{\mathsf{B}}$$

for some state $|\eta\rangle$ (depending on $|v\rangle, U, m$), such that if $r \in \mathbb{D}_m$ denotes the outcome of a standard-basis measurement of $|\eta\rangle$ then

$$\Pr(\Delta(r,\theta) \geq \varepsilon) \leq O\big(2^{-m}/\varepsilon\big) \tag{5.3.3}$$

for all $\varepsilon > 0$ [80, Chapter 5]. Let $m(n) = 9n$ and $P_n = \mathtt{PE}^{(U_n,m(n))}$. Since $(U_n)_n$ is space-uniform, the sequence $(P_n)_n$ is also space-uniform, as can be seen by inspection of the phase estimation algorithm.

Procedure 6 The circuit C_n

```
1: Initialize an $n$-qubit register $\mathsf{A}$ to the uniformly mixed state $2^{-n}I$.
2: Initialize an $m(n)$-qubit register $\mathsf{C}$ to $\left|0^{m(n)}\right\rangle$, and apply the phase estimation circuit
   $P_n$ with eigenvector register $\mathsf{A}$ and eigenvalue register $\mathsf{C}$.
3: Create a one-qubit register $\mathsf{D}$, and controlled on the state $|r\rangle$ of $\mathsf{C}$ where $r \in \mathbb{D}_{m(n)}$,
   construct the state $\sqrt{r}|0\rangle + \sqrt{1-r}|1\rangle$ in $\mathsf{D}$.
4: Measure $\mathsf{D}$ in the standard basis. If the measurement outcome is 0 then accept and
   output $\mathsf{A}$, otherwise reject.
```

The circuit C_n and its properties. The circuit C_n is described in Procedure 6; clearly $(C_n)_n$ is space-uniform. For a fixed $n \in \mathbb{N}$, when analyzing C_n we use the following notation. Let

$$U_n = \sum_{j=1}^{2^n} e^{2\pi i\theta_j}|v_j\rangle\langle v_j|$$

be an eigendecomposition of U_n where $2^{-3n} \leq \theta_j < 1-2^{-3n}$ for all j. (The phases $e^{2\pi i\theta_j}$ and eigenvectors $|v_j\rangle$ depend on n, but for notational clarity we leave this dependence implicit.) Let $t = \sum_j \theta_j$ denote the canonical evolution time for U_n, and let $|\eta_j\rangle = \sum_{r\in\mathbb{D}_{m(n)}} \alpha_{jr}\,|r\rangle$ be the $m(n)$-qubit state such that

$$P_n|v_j\rangle\left|0^{m(n)}\right\rangle = |v_j\rangle|\eta_j\rangle\ .$$

The state after Line 2 is $2^{-n}\sum_j v_j \otimes \eta_j$, so the state after Line 3 is

$$2^{-n}\sum_j v_j \otimes \left(\sum_r \alpha_{jr}|r\rangle\big(\sqrt{r}|0\rangle + \sqrt{1-r}|1\rangle\big)\right)\left(\sum_r \alpha^*_{jr}\langle r|\big(\sqrt{r}\langle 0| + \sqrt{1-r}\langle 1|\big)\right).$$

Define

$$\widetilde{\theta}_j = \sum_{r\in\mathbb{D}_{m(n)}} |\alpha_{jr}|^2\, r, \qquad\qquad \widetilde{t} = \sum_j \widetilde{\theta}_j\ .$$

Then

$$\Pr(C_n \text{ accepts}) = 2^{-n} \cdot \widetilde{t}, \tag{5.3.4}$$

and when accepting C_n outputs the state

$$\rho = \frac{1}{\widetilde{t}} \sum_j \widetilde{\theta}_j \cdot v_j.$$

Now we argue that

$$\left|\widetilde{\theta}_j - \theta_j\right| \leq 2^{-4n} \tag{5.3.5}$$

for all j, provided n is sufficiently large. Let $\varepsilon = \frac{1}{2} \cdot 2^{-4n}$. If r denotes the outcome of a standard-basis measurement of $|\eta_j\rangle$, then

$$\begin{aligned}\left|\widetilde{\theta}_j - \theta_j\right| &= |\mathbb{E}[r] - \theta_j| = |\mathbb{E}[r - \theta_j]| \leq \mathbb{E}|r - \theta_j| \\ &= \mathbb{E}[|r - \theta_j| \cdot \mathbb{1}_{|r-\theta_j| \leq \varepsilon}] + \mathbb{E}[|r - \theta_j| \cdot \mathbb{1}_{|r-\theta_j| > \varepsilon}] \\ &\leq \varepsilon + \Pr(|r - \theta_j| > \varepsilon) = \varepsilon + \Pr(\Delta(r, \theta_j) > \varepsilon),\end{aligned}$$

where the last equality holds because U_n is stable and $\varepsilon < 2^{-3n}$. By (5.3.3) it follows that

$$\left|\widetilde{\theta}_j - \theta_j\right| \leq \varepsilon + O\Big(2^{-m(n)}/\varepsilon\Big) = \varepsilon + O\big(2^{-5n}\big) \leq 2\varepsilon,$$

which establishes (5.3.5).

Proof that $(\rho_n)_n$ is in statePSPACE. Let σ_n denote the output state of the general quantum circuit D_n described in Procedure 7, and let $\ell(n) = 2^{4n}$. Observe that $(D_n)_n$ is space-uniform. Clearly

$$\sigma_n = \Pr(C_n \text{ rejects})^{\ell(n)} \cdot |0 \ldots 0\rangle\langle 0 \ldots 0| + \Big(1 - \Pr(C_n \text{ rejects})^{\ell(n)}\Big) \cdot \rho_n,$$

so by (5.3.4),

$$\begin{aligned}\mathrm{td}(\sigma_n, \rho_n) &\leq \Pr(C_n \text{ rejects})^{\ell(n)}(0 \ldots 0, \rho_n) \leq \big(1 - 2^{-n}\widetilde{t}\big)^{\ell(n)} \leq \exp\big(-2^{-n} \cdot \widetilde{t} \cdot \ell(n)\big) \\ &= \exp\big(-2^{3n} \cdot \widetilde{t}\big).\end{aligned}$$

By (5.3.5) and the fact that U_n is stable,

$$\widetilde{t} = \sum_j \widetilde{\theta}_j = \sum_j \Big(\theta_j - \Big(\theta_j - \widetilde{\theta}_j\Big)\Big) \geq \sum_j (2^{-3n} - 2^{-4n}) \geq \Omega(2^{-2n}), \tag{5.3.6}$$

so $\mathrm{td}(\sigma_n, \psi_n) \leq \exp(-\Omega(2^n)) \leq \exp\big(-n^{\omega(1)}\big)$ as desired.

The function f and its properties. Let C'_n be identical to C_n except that C'_n only outputs an accept/reject bit (as opposed to also outputting the register A). Then $(C'_n|0 \ldots 0\rangle)_n$ is in $\mathsf{statePSPACE}_0$, so by Lemma 5.1.6 and (5.3.4) there exists a PSPACE-computable function g such that for all $n \in \mathbb{N}$ it holds that $0 \leq g(1^n) \leq 1$ and

$$\big|g(1^n) - 2^{-n}\widetilde{t}\big| \leq 2^{-4n}.$$

Procedure 7 The circuit D_n

```
for $2^{4n}$ times do
    Execute $C_n$.
    if $C_n$ accepts then return the output state of $C_n$ and abort. end if
end for
return $|0\dots0\rangle$.
```

Let $f(1^n) = 2^n g(1^n)$, and observe that f is computable in PSPACE (since g is) and that

$$\left|f(1^n) - \widetilde{t}\right| = 2^n\left|g(1^n) - 2^{-n}\widetilde{t}\right| \leq 2^{-3n} \ . \tag{5.3.7}$$

By (5.3.5), (5.3.7) and the triangle inequality,

$$|f(1^n) - t| \leq \left|f(1^n) - \widetilde{t}\right| + \left|\widetilde{t} - t\right| \leq 2^{-3n} + \sum_j \left|\widetilde{\theta}_j - \theta_j\right| \leq 2^{-3n} + 2^n \cdot 2^{-4n} = 2 \cdot 2^{-3n},$$

so $f(1^n) \leq t + e^{-\Omega(n)}$ as required.

Fixing n, all that remains is to prove that ρ is a program state for $U = U_n$ with evolution time $f = f(1^n)$ and error $e^{-\Omega(n)}$. In other words, given an arbitrary n-qubit state $|\phi\rangle$, we would like to prove that

$$\mathrm{td}\Big(U\phi U^\dagger, \widetilde{U}\phi\widetilde{U}^\dagger\Big) \leq e^{-\Omega(n)} \tag{5.3.8}$$

for

$$\widetilde{U} = \exp(2\pi i \cdot f \cdot \rho) = \sum_j \exp\Big(2\pi i \cdot f\widetilde{\theta}_j/\widetilde{t}\Big)|v_j\rangle\langle v_j| \ .$$

By (1.6.3),

$$\mathrm{td}\Big(U\phi U^\dagger, \widetilde{U}\phi\widetilde{U}^\dagger\Big) \leq \Big\|U|\phi\rangle - \widetilde{U}|\phi\rangle\Big\| \leq \Big\|U - \widetilde{U}\Big\|.$$

Since $|v_1\rangle, \dots, |v_{2^n}\rangle$ are orthogonal and the operator norm equals the largest singular value,

$$\begin{aligned}
\Big\|U - \widetilde{U}\Big\| &= \left\|\sum_j \Big(\exp(2\pi i \cdot \theta_j) - \exp\Big(2\pi i \cdot f\widetilde{\theta}_j/\widetilde{t}\Big)\Big)|v_j\rangle\langle v_j|\right\| \\
&= \max_j \Big|\exp(2\pi i \cdot \theta_j) - \exp\Big(2\pi i \cdot f\widetilde{\theta}_j/\widetilde{t}\Big)\Big| \\
&\leq 2\pi \max_j \Big|\theta_j - f\widetilde{\theta}_j/\widetilde{t}\Big|,
\end{aligned}$$

where the last inequality holds because $\left|e^{ia} - e^{ib}\right| = \sqrt{2 - 2\cos(a-b)} \leq |a - b|$ for all $a, b \in \mathbb{R}$. Finally, for all j,

$$\left|\theta_j - \frac{f\widetilde{\theta}_j}{\widetilde{t}}\right| = \left|\frac{\theta_j\big(\widetilde{t} - f\big) + f\Big(\theta_j - \widetilde{\theta}_j\Big)}{\widetilde{t}}\right| \leq \frac{\theta_j\left|\widetilde{t} - f\right| + f\left|\theta_j - \widetilde{\theta}_j\right|}{|\widetilde{t}|} \leq \frac{\left|\widetilde{t} - f\right| + 2^n\left|\theta_j - \widetilde{\theta}_j\right|}{|\widetilde{t}|},$$

and by (5.3.5), (5.3.6), (5.3.7) it holds that

$$\frac{\left|\widetilde{t}-f\right|+2^n\left|\theta_j-\widetilde{\theta}_j\right|}{\left|\widetilde{t}\right|}\leq O\left(\frac{2^{-3n}+2^n\cdot 2^{-4n}}{2^{-2n}}\right)\leq e^{-\Omega(n)}$$

from which (5.3.8) follows.

Proof of Lemma 5.3.5 in the general case

Again, given n let $U_n=\sum_{j=1}^{2^n}e^{2\pi i\theta_j}|v_j\rangle\langle v_j|$ be an eigendecomposition of U_n where $0\leq\theta_j<1$ for all j. We now consider the case where U_n may not be stable for all n. To remedy this, we reduce to the stable case via the following claim:

Claim 5.3.7. *There exists a* PSPACE*-computable function* ϕ *such that for all* $n\in\mathbb{N}$ *it holds that* $0\leq\phi(1^n)\leq O(2^{-2n})$ *and the unitary* $e^{2\pi i\phi(1^n)}U_n$ *is stable.*

First we prove Lemma 5.3.5 assuming Claim 5.3.7, and then we prove Claim 5.3.7.

Proof of Lemma 5.3.5 assuming Claim 5.3.7. Let $U_n'=e^{2\pi i\phi(1^n)}U_n$. The family $(U_n')_n$ is space-uniform, because $(U_n)_n$ is space-uniform and $\phi(1^n)$ is PSPACE-computable.[4] Since U_n' is furthermore stable for all n, by the special case of Lemma 5.3.5 proved above there exists

- a PSPACE-computable function f such that $0\leq f(1^n)\leq 2^n$ and $f(1^n)\leq t_n'+e^{-\Omega(n)}$, where t_n' is the canonical evolution time for U_n',
- a sequence $(|\rho_n\rangle)_n\in\mathsf{statePSPACE}$,

such that for all n the state ρ_n is a program state for U_n' with evolution time $f(1^n)$ and error $e^{-\Omega(n)}$. Since U_n' and U_n differ only by a global phase, the state ρ_n is also a program state for U_n with evolution time $f(1^n)$ and error $e^{-\Omega(n)}$. Finally, letting $t_n=\sum_j\theta_j$ denote the canonical evolution time for U_n, we have

$$t_n'=\sum_j(\theta_j+\phi(1^n)-\lceil\theta_j+\phi(1^n)\rceil)\leq\sum_j(\theta_j+\phi(1^n))=\sum_j\theta_j+2^n\cdot\phi(1^n)\leq t_n+2^n\cdot O\big(2^{-2n}\big),$$

so $f(1^n)\leq t_n'+e^{-\Omega(n)}\leq t_n+e^{-\Omega(n)}$ as desired. □

Now we prove Claim 5.3.7. The function ϕ is defined relative to the circuit E_n described in Procedure 8. Here, $(P_n)_n$ denotes the family of phase estimation circuits (like described in the proof of the stable case of Lemma 5.3.5), where the eigenvalue register has $m(n)$ qubits for a sufficiently large polynomial m to be specified later. Given an implicit parameter n, let $\delta=2^{-3n}$ and $\varepsilon=2\cdot 2^{-3n}$.

Let $2^{-n}\sum_j v_j\otimes\eta_j$ denote the state of the circuit after Line 2 (where $|v_j\rangle,|\eta_j\rangle$ are defined as in the analysis of Procedure 6). Let $s_j\in\mathbb{D}_{m(n)}$ denote the outcome of a standard-basis measurement of $|\eta_j\rangle$; then

$$\Pr(E_n(r)\text{ rejects})=\mathbb{E}_{j\sim[2^n]}\Pr(\Delta(s_j,-r)\leq\varepsilon)\ .$$

[4] Technically, the phase $\exp(-2\pi i\phi(1^n))$ may not be implementable exactly using our assumed gate set; however it can be approximated with exponentially small error that does not alter the analysis. For clarity we assume that the phase $\exp(-2\pi i\phi(1^n))$ can be implemented exactly.

Procedure 8 The circuit E_n

```
Input: $r \in \mathbb{D}_{m(n)}$
1: Initialize an $n$-qubit registers $\mathsf{A}$ to the uniformly mixed state $2^{-n}I$.
2: Initialize an $m(n)$-qubit register $\mathsf{C}$ to $|0^{m(n)}\rangle$, and apply the phase estimation circuit $P_n$ with eigenvector register $\mathsf{A}$ and eigenvalue register $\mathsf{C}$.
3: Measure $\mathsf{C}$ in the standard basis, and let $s \in \mathbb{D}_{m(n)}$ denote the measurement outcome.
4: If $\Delta(s, -r) > \varepsilon$ then accept, otherwise reject.
```

Since $(P_n)_n$ is space-uniform, the sequence $(E_n|0\dots0\rangle)_n$ is in $\mathsf{statePSPACE}_0$, so by Lemma 5.1.6 and the above equality there exists a PSPACE-computable function h such that for all $n \in \mathbb{N}, r \in \mathbb{D}_{m(n)}$ it holds that $0 \leq h(1^n, r) \leq 1$ and

$$\left|h(1^n, r) - \mathbb{E}_{j\sim[2^n]}\Pr(\Delta(s_j, -r) \leq \varepsilon)\right| < 2^{-2n}. \tag{5.3.9}$$

Define

$$\phi(1^n) = \min\left\{r \in \mathbb{D}_{m(n)} : h(1^n, r) < 2 \cdot 2^{-2n}\right\}. \tag{5.3.10}$$

(It is not immediately clear that $\phi(1^n)$ is well defined, i.e. that there exists $r \in \mathbb{D}_{m(n)}$ such that $h(1^n, r) < 2 \cdot 2^{-2n}$, but we will see that this is the case.)

First we prove that $\phi(1^n)$ is well defined, at most $9 \cdot 2^{-2n}$, and PSPACE-computable, and then we prove that $U_n' = e^{2\pi i \phi(1^n)} U_n$ is stable, thus establishing Claim 5.3.7.

Proof that $\phi(1^n)$ is well-defined, at most $9 \cdot 2^{-2n}$, and PSPACE-computable. We first show by a counting argument that there exists an $r \in \mathbb{D}_{m(n)}$ such that $r \leq 9 \cdot 2^{-2n}$ and $\Delta(\theta_j, -r) > 2\varepsilon$ for all j. This holds because on the one hand, we have

$$\left|\left\{r \in \mathbb{D}_{m(n)} \mid r \leq 9 \cdot 2^{-2n}\right\}\right| \geq 9 \cdot 2^{m(n)-2n} .$$

On the other hand, we have

$$\begin{aligned}\left|\left\{r \in \mathbb{D}_{m(n)} \mid \exists j : \Delta(\theta_j, -r) \leq 2\varepsilon\right\}\right| &\leq \sum_{j=1}^{2^n}\left|\left\{r \in \mathbb{D}_{m(n)} \mid \Delta(\theta_j, -r) \leq 2\varepsilon\right\}\right| \\ &\leq \sum_{j=1}^{2^n}\left(2 \cdot 2\varepsilon \cdot 2^{m(n)} + 1\right) = 8 \cdot 2^{m(n)-2n} + 2^n < 9 \cdot 2^{m(n)-2n},\end{aligned}$$

where in the last inequality we take m to be a sufficiently large polynomial. This implies the existence of such an r.

We now prove that $h(1^n, r) < 2 \cdot 2^{-2n}$, which implies that $\phi(1^n)$ is well defined and at most $9 \cdot 2^{-2n}$ as required. By (5.3.9) it holds that

$$h(1^n, r) < \mathbb{E}_{j\sim[2^n]}\Pr(\Delta(s_j, -r) \leq \varepsilon) + 2^{-2n},$$

so it suffices to prove that $\Pr(\Delta(s_j, -r) \leq \varepsilon) < 2^{-2n}$ for all j. By the definition of r and the triangle inequality for Δ, the event $\Delta(s_j, -r) \leq \varepsilon$ implies the event

$$2\varepsilon < \Delta(\theta_j, -r) \leq \Delta(\theta_j, s_j) + \Delta(s_j, -r) \leq \Delta(\theta_j, s_j) + \varepsilon,$$

i.e. $\Delta(\theta_j, s_j) > \varepsilon$. So by (5.3.3),

$$\Pr(\Delta(s_j, -r) \leq \varepsilon) \leq \Pr(\Delta(\theta_j, s_j) > \varepsilon) \leq O\Big(2^{-m(n)}/\varepsilon\Big) < 2^{-2n},$$

where the last inequality follows by taking m to be a sufficiently large polynomial.

Finally, since h is PSPACE-computable, ϕ is as well. □

Proof that U'_n is stable. Since U'_n has eigenvalues $\exp(2\pi i(\theta_j + \phi(1^n)))$, the condition that U'_n is stable is equivalent to the condition that $\Delta(\theta_j + \phi(1^n), 0) \geq \delta$ for all $j \in [2^n]$, which in turn is equivalent to $\Delta(\theta_j, -\phi(1^n)) \geq \delta$ for all j. Suppose for contradiction that there exists a j^* such that $\Delta(\theta_{j^*}, -\phi(1^n)) < \delta$. By the definitions of h and $\phi(1^n)$ (i.e. (5.3.9) and (5.3.10)), we have

$$\begin{aligned}\Pr(\Delta(s_{j^*}, -\phi(1^n)) \leq \varepsilon) &\leq 2^n\, \mathbb{E}_{j\sim[2^n]} \Pr(\Delta(s_j, -\phi(1^n)) \leq \varepsilon) \leq 2^n\left(2^{-2n} + h(1^n, \phi(1^n))\right) \\ &\leq 2^n\left(2^{-2n} + 2\cdot 2^{-2n}\right) = e^{-\Omega(n)}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\Pr(\Delta(s_{j^*}, -\phi(1^n)) > \varepsilon) &\leq \Pr(\Delta(s_{j^*}, \theta_{j^*}) + \Delta(\theta_{j^*}, -\phi(1^n)) > \varepsilon) \leq \Pr(\Delta(s_{j^*}, \theta_{j^*}) > \varepsilon - \delta) \\ &= \Pr(\Delta(s_{j^*}, \theta_{j^*}) > 2^{-3n}) \leq O\Big(2^{-m(n)+3n}\Big) \leq e^{-\Omega(n)},\end{aligned}$$

where the first inequality is by the triangle inequality for Δ, the second is by the definition of j_*, the third is by the definitions of δ and ε, the fourth is by (5.3.3), and the last is by taking m to be a sufficiently large polynomial. Therefore

$$1 = \Pr(\Delta(s_{j^*}, -\phi(1^n)) \leq \varepsilon) + \Pr(\Delta(s_{j^*}, -\phi(1^n)) > \varepsilon) \leq e^{-\Omega(n)},$$

which gives the desired contradiction. □

Proof of Theorem 5.3.4

By Lemma 5.3.5 and the fact that U_n has canonical evolution time at most $\mathrm{poly}(n)$, there exists

- a PSPACE-computable function f such that $0 \leq f(1^n) \leq \mathrm{poly}(n)$ for all n, and
- a sequence $(\rho_n)_n \in \mathsf{statePSPACE}$,

such that for all n the state ρ_n is a program state for U_n with evolution time $f(1^n)$ and error $e^{-\Omega(n)}$.

Let $\varepsilon, \delta = 1/\mathrm{poly}(n)$; we will prove that $(U_n)n \in \mathsf{unitaryQIP}_{\varepsilon,\delta}(6)$, establishing the theorem. Let k be a sufficiently large polynomial to be chosen later, and let

$$\varphi_n = \rho_n^{\otimes k(n)} \otimes |f(1^n)\rangle\langle f(1^n)|.$$

It is easy to see that $(\varphi_n)_n \in \mathsf{statePSPACE}$, so by Theorem 1.4.3 there exists a $\mathsf{stateQIP}_{\varepsilon,\delta/2}(6)$ verifier $(V_n)_n$ for $(\varphi_n)_n$. The following describes a $\mathsf{unitaryQIP}_{\varepsilon,\delta}(6)$ verifier for $(U_n)_n$, as applied to an n-qubit input state $|\theta\rangle$: Simulate V_n, if V_n rejects then reject, and if V_n accepts and outputs a state σ then accept and output $\texttt{LMR}(\theta, \sigma)$. This verifier runs in polynomial time because V and $\texttt{LMR}$ do.

Completeness holds because a prover that simulates an honest prover for V is accepted with probability 1. Now we prove soundness, assuming for simplicity that n is sufficiently large. Consider an arbitrary prover (which may depend on the verifier's n-qubit input state $|\theta\rangle$) such that the verifier accepts with probability at least ε. By the definition of V_n, this means that the output state σ of V_n satisfies $\mathrm{td}(\sigma, \varphi) \leq \delta/2$ for $\varphi = \varphi_n$. Write $f = f(1^n)$ and $\rho = \rho_n$, and let $W = \exp(2\pi i \cdot f \cdot \rho)$. By the triangle inequality,

$$\mathrm{td}\Big(\texttt{LMR}(\theta, \sigma), U\theta U^\dagger\Big) \leq \mathrm{td}(\texttt{LMR}(\theta, \sigma), \texttt{LMR}(\theta, \varphi))\, \mathrm{td}\Big(\texttt{LMR}(\theta, \varphi), W\theta W^\dagger\Big) + \mathrm{td}\Big(W\theta W^\dagger, U\theta U^\dagger\Big),$$

by (1.6.1)

$$\mathrm{td}(\texttt{LMR}(\theta, \sigma), \texttt{LMR}(\theta, \varphi)) \leq \mathrm{td}(\sigma, \varphi) \leq \delta/2,$$

by the definition of φ and Theorem 5.3.3

$$\mathrm{td}\Big(\texttt{LMR}(\theta, \varphi), W\theta W^\dagger\Big) = \mathrm{td}\Big(\texttt{LMR}(\theta, \rho^{\otimes k}, f), W\theta W^\dagger\Big) \leq O(f^2/k),$$

and since ρ is a program state for U with evolution time f and error $e^{-\Omega(n)}$

$$\mathrm{td}\Big(W\theta W^\dagger, U\theta U^\dagger\Big) \leq e^{-\Omega(n)}.$$

It follows that

$$\mathrm{td}\Big(\texttt{LMR}(\theta, \sigma), U\theta U^\dagger\Big) \leq \delta/2 + O(f^2/k) + e^{-\Omega(n)},$$

and since $f \leq \mathrm{poly}(n)$ there exists a polynomial k such that this bound is at most δ.

5.3.2 Protocol for general unitaries, but with restricted inputs

Let $U = (U_n)_n$ denote a sequence of quantum circuits in $\mathsf{unitaryPSPACE}$, not necessarily with polynomial action (i.e. the unitaries can act nontrivially on the entire Hilbert space). We argue that there is a $\mathsf{unitaryQIP}$ protocol for U provided that the verifier also receives as input a *succinct description* of a polynomial-dimensional subspace S that contains the input state. By succinct description, we mean that there is a polynomial-space Turing machine M that on input 1^n, outputs the description of a polynomial-space quantum circuit R that on input $|0\rangle \otimes |\phi\rangle$ outputs $|0\rangle \otimes (I - \Pi_S)|\phi\rangle + |1\rangle \otimes \Pi_S|\phi\rangle$ (for all n-qubit states $|\phi\rangle$) where Π_S is the projection onto S. In other words, the Turing machine M succinctly describes a circuit R that "recognizes" states from the subspace S.

The protocol for applying U essentially reduces to using the protocol for polynomial-action unitary families from Section 5.3.1. Consider an input $(1^n, |\phi\rangle, M)$, where $|\phi\rangle$ is an n-qubit state and M is a succinct description of a polynomial-space quantum circuit R that recognizes a polynomial-dimensional subspace S that contains $|\phi\rangle$. Let S' be the $(n+1)$-qubit subspace spanned by $|\varphi\rangle|0\rangle$ for $|\varphi\rangle$ in S, and by $|\varphi\rangle|1\rangle$ for $|\varphi\rangle$ in the subspace U_nS. Let V be the $(n+1)$-qubit unitary defined by $V|\varphi\rangle|0\rangle = U_n|\varphi\rangle \otimes |1\rangle$ and $V|\varphi\rangle|1\rangle = U_n^\dagger|\varphi\rangle \otimes |0\rangle$ (for all n-qubit states $|\varphi\rangle$) and observe that S' is closed under action by V. Furthermore, since R and U_n are polynomial-space quantum circuits, there clearly exists a polynomial-space quantum circuit R' that recognizes S'. The verifier can run the protocol to synthesize the unitary V (on an input in S') which acts as $U_n \otimes I$ on the subspace $S \otimes |0\rangle$. The key to this reduction is to show that the program state corresponding to the unitary V can be generated in quantum polynomial space. We sketch an argument below.

The same phase-estimation-based approach can be used, except before applying phase estimation to the n-qubit uniformly mixed state $2^{-n}I$, an ancilla $|0\rangle$ qubit is adjoined and the polynomial-space circuit R' is run on the first $n+1$ qubits of $|0\rangle \otimes 2^{-n}I$. Then, the first qubit is measured and the circuit post-selects on it being in the state $|1\rangle$. In other words, the uniformly mixed state $2^{-n}I$ is projected to be supported only on the subspace S', and the resulting entangled state is

$$\frac{1}{\dim S'} \sum_k |w_k\rangle\langle w_k|$$

where $\{|w_k\rangle\}$ is a basis for the subspace S' in which V is diagonal (i.e. they are eigenvectors of V). All other steps of Procedure 6, Procedure 7, and Procedure 8 are the same. Thus, the resulting program states represent the unitary V restricted to the subspace S', and thus are program states for V. These program states can be generated in polynomial space because the post-selection probability of projecting $2^{-n}I$ to $\frac{1}{\dim S'} \sum_k |w_k\rangle\langle w_k|$ is at least 2^{-n} (assuming that the subspace S is at least one-dimensional).

5.4 Multiple entangled provers

The class QMIP is defined analogously to QIP, but with multiple provers who may share arbitrarily many entangled qubits, and similarly for the m-message variant $\mathsf{QMIP}(m)$. The following theorem characterizes this class:

Theorem 1.4.2 ([62, 84])**.** *$\mathsf{QMIP} = \mathsf{MIP}^* = \mathsf{RE}$.*

Similarly, we define classes $\mathsf{stateQMIP}$ and $\mathsf{unitaryQMIP}$ analogously to $\mathsf{stateQIP}$ and $\mathsf{unitaryQIP}$ respectively, but with multiple provers who may share arbitrarily many entangled qubits. We also define the following analogues of the class R of computable languages, where by "computable sequence of quantum circuits" we mean a sequence in which the description of the n'th circuit can be computed as a function of n:

Definition 5.4.1 (stateR)**.** stateR is the class of sequences of mixed states $(\rho_n)_n$ such that each ρ_n is a state on n qubits, and for every polynomial q there exists a computable family of general quantum circuits $(C_n)_n$ such that for all sufficiently large n, the circuit C_n takes no inputs and C_n outputs a mixed state σ such that $\mathrm{td}(\rho_n, \sigma) \le 1/q(n)$.

Definition 5.4.2 ($\mathsf{unitaryR}$)**.** $\mathsf{unitaryR}$ is the class of all computable sequences $(U_n)_n$ of unitary quantum circuits such that each U_n acts on n qubits.

We prove a statement identical to Theorem 1.4.3 but with stateR and $\mathsf{stateQMIP}$ in place of $\mathsf{statePSPACE}$ and $\mathsf{stateQIP}$ respectively, and we also prove the converse statement:

Theorem 1.4.4. *$\mathsf{stateR} = \mathsf{stateQMIP}(6)$.*

Proof. First we prove that $\mathsf{stateR} \subseteq \mathsf{stateQMIP}$. That is, given a state family $(\rho_n)_n$ in stateR and a polynomial q, we prove that $(\rho_n)_n$ is in $\mathsf{stateQMIP}[1/q, 1/q]$. The oracle from Theorem 2.3.1 is computable, and hence admits a QMIP verifier by Theorem 1.4.2 and the fact that $\mathsf{R} \subseteq \mathsf{RE}$. The rest of the proof is essentially the same as that of Theorem 1.4.3.

Now we prove that $\mathsf{stateQMIP} \subseteq \mathsf{stateR}$. Let $(\rho_n)_n \in \mathsf{stateQMIP}$ and let p be a polynomial. By the definition of stateR, it suffices to prove that there exists a computable family of general quantum circuits $(C_n)_n$ such that for all sufficiently large n, the circuit C_n takes no inputs and C_n outputs a mixed state σ such that $\mathrm{td}(\rho_n, \sigma) \le 1/p(n)$. The circuit C_n does the following, where V is a $\mathsf{stateQMIP}_{1/2,1/p}$ verifier:

1. Brute force over a discretization of the set of all provers for V_n, until finding a prover P that V_n accepts with probability at least $1/2$.
2. Output the state σ produced by $V_n \leftrightarrows P$ conditioned on accepting.

Such a prover P can be found because there exists an "honest" prover that V_n accepts with probability 1, and there exists an arbitrarily good approximation of the honest prover in the discretization of the set of provers.[5] Then $\mathrm{td}(\rho_n, \sigma) \leq 1/p(n)$ by the soundness guarantee of V. □

We also prove a statement identical to Theorem 5.3.4 but with unitaryR and unitaryQMIP in place of unitaryPSPACE and unitaryQIP respectively:

Theorem 5.4.3. *Let* $(U_n)_n \in \mathsf{unitaryR}$ *be a sequence of unitaries such that* U_n *has canonical evolution time at most* $\mathrm{poly}(n)$ *for all* n*. Then* $(U_n)_n$ *is in* $\mathsf{unitaryQMIP}(6)$*.*

Proof. The proof is essentially the same as that of Theorem 5.3.4, except that the natural analogue of Lemma 5.3.5 (i.e. with "stateR" in place of "statePSPACE" and with "computable" in place of "PSPACE-computable") holds trivially, and one should apply Theorem 1.4.4 in place of Theorem 1.4.3. □

[5] For comparison, the fact that $\mathsf{QMIP} \subseteq \mathsf{RE}$ also follows by brute-forcing over a discretization of the set of all provers. But unlike the state synthesis algorithm described above, this $\mathsf{QMIP} \subseteq \mathsf{RE}$ algorithm is not guaranteed to terminate, which is why Theorem 1.4.4 is not directly analogous to Theorem 1.4.2.

Appendix A

Low-depth quantum circuit implementations of basic tasks

In Appendix A.1 we show that QNC^1 circuits can efficiently simulate $\mathsf{QAC}^0_\mathsf{f}$ circuits, and in Appendix A.2 we show that $\mathsf{QAC}^0_\mathsf{f}$ circuits can efficiently perform certain operations that we take for granted in QNC circuits.

A.1 QNC simulation of QAC_f circuits

Recall that n-qubit *restricted fanout* maps $\left|b, 0^{n-1}\right\rangle$ to $|b^n\rangle$ for $b \in \{0, 1\}$.

Lemma A.1.1 (Green et al. [49])**.** *The n-qubit restricted fanout transformation can be implemented by a size-$(n-1)$, depth-$\lceil \log n \rceil$ circuit consisting of CNOT gates with no ancillae.*

Proof. The idea is illustrated in Fig. 15. On input $\left|b, 0^{n-1}\right\rangle$ where $b \in \{0, 1\}$, for $1 \le k < \lceil \log n \rceil$ the k'th layer maps $\left|b^{2^{k-1}}, 0 \ldots 0\right\rangle$ to $\left|b^{2^k}, 0 \ldots 0\right\rangle$ using 2^{k-1} CNOT gates, and similarly the $\lceil \log n \rceil$'th layer makes $n - 2^{\lceil \log n \rceil - 1}$ additional copies of b. □

Lemma 1.3.7. *Every n-qubit, depth-d QAC_f circuit can be cleanly simulated by an $O(n)$-qubit, depth-$O(d \log n)$, size-$O(dn)$ QNC circuit.*

Proof. An n-qubit generalized Toffoli gate can be cleanly simulated by a size-$O(n)$, depth-$O(\log n)$ QNC circuit with $O(n)$ ancillae. This follows by simulating a log-depth DeMorgan formula for the AND function (i.e. the circuit whose graph is a balanced binary tree of 2-bit AND gates), with one ancilla qubit allocated to store the value of each gate in the DeMorgan formula, and uncomputing the garbage using Fig. 1. If C computes restricted

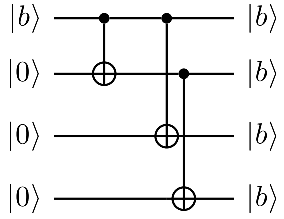

Figure 15: Four-qubit restricted fanout.

fanout then the circuit in Fig. 1 computes fanout [49], so by Lemma A.1.1 it follows that n-qubit fanout can also be cleanly implemented by a size-$O(n)$, depth-$O(\log n)$ QNC circuit with $O(n)$ ancillae.

A general n-qubit, depth-1 $\mathsf{QAC_f}$ circuit can be written as $\bigotimes_j G_j$, where each G_j is a k_j-qubit gate such that $\sum_j k_j \leq n$, and if $k_j > 1$ then G_j is either a generalized Toffoli or fanout gate. It follows that $\bigotimes_j G_j$ can be cleanly simulated by a QNC circuit where the size and number of ancillae are $O\Big(\sum_j k_j\Big) \leq O(n)$ and the depth is $O(\max_j \log k_j) \leq O\Big(\log \sum_j k_j\Big) \leq O(\log n)$. The lemma follows by successively implementing each layer of a $\mathsf{QAC_f}$ circuit in this way, reusing the same ancillae to simulate each layer. □

A.2 Properties of $\mathsf{QAC_f}$ circuits

Lemma A.2.1. *There is a uniform family of $O(mn\log n)$-qubit $\mathsf{QAC_f^0}$ circuits $(C_{n,m})_{n,m}$, where $C_{n,m}$ takes as input a $(\log n)$-qubit register K and m-qubit registers $\mathsf{A}_0, \ldots, \mathsf{A}_{n-1}, \mathsf{B}$, and $C_{n,m}$ cleanly swaps A_k and B controlled on the classical state $|k\rangle_{\mathsf{K}}$.*

Proof. We can assume without loss of generality that $m = 1$, because then the general case follows by swapping the i'th qubits of A_k and B for all i in parallel. By linearity we may assume that the input is a standard basis state

$$|k\rangle_{\mathsf{K}} |x_0\rangle_{\mathsf{A}_0} \cdots |x_{n-1}\rangle_{\mathsf{A}_{n-1}} |y\rangle_{\mathsf{B}}.$$

For now assume that y is promised to be 0^n. First compute $x_k = \bigvee_{j=0}^{n-1} (\mathbb{1}_{j=k} \wedge x_j)$ in B, using that $\mathsf{QAC_f^0}$ circuits can simulate $\mathsf{AC^0}$ circuits; note that comparing j and k requires $O(\log n)$ qubits for any given value of j. Then controlled on the state $|x_k\rangle_{\mathsf{B}}$, for all $j < n$ in parallel (using fanout) XOR the bit $\mathbb{1}_{j=k} \wedge x_k$ into A_j.

For the general case where y might not be 0^n, let C be an n-qubit register in the ancillae. First swap A_k and C as described above, then swap B and A_k as described above, and finally swap C and B. □

Lemma A.2.2. *If C is an n-qubit, size-s, depth-d $\mathsf{QAC_f}$ circuit then* ctrl-C *can be cleanly implemented by an $O(n)$-qubit, size-$O(s)$, depth-$O(d)$ $\mathsf{QAC_f}$ circuit.*

Proof. Controlled on a bit $b \in \{0, 1\}$, each gate in a $\mathsf{QAC_f}$ circuit can be implemented controlled on b as follows. A k-qubit generalized Toffoli gate controlled on b is equivalent to a $(k+1)$-qubit generalized Toffoli gate, fanning out a bit c controlled on b is equivalent to fanning out bc, and applying a one-qubit gate controlled on b can be done trivially. The result follows by making n copies of b, and using these copies to implement all gates in a given layer of C in parallel controlled on b, where the same ancillae are reused in simulations of successive layers of C. □

Appendix B

Proof of Lemma 2.2.2

Recall that in Section 2.2.1 we defined $\alpha = 0.35$ and

$$|p_{\eta,C}\rangle = C \cdot 2^{-n/2} \sum_{x\in\{0,1\}^n} \mathrm{sgnRe}(\langle\eta|C|x\rangle)|x\rangle$$

for a Clifford unitary C and vector $|\eta\rangle \in \left(\mathbb{C}^2\right)^{\otimes n}$. We establish the following fact:

Lemma 2.2.2 ([59])**.** *For all states $|\eta\rangle$ there exists a Clifford unitary C such that* $\mathrm{Re}(\langle\eta|p_{\eta,C}\rangle) \geq$ *α.*

Eq. (A.22) of Irani et al. [59]—where their $|\tau\rangle$ equals our $|\eta\rangle$, their γ can be set to 0.24999, and their d equals 2^n—implies that

$$\Pr\Big(\|\mathrm{Re}(C|\eta\rangle)\|_1 \geq \sqrt{0.24999 \cdot 2^n}\Big) > 0$$

for a random Clifford unitary C. Therefore there exists a fixed Clifford unitary C such that $2^{-n/2}\left\|\mathrm{Re}\left(C^\dagger|\eta\rangle\right)\right\|_1 \geq 0.4999$. Finally it follows from the definition of $|p_{\eta,C}\rangle$ that

$$\mathrm{Re}(\langle\eta|p_{\eta,C}\rangle) = 2^{-n/2}\sum_x |\mathrm{Re}(\langle\eta|C|x\rangle)| = 2^{-n/2}\left\|\mathrm{Re}(C^\dagger|\eta\rangle)\right\|_1,$$

implying that Lemma 2.2.2 holds with $\alpha = 0.4999$.

We instead define $\alpha = 0.35$ because we believe that there is a typo in Irani et al. [59], and that the right side of their Eq. (A.22) should be $1/2 - 4\gamma$ instead of $1/2 - 2\gamma$. So in the above analysis we should actually set γ to be slightly less than $1/8$, and so the value of α should be slightly less than $\sqrt{1/8} \approx 0.354$. The exact value of α is not important for our main results however.

Our disagreement with the argument in Irani et al. [59] is as follows. We will use their notation; in particular they assign a different meaning to the variable α than we have done. First—and this part is actually an understatement by Irani et al., not an error—in Eq. (A.13) the expression $\sqrt{(2^n+1)/(2\alpha)}$ can trivially be replaced by $\sqrt{(2^n+1)/(4\alpha)}$, and so Eq. (A.15) can be replaced by "$\geq 1 - 1/(2\alpha)$". Applying this strengthening of Eq. (A.15) with $\alpha = 1/(16\gamma)$ implies that $\Pr\left(\||\psi\rangle\|_1 \geq 2\sqrt{\gamma 2^n}\right) \geq 1 - 8\gamma$, where $|\psi\rangle$ is as defined in Lemma A.5 of Irani et al.

Write $|\psi\rangle = |a\rangle + i|b\rangle$ where $|a\rangle, |b\rangle \in \mathbb{R}^{2^n}$. Then

$$\Pr\Big(\||a\rangle\|_1 \geq \sqrt{\gamma 2^n}\Big) \geq \Pr\Big(\||a\rangle\|_1 \geq \sqrt{\gamma 2^n} \,\Big|\, \||\psi\rangle\|_1 \geq 2\sqrt{\gamma 2^n}\Big) \Pr\Big(\||\psi\rangle\|_1 \geq 2\sqrt{\gamma 2^n}\Big).$$

By the triangle inequality $\||\psi\rangle\|_1 \leq \||a\rangle\|_1 + \||b\rangle\|_1$, so conditioned on $\||\psi\rangle\|_1 \geq 2\sqrt{\gamma 2^n}$ either $\||a\rangle\|_1 \geq \sqrt{\gamma 2^n}$ or $\||b\rangle\|_1 \geq \sqrt{\gamma 2^n}$ (or both). Furthermore $\||a\rangle\|_1$ and $\||b\rangle\|_1$ are identically distributed conditioned on any value of $\||\psi\rangle\|_1$, because applying a global phase of i to $|\psi\rangle$ has the effect of swapping $\||a\rangle\|_1$ and $\||b\rangle\|_1$ without changing $\||\psi\rangle\|_1$. Therefore

$$\Pr\Big(\||a\rangle\|_1 \geq \sqrt{\gamma 2^n} \,\Big|\, \||\psi\rangle\|_1 \geq 2\sqrt{\gamma 2^n}\Big) \geq 1/2$$

and so $\Pr\big(\||a\rangle\|_1 \geq \sqrt{\gamma 2^n}\big) \geq \frac{1}{2}(1 - 8\gamma) = 1/2 - 4\gamma$.

Bibliography

[1] Scott Aaronson. "Open problems related to quantum query complexity". Sec. 6. 2021. URL: https://www.scottaaronson.com/papers/open.pdf (p. 5).

[2] Scott Aaronson. "Quantum copy-protection and quantum money". In: *CCC*. 2009, pp. 229–242. DOI: 10.1109/CCC.2009.42. arXiv: 1110.5353 (pp. 2, 5).

[3] Scott Aaronson. "The complexity of quantum states and transformations: from quantum money to black holes". 2016. arXiv: 1607.05256 (pp. 2, 4–6, 19).

[4] Scott Aaronson and Andrew Drucker. "A full characterization of quantum advice". In: *SIAM J. Comput.* 43.3 (2014), pp. 1131–1183. DOI: 10.1137/110856939. URL: https://doi.org/10.1137/110856939 (p. 5).

[5] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". In: *Phys. Rev. A* 70.5 (2004), p. 052328. DOI: PhysRevA.70.052328. arXiv: quant-ph/0406196 (p. 31).

[6] Scott Aaronson and Greg Kuperberg. "Quantum versus classical proofs and advice". In: *Theory Comput.* 3.7 (2007), pp. 129–157. DOI: 10.4086/toc.2007.v003a007. arXiv: quant-ph/0604056 (p. 4).

[7] Dorit Aharonov and Amnon Ta-Shma. "Adiabatic quantum state generation and statistical zero knowledge". In: *STOC*. 2003, pp. 20–29. DOI: 10.1145/780542.780546 (p. 2).

[8] Noga Alon, Martin Dietzfelbinger, Peter Bro Miltersen, Erez Petrank, and Gábor Tardos. "Linear hash functions". In: *Journal of the ACM (JACM)* 46.5 (1999), pp. 667–683. DOI: 10.1145/324133.324179 (p. 36).

[9] Noga Alon and Joel Spencer. *The Probabilistic Method*. 4th ed. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, 2016. Chap. The Probabilistic Lens: Turán's Theorem (p. 66).

[10] Andris Ambainis. "Quantum lower bounds by quantum arguments". In: *J. Comput. System Sci.* 64.4 (2002), pp. 750–767. DOI: 10.1006/jcss.2002.1826. arXiv: quant-ph/0002066 (pp. 8, 43).

[11] Anurag Anshu and Srinivasan Arunachalam. "A survey on the complexity of learning quantum states". In: (2023). arXiv: 2305.20069 (p. 25).

[12] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009 (pp. 9, 26).

[13] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. "Proof verification and the hardness of approximation problems". In: *J. ACM* 45.3 (1998), pp. 501–555. DOI: 10.1145/278298.278306 (p. 20).

[14] Yosi Atia and Dorit Aharonov. “Fast-forwarding of Hamiltonians and exponentially precise measurements”. In: *Nat Commun* 8.1 (2017), pp. 1–9. DOI: 10.1038/s41467-017-01637-7. arXiv: 1610.09619 (p. 23).

[15] László Babai, Lance Fortnow, and Carsten Lund. “Non-deterministic exponential time has two-prover interactive protocols”. In: *Computational complexity* 1 (1991), pp. 3–40. DOI: 10.1007/BF01200056 (p. 20).

[16] Zongbo Bao and Penghui Yao. “Nearly Optimal Algorithms for Testing and Learning Quantum Junta Channels”. In: (2023). arXiv: 2305.12097 (p. 25).

[17] Adriano Barenco, Charles H Bennett, Richard Cleve, David P DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A Smolin, and Harald Weinfurter. “Elementary gates for quantum computation”. In: *Physical review A* 52.5 (1995), p. 3457. arXiv: quant-ph/9503016 (p. 18).

[18] James Bartusek, Dakshita Khurana, and Akshayaram Srinivasan. “Secure Computation with Shared EPR Pairs (Or: How to Teleport in Zero-Knowledge)”. In: (2023). arXiv: 2304.10480 (p. 22).

[19] Paul Beame. *A switching lemma primer*. Tech. rep. University of Toronto, 1994. URL: http://www.cs.utoronto.ca/~toni/Courses/Complexity2015/handouts/primer.pdf (p. 14).

[20] Adam Bene Watts, Robin Kothari, Luke Schaeffer, and Avishay Tal. “Exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits”. In: *STOC*. 2019, pp. 515–526. DOI: 10.1145/3313276.3316404. arXiv: 1906.08890 [quant-ph] (p. 64).

[21] Debajyoti Bera. “A lower bound method for quantum circuits”. In: *Inform. Process. Lett.* 111.15 (2011), pp. 723–726. DOI: 10.1016/j.ipl.2011.05.002 (p. 15).

[22] Debajyoti Bera, Frederic Green, and Steven Homer. “Small depth quantum circuits”. In: *ACM SIGACT News* 38.2 (2007), pp. 35–50. DOI: 10.1145/1272729.1272739 (p. 13).

[23] Dominic W. Berry, Andrew M. Childs, and Robin Kothari. “Hamiltonian simulation with nearly optimal dependence on all parameters”. In: *FOCS*. 2015, pp. 792–809. DOI: 10.1109/FOCS.2015.54. arXiv: 1501.01715 (pp. 2, 7).

[24] John Bostanci, Yuval Efron, Tony Metger, Alexander Poremba, Luowen Qian, and Henry Yuen. “Unitary Complexity and the Uhlmann Transformation Problem”. In: (2023). arXiv: 2306.13073 (pp. 24, 25, 83).

[25] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013. DOI: 10.1093/acprof:oso/9780199535255.001.0001 (p. 55).

[26] Adam Bouland and Tudor Giurgică-Tiron. “Efficient Universal Quantum Compilation: An Inverse-free Solovay-Kitaev Algorithm”. In: (2021). arXiv: 2112.02040 (p. 10).

[27] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. “Quantum amplitude amplification and estimation”. In: *Quantum computation and information*. Vol. 305. Contemp. Math. Amer. Math. Soc., 2002, pp. 53–74. DOI: 10.1090/conm/305/05215. arXiv: quant-ph/0005055 (p. 42).

[28] Harry Buhrman, Richard Cleve, Michal Koucký, Bruno Loff, and Florian Speelman. "Computing with a full memory: catalytic space". In: *STOC*. 2014, pp. 857–866. DOI: 10.1145/2591796.2591874. URL: https://doi.org/10.1145/2591796.2591874 (p. 25).

[29] Stephen S Bullock, Dianne P O'Leary, and Gavin K Brennen. "Asymptotically optimal quantum circuits for d-level systems". In: *Physical review letters* 94.23 (2005), p. 230502. arXiv: 0410116 (p. 18).

[30] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J Coles. "Variational quantum algorithms". In: *Nat. Rev. Phys.* 3.9 (2021), pp. 625–644. DOI: 10.1038/s42254-021-00348-9. arXiv: 2012.09265 (p. 2).

[31] Thomas Chen, Shivam Nadimpalli, and Henry Yuen. "Testing and learning quantum juntas nearly optimally". In: *SODA*. 2023, pp. 1163–1185. DOI: 10.1137/1.9781611977554. arXiv: 2207.05898 (p. 25).

[32] Andrew M. Childs and Nathan Wiebe. "Hamiltonian simulation using linear combinations of unitary operations". In: *Quantum Inf. Comput.* 12.11-12 (2012), pp. 901–924. DOI: 10.5555/2481569.2481570. arXiv: 1202.5822 (pp. 2, 7).

[33] B David Clader, Alexander M Dalzell, Nikitas Stamatopoulos, Grant Salton, Mario Berta, and William J Zeng. "Quantum resources required to block-encode a matrix of classical data". In: *IEEE Transactions on Quantum Engineering* 3 (2022), pp. 1–23. arXiv: 2206.03505 (p. 16).

[34] Léo Colisson, Garazi Muguruza, and Florian Speelman. "Oblivious Transfer from Zero-Knowledge Proofs, or How to Achieve Round-Optimal Quantum Oblivious Transfer and Zero-Knowledge Proofs on Quantum States". In: (2023). arXiv: 2303.01476 (p. 22).

[35] Quantiki contributors. *The Church of the larger Hilbert space*. URL: https://quantiki.org/wiki/church-larger-hilbert-space (p. 2).

[36] Quantiki contributors. *The no-cloning theorem*. URL: https://www.quantiki.org/wiki/no-cloning-theorem (p. 3).

[37] G Mauro D'Ariano, Matteo GA Paris, and Massimiliano F Sacchi. "Quantum tomography". In: *Advances in imaging and electron physics* 128 (2003), pp. 206–309 (p. 2).

[38] Christopher M. Dawson and Michael A. Nielsen. "The Solovay–Kitaev algorithm". In: *Quantum Inf. Comput.* 6.1 (2006), pp. 81–95. arXiv: quant-ph/0505030 (p. 10).

[39] Hugo Delavenne, François Le Gall, Yupan Liu, and Masayuki Miyamoto. "Quantum Merlin-Arthur proof systems for synthesizing quantum states". In: (2023). arXiv: 2303.01877 (pp. 22, 24).

[40] D. Deutsch. "Quantum theory, the Church-Turing principle and the universal quantum computer". In: *Proc. Roy. Soc. London Ser. A* 400.1818 (1985), pp. 97–117 (p. 1).

[41] Mark Ettinger, Peter Høyer, and Emanuel Knill. "The quantum query complexity of the hidden subgroup problem is polynomial". In: *Inform. Process. Lett.* 91.1 (2004), pp. 43–48. DOI: 10.1016/j.ipl.2004.01.024. arXiv: quant-ph/0401083 (p. 5).

[42] Maosen Fang, Stephen Fenner, Frederic Green, Steven Homer, and Yong Zhang. "Quantum lower bounds for fanout". In: *Quantum Inf. Comput.* 6.1 (2006), pp. 46–57. arXiv: quant-ph/0312208 (pp. 12, 15, 45).

[43] Omar Fawzi, Nicolas Flammarion, Aurélien Garivier, and Aadil Oufkir. "Quantum Channel Certification with Incoherent Strategies". In: (2023). arXiv: 2303.01188 (p. 25).

[44] Richard P. Feynman. "Simulating physics with computers". In: vol. 21. 6-7. Physics of computation, Part II (Dedham, Mass., 1981). 1982, pp. 467–488. DOI: 10.1007/BF02650179. URL: https://doi.org/10.1007/BF02650179 (pp. 1, 2).

[45] Dmitry Gavinsky, Shachar Lovett, Michael Saks, and Srikanth Srinivasan. "A tail bound for read-k families of functions". In: *Random Structures Algorithms* 47.1 (2015), pp. 99–108. DOI: 10.1002/rsa.20532. arXiv: 1205.1478 [cs.DM] (pp. 50, 51).

[46] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum random access memory". In: *Phys. Rev. Lett.* 100.16 (2008), p. 160501. DOI: 10.1103/PhysRevLett.100.160501. arXiv: 0708.1879 (p. 8).

[47] 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". 2020. arXiv: 2007.04246 [quant-ph] (p. 13).

[48] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. "The knowledge complexity of interactive proof systems". In: *SIAM J. Comput.* 18.1 (1989), pp. 186–208. DOI: 10.1137/0218012 (p. 20).

[49] Frederic Green, Steven Homer, Cristopher Moore, and Christopher Pollett. "Counting, fanout, and the complexity of quantum ACC". In: *Quantum Inf. Comput.* 2.1 (2002), pp. 35–65. arXiv: quant-ph/0106017 (pp. 12–14, 31, 35, 45, 95, 96).

[50] Lov Grover and Terry Rudolph. "Creating superpositions that correspond to efficiently integrable probability distributions". In: (2002). arXiv: quant-ph/0208112 (p. 6).

[51] Kaiwen Gui, Alexander M Dalzell, Alessandro Achille, Martin Suchara, and Frederic T Chong. "Spacetime-Efficient Low-Depth Quantum State Preparation with Applications". In: (2023). arXiv: 2303.02131 (p. 17).

[52] 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". 2020. arXiv: 2007.00662 [quant-ph] (p. 13).

[53] Jeongwan Haah, Robin Kothari, Ryan O'Donnell, and Ewin Tang. "Query-optimal estimation of unitary channels in diamond distance". In: (2023). arXiv: 2302.14066 (p. 25).

[54] Daniel Harlow and Patrick Hayden. "Quantum computation vs. firewalls". In: *J. High Energy Phys.* 6 (2013), 085, front matter+55. DOI: 10.1007/JHEP06(2013)085. URL: https://doi.org/10.1007/JHEP06(2013)085 (pp. 2, 5).

[55] Aram W. Harrow, Benjamin Recht, and Isaac L. Chuang. "Efficient discrete approximations of quantum gates". In: vol. 43. 9. Quantum information theory. 2002, pp. 4445–4451. DOI: 10.1063/1.1495899. arXiv: quant-ph/0111031 (p. 10).

[56] Johan Håstad. “Almost optimal lower bounds for small depth circuits”. In: *STOC*. 1986, pp. 6–20. DOI: 10.1145/12130.12132 (pp. 13–15).

[57] Peter Høyer and Robert Špalek. “Quantum fan-out is powerful”. In: *Theory Comput.* 1.5 (2005), pp. 81–103. DOI: 10.4086/toc.2005.v001a005 (pp. 13, 19).

[58] Sándor Imre and Ferenc Balázs. *Quantum Computing and Communications: an engineering approach*. John Wiley & Sons, 2005. Chap. 7. DOI: 10.1002/9780470869048 (p. 41).

[59] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. “Quantum search-to-decision reductions and the state synthesis problem”. In: *CCC*. Vol. 234. 2022, 5:1–5:19. DOI: 10.4230/lipics.ccc.2022.5. arXiv: 2111.02999 (pp. 6, 7, 30, 97).

[60] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. “QIP = PSPACE”. In: *J. ACM* 58.6 (2011), pp. 1–27. DOI: 10.1145/2049697.2049704. arXiv: 0907.4737 (pp. 21, 78).

[61] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. “Pseudorandom quantum states”. In: *CRYPTO*. 2018, pp. 126–152. DOI: 10.1007/978-3-319-96878-0_5 (p. 2).

[62] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. “MIP* = RE”. 2020. arXiv: 2001.04383 (pp. 21, 93).

[63] Yifan Jia and Michael M Wolf. “Hay from the haystack: explicit examples of exponential quantum circuit complexity”. In: *Communications in Mathematical Physics* (2023), pp. 1–16. DOI: 10.1007/s00220-023-04720-x. arXiv: 2205.06977 (p. 11).

[64] Stasys Jukna. *Boolean function complexity*. Vol. 27. Algorithms and Combinatorics. Advances and frontiers. Springer, Heidelberg, 2012. DOI: 10.1007/978-3-642-24508-4 (pp. 9, 11, 13).

[65] Alastair Kay. *Tutorial on the Quantikz package*. 2020. DOI: 10.17637/rh.7000520. arXiv: 1809.03842 [quant-ph] (p. iii).

[66] Shelby Kimmel, Cedric Yen-Yu Lin, Guang Hao Low, Maris Ozols, and Theodore J. Yoder. “Hamiltonian simulation with optimal sample complexity”. In: *npj Quantum Inf* 3.13 (2017), pp. 1–7. DOI: 10.1038/s41534-017-0013-7. arXiv: 1608.00281 (pp. 23, 84).

[67] Emanuel Knill. “Approximation by quantum circuits”. In: (1995). arXiv: quant-ph/9508006 (pp. 16, 18).

[68] Dirk P Kroese, Tim Brereton, Thomas Taimre, and Zdravko I Botev. “Why the Monte Carlo method is so important today”. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 6.6 (2014), pp. 386–392 (p. 1).

[69] Ajay Kumar and Sunita Garhwal. “State-of-the-art survey of quantum cryptography”. In: *Arch. Comput. Methods Eng.* 28.5 (2021), pp. 3831–3868. DOI: 10.1007/s11831-021-09561-2. URL: https://doi.org/10.1007/s11831-021-09561-2 (p. 2).

[70] François Le Gall, Masayuki Miyamoto, and Harumichi Nishimura. “Distributed Merlin-Arthur Synthesis of Quantum States and Its Applications”. In: (2022). arXiv: 2210.01389 (p. 25).

[71] Daniel A. Lidar and Todd A. Brun. *Quantum Error Correction*. Cambridge University Press, 2013. URL: www.cambridge.org/9780521897877 (p. 2).

[72] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. "Quantum principal component analysis". In: *Nature Phys* 10.9 (2014), pp. 631–633. DOI: 10.1038/nphys3029. arXiv: 1307.0401 (pp. 23, 84).

[73] Carsten Lund, Lance Fortnow, Howard Karloff, and Noam Nisan. "Algebraic methods for interactive proof systems". In: *J. ACM* 39.4 (1992), pp. 859–868. DOI: 10.1145/146585.146605 (p. 20).

[74] Oleg Lupanov. "On a method of circuit synthesis". In: *Izvestia VUZ* 1 (1958), pp. 120–140. DOI: 10.2307/2271493 (p. 16).

[75] Fermi Ma. Personal communication. 2023 (p. 30).

[76] Tony Metger and Henry Yuen. "stateQIP= statePSPACE". 2023. arXiv: 2301.07730 (pp. 21, 74, 76, 79, 85).

[77] Abel Molina and John Watrous. "Revisiting the simulation of quantum Turing machines by quantum circuits". In: *Proceedings of the Royal Society A* 475.2226 (2019), p. 20180767. arXiv: 1808.01701 (p. 10).

[78] Rajeev Motwani and Prabhakar Raghavan. "Randomized algorithms". In: *ACM Computing Surveys (CSUR)* 28.1 (1996), pp. 33–37 (p. 1).

[79] Ashwin Nayak. "Inverting a permutation is as hard as unordered search". In: *Theory Comput.* 7 (2011), pp. 19–25. DOI: 10.4086/toc.2011.v007a002. arXiv: 1007.2899 (pp. 8, 43).

[80] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667 (pp. 8, 10, 17, 18, 26, 28, 39, 42, 75, 86).

[81] Daniel Padé, Stephen Fenner, Daniel Grier, and Thomas Thierauf. "Depth-2 QAC circuits cannot simulate quantum parity". 2020. arXiv: 2005.12169 [quant-ph] (p. 15).

[82] Natalie Parham. "On the Power and Limitations of Shallow Quantum Circuits". MA thesis. 2022. URL: http://hdl.handle.net/10012/18702 (p. 12).

[83] Alexander Razborov. "Lower bounds on the size of bounded depth circuits over a complete basis with logical addition". In: *Math. Notes* 41.4 (1987), pp. 333–338. DOI: 10.1007/BF01137685 (p. 15).

[84] Ben W. Reichardt, Falk Unger, and Umesh Vazirani. "A Classical Leash for a Quantum System: Command of Quantum Systems via Rigidity of CHSH Games". In: *ITCS*. 2013, pp. 321–322. DOI: 10.1145/2422436.2422473. arXiv: 1209.0448 (pp. 21, 93).

[85] Gregory Rosenthal. "Bounds on the QAC^0 Complexity of Approximating Parity". In: *ITCS*. Vol. 185. 2021, 32:1–32:20. DOI: 10.4230/LIPIcs.ITCS.2021.32. arXiv: 2008.07470 (p. 25).

[86] Gregory Rosenthal. "Efficient Quantum State Synthesis with One Query". 2023. arXiv: 2306.01723 (pp. 25, 74).

[87] Gregory Rosenthal. "Query and Depth Upper Bounds for Quantum Unitaries via Grover Search". 2021. arXiv: 2111.07992 (p. 25).

[88] Gregory Rosenthal and Henry Yuen. "Interactive Proofs for Synthesizing Quantum States and Unitaries". In: *ITCS*. Vol. 215. 2022, 112:1–112:4. DOI: 10.4230/LIPIcs.ITCS.2022.112. arXiv: 2108.07192 (pp. 25, 74, 76, 80, 85).

[89] Adi Shamir. "IP = PSPACE". In: *J. ACM* 39.4 (1992), pp. 869–877. DOI: 10.1145/146585.146609 (p. 20).

[90] Claude Shannon. "The synthesis of two-terminal switching circuits". In: *Bell System Tech. J.* 28 (1949), pp. 59–98. DOI: 10.1002/j.1538-7305.1949.tb03624.x (pp. 9, 16).

[91] Adrian She and Henry Yuen. "Unitary property testing lower bounds by polynomials". In: (2022). arXiv: 2210.05885 (p. 25).

[92] Peter W Shor. "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer". In: *SIAM review* 41.2 (1999), pp. 303–332 (p. 1).

[93] Roman Smolensky. "Algebraic methods in the theory of lower bounds for Boolean circuit complexity". In: *STOC*. 1987, pp. 77–82. DOI: 10.1145/28395.28404 (p. 15).

[94] Xiaoming Sun, Guojing Tian, Shuai Yang, Pei Yuan, and Shengyu Zhang. "Asymptotically optimal circuit depth for quantum state preparation and general unitary synthesis". In: *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* (2023). DOI: 10.1109/TCAD.2023.3244885. arXiv: 2108.06150 (pp. 16–18).

[95] Brian Swingle. "Unscrambling the physics of out-of-time-order correlators". In: *Nature Phys* 14.10 (2018), pp. 988–990. DOI: 10.1038/s41567-018-0295-5 (p. 2).

[96] Yasuhiro Takahashi and Seiichiro Tani. "Collapse of the hierarchy of constant-depth exact quantum circuits". In: *Comput. Complexity* 25.4 (2016), pp. 849–881. DOI: 10.1007/s00037-016-0140-0. arXiv: 1112.6063 (pp. 13, 19).

[97] A. M. Turing. "On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction". In: *Proc. London Math. Soc. (2)* 43.7 (1937), pp. 544–546. DOI: 10.1112/plms/s2-43.6.544. URL: https://doi.org/10.1112/plms/s2-43.6.544 (p. 1).

[98] Leslie G Valiant. "Exponential lower bounds for restricted monotone circuits". In: *Proceedings of the fifteenth annual ACM symposium on Theory of computing*. 1983, pp. 110–117. DOI: 10.1145/800061.808739 (p. 13).

[99] Thomas Vidick and John Watrous. "Quantum Proofs". In: *Found. Trends Theor. Comput. Sci.* 11.1-2 (2016), pp. 1–215. DOI: 10.1561/0400000068. arXiv: 1610.01664 (pp. 21, 78).

[100] John Watrous. "PSPACE has constant-round quantum interactive proof systems". In: *Theoret. Comput. Sci.* 292.3 (2003), pp. 575–588. DOI: 10.1016/S0304-3975(01)00375-9 (pp. 21, 78).

[101] John Watrous. "On the complexity of simulating space-bounded quantum computations". In: *Comput. Complexity* 12.1–2 (2003), pp. 48–84. DOI: 10.1007/s00037-003-0177-8 (p. 76).

[102] Nathan Wiebe. Personal communication. 2021 (pp. 39, 41).

[103] Wikipedia contributors. *n-sphere. Recurrences.* URL: https://en.wikipedia.org/wiki/N-sphere#Recurrences (p. 71).

[104] A Chi-Chih Yao. "Quantum circuit complexity". In: *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*. IEEE. 1993, pp. 352–361. DOI: 10.1109/SFCS.1993.366852 (p. 10).

[105] Pei Yuan and Shengyu Zhang. "Optimal (controlled) quantum state preparation and improved unitary synthesis by quantum circuits with any number of ancillary qubits". In: *Quantum* 7 (2023), p. 956. DOI: 10.22331/q-2023-03-20-956. arXiv: 2202.11302 (pp. 17, 18).

[106] Xiao-Ming Zhang, Tongyang Li, and Xiao Yuan. "Quantum state preparation with optimal circuit depth: Implementations and applications". In: *Physical Review Letters* 129.23 (2022), p. 230504. DOI: 10.1103/PhysRevLett.129.230504. arXiv: 2201.11495 (p. 16).