Lower bounds on the non-Clifford resources for quantum computations

Lower bounds on the non-Clifford resources for quantum computations

Michael Beverland[1], Earl Campbell[2], Mark Howard[3], and Vadym Kliuchnikov[1]

[1]QuArc, Microsoft Quantum, Redmond, Washington, US

[2]Department of Physics and Astronomy, University of Sheffield, Sheffield, UK

[3]School of Mathematics, Statistics & Applied Mathematics, NUI Galway, Ireland

November 26, 2019

Treating stabilizer operations as free, we establish lower bounds on the number of resource states, also known as magic states, needed to perform various quantum computing tasks. Our bounds apply to adaptive computations using measurements with an arbitrary number of stabilizer ancillas. We consider (1) resource state conversion, (2) single-qubit unitary synthesis, and (3) computational subroutines including the quantum adder and the multiply-controlled Z gate.

To prove our resource conversion bounds we introduce two new monotones, the stabilizer nullity and the dyadic monotone, and make use of the already-known stabilizer extent. We consider conversions that borrow resource states, known as catalyst states, and return them at the end of the algorithm. We show that catalysis is necessary for many conversions and introduce new catalytic conversions, some of which are optimal.

By finding a canonical form for post-selected stabilizer computations, we show that approximating a single-qubit unitary to within diamond-norm precision ε requires at least $1/7 \cdot \log_2(1/\varepsilon) - 4/3$ T-states on average. This is the first lower bound that applies to synthesis protocols using fall-back, mixing techniques, and where the number of ancillas used can depend on ε.

Up to multiplicative factors, we optimally lower bound the number of T or CCZ states needed to implement the ubiquitous modular adder and multiply-controlled-Z operations. When the probability of Pauli measurement outcomes is $1/2$, some of our bounds become tight to within a small additive constant.

Contents

1 Introduction and high-level overview **3**
- 1.1 Monotones under stabilizer operations . . . 3
- 1.2 Resource state conversion . . . 4
- 1.3 Computational tasks . . . 4
- 1.4 Unitary synthesis . . . 5
- 1.5 Measurement with probability 1/2 . . . 6

2 Some basic techniques **7**
- 2.1 Stabilizer nullity . . . 7
- 2.2 Stabilizer extent . . . 8
- 2.3 Catalysis . . . 9

3 Conversion between resource states **11**
- 3.1 Phase polynomial protocols . . . 12
- 3.2 One-bit adder conversion protocols . . . 13
- 3.3 Conversion bounds . . . 16

4 Computational task lower bounds **17**
- 4.1 Lower bounds for the C^nZ gate . . . 18
- 4.2 Lower bounds for the modular adder . . . 19

5 Lower bounds for approximate unitary synthesis **21**
- 5.1 Approximate unitary synthesis with and without post-selection . . . 22
- 5.2 Lower bounds with $|CS\rangle$ and $|CCZ\rangle$ resource states . . . 24
- 5.3 Lower bounds with $|T\rangle$ resource states . . . 25

6 Tighter lower bounds with measurement probabilities one half **27**
- 6.1 Lower bound with CCZ gates for C^nZ gate . . . 27
- 6.2 Lower bounds for the modular adder . . . 30
- 6.3 Lower bounds for resource state conversion . . . 31

7 Conclusion and open problems **32**

8 Acknowledgements **33**

A Appendices **37**
- A.1 Generic circuits for injecting diagonal gates . . . 37
- A.2 Reducing the cost of unitary synthesis using $\sqrt{T}$ gates . . . 38
- A.3 Explicit circuits for some common resource conversions . . . 40
- A.4 Extent values . . . 42
- A.5 Further details on phase polynomial protocols . . . 43
- A.6 Lower bound reduction from the modular adder to the multiply-controlled Z state 46
- A.7 Canonical form for post-selected stabilizer computations . . . 46
- A.8 Conversion protocols for dyadic rational powers of T gate . . . 51
- A.9 Overview of some definitions and results from Number Theory . . . 53
- A.10 Some properties of dyadic monotone μ_2 . . . 59
- A.11 General lower bounds for approximate unitary synthesis . . . 60

1 Introduction and high-level overview

Many promising architectures for universal fault-tolerant quantum computing [20, 34] perform computation by applying *stabilizer operations* to carefully prepared *resource states* known as magic states [6, 13, 38]. The stabilizer operations, which consist of Clifford gates, preparation of stabilizer ancilla states, and measurements in the Pauli basis, tend to be relatively easy to implement in these architectures. On the other hand, due to restrictions imposed by error correction [3, 7, 18, 40], the resource states tend to be produced by hefty distillation protocols [5, 12, 27, 33], that dominate the space-time overhead of the overall computation. It is therefore very natural and practically motivated to ask:

What is the minimum number of copies of a particular resource state that must be consumed to perform a given computational task using an arbitrary number of stabilizer operations?

We address this question by providing lower bounds for a number of computational tasks.

In the early days of quantum computing research, upper bounds for the resources required to implement compelling algorithms were crucial to motivate the development of scalable quantum computing hardware. Today, lower bounds are arguably more important since they can identify opportunities for further optimization. Unfortunately, lower bounds are infamously elusive: where an upper bound of resource requirements can be obtained by identifying an explicit algorithm, proving that no algorithm exists with certain properties can be very difficult.

Good lower bounds have however been forthcoming for some models of quantum computation, such as the two-qubit gate cost in the absence of measurement considering single-qubit gates as free [31, 37, 45–47]. For example, the multiply controlled phase operation needs at least a linear number of two-qubit gates [2]. These results are not tailored to the fault-tolerant setting where two-qubit CNOT gates and measurements are much cheaper than single-qubit non-Clifford gates, leading us to seek new theoretical tools.

We provide lower bounds for the production of particular target states, the implementation of important subroutines such as the adder and the multiply controlled phase operation, as well as approximating an arbitrary unitary to a desired precision. Our bounds significantly strengthen the best that were previously known and in some cases are the first non-trivial bounds that apply. We give separate bounds in terms of a variety of the most common basic resource states, and map out how these basic states themselves can be converted into one another.

1.1 Monotones under stabilizer operations

In Section 2, we introduce the *stabilizer nullity*, a function $\nu(|\psi\rangle)$ of any pure state $|\psi\rangle$ that is non-increasing under stabilizer operations. The stabilizer nullity is surprisingly powerful given its simplicity: it is the number of qubits that $|\psi\rangle$ is hosted in, minus the number of independent Pauli operators that stabilize $|\psi\rangle$. It is easy to see that $\nu = 0$ for any stabilizer state. We also leverage a previously known monotone called the *stabilizer extent* [4, 42] (see Section 2.2) which is also non-increasing under stabilizer operations. These monotones allow us to bound some state preparation tasks, for example n copies of the state $|\psi\rangle$ cannot be sufficient to produce a target state $|\text{tar.}\rangle$ if $\nu(|\psi\rangle^{\otimes n}) < \nu(|\text{tar.}\rangle)$. We write this as

$$\nu(|\psi\rangle^{\otimes n}) < \nu(|\text{tar.}\rangle) \quad \text{implies} \quad |\psi\rangle^{\otimes n} \not\longrightarrow |\text{tar.}\rangle.$$

An important factor in understanding the limitations of stabilizer operations is that their power can be increased not only by consuming resource states, but also by borrowing other resource states, known as *catalyst states* and returning them unchanged at the end of the algorithm, that is

$$|A\rangle \not\longrightarrow |B\rangle \text{ but } |A\rangle|\text{cat.}\rangle \longrightarrow |B\rangle|\text{cat.}\rangle, \text{ written as } |A\rangle \xrightarrow{|\text{cat.}\rangle} |B\rangle.$$

In Section 2.3, after establishing a general no-go theorem, we give several examples of resource state conversion where catalysis is necessary and sufficient. Similar results have been proven in the past for a more restricted set of scenarios e.g. [8, 39, 44].

1.2 Resource state conversion

Standard choices of which basic resource states are consumed by an algorithm could include the T-state $|T\rangle := T|+\rangle$, the CCZ-state $|CCZ\rangle := CCZ|+\rangle^{\otimes 3}$, the $\sqrt{T}$-state $|\sqrt{T}\rangle := \sqrt{T}|+\rangle$, along with many others. Before we turn to resource lower bounds for computational tasks such as implementing arbitrary unitaries, we should first consider how various basic resource states relate to one another, which is the focus of Section 3. Thankfully, the stabilizer nullity is *additive*, such that $\nu(|\psi\rangle|\phi\rangle) = \nu(|\psi\rangle) + \nu(|\phi\rangle)$ for all $|\psi\rangle$ and $|\phi\rangle$, which allows us to say even more. For example, we can rule out catalyzed conversions since $\nu(|A\rangle) < \nu(|B\rangle)$ implies that $\nu(|A\rangle|\text{cat.}\rangle) < \nu(|B\rangle|\text{cat.}\rangle)$ for any catalyzing state $|\text{cat.}\rangle$. Moreover, tensor powers of states simplify, allowing us to make asymptotic implications, i.e.,

$$\nu(|A\rangle) < r \cdot \nu(|B\rangle) \quad \text{implies} \quad |A\rangle^{\otimes n} \not\Longrightarrow |B\rangle^{\otimes \lceil rn \rceil} \quad \forall n.$$

Here, the double arrow indicates that *even with an arbitrary catalyst state* $\lceil rn \rceil$ copies of $|B\rangle$ cannot be produced from n copies of $|A\rangle$ using stabilizer operations.

These state conversion bounds and algorithms put our computational task lower bounds on more solid footing by allowing us to analyze the cost in terms of different input resource states. We also foresee our conversion results being useful in a much broader context, such as allowing a meaningful comparison of protocols that distill T-states with protocols that distill CCZ-states. For example, two T-states can be produced from a single CCZ-state, and therefore a distillation protocol A for $|CCZ\rangle$ outperforms a distillation protocol B for $|T\rangle$ if the protocol formed from converting the output of A into T-states outperforms B.

The bounds we obtain show there is a conversion gap: starting with n copies of $|T\rangle$, and applying the best possible $|T\rangle$ to $|CCZ\rangle$ conversion followed by the best possible $|CCZ\rangle$ to $|T\rangle$ conversion will yield fewer than n copies of $|T\rangle$. This gap survives in the asymptotic limit.

In Table 1 and Table 2 at the end of Section 3 we summarize many of our conversion bounds along with the most efficient known conversion algorithms (some of which were previously known, some of which we introduce in Section 3.1 and Section 3.2). Many of the conversion algorithms do not match the bounds suggesting that more efficient algorithms remain to be found.

1.3 Computational tasks

We have outlined how monotones can help bound the resources required to produce a particular target state. In Section 4 we bootstrap these techniques to lower bound the resources required to perform certain computational tasks.

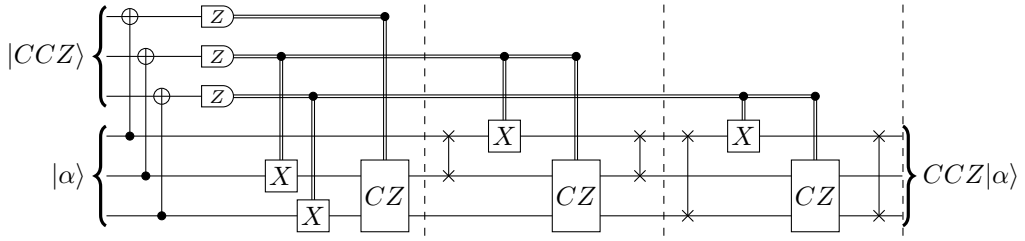


Figure 1: Implementation of the CCZ gate via injection of the $|CCZ\rangle$ state.

To exactly implement a unitary U, note that a lower bound of the number of copies of $|\psi\rangle$ needed to produce a state $U|S\rangle$, where $|S\rangle$ is a stabilizer state, also serves as a lower bound for applying U. We use this strategy in Section 4.1 to study the multiply controlled Z gate C^nZ, which is a key component of many important algorithms, including part of the reflection step in Grover's search [26]. By calculating the stabilizer nullity of the state $|C^{n-1}Z\rangle = C^{n-1}Z|+\rangle^{\otimes n}$ we straightforwardly show that for $n \geq 3$ it is not possible to apply the multiply controlled Z gate $C^{n-1}Z$ with stabilizer operations consuming fewer than n copies of $|T\rangle$, or $n/2$ copies of $|CS\rangle$, or $n/3$ copies of $|CCZ\rangle$. For comparison, the most efficient known algorithm [32] produces $C^{n-1}Z$ with $n-2$ copies of $|CCZ\rangle$.

It is also useful to consider catalysis as a proof technique when establishing lower bounds for computational tasks. For example, suppose U maps a state $|S\rangle|\Psi\rangle$ to a state $|\Phi\rangle|\Psi\rangle$ for stabilizer state $|S\rangle$ and non-stabilizer states $|\Phi\rangle$ and $|\Psi\rangle$. Then, a resource lower bound for catalytically producing the state $|\Phi\rangle$ must also serve as a lower bound for implementing U. In Section 4.2, we use this strategy to lower bound one of the most fundamental quantum arithmetic operations: the adder circuit, which acts on n-qubit basis states as $A(|i\rangle|j\rangle) = |i\rangle|i+j\rangle$ with $i+j$ evaluated modulo 2^n. The key is that the modular adder circuit acts on the input $|+\rangle^{\otimes n}|QFT_n\rangle$ to produce $|QFT_n^*\rangle|QFT_n\rangle$, where $|QFT_n\rangle$ is sometimes known as the quantum Fourier state, which becomes $|QFT_n^*\rangle$ under complex conjugation of coefficients in the computational basis. Crucially, we find that $\nu(|QFT_n\rangle) = \nu(|QFT_n^*\rangle) = n-2$ implying the adder circuit cannot be implemented with fewer than $n-2$ copies of $|T\rangle$, or $(n-2)/2$ copies of $|CS\rangle$, or $(n-2)/3$ copies of $|CCZ\rangle$. The most efficient known implementation of a modular adder uses $n-1$ copies of $|CCZ\rangle$ state [21].[1]

There are diagonal unitaries for which the cost of the resource state $U|+\rangle^{\otimes n}$ is the same as the cost of implementing the unitary itself. In particular, this is the case for all diagonal unitaries from the third level of the Clifford hierarchy, as can be seen from the state injection protocol described in Appendix A.1.

1.4 Unitary synthesis

In Section 5 we consider the number of resource states required to approximate an arbitrary single-qubit unitary to within diamond-norm precision ε using stabilizer operations. Although the importance of unitary synthesis has been long recognized, less is known regarding synthesis

[1] After the first posting of this paper, Craig Gidney [22] showed that the state $|C^nZ\rangle$ can be produced using the n-qubit modular adder. We reproduce his argument in Appendix A.6 for completeness. Using our (slightly stronger) bounds for $|C^{n-1}Z\rangle$ the adder circuit cannot be implemented with fewer than $n+1$ copies of $|T\rangle$, or $(n+1)/2$ copies of $|CS\rangle$, or $(n+1)/3$ copies of $|CCZ\rangle$.

strategies exploiting measurements, classical feed-forward and ancilla qubits. Crucially our lower bounds apply to synthesis algorithms in this general setting.

Loosely, our proof strategy is to select a target unitary U that can just be resolved from the identity given the required precision ε, and lower bound the resources required to produce the state $U|0\rangle$. As ε becomes small, $U|0\rangle$ approaches (but never quite reaches) the basis state $|0\rangle$. Unfortunately, the associated resource requirement divergence is not captured by either of the monotones we have discussed as they do not diverge for states approaching $|0\rangle$.

To achieve the required lower bound, we find a canonical form for stabilizer circuits applied to resource states (which may be of independent interest - see Theorem 5.3), and turn to a number theoretic approach. First, note that if the state $U|0\rangle$ is measured, if ε is small then the probability $|\langle 1|U|0\rangle|^2$ must be finite but very close to zero. Second, we use the canonical form to show that any single-qubit state produced using stabilizer circuits on a fixed number of resource states can only have a discrete set of allowed measurement probabilities, irrespective of the length of the stabilizer circuit and the number of stabilizer ancillas. This establishes a bound since producing a state $U|0\rangle$ with sufficiently small $|\langle 1|U|0\rangle|^2$ requires a sufficiently large number of resource states.

Our unitary synthesis results do not hold when a catalyst state is allowed, in contrast to those bounds proven with the stabilizer nullity due to its additive property. Another complication is that the number of resource states consumed by a protocol is actually a random variable, which can depend on the sequence of measurement outcomes obtained during the protocol. Our previous bounds held for every possible sequence of measurement outcomes, but for unitary synthesis we have to account for this subtlety. Let $\mathcal{N}_{|\Psi\rangle}(U,\varepsilon)$ be the number of copies of $|T\rangle$ consumed by a stabilizer circuit that approximates a unitary U to within diamond-norm precision ε. We show that there exist (diagonal) target unitaries U such that the expectation of $\mathcal{N}_{|\Psi\rangle}(U,\varepsilon)$ satisfies

$$\mathbf{E}\mathcal{N}_{|T\rangle}(U,\varepsilon) \geq \frac{1}{7}\log_2(1/\varepsilon) - \frac{4}{3}.$$

The best existing algorithm requires at most $2\log_2(1/\varepsilon)+O(\log(\log 1/\varepsilon))$ T-gates to implement any diagonal unitary, which comes from a combination of Refs [9, 28] with [43]. Our bound above is inferred from Theorem 5.2, which is a stronger but more nuanced result. We also have results that apply to stabilizer circuits with post selection in Section 5.1 and Section 5.2.

1.5 Measurement with probability 1/2

We can further tighten some lower bounds in a common restricted setting [21, 23, 25, 32] where arbitrary single-qubit Pauli measurements are not permitted, but only those measurements with outcomes which occur with probability one half. In Section 6, we introduce a quantity similar to the stabilizer nullity, which we call the *dyadic monotone* $\mu_2(|\psi\rangle)$ to prove a number of known subroutines exhibit nearly optimal resource consumption in this setting, narrowing the search for future algorithm improvements. The dyadic monotone requires that we restrict to states (including catalyst states) which can be written in the computational basis with coefficients that are integer combinations of $\exp(i\pi j/2^d)/2^k$ for integers d, j, k. This includes stabilizer states as well as resource states associated with all higher levels of the Clifford hierarchy [25]. Among other results we show that in this setting, the well-known circuit [32] which implements the multiply-controlled-Z gate $C^{n-1}Z$ using $n-2$ copies of $|CCZ\rangle$ is optimal.

In addition, we show that $n-2$ copies of $|CCZ\rangle$ are required to implement modular adder. The best known modular adder circuit [21] uses just one more $|CCZ\rangle$ gate. [2]

2 Some basic techniques

In this section we present some general techniques that are used throughout the paper, and defer our more specialized techniques to later sections and appendices. In Section 2.1 we introduce a number of properties of quantum states, including a simple but surprisingly powerful monotone under stabilizer operations which we call the stabilizer nullity. In Section 2.2 we review another monotone under stabilizer operations known as the stabilizer extent which was recently introduced [4, 42]. In Section 2.3 we show that the number of resources required to accomplish a computational goal can depend upon whether or not an additional catalyzing resource state is allowed which is returned unchanged at the end of an algorithm.

2.1 Stabilizer nullity

Let us first recall the definition of a stabilizer state and introduce a slight generalization of it.

Definition 2.1. *Let $|\psi\rangle$ be a non-zero n-qubit state.* The stabilizer *of $|\psi\rangle$, denoted* $\mathrm{Stab}|\psi\rangle$*, is the sub-group of the Pauli group $\mathcal{P}_n$ on n qubits for which $|\psi\rangle$ is a $+1$ eigenstate, that is* $\mathrm{Stab}\,|\psi\rangle = \{P \in \mathcal{P}_n : P|\psi\rangle = |\psi\rangle\}$*. The states for which the size of the stabilizer is 2^n are called* stabilizer states*. States for which the stabilizer contains only the identity matrix are said to have a* trivial stabilizer*. If Pauli P is in* $\mathrm{Stab}\,|\psi\rangle$*, we say that P stabilizes $|\psi\rangle$.*

Note that $\mathrm{Stab}(|\psi\rangle)$ can not contain $-I$. In addition, note that all Pauli group elements contained in $\mathrm{Stab}(|\psi\rangle)$ commute with each other and are Hermitian matrices. The size of the stabilizer of any state is equal to some power of two. For any Clifford unitary C, the size of $\mathrm{Stab}|\psi\rangle$ is always equal to the size of $\mathrm{Stab}(C|\psi\rangle)$. The size of the stabilizer is also multiplicative for the tensor product of states, that is $|\mathrm{Stab}(|\psi\rangle|\phi\rangle)| = |\mathrm{Stab}\,|\psi\rangle| \cdot |\mathrm{Stab}\,|\phi\rangle|$. A key quantity that we use throughout the paper is simply related to $\mathrm{Stab}(|\psi\rangle)$.

Definition 2.2 (Stabilizer nullity)**.** *Let $|\psi\rangle$ be a non-zero n-qubit state.* **The stabilizer nullity** *of $|\psi\rangle$ is $\nu(|\psi\rangle) = n - \log_2 |\,\mathrm{Stab}|\psi\rangle|$.*

Let us next see that the stabilizer nullity is non-increasing when multiple-qubit Pauli measurements are applied.

Proposition 2.3. *Let $|\psi\rangle$ be a non-zero n-qubit state and let P be an n-qubit Pauli matrix and suppose that the probability of a $+1$ outcome when measuring P on $|\psi\rangle$ is non-zero. Then there are two alternatives for the state $|\phi\rangle$ after the measurement: either $|\mathrm{Stab}|\phi\rangle| = |\mathrm{Stab}|\psi\rangle|$, or $|\mathrm{Stab}|\phi\rangle| \geq 2|\mathrm{Stab}|\psi\rangle|$, both of which satisfy $\nu(|\phi\rangle) \leq \nu(|\psi\rangle)$.*

Proof. First consider the simple case when P is in $\mathrm{Stab}|\psi\rangle$. In this case, the "+1" measurement outcome occurs with probability 1 and $|\psi\rangle$ is unchanged. When P is not in $\mathrm{Stab}|\psi\rangle$ we consider

[2] This bound becomes tight in light of Craig Gidney's blog post [22] which showed that the n-qubit modular adder can be used to produce the state $|C^nZ\rangle$, which requires at least $n-1$ copies of $|CCZ\rangle$ in this setting.

two alternatives. The first alternative is that P commutes with all elements of $\mathrm{Stab}|\psi\rangle$, then $\mathrm{Stab}|\phi\rangle$ contains $\mathrm{Stab}|\psi\rangle \cup P\,\mathrm{Stab}|\psi\rangle$ and its size is at least double that of $\mathrm{Stab}|\psi\rangle$. The second alternative is that P anti-commutes with some element Q from $\mathrm{Stab}|\psi\rangle$. In this case, we will see that the size of the stabilizer does not change as a result of the measurement. First note that in this case the probability of the $+1$ measurement outcome is $1/2$, because the probability of the $+1$ outcome is $\langle\psi|(I+P)|\psi\rangle/2$ and equal to $\langle\psi|Q(I+P)Q|\psi\rangle/2 = \langle\psi|(I-P)|\psi\rangle/2$ which is the probability of the -1 outcome, where we have used $Q|\psi\rangle = |\psi\rangle$ and $QPQ = -P$. Therefore $|\phi\rangle = (I+P)/\sqrt{2}|\psi\rangle$, where we have fixed the normalization such that $\langle\phi|\phi\rangle = \langle\psi|\psi\rangle$. Using that Q stabilizes $|\psi\rangle$ we also see that, $|\phi\rangle = (I+PQ)/\sqrt{2}|\psi\rangle$. Finally we observe that $(I+PQ)/\sqrt{2}$ is a Clifford unitary equal to $\exp(i\pi P'/4)$ for Hermitian Pauli matrix $P' = -iPQ$. As $|\phi\rangle$ and $|\psi\rangle$ differ by a Clifford, we conclude that $\mathrm{Stab}|\psi\rangle$ and $\mathrm{Stab}|\phi\rangle$ are the same size. □

One might wonder, if the second alternative in the proposition statement above should be $|\mathrm{Stab}|\phi\rangle| = 2|\mathrm{Stab}|\psi\rangle|$ instead of $|\mathrm{Stab}|\phi\rangle| \geq 2|\mathrm{Stab}|\psi\rangle|$. Here is an example that shows that the size of the stabilizer can more than double after one measurement. Consider an initial state $|\psi\rangle = |T\rangle|T\rangle$ states and measure the $+1$ outcome of $-Z\otimes Z$. The resource state $|T\rangle$ has a trivial stabilizer and so do its tensor powers, such that $|\,\mathrm{Stab}|\psi\rangle| = 1$. The result of the measurement is $|\phi\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ which is a two-qubit stabilizer state which therefore has $|\,\mathrm{Stab}|\phi\rangle| = 4$.

By Prop. 2.3 and the other aforementioned properties of $\mathrm{Stab}|\psi\rangle$, we see that the stabilizer nullity ν is invariant under Clifford unitaries, is non-increasing under Pauli measurements, and is additive under the tensor product. Moreover, as $\nu(|\psi\rangle) = 0$ when $|\psi\rangle$ is a stabilizer state, the stabilizer nullity is invariant under the inclusion or removal of stabilizer states. Another useful definition is the Pauli spectrum.

Definition 2.4 (Pauli spectrum)**.** *Let $|\psi\rangle$ be a non-zero n-qubit state.* The Pauli spectrum $\mathrm{Spec}|\psi\rangle$ *of $|\psi\rangle$ is:*

$$\mathrm{Spec}\,|\psi\rangle = \left\{ \frac{|\langle\psi|P|\psi\rangle|}{\langle\psi|\psi\rangle}, \quad \forall\ P \in \{I,X,Y,Z\}^{\otimes n} \right\}. \tag{1}$$

The Pauli spectrum is a list of 4^n real numbers each between 0 and 1 which is invariant under Clifford gates. Consider the following example

Proposition 2.5. *The Pauli spectrum of the state $|\theta\rangle = (|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$ is $\{\cos\theta, \sin\theta, 0\}$. The state $|\theta\rangle$ is therefore a stabilizer state only for $\theta = m\ \pi/2$ for some integer m.*

Proof. This follows from direct verification of the Pauli spectrum of $(|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$. □

We will make further use of the Pauli spectrum later in the paper, but for now note that the number of 1s in the Pauli spectrum of $|\psi\rangle$ is $|\,\mathrm{Stab}|\psi\rangle|$.

2.2 Stabilizer extent

In Definition 2.2 we introduced the stabilizer nullity ν which is additive and monotonic under stabilizer operations. Another monotone under stabilizer operations known as the *stabilizer extent* ξ was recently introduced [4, 42]. The stabilizer extent has a number of other very desirable properties shown in [42].

Definition 2.6 (stabilizer extent)**.** *For an arbitrary pure state $|\psi\rangle$,* **the stabilizer extent**, *denoted $\xi(|\psi\rangle)$, is*

$$\xi(|\psi\rangle) = \min ||(c_1, \ldots, c_k)||_1^2 \ s.t. \ |\psi\rangle = \sum_{\alpha=1}^{k} c_\alpha |\phi_\alpha\rangle. \tag{2}$$

where the minimization is over all complex linear combinations of stabilizer states $\{|\phi_\alpha\rangle\}$.

It is clearly submultiplicative but for many interesting cases it has been proven to be strictly multiplicative. More precisely,

Lemma 2.7. *The stabilizer extent is multiplicative with respect to a given set of states $\{|\psi_1\rangle, |\psi_2\rangle, \ldots |\psi_\ell\rangle\}$, such that, $\xi(|\psi_1\rangle|\psi_2\rangle \ldots |\psi_\ell\rangle) = \prod_{j=1}^{\ell} \xi(|\psi_j\rangle)$, if for each state $|\psi_j\rangle$ at least one of the following conditions is satisfied ($|\phi_j\rangle$ is always a stabilizer state):*

1. *$|\psi_j\rangle$ is a state of at most three qubits,*
2. *There exist states $|\omega_j\rangle$ and $|\phi_j\rangle$ such that $\xi(|\psi_j\rangle) = \frac{|\langle\psi_j|\omega_j\rangle|^2}{\max_{\phi_j} |\langle\omega_j|\phi_j\rangle|^2}$ **and** $|\langle\omega_j|\phi_j\rangle|^2 \geq 1/4$.*

This may actually hold more generally as we do not know of any counterexamples to multiplicativity of the stabilizer extent.

2.3 Catalysis

When considering the action of a sequence of stabilizer operations, it is important to consider scenarios in which another resource state is present which is returned at the end of the sequence. As the additional resource state is not consumed, we refer to it as a *catalyst*. By considering restrictions of the entries of density matrices, we prove here that a broad class of resource state conversions are impossible without catalysis, but can be achieved with catalysis. Similar results have been proven in the past for a more restricted set of scenarios [8, 39, 44].

It will be useful to recall some standard number fields:

$$\begin{aligned}
\mathbb{Q}(i) &= \{a_0 + ia_1 : a_0, a_1 \in \mathbb{Q}\}, \\
\mathbb{Q}(\zeta_8) &= \left\{a_0 + \zeta_8 a_1 + \ldots + \zeta_8^3 a_3 : a_k \in \mathbb{Q}\right\}, \zeta_8 = \exp(2\pi i/8), \\
\mathbb{Q}(\zeta_{16}) &= \left\{a_0 + \zeta_{16} a_1 + \ldots + \zeta_{16}^7 a_7 : a_k \in \mathbb{Q}\right\}, \zeta_{16} = \exp(2\pi i/16),
\end{aligned}$$

where $\mathbb{Q}$ is the set of rational numbers. As fields, each of these sets is closed under addition, multiplication, negation and taking the inverse. Note that $\mathbb{Q}(i)$ is a subset of both $\mathbb{Q}(\zeta_8)$ and $\mathbb{Q}(\zeta_{16})$. It is straightforward to verify that in the computational basis $|CS\rangle$ and $|CCZ\rangle$ states have density matrices with all entries in $\mathbb{Q}(i)$, but that the density matrix for the $|T\rangle$ state has some entries outside of $\mathbb{Q}(i)$ – instead all its entries are in $\mathbb{Q}(\zeta_8)$. More generally, the density matrix for $|T\rangle^{\otimes k} \otimes |0\rangle^{n-k}$ is given by:

$$\frac{1}{2^k} \sum_{a,b \in \{0,1\}^k} \zeta_8^{\text{weight}(a) - \text{weight}(b)} |a\rangle\langle b| \otimes (|0\rangle\langle 0|)^{\otimes n-k}.$$

where weight(a) is the Hamming weight of the bit string a. Similarly, we observe that the density matrix of $\left|\sqrt{T}\right\rangle^{\otimes k} \otimes |0\rangle^{n-k}$ is given by:

$$\frac{1}{2^k} \sum_{a,b \in \{0,1\}^k} \zeta_{16}^{\text{weight}(a)-\text{weight}(b)} |a\rangle\langle b| \otimes (|0\rangle\langle 0|)^{\otimes n-k}.$$

Consider the following theorem, which rules out a number of uncatalysed resource state conversions.

Theorem 2.8. *Let F be a number field which contains $\mathbb{Q}(i)$ and which is closed under complex conjugation. Any stabilizer circuit applied to a density matrix with all entries in F produces a density matrix with all entries in F, with both density matrices written in the computational basis.*

For example, no stabilizer circuit on any number of $|CS\rangle$ or $|CCZ\rangle$ states (which have density matrices with all entries in $\mathbb{Q}(i)$) can be used to produce a $|T\rangle$ state (which has a density matrix with all entries in $\mathbb{Q}(\zeta_8)$). Similarly, no stabilizer circuit on any number of $|T\rangle$ states can be used to produce a $|\sqrt{T}\rangle$ state (with entries in $\mathbb{Q}(\zeta_{16})$).

Proof. Suppose our stabilizer circuit acts upon n qubits initially in the $|0\rangle$ state. Clearly the density matrix $\rho_{\text{initial}} = (|0\rangle\langle 0|)^{\otimes n}$ has entries over $\mathbb{Q}$. We point out that all Clifford unitaries can be written as matrices with entries over $\mathbb{Q}(i)$, and therefore as matrices with entries over F. Explicitly, the Clifford group is generated by H, CZ and S

$$\begin{aligned} H &= \frac{1}{1+i}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \\ S &: |0\rangle \mapsto |0\rangle,\ |1\rangle \mapsto i|1\rangle, \\ CZ &: |ab\rangle \mapsto (-1)^{a \wedge b} |ab\rangle. \end{aligned}$$

Given that any gate U in the circuit is a tensor product of a unitary with entries over F and I and ρ has entries over F the product $U\rho U^\dagger$ is a density matrix with entries over F. Therefore applying the gates in the circuit preserves required property.

Note that measurement with or without post-selection can be described as:

$$\begin{aligned} \rho &\mapsto \frac{P\rho P}{\text{Tr}\rho P}, \\ \rho &\mapsto \sum_{P \in \mathcal{P}} P\rho P. \end{aligned}$$

The projectors P above correspond to measurement in the computational basis and therefore can be written as matrices with entries over $\mathbb{Q}(i)$ and therefore over F. The product of matrices over F is a matrix over F. The trace of a matrix over F is also in F by the definition of a field. The quotient of a matrix over F and an element of F is again a matrix over F because any field is closed under the division operation. This completes the proof. □

Importantly, the no-go results of Theorem 2.8 can be evaded by including a catalyst state. For example in Figure 2 we show how a $|CS\rangle$ state can be used to produce a $|T\rangle$ state by using an additional $|T\rangle$ state which is not consumed. Some examples of catalytic conversion have

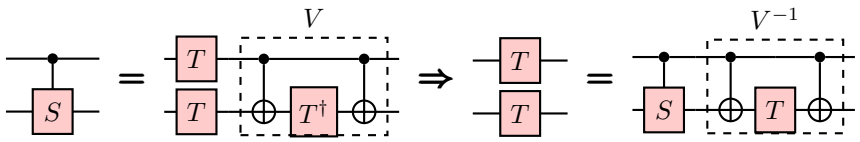


Figure 2: Upon applying the rightmost circuit to the $|+\rangle|+\rangle$ state, $|CS\rangle$ is converted into two copies of $|T\rangle$, using an additional T-gate. In terms of resource states $|CS\rangle|T\rangle \longrightarrow |T\rangle|T\rangle$, or equivalently $|CS\rangle \implies |T\rangle$. From Theorem 2.8, the conversion of $|CS\rangle$ into $|T\rangle$ would be impossible without catalysis. The leftmost gate identity is the standard implementation of a CS-gate from T-gates [2].

been noted before [8, 44], and this particular example is Clifford equivalent to that in [8]. In Section 3 we introduce two new catalytic conversion families.

Clearly we must distinguish scenarios in which catalysts are allowed from those in which they are not allowed. Throughout the remainder of the paper we use the following notation of single and double arrows.

Definition 2.9 (Conversion notation)**.** *The equation $|A\rangle \to |B\rangle$ indicates that resource state $|A\rangle$ can be converted into resource state $|B\rangle$ with stabilizer operations in the absence of a catalyst. On the other hand, $|A\rangle \overset{|C\rangle}{\Longrightarrow} |B\rangle$, which is equivalent to $|A\rangle|C\rangle \to |B\rangle|C\rangle$, indicates the conversion can proceed with the use of a catalyst $|C\rangle$ (which we sometimes omit above the arrow). When a process is impossible, we strike through the arrow, for example $|A\rangle \not\Longrightarrow |B\rangle$ signifies that $|A\rangle$ cannot be converted to $|B\rangle$ by stabilizer operations even in the presence of an arbitrary catalyst. In cases involving multiple copies of a given state such as $|A\rangle^{\otimes 2} \overset{|C\rangle}{\Longrightarrow} |B\rangle$, we sometimes write $2|A\rangle \overset{|C\rangle}{\Longrightarrow} |B\rangle$ to avoid clutter.*

3 Conversion between resource states

In this section we collect several results on the inter-conversion of resource states. These state conversion bounds and algorithms put the computational task lower bounds in later sections on more solid footing by allowing us to analyze the cost in terms of different input resource states. We also foresee our conversion results being useful in a much broader context, such as allowing a meaningful comparison of protocols that distill different types of resource states.

We saw in Section 2.3 that some resource conversions are impossible without access to a non-consumable resource often called a catalyst. Here we give two families of catalyzed conversion circuits, generalizing the previously known examples [8, 21, 44]. First, in Section 3.1 we introduce a general set of techniques for catalytic conversion of Clifford magic states. Second, in Section 3.2 we specify the use of adder circuits to perform catalysis, building on ideas of Gidney [21]. Finally, in Section 3.3, we utilize the monotones discussed in Section 2 to bound the optimal rates for conversion of resource states. One interesting observation is that although many pairs of resource states can be exactly converted into one another, it is impossible to do so without loss, even asymptotically. This complements the recent work [48] which applies to odd-prime qudits but not qubits.

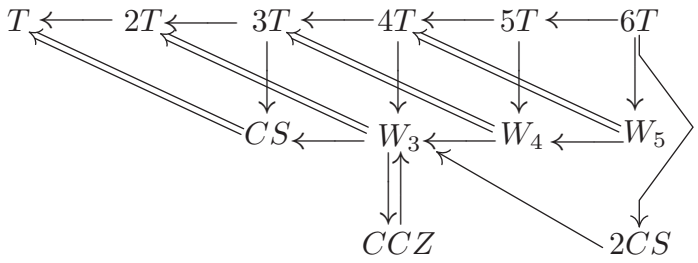


Figure 3: Phase polynomial conversion protocols for states of the form $|U\rangle := U|+\rangle^{\otimes n}$ where U is any diagonal unitary in the 3rd level of the Clifford hierarchy. The single arrow $\longrightarrow$ shows when a conversion can be directly realized, whereas a double arrow $\Longrightarrow$ indicates catalysis is used (and necessary). A subclass of n-qubit Clifford magic states are denoted $|W_n\rangle$ and arise from the diagonal unitary $W_n = \sum_x \exp(i\pi g(x)/4)|x\rangle\langle x|$, $g(x) = (\oplus_{i=1}^n x_i) + \sum_{i=1}^n x_i$, where $\oplus$ denotes addition modulo 2.

3.1 Phase polynomial protocols

Here we introduce a general set of techniques for catalytic conversion of Clifford magic states. Our main results are summarized in Figure 3. Recall that for any diagonal unitary U in the 3rd level of the Clifford hierarchy, the resource state $|U\rangle := U|+\rangle^{\otimes n}$ can be used to deterministically apply U and is known as a *Clifford magic state*. The unitary U can always be implemented using CNOT, S and T gates [1, 29]. The Clifford hierarchy is nested, so that the Clifford group (the 2nd level) is contained within the 3rd level. We have the following result

Theorem 3.1. *Let $|U\rangle = U|+\rangle^{\otimes n}$ be an n-qubit magic state for a diagonal unitary U from the 3rd level of the Clifford hierarchy, and let $\tau(U)$ be the minimum number of T gates needed to implement U using the gate set $\{CNOT, S, T\}$. The following resource conversion is possible*

$$|U\rangle \xrightarrow{|T\rangle^{\otimes \tau(U)-\nu(|U\rangle)}} |T\rangle^{\otimes 2\nu(|U\rangle)-\tau(U)}. \tag{3}$$

In the theorem, we follow the conversion notation of Definition 2.9 and use ν that was defined earlier as the stabilizer nullity (recall Definition 2.2). The proof of this theorem can be found in Appendix A.5.

An interesting family of ($n > 1$-qubit) unitaries that we call W_n are

$$W_n = \sum_x \exp(i\pi g(x)/4)|x\rangle\langle x|, \text{ with } g(x) = (\oplus_{i=1}^n x_i) + \sum_{i=1}^n x_i, \tag{4}$$

where the $\oplus$ sum is performed modulo 2. The corresponding Clifford magic state $|W_n\rangle = W_n|+\rangle^{\otimes n}$, when expressed as a density matrix $\rho = |W_n\rangle\langle W_n|$, has entries in $\mathbb{Q}[i]$ by virtue of the fact that $g(x) \equiv 0 \mod 2$ for all x. By Theorem 2.8 this implies that no T-states can be derived from $|W_n\rangle$ in the absence of a catalyst. In Figure 4 we give an explicit circuit for converting $|W_n\rangle$ to $|T\rangle^{\otimes n-1}$ using catalysis. This matches Theorem 3.1 by virtue of the following lemma, which is proved in Appendix A.5.

Lemma A.5. $\tau(W_n) = n + 1$.

We also have that the W_n states can be reduced in the sense that

$$|W_n\rangle \to |W_{n-1}\rangle. \tag{5}$$

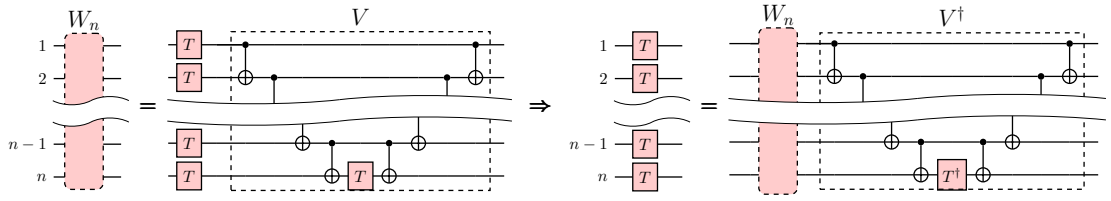


Figure 4: The first circuit identity can be verified explicitly, and follows from reasoning presented in [1]. By applying the rightmost circuit to the $|+\rangle^{\otimes n}$ state, $|W_n\rangle$ is converted into n copies of $|T\rangle$, using an additional $T^\dagger$-gate. In terms of resource states $|W_n\rangle|T\rangle \longrightarrow |T\rangle^{\otimes n}$, or equivalently $|W_n\rangle \implies |T\rangle^{\otimes n-1}$. The leftmost circuit identity depicts how the unitary W_n-gate can be implemented using a minimal (i.e., $\tau(W_n) = n+1$) number of T-gates.

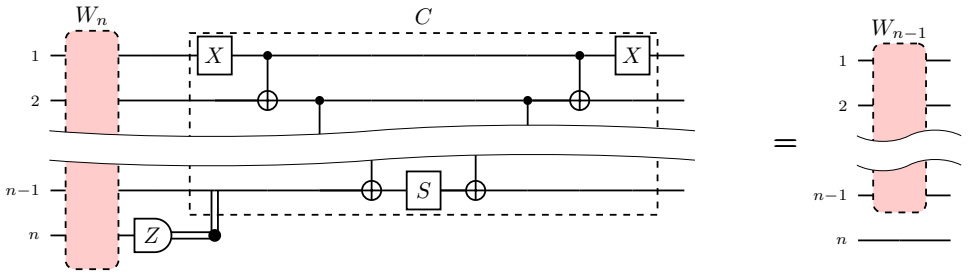


Figure 5: By applying the circuit on the left to the $|+\rangle^{\otimes n}$ state, $|W_n\rangle$ is converted into $|W_{n-1}\rangle$, where, depending on the Z measurement result, a Clifford correction C as in (6) may be required. This equality can be understood as follows. First note that the gate is symmetric with respect to permutations of qubits, so measuring the last qubit is equivalent to meausring the first. Since a pair of CNOT gates are applied in W_n controlled on the first qubit, if the outcome is zero, then those CNOTs can be removed, and W_{n-1} is applied directly. On the other hand, if the outcome is one, then instead of W_{n-1}, we have W_{n-1} sandwiched between a pair of X gates applied to the target of those CNOTs. This can be fixed by the content of the dashed box, which can be verified by taking the product of the gate which is applied with W_{n-1}, propagating the Xs through the circuit and making use of $XTX^\dagger = T^\dagger$ before cancelling adjacent CNOTs.

This conversion is achieved by first measuring the last qubit in the computational basis as shown in Figure 5. If one obtains the "0" outcome, we immediately have the state $|W_{n-1}\rangle|0\rangle$ and so just discard the last qubit. In the case of a "1" outcome, then a Clifford correction

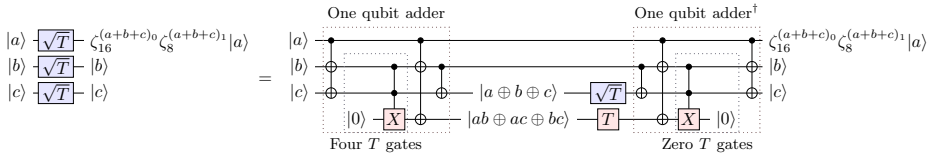
$$C = \sum_x i^{1 \oplus_{j=1}^n x_j} |x\rangle\langle x|, \tag{6}$$

is required. Performing this correction and discarding the last qubit we again obtain $|W_{n-1}\rangle$.

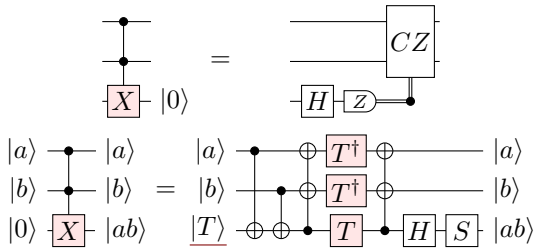
3.2 One-bit adder conversion protocols

In this subsection, we present a class of protocols that use catalysis to convert resource states for the third level of the Clifford hierarchy (i.e. Clifford magic states) to resource states for the higher levels of the Clifford hierarchy. It is beneficial to apply some of the protocols directly at the gate level too. The main building block for the protocols in this subsection is the circuit shown on Figure 6a, which is a special case of an idea described on Page 4 in [21]. This circuit implements three $\sqrt{T}$ gates (which are in the 4^{th} level of the Clifford heirarchy) using one $\sqrt{T}$

gate along with a few gates from the third level of the Clifford hierarchy. The key difference in our approach from that in Ref. [21] is that to scale this small example to parallel rotations on an n-qubit register, we use recursion, whereas in Ref. [21] a Hamming weight generalization is used. Compared with the Hamming weight construction, our recursive construction amortizes the cost of the correction operations associated with injecting gates from higher levels of the Clifford hierarchy. Later, in Section 6.3, we show that our construction is asymptotically optimal under the assumption that only measurements with probability one-half are used.



(a) Three $\sqrt{T}$ gates can be applied using a circuit with just one $\sqrt{T}$ gate and other gates in the third level of the Clifford hierarchy. This uses the Hamming weight register idea from [21] along with the adder from Figure 4 in [21].



(b) Simplified circuits for the Toffolli (i.e. the doubly-controlled-X gate) when the target qubit ends in the $|0\rangle$ state, and when the target qubit starts in the $|0\rangle$ state. See Figure 3 in [21]. Using the first of these circuits as a subroutine, Figure 6a implements three $\sqrt{T}$ gates using one $\sqrt{T}$ gate, one T gate and one CCX gate. Additionally making use of the second of these circuits, Figure 6a implements three $\sqrt{T}$ gates using one $\sqrt{T}$ gate and five T gates.

Figure 6: Circuits for applying three $\sqrt{T}$ gates using five T gates and one $\sqrt{T}$ gate.

To understand the circuit in Figure 6a, first note that the gate $\exp\left(i\theta|1\rangle\langle 1|\right)^{\otimes n}$ acting on an n-qubit register in the computational basis state $|w\rangle$ gives $e^{i\theta\cdot \mathrm{hw}(w)}|w\rangle$, where $\mathrm{hw}(w)$ is the Hamming weight of the bit string w. Therefore an alternative way of applying the gate $\exp\left(i\theta|1\rangle\langle 1|\right)^{\otimes n}$ is to compute the binary representation of $\mathrm{hw}(w)$ and store it in a quantum register $|x_k \ldots x_0\rangle$, and for j from 0 to k apply $\exp\left(i2^j\theta|1\rangle\langle 1|\right)$ to qubit j in the register. In Figure 6a we use the adder circuit shown in Figure 4 in [21] to compute the Hamming weight of the bit string a, b, c. For bit strings of length three the Hamming weight can be represented using two bits. The lower bit is the parity $a \oplus b \oplus c$ and the higher bit is the majority function $ab \oplus bc \oplus ac$. These are exactly the values computed by the adder. An important efficiency gain comes from the observation illustrated in the first circuit in Figure 6b that the one qubit adder can be un-computed by using Clifford gates and single qubit Pauli measurements only [21]. With this trick, the circuit shown in Figure 6a applies three $\sqrt{T}$ gates using only one $\sqrt{T}$ gate, and either one T gate and one CCX gate, or five T gates if the second circuit in Figure 6b is used.

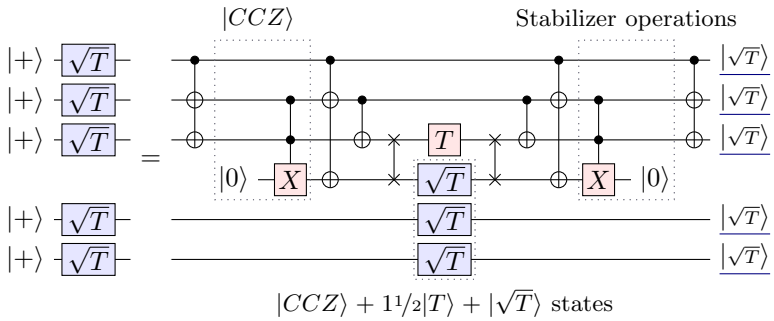


Figure 7: Conversion $k|CCZ\rangle + (k + 1/2)|T\rangle \xrightarrow{|\sqrt{T}\rangle} 2k|\sqrt{T}\rangle$ for $k = 2$.

Figure 6a can be adapted to form resource state conversion protocols. For example the protocol $|\sqrt{T}\rangle + 5.5|T\rangle \to 3|\sqrt{T}\rangle$ follows directly when $|+\rangle$ states are fed into Figure 6a when the $\sqrt{T}$ gate is implemented by injection of $\left|\sqrt{T}\right\rangle$, and when the third Clifford level gates are implemented with $|T\rangle$ resource states. We need to use 5 $|T\rangle$ states to implement the first CCX gate and the T gate in Figure 6a. The $\sqrt{T}$ gate is implemented with the injection circuit which requires an additional T gate correction 50% of the time, which adds 1/2 to the number of $|T\rangle$ states consumed (on average).

The extra T gate can be amortized to give the conversion

$$k|CCZ\rangle + (k + 1/2)|T\rangle \xrightarrow{|\sqrt{T}\rangle} 2k|\sqrt{T}\rangle$$

valid for any positive integer k. We see that asymptotically $|\sqrt{T}\rangle$ state costs half of $|T\rangle$ state plus half of $|CCZ\rangle$ state. Using the circuit on Figure 6a we can reduce the parallel application of $2k + 1$ $\sqrt{T}$ gates to the parallel application of $2k - 1$ $\sqrt{T}$ gates for any positive integer k. We use the circuit on Figure 6a to execute first three out of $2k+1$ $\sqrt{T}$ gates by only using one $\sqrt{T}$ gate. Then we observe that the rest of the $2k - 2$ $\sqrt{T}$ gates can be executed in parallel with the newly introduced $\sqrt{T}$ gate. Figure 7 shows how to reduce the parallel application of five $\sqrt{T}$ gates to the parallel application of three $\sqrt{T}$ gates. We also note that all above results can be applied to conversion of $|\sqrt{T}^3\rangle$ states and application of $\sqrt{T}^3$ gates. The cost of applying $\sqrt{T}^3$ gate is the same as the cost of applying $\sqrt{T}$. This is an important observation for single qubit circuit synthesis applications.

A similar idea leads to the lower cost of applying many $T^{j/2^{d-2}} = \exp(\pi ij/2^d|1\rangle\langle 1|)$ for positive integer $d \geq 2$ and an odd j proved in Appendix A.8.

Theorem A.16. *Let $k, d \geq 1$ be positive integers and let j be an odd integer and let $a_{d,k} = 2^{d-1}(k-1)+2$. Then $a_{d,k}$ gates $\exp(\pi ij/2^d|1\rangle\langle 1|)$ can be executed in parallel by using stabilizer operations with measurements that have probability 50%, $b_{d,k} = (2^{d-1}-1)(k-1)+d-1$ copies of $|CCZ\rangle$ state and using one copy of each of the states $|\pi j/2^d\rangle, |\pi j/2^{d-1}\rangle, \ldots, |\pi j/2^2\rangle$ as a catalyst. Asymptotically, the state $|\pi j/2^d\rangle$ is produced using $1 - 1/2^{d-1}$ $|CCZ\rangle$ states.*

Note that the number of $|CCZ\rangle$ states used by protocols described in the theorem above is asymptotically the same as the lower bounds established later in Section 6:

Lemma 6.9. *Consider a protocol that uses stabilizer operations with measurements probability* 50%*,* $|CCZ\rangle$ *states and a multi-qubit state as a catalyst. The catalyst has entries in* $\mathcal{R}_{d'}$ *for some* d'*. Suppose that such circuit uses* k $|CCZ\rangle$ *states and produces* n *states* $|\pi j/2^d\rangle$ *for odd* j *and integer* $d \geq 2$*, then* $k \geq n(1 - 1/2^{d-1})$*. Asymptotically, at least* $1 - 1/2^{d-1}$ *copies of* $|CCZ\rangle$ *state are needed to produce state* $|\pi j/2^d\rangle$*.*

These protocols are useful for reducing the cost of approximate unitary synthesis, as described in Appendix A.2.

3.3 Conversion bounds

Suppose we can identify a monotone $\mathcal{M}$ such that $\mathcal{M}(|\psi\rangle)$ is real for any state $|\psi\rangle$, and is non-increasing under stabilizer operations. We say such a function $\mathcal{M}$ is a *monotone*, and can use it to bound conversion processes since for example, a resource state $|A\rangle$ cannot be used to produce a resource state $|B\rangle$ with stabilizer operations if $\mathcal{M}(|A\rangle) < \mathcal{M}(|B\rangle)$, i.e.,

$$\mathcal{M}(|A\rangle) < \mathcal{M}(|B\rangle) \quad \text{implies} \quad |A\rangle \not\longrightarrow |B\rangle.$$

If the monotone is also *additive*, such that $\mathcal{M}(|\psi\rangle \otimes |\phi\rangle) = \mathcal{M}(|\psi\rangle) + \mathcal{M}(|\phi\rangle)$ for all $|\psi\rangle$ and $|\phi\rangle$, then we can say even more. For example we can rule out catalyzed conversions since $\mathcal{M}(|A\rangle) < \mathcal{M}(|B\rangle)$ implies that $\mathcal{M}(|A\rangle \otimes |\text{cat.}\rangle) < \mathcal{M}(|B\rangle \otimes |\text{cat.}\rangle)$ for any catalyzing state $|\text{cat.}\rangle$. Tensor powers of states simplify, allowing us to make asymptotic implications, i.e.,

$$\mathcal{M}(|A\rangle) < \alpha \cdot \mathcal{M}(|B\rangle) \quad \text{implies} \quad |A\rangle^{\otimes n} \not\Longrightarrow |B\rangle^{\otimes \alpha n} \quad \forall n.$$

using the arrow notation described in Definition 2.9. In other words this would put an upper bound of α on the catalytic rate of conversion from $|A\rangle$ to $|B\rangle$. Note that equivalent implications hold if the monotone is *multiplicative* rather than additive, i.e., if $\mathcal{M}(|\psi\rangle \otimes |\phi\rangle) = \mathcal{M}(|\psi\rangle) \cdot \mathcal{M}(|\phi\rangle)$ for all $|\psi\rangle$ and $|\phi\rangle$.

For example, consider the states $|T\rangle$ and $|CCZ\rangle$ for which the best known conversion algorithms are:

$$\begin{aligned} 4|T\rangle &\rightarrow |CCZ\rangle, \\ |CCZ\rangle &\overset{|T\rangle}{\Longrightarrow} 2|T\rangle. \end{aligned}$$

Clearly these algorithms would have loss if feeding the output of one into the other. The best possible conversion algorithms have (for any n and any catalyst) the minimum r and maximum r' in

$$\begin{aligned} rn|T\rangle &\Rightarrow n|CCZ\rangle, \\ n|CCZ\rangle &\Rightarrow r'n|T\rangle. \end{aligned}$$

It is straightforward to compute the stabilizer nullity values $\nu(|T\rangle) = 1$ and $\nu(|CCZ\rangle) = 3$. As described above, the fact that ν is additive immediately implies $r \geq 3$ and $r' \leq 3$. It is also possible to compute the extent values $\xi(|T\rangle) = (\sec \pi/8)^2 = 1.17157$ and $\xi(|CCZ\rangle) = 16/9 = 1.77778$. Moreover, $\log \xi$ is an additive monotone with respect to collections of $|T\rangle$ states

and $|CCZ\rangle$ states (which satisfy Lemma 2.7)[3] and therefore $r \geq \log[1.77778]/\log[1.17157] = 3.63356$ and $r' \leq \log[1.77778]/\log[1.17157] = 3.63356$. We therefore have that $r \geq 3.63356$ and $r' \leq 3$. From these bounds we see there is a gap: the best possible algorithm would require at least 3.63 $|T\rangle$ states to produce a $|CSS\rangle$ state, which can then be converted back into at most 3 $|T\rangle$ states. In Table 1 and Table 2 we show these conversion bounds along with those for many other pairs of states.

$\|\psi\rangle$	Best algo. (lower bound) [Ref.] $rn\|T\rangle \Rightarrow n\|\psi\rangle$	Best algo. (upper bound) [Ref.] $n\|\psi\rangle \Rightarrow r'n\|T\rangle$
$\left\|\sqrt{T}\right\rangle$	2.5 (1) [Fig. 7]	0.25 (0.754933*)
$\|T\rangle$	1 (1)	1 (1)
$\|CS\rangle = \|W_2\rangle$	3 (2.96818*) [Fig. 2]	1 (2) [Fig. 2]
$\|CCS\rangle$	7 (4.53328*) [32]	0.5 (3) [Prop. A.3]
$\|C^3S\rangle$	11 (4) [32]	0.25 (3.82743*) [Prop. A.3]
$\|CCZ\rangle$	4 (3.63356*) [32]	2 (3) [23]
$\|C^3Z\rangle$	6 (5.12122*) [32]	1 (4) [Tab. 2]
$\|C^4Z\rangle$	12 (5) [32]	0.5 (3.8233*) [Tab. 2]
$\|CCZ_{123,145}\rangle$	8 (5) [Tab. 2]	2 (4.37739*) [Tab. 2]
$\|W_3\rangle$	4 (3.63356*) [Tab. 2]	2 (3) [Tab. 2]
$\|W_4\rangle$	5 (4.99907*) [Fig. 4]	3 (4) [Fig. 4]
$\|W_5\rangle$	6 (5.93637*) [Fig. 4]	4 (5) [Fig. 4]

Table 1: Catalytic conversion rates to and from $|T\rangle$ states. In the first column, the produced or consumed state is specified. The second and third columns list the conversion rates (r to consume, and r' to produce) for the best known algorithm, along with the tightest bound implied by stabilizer extent or nullity in parenthesis. The references to particular results from Appendix A.3 are provided in square brackets. Note that bounds from the stabilizer extent, marked here by an asterisk, are not known to hold for arbitrary catalysts since the stabilizer extent is currently not known to be multiplicative for all states. The results that reference Table 2 are direct consequence of corresponding result from the table together with the inter-conversion between $|T\rangle$ and $|CCZ\rangle$. The values of the extent are calculated in Appendix A.4.

4 Computational task lower bounds

In the previous section we established bounds on the resources required to produce specific states. In this section we lower bound the non stabilizer resources needed to implement some important computational tasks. Specifically, we consider the multiply controlled Z gate in Section 4.1 and the modular adder in Section 4.2.

Our strategy to lower bound the number of copies of a resource state $|\psi\rangle$ needed to implement a unitary U (where U corresponds to some computational task) is to bootstrap bounds on the resources required to produce specific states. For example, note that a lower bound of

[3]Note that the bounds from the extent are only guaranteed to hold for catalysts which satisfy Lemma 2.7, but if (as conjectured) the extent is multiplicative for all states then it will hold in general.

$\|\psi\rangle$	Best algo. (lower bound) $rn\|CCZ\rangle \Rightarrow n\|\psi\rangle$	Best algo. (upper bound) $n\|\psi\rangle \Rightarrow r'n\|CCZ\rangle$
$\left\|\sqrt{T}\right\rangle$	0.75 (0.33333) [Fig. 7]	0.0625 (0.207767*) [Tab. 1]
$\|T\rangle$	0.5 (0.33333) [23]	0.25 (0.275212*) [32]
$\|CS\rangle = \|W_2\rangle$	1 (0.81688*) [Fig. 5, Fig. 13a]	0.5 (0.66666) [Fig. 13b]
$\|CCS\rangle$	2 (1.24763*) [32]	0.25 (1) [Prop. A.3]
$\|C^3S\rangle$	3 (1.33333) [32]	0.125 (1.05336*) [Prop. A.3]
$\|CCZ\rangle$	1(1)	1(1)
$\|C^3Z\rangle$	2 (1.40942*) [32]	0.5 (1.33333) [Prop. A.3]
$\|C^4Z\rangle$	3 (1.66667) [32]	0.25 (1.05336*) [Prop. A.3]
$\|CCZ_{123,145}\rangle$	2 (1.66667) [Fig. 14]	1 (1.20471*) [Fig. 14]
$\|W_3\rangle$	1 (1) [30]	1 (1) [30]
$\|W_4\rangle$	2.5 (1.3758*) [Tab. 1]	1 (1.33333) [Fig. 5]
$\|W_5\rangle$	3 (1.66667) [Tab. 1]	1 (1.63376*) [Fig. 5]

Table 2: Catalytic conversion rates to and from $|CCZ\rangle$ states. In the first column, the produced or consumed state is specified. The second and third columns list the conversion rates (r to consume, and r' to produce) for the best known algorithm, along with the tightest bound implied by stabilizer extent or nullity in parenthesis. The references to particular results from Appendix A.3 are provided in square brackets. Note that bounds from the stabilizer extent, marked here by an asterisk, are not known to hold for arbitrary catalysts since the stabilizer extent is currently not known to be multiplicative for all states. The results that reference Table 1 are direct consequence of corresponding result from the table together with the inter-conversion between $|T\rangle$ and $|CCZ\rangle$. The values of the extent are calculated in Appendix A.4.

the number of copies of $|\psi\rangle$ needed to produce a *state* $U|S\rangle$, where $|S\rangle$ is a stabilizer state, also serves as a lower bound for applying U. It is also useful to consider catalysis when establishing lower bounds for computational tasks. For example, suppose U maps the state $|\Psi\rangle|S\rangle$ to a state $|\Psi\rangle|\Phi\rangle$ for some non-stabilizer states $|\Phi\rangle$, $|\Psi\rangle$, then the number of copies of $|\psi\rangle$ needed to (catalytically) produce $|\Phi\rangle$ also serves as a lower bound for applying U.

4.1 Lower bounds for the C^nZ gate

The multiply controlled Z gate C^nZ is a key component of many important algorithms, for example to implement the reflection step in Grover's search [26]. We can lower bound the resources required to implement C^nZ as follows:

Proposition 4.1. *For $n \geq 3$, it is not possible to apply the multiply controlled Z gate $C^{n-1}Z$ or produce the state $\left|C^{n-1}Z\right\rangle = C^{n-1}Z|+\rangle^{\otimes n}$ by Clifford gates and measurements using fewer than n $|T\rangle$ states, or $n/2$ $|CS\rangle$ states, or $n/3$ $|CCZ\rangle$ states.*

Proof. First note that proving that a bound holds for the state $\left|C^{n-1}Z\right\rangle$ implies that it holds for the gate $C^{n-1}Z$. The proof for each of the bounds is then very straightforward: we simply show that the stabilizer nullity of the input state is smaller than the output state unless the bound is satisfied. Direct verification shows that $\nu(|T\rangle) = 1, \nu(|CS\rangle) = 2$, and $\nu(|CCZ\rangle) = 3$.

Finally, it is clear that $\nu(|C^{n-1}Z\rangle) = n$ for all $n \geq 3$ from Proposition 4.2 since we see that no non-trivial Pauli operator has expectation value $+1$ for the state $|CCZ\rangle$. □

Proposition 4.2. *For all $n \geq 3$, the Pauli spectrum of the state $|C^{n-1}Z\rangle = C^{n-1}Z|+\rangle^{\otimes n}$ has values (and multiplicities):* 1 (1); 0 $(-1 + 2^{n-1} + 2^{2n-1})$; $1 - 2^{2-n}$ $(2^n - 1)$; 2^{2-n} $(1 - 3 \cdot 2^{n-1} + 2^{2n-1})$.

Proof. Consider the multiply controlled Z state $|C^{n-1}Z\rangle$, defined as

$$\left|C^{n-1}Z\right\rangle = C^{n-1}Z|+\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{b\in\{0,1\}^n} (-1)^{b_1 \cdot b_2 \cdot \cdots \cdot b_n} |b\rangle.$$

We are interested in the set of Pauli expectation values $\langle C^{n-1}Z|X^x Z^z|C^{n-1}Z\rangle$ for arbitrary bit strings x and z. Explicit calculation shows

$$\begin{aligned} 2^n \langle C^{n-1}Z|X^x Z^z|C^{n-1}Z\rangle &= \sum_{b,b'\in\{0,1\}^n} (-1)^{b_1 \cdot b_2 \cdot \cdots \cdot b_n} (-1)^{b'_1 \cdot b'_2 \cdot \cdots \cdot b'_n} \langle b'|X^x Z^z|b\rangle, \\ &= \sum_{b,b'\in\{0,1\}^n} (-1)^{b_1 \cdot b_2 \cdot \cdots \cdot b_n} (-1)^{b'_1 \cdot b'_2 \cdot \cdots \cdot b'_n} (-1)^{z\cdot b} \langle b' + x|b\rangle, \\ &= \sum_{b\in\{0,1\}^n} (-1)^{b_1 \cdot b_2 \cdot \cdots \cdot b_n} (-1)^{(b_1+x_1)\cdot(b_2+x_2)\cdot \cdots \cdot (b_n+x_n)} (-1)^{z\cdot b}. \end{aligned}$$

When $x = 0^n$, we see that the sum simplifies to $\sum_{b\in\{0,1\}^n} (-1)^{z\cdot b}$, which is 2^n for $z = 0^n$, and 0 for any other z. For $x \neq 0^n$, note that the terms in the sum over b differ from $\sum_{b\in\{0,1\}^n} (-1)^{z\cdot b}$ only for $b = 1^n$ and $b = 1^n + x$. Therefore,

$$\begin{aligned} 2^n \langle C^{n-1}Z|X^x Z^z|C^{n-1}Z\rangle &= \sum_{b\in\{0,1\}^n} (-1)^{b_1 \cdot b_2 \cdot \cdots \cdot b_n} (-1)^{(b_1+x_1)\cdot(b_2+x_2)\cdot \cdots \cdot (b_n+x_n)} (-1)^{z\cdot b}, \\ &= -2(-1)^{z\cdot 1^n} - 2(-1)^{z\cdot(1^n+x)} + \sum_{b\in\{0,1\}^n} (-1)^{z\cdot b}. \end{aligned}$$

When $z = 0^n$, this is simply $2^n - 4$. When $z \neq 0^n$, it is $-2(-1)^{z\cdot 1^n} - 2(-1)^{z\cdot(1^n+x)}$. Summarizing,

$$|\langle C^{n-1}Z|X^x Z^z|C^{n-1}Z\rangle| = \begin{cases} 1 & \text{if } x = 0^n \text{ and } z = 0^n, \\ 0 & \text{if } x \cdot z \text{ is odd, or if } x = 0^n \text{ and } z \neq 0^n, \\ 1 - 2^{2-n} & \text{if } x \neq 0^n \text{ and } z = 0^n, \\ 2^{2-n} & \text{if } x \neq 0^n \text{ and } z \neq 0^n \text{ and } x \cdot z \text{ is even.} \end{cases} \tag{7}$$

We can count the number of each subset of binary vectors x and z to find the multiplicities. □

4.2 Lower bounds for the modular adder

The adder circuit is one of the most fundamental quantum arithmetic operations, which implements addition on a pair of registers in superposition. We can lower bound the resources required to implement it as follows:[4]

[4] After the first posting of this paper, Craig Gidney [22] showed that the state $|C^n Z\rangle$ can be produced using the n-qubit modular adder. We reproduce his argument in Appendix A.6 for completeness. Using our (slightly stronger) bounds for $\left|C^{n-1}Z\right\rangle$ the adder circuit cannot be implemented with fewer than $n + 1$ copies of $|T\rangle$, or $(n + 1)/2$ copies of $|CS\rangle$, or $(n + 1)/3$ copies of $|CCZ\rangle$.

Proposition 4.3. *An adder circuit on two n-qubit registers acts on basis states as*

$$A(|i\rangle|j\rangle) = |i\rangle|i+j\rangle$$

with $i+j$ evaluated modulo 2^n. It is not possible to implement the adder circuit with Clifford gates and measurements using fewer than $n-2$ $|T\rangle$ states, $(n-2)/2$ $|CS\rangle$ states or $(n-2)/3$ $|CCZ\rangle$ states.

Proof. The proof proceeds in two steps. First we show that the adder circuit A acting on the n-qubit quantum Fourier state $\left|QFT_n^b\right\rangle$ (defined below) and the stabilizer state $|+\rangle^{\otimes n}$ has the action $A(|+\rangle^{\otimes n}\left|QFT_n^b\right\rangle) = \left|QFT_n^{-b}\right\rangle\left|QFT_n^b\right\rangle$. This tells us that if A is implemented by a set of Clifford gates and Pauli measurements along with some input resource state $|\psi\rangle$, it must be that $\nu(|\psi\rangle\left|QFT_n^b\right\rangle) \geq \nu(\left|QFT_n^{-b}\right\rangle\left|QFT_n^b\right\rangle)$, and hence $\nu(|\psi\rangle) \geq \nu(\left|QFT_n^{-b}\right\rangle)$ by the additive property of the stabilizer nullity. Second we show that $\nu(\left|QFT_n^{-1}\right\rangle) = n-2$, which then directly implies our bounds since if the bounds are not satisfied, $\nu(|\psi\rangle) \geq \nu(\left|QFT_n^{-1}\right\rangle)$ would not be satisfied.

Given this proof structure, it remains to show that

$$A(|+\rangle^{\otimes n}\left|QFT_n^b\right\rangle) = \left|QFT_n^{-b}\right\rangle\left|QFT_n^b\right\rangle,$$

and that $\nu(\left|QFT^{-1}\right\rangle) = n-2$. First we recall the family of quantum Fourier states for each integer $a = 0, 1, \ldots, 2^n-1$:

$$|QFT_n^a\rangle = \frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}\exp\left[\frac{i2\pi ay}{2^n}\right]|y\rangle = \otimes_{k=1}^{n}\frac{|0\rangle + e^{i2\pi a/2^k}|1\rangle}{\sqrt{2}}, \tag{8}$$

where $|y\rangle$ is an n-qubit basis state (with y expressed in binary), and note that $\left|QFT_n^0\right\rangle = |+\rangle^{\otimes n}$, and $|QFT_n^a\rangle = |a+2^n\rangle$. Consider applying the adder to a pair of such states:

$$\begin{aligned} A\left(|QFT_n^a\rangle\left|QFT_n^b\right\rangle\right) &= \frac{1}{2^n}\sum_{y=0}^{2^n-1}\sum_{z=0}^{2^n-1}\exp\left[\frac{i2\pi(ay+bz)}{2^n}\right]|y\rangle|z+y\rangle, \\ &= \frac{1}{2^n}\sum_{y=0}^{2^n-1}\sum_{x=0}^{2^n-1}\exp\left[\frac{i2\pi(ay+b(x-y))}{2^n}\right]|y\rangle|x\rangle, \\ &= \left|QFT_n^{a-b}\right\rangle\left|QFT_n^b\right\rangle. \end{aligned}$$

By taking $a=0$, we have $A(|+^{\otimes n}\rangle\left|QFT_n^b\right\rangle) = \left|QFT_n^{-b}\right\rangle\left|QFT_n^b\right\rangle$ as required.

Finally, to calculate the stabilizer nullity of $\left|QFT_n^{-b}\right\rangle$, we use the tensor product decomposition in Eq. (8), and note that $\nu(\left|QFT_n^{-b}\right\rangle)$ is the sum of the stabilizer nullity for each state in the tensor product. When $b=1$, from prop. 4.3 the first two states in the tensor decomposition have $\nu=0$, whereas the remaining the $n-2$ states have $\nu=1$, such that $\nu(\left|QFT_n^1\right\rangle) = n-2$. The bounds are then implied from the fact that $\nu(|T\rangle)=1$, $\nu(|CS\rangle)=2$ and $\nu(|CCZ\rangle)=3$. □

The calculation of $\nu(\left|QFT_n^1\right\rangle) = n-2$ that we performed in the proof above also implies that the Quantum Fourier Transform on n qubits can not be performed using fewer than $n-2$ copies of $|T\rangle$.

5 Lower bounds for approximate unitary synthesis

In this section, we lower bound the number of resource states needed to approximate an arbitrary single-qubit unitary using Clifford gates and Pauli measurements. Unlike the previously-known lower bounds, our bounds: (1) allow for Pauli measurements; (2) allow measurement outcomes to affect the subsequent parts of the protocol; and (3) do not depend on the number of ancillary qubits used in the protocol.

There are some subtleties to be addressed when analyzing a protocol containing measurements that can affect the operations applied in subsequent parts of the protocol. In particular, the state the protocol outputs and the number of resource states it consumes are random variables, which depend on the sequence of measurement outcomes obtained. The following definition is convenient for formulating lower bounds in this setting.

Definition 5.1. *Consider a protocol with measurement outcomes that can affect subsequent parts of the protocol. Fixing a sequence of measurement outcomes in the protocol specifies an associated post-selected quantum circuit. Every input state to such a protocol defines a probability distribution on the set of all measurement outcomes and on their associated post-selected quantum circuits. We say that the protocol has some property P with probability at least p if, for all states input to the protocol, a sample drawn from the distribution of post-selected quantum circuits has the property P with probability at least p.*

For example, the property P above could be *the number of $|T\rangle$ states consumed is at least M*. The primary goal of this section is to establish the following result:

Theorem 5.2. *Consider a protocol that uses $\mathcal{N}_{|\Psi\rangle}(U,\varepsilon)$ copies of the resource state $|\Psi\rangle$ and stabilizer operations to approximate a one-qubit unitary U to within precision ε (measured by the diamond norm). For any positive $C > 1$ and $\varepsilon < 1/(2^8 C)$ there exists a unitary U such that the following inequalities must hold*

$$\begin{aligned}
\mathcal{N}_{|T\rangle}(U,\varepsilon) &\geq \frac{1}{6}\log_2(1/\varepsilon) - \frac{1}{6}\log_2(C) - 1, \\
\mathcal{N}_{|CCZ\rangle}(U,\varepsilon) &\geq \frac{1}{8}\log_2(1/\varepsilon) - \frac{1}{8}\log_2(C) - \frac{3}{4}, \\
\mathcal{N}_{|CS\rangle}(U,\varepsilon) &\geq \frac{1}{6}\log_2(1/\varepsilon) - \frac{1}{6}\log_2(C) - 1.
\end{aligned}$$

with probability at least $(C-1)/C$. In particular, this is the case for all unitaries U such that $2\sqrt{C\varepsilon} \leq |\langle 0|U|1\rangle|^2 \leq 6\sqrt{C\varepsilon}$.

The bounds in Theorem 5.2 directly imply related bounds on the average case, such as:

$$\mathbf{E}\mathcal{N}_{|T\rangle}(U,\varepsilon) \geq \frac{C-1}{C}\left(\frac{1}{6}\log_2(1/\varepsilon) - \frac{1}{6}\log_2(C) - 1\right).$$

In the rest of this section we put together the pieces to prove Theorem 5.2. Our strategy to lower bound the number of resource states needed to approximate unitary U to within diamond-norm precision ε is to establish a relation between this and lower bounds on approximating the state $U|1\rangle$ to within trace norm ε'. Unfortunately, the associated resource requirement divergence is not captured by either the nullity or extent monotones we have discussed as they do not diverge for states approaching $|0\rangle$. Our unitary synthesis results do not

hold when a catalyst state is allowed, in contrast to those bounds proven with the stabilizer nullity due to its additive property. We make the relation between approximating the unitary U to within diamond-norm precision ε and approximating the state $U|1\rangle$ to within trace norm ε' concrete in Section 5.1, and then prove lower bounds for state approximation using different resource states in Section 5.2 and Section 5.3. Before this, we present a theorem which we use to prove our lower bounds apply even with an arbitrary number of additional stabilizer ancillas:

Theorem 5.3. *Consider a post-selected stabilizer circuit with input $|\psi_{\text{in}}\rangle$ and output $|\psi_{\text{out}}\rangle$, where $|\psi_{\text{in}}\rangle$ is defined on no fewer qubits than $|\psi_{\text{out}}\rangle$. Then there exists a set of $k = \nu(|\psi_{\text{in}}\rangle) - \nu(|\psi_{\text{out}}\rangle)$ independent commuting Pauli operators $P_1, \ldots, P_k$ and a Clifford unitary C such that*

$$|\psi_{\text{out}}\rangle \otimes |S\rangle \propto C M_{P_1} \ldots M_{P_k} |\psi_{\text{in}}\rangle,$$

where $|S\rangle$ is a stabilizer state and where M_P is the projector on the $+1$ eigenspace of P.

From Theorem 5.3, without loss of generality we can assume that there are only commuting measurements in the protocol and no ancillary qubits which simplifies our analysis. However, note that this canonical form works for post-selected measurements. We highlight this theorem here because we expect that it may be of broader application and interest. We defer the proof to Appendix A.7.

5.1 Approximate unitary synthesis with and without post-selection

Our starting point addresses the order of taking averages for a protocol with measurement outcomes that can affect subsequent parts of the protocol. In particular, the following lemma shows that a protocol that has an average output density matrix which is close to a desired state also has, on average, an output density matrix which is close to the desired state *on individual runs of the protocol*.

Lemma 5.4. *Consider a protocol that, when averaged over measurement outcomes, produces a density matrix ρ that has fidelity $\langle\psi|\rho|\psi\rangle$ at least $1-\delta$ with a pure state $|\psi\rangle$. Then, for any $C > 1$, with probability at least $(C-1)/C$ the fidelity between $|\psi\rangle$ and the protocol's output is at least $1 - C\delta$ following the convention of Definition 5.1.*

Proof. Suppose the protocol has N possible sequences of measurement outcomes. Let p_k be the probability of the k^{th} sequence of measurement outcomes occurring, and let ρ_k be the normalized density matrix of the output register for that sequence.

For fixed $C > 1$ we split the set of all fixed sequences of measurement outcomes into two subsets, S and its complement $\overline{S}$. The set S contains sequences that output good approximations of $|\psi\rangle$ such that for $k \in S$, $\langle\psi|\rho_k|\psi\rangle \geq 1 - C\delta$, and $\overline{S}$ contains sequences that output worse approximations, such that for $k \in \overline{S}$, $\langle\psi|\rho_k|\psi\rangle < 1 - C\delta$. Because the overall average output ρ has fidelity at least $1-\delta$ with $|\psi\rangle$, the probability p_S of all outcomes leading to a good approximation can not be small. More explicitly, let ρ_S and $\rho_{\overline{S}}$ be the normalized density matrices corresponding to averaging over the subsets S and $\overline{S}$ respectively:

$$\rho_S \propto \sum_{k \in S} p_k \rho_k \text{ and } \rho_{\overline{S}} \propto \sum_{k \in \overline{S}} p_k \rho_k.$$

The density matrix of the output is then $\rho = p_S \rho_S + (1 - p_S)\rho_{\overline{S}}$. By construction $\langle\psi|\rho_{\overline{S}}|\psi\rangle < 1 - C\delta$, therefore

$$1 - \delta \leq \langle\psi|\rho|\psi\rangle = p_S\langle\psi|\rho_S|\psi\rangle + (1 - p_S)\langle\psi|\rho_{\overline{S}}|\psi\rangle \leq p_S + (1 - p_S)(1 - C\delta).$$

By solving the inequality $1 - \delta \leq p_S + (1 - p_S)(1 - C\delta)$ we derive the required lower bound on p_S. □

Thus far we have used fidelity to compare a state and its approximation, but we wish to deduce something about the diamond norm distance between channels. We can give bounds in both directions between the trace distance and the fidelity $\sqrt{\langle\psi|\rho|\psi\rangle}$ using the Fuchs–van de Graaf inequalities:

$$\sqrt{\langle\psi|\rho|\psi\rangle} \geq 1 - \frac{1}{2}\||\psi\rangle\langle\psi| - \rho\|_1, \tag{9}$$

$$\||\psi\rangle\langle\psi| - \rho\|_1 \leq 2\sqrt{1 - \langle\psi|\rho|\psi\rangle}. \tag{10}$$

From the second of these inequalities and from Lemma 5.4, the following is implied: *Consider a protocol which, when averaged over measurement outcomes, produces a density matrix ρ that has fidelity at least $1-\delta$ with a pure state $|\psi\rangle$. Then, for any $C > 1$, with probability at least $(C-1)/C$ the trace distance between $|\psi\rangle$ and the protocol's output is at most $2\sqrt{C\delta}$.* Note that the square root is necessary, as exemplified by randomized protocols [9, 10, 14, 28]. A corollary of these protocols is approximate state preparation protocols that achieve trace distance $\sim \delta$ by randomly choosing between different deterministic state preparation procedures, each with trace distance $\sim \sqrt{\delta}$.

The next lemma establishes connection between the lower bounds for state preparation protocols with post-selection and lower bounds for non-post-selected protocols for approximating unitaries.

Lemma 5.5. *Consider a protocol that uses $\mathcal{N}_{|\Psi\rangle}(U, \varepsilon)$ copies of the resource state $|\Psi\rangle$ and stabilizer operations to approximate a one-qubit unitary U to within precision ε (measured by the diamond norm). For any $C > 1$, let N be the minimum number of copies of a resource state $|\Psi\rangle$ needed to approximate the state $|\psi\rangle = U|1\rangle$ to trace distance $2\sqrt{C\varepsilon}$ with any protocol composed of stabilizer operations and post-selection. Then $\mathcal{N}_{|\Psi\rangle}(U, \varepsilon) \geq N$ with probability at least $(C-1)/C$, following the convention of Definition 5.1.*

Proof. Given a protocol that uses $\mathcal{N}_{|\Psi\rangle}(U, \varepsilon)$ copies of $|\Psi\rangle$ to approximate U to diamond-norm precision ε, we could approximate the state $|\psi\rangle = U|1\rangle$ to within trace distance ε with $\mathcal{N}_{|\Psi\rangle}(U, \varepsilon)$ copies of $|\Psi\rangle$. By Fuchs-van de Graaf inequality (9), our protocol approximates $|\psi\rangle$ with fidelity at least $1 - \varepsilon/2$. We now have a statement regarding the fidelity of the protocol, averaged over all the protocol's possible measurement sequences, and we wish to connect this to post-selected protocols. By direct application of Lemma 5.4, the fidelity between the output of this protocol and $|\psi\rangle$ is at least $1 - C\varepsilon$ with probability at least $(C-1)/C$, following the convention of Definition 5.1. Finally, by Fuchs-van de Graaf inequality (10), the output density matrix ρ is within trace distance $2\sqrt{C\varepsilon}$ with probability at least $(C-1)/C$. Therefore $\mathcal{N}_{|\Psi\rangle}(U, \varepsilon) \geq N$ with probability at least $(C-1)/C$. □

In the next sub-sections we establish lower bounds on the number of $|T\rangle$ and $|CS\rangle$ states needed to approximate one qubit states when using post-selected stabilizer operations. We first establish the lower bounds involving $|CS\rangle$ because it is simpler and illustrates main ideas used for the lower bound in terms of $|T\rangle$ states.

5.2 Lower bounds with $|CS\rangle$ and $|CCZ\rangle$ resource states

We start by establishing approximation lower bound using $|CS\rangle$ states because it is the simplest case sufficient to illustrate the main proof techniques. The aim of this subsection is to prove the following result:

Lemma 5.6. *Let $N_{|CS\rangle}(|\psi\rangle, \varepsilon)$ be the minimum number of $|CS\rangle$ resource states required to approximate the one-qubit state $|\psi\rangle$ to within trace distance ε using stabilizer operations and post-selection. When $\varepsilon < 1/8$, there exists a state $|\psi\rangle$ such that $N_{|CS\rangle}(|\psi\rangle, \varepsilon) \geq 1/3 \cdot \log_2(1/\varepsilon) - 2/3$. For example, this is the case for all states such that $\varepsilon < |\langle\psi|0\rangle|^2 < 3\varepsilon$.*

Proof. Our proof has two main parts. Firstly, we note that the existence of a protocol that uses n copies of $|CS\rangle$ to approximately prepare a state $|\psi\rangle$ to within trace distance ε, where the target state satisfies $\varepsilon < |\langle\psi|0\rangle|^2 < 3\varepsilon$, implies that there must be a set of $k \leq 2n$ commuting Pauli operators which, when measured on the input state $|CS\rangle^{\otimes n}$, have a probability of all giving $+1$ outcomes in the interval $(0, 4\varepsilon)$. Secondly, we observe that the probability of a joint measurement of any $k \leq 2n$ commuting Pauli operators on the input state $|CS\rangle^{\otimes n}$ can either be zero, or must be at least $1/2^{k+n}$. We then conclude that $4\varepsilon \geq 1/2^{n+k} \geq 1/2^{3n}$ and therefore $N_{|CS\rangle}(|\psi\rangle, \varepsilon) \geq 1/3 \cdot \log_2(1/(\varepsilon)) - 2/3$.

Consider $|\psi\rangle$ such that $\varepsilon < |\langle\psi|0\rangle|^2 < 3\varepsilon$ and assume that the first qubit is the output qubit of the protocol. Let ρ be the density matrix of the output qubit. By Theorem 5.3, we can write the approximate preparation of ρ in terms of a Clifford unitary C and a set of $k - 1 = \nu(|CS\rangle^{\otimes n}) - \nu(|\Psi_{\text{out}}\rangle)$ independent commuting Pauli operators $P_1, \ldots, P_{k-1}$. Let us define $P_k = C^\dagger Z_1 C$ and show that P_k commutes with $P_1, \ldots, P_{k-1}$. Recall that if P_k anti-commutes with one of $P_1, \ldots, P_{k-1}$, this implies that $p' = 1/2$, where $p' = \text{Tr}(|0\rangle\langle 0|\rho)$ is the probability of getting a $+1$ measurement of Z_1. Next we estimate this probability based on the precision requirement $\||\psi\rangle\langle\psi| - \rho\|_1 \leq \varepsilon$. Note that p' satisfies the inequality:

$$\left||\langle 0|\psi\rangle|^2 - p'\right| = |\text{Tr}\left(|0\rangle\langle 0|\psi\rangle\langle\psi|\right) - \text{Tr}(|0\rangle\langle 0|\rho)| \leq \||\psi\rangle\langle\psi| - \rho\|_1 \leq \varepsilon,$$

where we have used the inequality $|\text{Tr}AB| \leq \|A\|_\infty \|B\|_1$, and that $\||0\rangle\langle 0|\|_\infty = 1$. This implies that the probability p' must belong to the interval $(0, 4\varepsilon)$. The condition $\varepsilon < 1/8$ implies that $p' \in (0, 1/2)$ and therefore P_k must commute with $P_1, \ldots, P_{k-1}$. Next we show that $k \leq 2n$, by showing that $\nu(|\Psi_{\text{out}}\rangle) \geq 1$. If $\nu(|\Psi_{\text{out}}\rangle) = 0$ this means that the output state in a stabilizer state and this would imply that probability of measuring $|0\rangle$ on output qubit must be $0, 1$ or $1/2$ which is ruled out by our estimate $p' \in (0, 4\varepsilon)$. The joint probability of measuring $P_1, \ldots, P_k$ is non-zero and less than the conditional probability p' and therefore also belongs to interval $(0, 4\varepsilon)$, as required.

Next we show that if the joint probability of measuring any k commuting Pauli operators $P_1, \ldots, P_k$ is non-zero, then it must be at least $1/2^{n+k}$. Consider

$$\langle CS|^{\otimes n} \prod_{j=1}^{k} \frac{(I + P_j)}{2} |CS\rangle^{\otimes n} = \frac{1}{2^k} \sum_{P \in \langle P_1, \ldots, P_k\rangle} \langle CS|^{\otimes n} P |CS\rangle^{\otimes n}.$$

The Pauli expectations of $|CS\rangle$ can only be 0, 1 or $\pm 1/2$. Therefore, the value of the expression above can always be written as $a/2^{k+n}$ for some non-negative integer a and its smallest non-zero value is $1/2^{k+n}$. □

The key to generalizing the above result from $|CS\rangle$ states to an arbitrary k-qubit resource state $|\Psi\rangle$ is to establish a lower bound on the quantity:

$$\frac{1}{2^m} \sum_{P \in \langle P_1, \ldots, P_m \rangle} \langle \Psi |^{\otimes n} P |\Psi\rangle^{\otimes n}, \tag{11}$$

where $\{P_1, \ldots, P_m\}$ are independent commuting Pauli operators and $m \leq k \cdot n$. Note that replacing p with one in the statement of the lemma leads to a slightly weaker lower bound that does not require the knowledge of p. For example, it is not too difficult to generalize the above result to use $|CCZ\rangle$ states in place of $|CS\rangle$ states, because their Pauli expectations also take values 0, 1 and $\pm 1/2$. The resulting lemma is

Lemma 5.7. *Let $N_{|CCZ\rangle}(|\psi\rangle, \varepsilon)$ be the minimum number of $|CCZ\rangle$ resource states required to approximate the one-qubit state $|\psi\rangle$ to within trace distance ε using stabilizer operations and post-selection probability p. When $\varepsilon < 1/8$, there exists a state $|\psi\rangle$ such that $N_{|CCZ\rangle}(|\psi\rangle, \varepsilon) \geq {}^1/_4 \cdot \log_2(1/\varepsilon) - {}^1/_2$.*

5.3 Lower bounds with $|T\rangle$ resource states

The goal of this subsection is to establish the lower bound on the probability of a sequence of measurements of k independent commuting Pauli operators on input state $|T\rangle^{\otimes n}$ for $k \leq n$ and then find the lower bound on the number of $|T\rangle$ states needed to approximate single a qubit unitary. The following result is the missing piece needed to generalize Lemma 5.6.

Proposition 5.8. *Let $\{P_1, \ldots, P_k\}$ be independent commuting Pauli operators and let the probability of measuring the $+1$ eigenvalue of each be*

$$p = \frac{1}{2^k} \sum_{P \in \langle P_1, \ldots, P_k \rangle} \langle T |^{\otimes n} P |T\rangle^{\otimes n}. \tag{12}$$

If the value of p is non-zero, then $p \geq \frac{1}{2^{2k+n}}$.

Before proceeding we need to introduce several concepts we are going to use in the proof [16]. Consider the following set:

$$\mathcal{R} = \left\{ \frac{a + bi + \sqrt{2}(c + di)}{2^j} \; : \; \text{for } a, b, c, d, j \text{ integers} \right\}.$$

Note that the set $\mathcal{R}$ is closed under addition, negation and multiplication, and contains 0 and 1. Thus, the set $\mathcal{R}$ is an example of a ring. Also note that the set $\mathcal{R}$ is closed under complex conjugation.

Note that the state $|T\rangle$ can can be written as a vector with entries in $\mathcal{R}$ as $(\sqrt{2}/2, (1+i)/2)$. Similarly, all Pauli operators can be written as matrices with entries in $\mathcal{R}$. For this reason, p defined in Equation (12) also belongs to $\mathcal{R}$. Moreover, as a real number, we can write $p = (a_p + c_p\sqrt{2})/2^k$ for some integers a_p, c_p, k. We cannot directly use the approach of lower

bounding p directly that we used in Sec. 5.2, because $\sqrt{2}$ is an irrational number and $a_p + c_p\sqrt{2}$ can be made arbitrary small. To address this new complication, we use the bullet map that preserves $\mathcal{R}$ and is similar to complex conjugation:

$$\left(\frac{a + bi + \sqrt{2}(c + di)}{2^k}\right)^\bullet = \left(\frac{a + bi - \sqrt{2}(c + di)}{2^k}\right).$$

One can directly check that for arbitrary elements of r_1 and r_2 of $\mathcal{R}$, the following holds:

$$(r_1 + r_2)^\bullet = r_1^\bullet + r_2^\bullet, \tag{13}$$
$$(r_1 \cdot r_2)^\bullet = r_1^\bullet \cdot r_2^\bullet, \tag{14}$$
$$(r_1^\bullet)^* = (r_1^*)^\bullet. \tag{15}$$

In addition, the map $(\cdot)^\bullet$ helps us convert numbers of the form $(a + c\sqrt{2})/2^k$ into numbers of the form $d/2^k$ because:

$$(a + c\sqrt{2})(a + c\sqrt{2})^\bullet = a^2 - 2c^2 \tag{16}$$

Now we are ready to prove the proposition:

Proof of Proposition 5.8. We will show that if p is non-zero, then $p^\bullet$ belongs to the interval $(0, 1]$ and $pp^\bullet$ is a non-negative number of the form $n_p/2^{2k+n}$ for some integer n_p. This implies that the smallest non-zero value of $p = (n_p/2^{2k+n})/p^\bullet$ is at least $1/2^{2k+n}$.

First note that the Pauli expectations of $|T\rangle$ can only be 0, 1 or $\pm 1/\sqrt{2}$. For this reason, p must be a number of the form $(a_p + c_p\sqrt{2})\sqrt{2}^n/2^k$. Using (13), (14), (15) and (16) we see that:

$$p^\bullet = \frac{1}{2^n} \sum_{P \in \langle P_1, \ldots, P_n \rangle} \langle T^\bullet|^{\otimes n} P |T^\bullet\rangle^{\otimes n} \text{ where } |T^\bullet\rangle = (-\sqrt{2}/2, (1+i)/2).$$

Therefore $p^\bullet$ is the probability of measuring a projector on the state $|T^\bullet\rangle^{\otimes n}$ and must be less or equal to one. By definition of $(\cdot)^\bullet$, $p^\bullet$ can be zero if and only if p is zero. We conclude that $p^\bullet$ belongs to the interval $(0, 1]$ as required.

Finally let us compute

$$pp^\bullet = (a_p^2 - 2c_p^2)(-1)^n/2^{2k+n} = n_p/2^{2k+n} \text{ for some integer } n_p,$$

as required. □

Using the same techniques as in the proof of Lemma 5.6 we get the following result:

Lemma 5.9. *Let $N_{|T\rangle}(|\psi\rangle, \varepsilon)$ be the minimum number of $|T\rangle$ resource states required to approximate the one-qubit state $|\psi\rangle$ to within trace distance ε using stabilizer operations. When $\varepsilon < 1/8$, there exists a state $|\psi\rangle$ such that $N_{|T\rangle}(|\psi\rangle, \varepsilon) \geq 1/3 \cdot \log_2(1/\varepsilon) - 2/3$. For example, this is the case for all states such that $\varepsilon < |\langle\psi|0\rangle|^2 < 3\varepsilon$.*

We omit the proof here because it is very similar to the proof of Lemma 5.6. These can be generalized further to include states like $|\sqrt{T}\rangle$, $|\sqrt{T}^3\rangle$ as shown in Theorem A.30 in the Appendix and other roots of T using methods described in Appendix A.11 using the dyadic monotone introduced in the next section.

Proof of Theorem 5.2. First note that setting $p = 1$ on the right hand side of the inequalities in Lemma 5.6, Lemma 5.7 and Lemma 5.9 form new (weaker) inequalities which hold for all p. Then apply Lemma 5.5 to each of these inequalities. □

6 Tighter lower bounds with measurement probabilities one half

The goal of this section is to introduce a quantity similar to the stabilizer nullity $\nu(|\psi\rangle)$ that lets us establish stronger lower bounds on the number of resource states needed for certain tasks. The drawback is that these tighter bounds are not for completely arbitrary sequences of Clifford gates and Pauli measurements, but only those in which each measurement outcome occurs with probability half. However, as so many of the known circuits are of this class, we foresee these bounds being of interest and expect them to encourage researchers to turn to more rich classes of circuits to evade them. In what follows, we first show that the well-known circuit [32] to implement the multiply-controlled-Z gate using $|CCZ\rangle$ states is optimal with probability half measurements. We then show that the best-known circuit for the modular adder [21] using $|CCZ\rangle$ states with probability half measurements uses the number of $|CCZ\rangle$ states that differs by one from the lower bound.

6.1 Lower bound with CCZ gates for C^nZ gate

Consider quantum states which, when written in the computational basis, have entries in the following set:

$$\mathbb{Z}[i, 1/2] = \left\{ \frac{a + ib}{2^k} : a, b, k \in \mathbb{Z} \right\}.$$

Indeed, $|C^nZ\rangle$ can we written as vectors with entries in the above set. Note that the set $\mathbb{Z}[i, 1/2]$ is a ring since it is closed under addition, negation, multiplication, and contains 0 and 1.

Observe that if a state $|\psi\rangle$ has entries in $\mathbb{Z}[i, 1/2]$ then for any Hermitian multi-qubit Pauli operator P, the expectation $\langle\psi|P|\psi\rangle$ can be written as $a/2^k$ for integers a, k. The expectation is in $\mathbb{Z}[i, 1/2]$ because the entries of the Pauli matrices are in $\mathbb{Z}[i, 1/2]$ and $\mathbb{Z}[i, 1/2]$ is closed under complex conjugation. The expectation is also a real number and all the real numbers in $\mathbb{Z}[i, 1/2]$ are of the form $a/2^k$ for integers a, k. Note that for stabilizer states Pauli expectations can only be ± 1 and 0. Roughly speaking, the power of 2 in the denominator of the Pauli expectation lets us capture how non-stabilizer the state is. Next we develop this intuition more rigorously.

First we need a more rigorous way to talk about the power of 2 in the denominator. Let q be a non-zero rational number. It can be written as a product of integer powers of prime numbers in a unique way:

$$q = \pm 2^k \cdot p_1^{k(1)} \cdots p_m^{k(m)},\ p_k \text{ are odd primes}, k, k(1), \ldots, k(m) \text{ are integers}$$

Let us define $v_2(q)$ to be k. Note that function v_2 is somewhat similar to $\log|\cdot|$ in that $v_2(q_1q_2) = v_2(q_1) + v_2(q_2)$, $v_2(\pm 1) = 0$ and $v_2(q) = v_2(-q)$. For odd integer a and integer k the value is $v_2\Big(a/2^k\Big) = -k$. Note also that v_2 is always non-negative for integer arguments. It is convenient to extend v_2 to all rational numbers, by defining $v_2(0) = +\infty$. Note that with this extension the multiplicative property still holds. Now we are ready to define the quantity of interest.

Definition 6.1 (Dyadic monotone)**.** *Let $|\psi\rangle$ be an n-qubit state with entries in $\mathbb{Z}[i, 1/2]$, the* dyadic monotone *is*

$$\mu_2|\psi\rangle = \max\{-v_2(\langle\psi|P|\psi\rangle) : P \in \{I, X, Y, Z\}^{\otimes n}\}.$$

The dyadic monotone is essentially the maximum power of two in the denominator over the Pauli spectrum (the set of all Pauli expectations). It is invariant under Clifford unitaries because they map the set of all multi-qubit Pauli matrices to the set of all Pauli matrices up to a sign and v_2 is insensitive to the sign of its argument. In addition, Clifford unitaries map states with entries in $\mathbb{Z}[i, 1/2]$ to states with entries in $\mathbb{Z}[i, 1/2]$, because all Clifford unitaries can be written as matrices with entries in $\mathbb{Z}[i, 1/2]$, up to a global phase.

Similarly to the stabilizer nullity ν, the dyadic monotone μ_2 behaves nicely under taking tensor products.

Proposition 6.2. *Let $|\phi\rangle$ and $|\psi\rangle$ be states with entries in $\mathbb{Z}[i, 1/2]$, then*

$$\mu_2(|\phi\rangle \otimes |\psi\rangle) = \mu_2|\phi\rangle + \mu_2|\psi\rangle.$$

Proof. The result follows from the fact that for Pauli matrices P and Q such that the expectations $\langle\phi|P|\phi\rangle$ and $\langle\psi|Q|\psi\rangle$ are non-zero it is the case that:

$$v_2(\langle\phi| \otimes \langle\psi|(P \otimes Q)|\phi\rangle \otimes |\psi\rangle) = v_2(\langle\phi|P|\phi\rangle) + v_2(\langle\psi|Q|\psi\rangle).$$

□

Another important property is that the dyadic monotone is minimal for stabilizer states:

Proposition 6.3. *Let $|\phi\rangle$ be a state $\mathbb{Z}[i, 1/2]$, then $\mu_2|\psi\rangle \geq 0$, with equality achieved if and only if $|\psi\rangle$ is a stabilizer state.*

Proof. Consider a non-zero Pauli expectation $\langle\psi|P|\psi\rangle$ and write it as $a/2^k$ for some odd integer a. Note that k must be non-negative because $|\langle\psi|P|\psi\rangle| \leq 1$. This shows that $\mu_2|\psi\rangle \geq 0$. For stabilizer states, the only non-zero expectations can be ± 1 and therefor μ_2 is zero. It remains to show that $\mu_2(|\psi\rangle) = 0$ implies that $|\psi\rangle$ is stabilizer state. First note that $\mu_2(|\psi\rangle) = 0$ implies that all non-zero Pauli expectations are odd integers. Together with the condition $|\langle\psi|P|\psi\rangle| \leq 1$ this implies that the expectations can only be ± 1, in other words either P or $-P$ is in $\mathrm{Stab}|\psi\rangle$. Suppose that $|\psi\rangle$ is an n-qubit state and let us compute the size of $\mathrm{Stab}|\psi\rangle$. Note that the set $\{I, X, Y, Z\}^{\otimes n}$ is an orthogonal basis of the space of matrices with respect to the inner product $\langle A, B\rangle = \mathrm{Tr}AB^\dagger$. The norm squared of the density matrix $|\psi\rangle\langle\psi|$ is given by the following expression:

$$1 = \langle\psi|\psi\rangle^2 = \frac{1}{2^n} \sum_{P \in \{I,X,Y,Z\}^{\otimes n}} |\mathrm{Tr}(|\psi\rangle\langle\psi|P)|^2,$$

which implies that the size of $\mathrm{Stab}|\psi\rangle$ is 2^n and therefore that $|\psi\rangle$ is a stabilizer state. □

Now we show that Pauli measurements with probability half take states with entries in $\mathbb{Z}[i, 1/2]$ to states with entries in $\mathbb{Z}[i, 1/2]$ (allowing the dyadic monotone to be evaluated). Such measurements are used in magic-state injection protocols and play an important role in reducing state preparation using non-Clifford gates to state preparation using resource states, Clifford unitaries and Pauli measurements. Measuring a ± 1 eigenvalue of a Pauli observable P with probability $1/2$ is equivalent to multiplying the state by the matrix $(I \pm P)/\sqrt{2}$ which is equal to $(1 + i)(I \pm P)/2$ up to a global phase. The matrix $(1 + i)(I \pm P)/2$ has entries in the ring $\mathbb{Z}[i, 1/2]$ and therefore the resulting state will also have entries in $\mathbb{Z}[i, 1/2]$.

Next we show that μ_2 is non-increasing under these measurements. To do this, we need another property of the function v_2 given by the following proposition:

Proposition 6.4. *For arbitrary rational numbers a, b the following inequality holds*

$$v_2(a_1 + a_2) \geq \min(v_2(a_1), v_2(a_2)). \tag{17}$$

Proof. Let us first prove the inequality for non-zero a_1, a_2. Rewrite $a_j = 2^{k_j} p_j/q_j$ for integer k_j and odd integers p_j and q_j such that

$$a_1 + a_2 = 2^{\min(k_1,k_2)}\Big(2^{k_1-\min(k_1,k_2)}p_1q_2 + 2^{k_2-\min(k_1,k_2)}p_2q_1\Big)/q_1q_2.$$

Since q_1 and q_2 are odd, $v_2(a_1 + a_2)$ is equal to

$$\min(k_1, k_2) + v_2\Big(2^{k_1-\min(k_1,k_2)}p_1q_2 + 2^{k_2-\min(k_1,k_2)}p_2q_1\Big)$$

by the multiplicative property of v_2. Since $2^{k_1-\min(k_1,k_2)}p_1q_2 + 2^{k_2-\min(k_1,k_2)}p_2q_1$ is an integer, its value of ν_2 is non-negative. The case when at least one of a_j is zero follows from the fact $\min(x, +\infty) = x$. This concludes the proof of the inequality. □

Now we are ready to prove desired result:

Proposition 6.5. *Let $|\psi\rangle$ be a state with entries in $\mathbb{Z}[i, 1/2]$, let P be a Pauli observable such that measuring its eigenvalue $+1$ has probability $1/2$ and let $|\psi_+\rangle$ be the normalized result of that measurement. Then $\mu_2|\psi\rangle \geq \mu_2|\psi_+\rangle$.*

Proof. Let us bound the value of v_2 for some Pauli operator Q evaluated on the expectation $\langle\psi_+|Q|\psi_+\rangle$. The normalized state is $|\psi_+\rangle = \frac{(I+P)}{\sqrt{2}}|\psi\rangle$. The expectation of Q is therefore equal to:

$$\langle\psi_+|Q|\psi_+\rangle = \langle\psi|(I + P)Q(I + P)|\psi\rangle/2.$$

If P and Q anti-commute, the expectation is zero and does not contribute to the calculation of μ_2. When P and Q commute, the expectation is equal to $\langle\psi|Q|\psi\rangle + \langle\psi|PQ|\psi\rangle$. Next we use inequality $v_2(a + b) \geq \min(v_2(a), v_2(b))$, to see that:

$$v_2(\langle\psi|Q|\psi\rangle + \langle\psi|PQ|\psi\rangle) \geq \min(\langle\psi|Q|\psi\rangle, \langle\psi|PQ|\psi\rangle) \geq -\mu_2|\psi\rangle.$$

We have upper-bounded $-v_2(\langle\psi_+|Q|\psi_+\rangle)$ by $\mu_2|\psi\rangle$ as required. □

Now we use these techniques to show the optimality of the well-known circuit [32] to implement the multiply-controlled-Z gate using stabilizer operations with measurement probablities half and $|CCZ\rangle$ magic states.

Lemma 6.6. *At least $n - 2$ $|CCZ\rangle$ states are needed to implement the n-qubit multiply controlled Z gate $C^{n-1}Z$ by using stabilizer operations with measurement probabilities one half. The optimal circuit follows from the construction for multiply-controlled unitaries described in [32].*

Proof. The circuit for $C^{n-1}Z$ that follows from [32] uses $n-2$ CCZ gates. By applying that circuit to $|+\rangle^{\otimes n}$ we can prepare $|C^{n-1}Z\rangle$. If there existed a circuit that used k CCZ gates for $k < n-2$, we would be able to prepare states $|C^{n-1}Z\rangle$ starting from k $|CCZ\rangle$ states and then using Clifford unitaries and Pauli observable measurements with probability half. Let us show that this is impossible. Indeed for the input state we would have value $\mu_2(|CCZ\rangle^{\otimes k}) = k$. For the output state we would have $\mu_2(|C^{n-1}Z\rangle) = n-2$. This follows from the calculation of Pauli spectrum of $|C^nZ\rangle$ in Proposition 4.2. We have shown above that μ_2 is non-increasing when we apply Clifford unitaries and measurements with probability $1/2$, therefore $k \geq n-2$ which concludes the proof. □

6.2 Lower bounds for the modular adder

To establish lower bounds for adder circuits we will use the fact that adder can create a complex conjugate copy of a Fourier state. Our strategy is to generalize μ_2 to be defined on a wider set of states including Fourier states. This is achieved by extending the domain of v_2 to a wider set of values. We postpone all the details of the construction of the generalization of v_2 to Appendix A.9. Instead we list and discuss all the properties of v_2 needed for the lower-bound proof and prove the lower bound for the adder using them. The properties are then proved in the appendix.

In the previous section, to establish the lower bounds we needed to define rings over which we can write coordinates of $|C^nZ\rangle$ states. We will need to define the rings we can use to write down the coordinates of Fourier states. We extend the domain of μ_2 to the union of the following family of sets:

$$\mathcal{R}_d = \mathbb{Z}\Big[\exp(i\pi/2^d), 1/2\Big] = \left\{ \frac{1}{2^k} \sum_{j=0}^{2^d-1} a_j \exp(i\pi j/2^d) : \text{ where } a_j, k \text{ are integers} \right\}.$$

Note that each of the sets $\mathcal{R}_d$ is closed under addition, negation, multiplication and therefore each of $\mathcal{R}_d$ is an example of a ring. In addition, ring $\mathcal{R}_d$ is closed under taking complex conjugate. Note also that $\mathcal{R}_1$ is exactly the ring $\mathbb{Z}[i, 1/2]$ and $\mathcal{R}_d \subset \mathcal{R}_{d+1}$ for all positive d.

After we defined the rings, we extend the domain of function v_2 so it is defined on values of Pauli expectations of Fourier states. For this reason, v_2 must be defined at least on the real subsets of $\mathcal{R}_d$. The proof of the lower bound for multiply-controlled-Z gate relied on additivity for a tensor product of states and monotonicity under measurements with probability $1/2$ of dyadic monotone μ_2. In turn, our proofs of the mentioned properties of dyadic monotone μ_2 relied on the following two properties of v_2:

- $v_2(a \cdot b) = v_2(a) + v_2(b)$
- $v_2(a + b) \geq \min(v_2(a), v_2(b))$

Above properties also hold for our extension of v_2. We will also need to know some explicit values of v_2 to compute μ_2 for Fourier states:

$$\text{For all odd integers } k, \text{ integers } d \geq 2 : v_2\Big(\cos(\pi k/2^d)\Big) = v_2\Big(\sin(\pi k/2^d)\Big) = \frac{1}{2^{d-1}} - 1. \quad (18)$$

For example, using above we see that $\mu_2|T\rangle = 1/2$ because $v_2(1/\sqrt{2}) = -1/2$. We can immediately conclude that C^nZ gate requires at least $2(n-2)$ $|T\rangle$ states. Next we proceed to calculate μ_2 for Fourier states:

Proposition 6.7. *Consider Fourier state*

$$|QFT_n^a\rangle = \sum_{y=0}^{2^n-1} \exp\left[\frac{i2\pi ay}{2^n}\right] |y\rangle = \otimes_{k=1}^{n} \left(|0\rangle + e^{i2\pi a/2^k}|1\rangle\right),$$

For all odd a, $\mu_2|QFT_n^a\rangle = n - 3 + (1/2)^{n-2}$.

Proof. Recall that Pauli expectations of $(|0\rangle + e^{i2\pi a/2^k}|1\rangle)/\sqrt{2}$ are

$$\{0, \cos\Big(2\pi a/2^k\Big), \sin\Big(2\pi a/2^k\Big)\}.$$

For this reason, for $k \geq 2$ we have:

$$\mu_2\Big((|0\rangle + e^{i2\pi a/2^k}|1\rangle)/\sqrt{2}\Big) = v_2\Big(\sin(\pi a/2^{k-1})\Big) = 1 - 1/2^{k-2}$$

Using multiplicative property of μ_2 we get:

$$\mu_2(|QFT_n^a\rangle) = \sum_{k=2}^{n}\Big(1 - 1/2^{k-2}\Big) = n - 3 + 1/2^{n-2}$$

□

Above leads to the following lower bound on the number of $|CCZ\rangle$ states needed to implement the modular adder:[5]

Lemma 6.8. *At least $n-2$ $|CCZ\rangle$ states are needed to implement the n-qubit modular adder for $n \geq 3$ by using stabilizer operations with measurement probabilities one half.*

Proof. Recall that by applying a circuit for modular adder to $|+\rangle^{\otimes n} \otimes |QFT_n^1\rangle$ we can create a state $|QFT_n^{-1}\rangle \otimes |QFT_n^1\rangle$. If there existed a circuit that used k CCZ gates for $k < n-2$, we would be able to prepare states $|QFT_n^{-1}\rangle$ starting from k $|CCZ\rangle$ and then using Clifford unitaries and Pauli observable measurements with probability half by using $|QFT_n^1\rangle$ as a catalyst. Let us show that this this impossible. Indeed for the input state we would have value μ_2 equal to $k + \mu_2|QFT_n^1\rangle$ and for the output state we would have $\mu_2|QFT_n^{-1}\rangle + \mu_2|QFT_n^1\rangle$. We know that μ_2 is non-increasing when we apply Clifford unitaries and measurements with probability $1/2$, therefore $k \geq \mu_2|QFT_n^{-1}\rangle = n-3+(1/2)^{n-2}$ which implies that $k \geq n-2$. □

The best known [21] modular adder construction uses $n-1$ $|CCZ\rangle$ states, therefore our bound is one $|CCZ\rangle$ state short of the optimum. Using the same techniques we can derive a lower bound of $2n-5$ $|T\rangle$ states for $n \geq 3$. This lower bound multiplicative constant is twice less than the best known construction.

It is also possible to show that the extension of μ_2 to the union of $\mathcal{R}_d$ in non-negative and that its equality to zero implies that its argument is a stabilizer state. We defer prove of this fact to Proposition A.26 in the Appendix.

6.3 Lower bounds for resource state conversion

In Section 3.2 and Appendix A.8, we have introduced protocols for catalysis assisted conversion of $|CCZ\rangle$ states into states $|\pi j/2^d\rangle$. We have found that for odd j and integer $d \geq 2$, asymptotically, one can create one $|\pi j/2^d\rangle$ state at the cost of $1-1/2^{d-1}$ $|CCZ\rangle$ states. Using the dyadic monotone we can show that this is optimal when only Pauli measurements with probability 50% are allowed.

[5] After the first posting of this paper, Craig Gidney [22] showed that the state $|C^nZ\rangle$ can be produced using the n-qubit modular adder. We reproduce his argument in Appendix A.6 for completeness. The requires at least $n-1$ copies of $|CCZ\rangle$ in this setting, which gives a tight lower bound of $n-1$ copies of CCZ to implement the modular adder for a pair of n-qubit states.

Lemma 6.9. *Consider a protocol that uses stabilizer operations with measurements probability* 50%, $|CCZ\rangle$ *states and a multi-qubit state as a catalyst. The catalyst has entries in* $\mathcal{R}_{d'}$ *for some* d'*. Suppose that such circuit uses* k $|CCZ\rangle$ *states and produces* n *states* $|\pi j/2^d\rangle$ *for odd* j *and integer* $d \geq 2$*, then* $k \geq n(1 - 1/2^{d-1})$*. Asymptotically, at least* $1 - 1/2^{d-1}$ *copies of* $|CCZ\rangle$ *state are needed to produce state* $|\pi j/2^d\rangle$*.*

Proof. Let $|\text{cat}\rangle$ be a state used as a catalyst, then μ_2 for the input of our protocol is $\mu_2|\text{cat}\rangle + k$ and for the output the value of μ_2 is $n(1 - 1/2^{d-1}) + \mu_2|\text{cat}\rangle$. This is because for odd j and integer $d \geq 2$, $\mu_2|\pi j/2^d\rangle = 1 - 1/2^{d-1}$. Above implies that $k \geq n(1 - 1/2^{d-1})$. □

It is possible to show the monotonicity of μ_2 for a wider range of measurements, namely the Pauli measurements map the state defined over $\mathcal{R}_d$ to the state defined over $\mathcal{R}_d$. We provide more details on this in Proposition A.27 in the appendix.

$\|\psi\rangle$	Best algo. (lower bound) $rn\|CCZ\rangle \Rightarrow n\|\psi\rangle$	Best algo. (upper bound) $n\|\psi\rangle \Rightarrow r'n\|CCZ\rangle$
$\left\|\sqrt{T}\right\rangle$	0.75 (0.33333, 0.75†) [Fig. 7]	0.0625 (0.207767*) [Tab. 1]
$\|T\rangle$	0.5 (0.33333, 0.5†) [23]	0.25 (0.275212*) [32]
$\|CS\rangle = \|W_2\rangle$	1 (0.81688*,1†) [Fig. 5, Fig. 13a]	0.5 (0.66666) [Fig. 13b]
$\|CCS\rangle$	2 (1.24763*,2†) [32]	0.25 (1) [Prop. A.3]
$\|C^3S\rangle$	3 (1.33333,3†) [32]	0.125 (1.05336*) [Prop. A.3]
$\|CCZ\rangle$	1(1)	1(1)
$\|C^3Z\rangle$	2 (1.40942*,2†) [32]	0.5 (1.33333) [Prop. A.3]
$\|C^4Z\rangle$	3 (1.66667,3†) [32]	0.25 (1.05336*) [Prop. A.3]
$\|CCZ_{123,145}\rangle$	2 (1.66667,2†) [Fig. 14]	1 (1.20471*) [Fig. 14]
$\|W_3\rangle$	1 (1) [30]	1 (1) [30]
$\|W_4\rangle$	2.5 (1.3758*,2†) [Tab. 1]	1 (1.33333) [Fig. 5]
$\|W_5\rangle$	3 (1.66667,2†) [Tab. 1]	1 (1.63376*) [Fig. 5]

Table 3: Catalytic conversion rates to and from $|CCZ\rangle$ states. This is an extended version of Table 2 that includes bounds based on dyadic monotone μ_2. In the first column, the produced or consumed state is specified. The second and third columns list the conversion rates (r to consume, and r' to produce) for the best known algorithm, along with the tightest bound implied by stabilizer extent or nullity in parenthesis. The references are provided in square brackets. The bounds from the stabilizer extent, marked here by an asterisk, are not known to hold for arbitrary catalysts since the stabilizer extent is currently not known to be multiplicative for all states. The bounds from the dyadic monotone, marked here by †, hold only for protocols that use measurements with outcome probabilities one half and for catalysts for which μ_2 is defined. The results that reference Table 1 are direct consequence of corresponding result from the table together with the inter-conversion between $|T\rangle$ and $|CCZ\rangle$.

7 Conclusion and open problems

We have presented a number of resource lower bounds for a variety of scenarios including resource state conversion, unitary synthesis, and computational tasks. To do so, we have

introduced a number of new tools, most notably the monotones that we call the stabilizer nullity and the dyadic monotone, along with a canonical form for post-selected stabilizer circuits. We anticipate that these tools can be used much more broadly, and for example expect the following to be fruitful applications:

- Lower bounds for the multiply-controlled adder, used in multiplication,
- Lower bounds for the hamming weight-one state preparation,
- Lower bounds for the hamming weight computation circuit,
- Lower bounds for small circuits, such as the quantum Fourier transform on small number of qubits.

There are a number of other questions which are raised by this work, which we feel are also deserving of further study:

1. For what set of states is the stabilizer extent multiplicative? Although it is multiplicative for all the states that we apply it to, it is not known to be multiplicative for all states, such that not all of our inter-conversion bounds apply in the presence of arbitrary catalysts.
2. We have found that the exact inter-conversion of stabilizer states is unavoidably lossy, even in the asymptotic limit. In the setting of entanglement theory, exactly converting between different types of entangled state is not possible, but upon relaxing the exact requirement, loss free inter-conversion is possible in entanglement theory. It would be interesting to extend the unavoidably lossy resource inter-conversion results to the inexact setting. This has been done for odd-prime qudits [48], but is not for qubits.
3. Is there a more efficient algorithm for the quantum adder which is outside the setting of probability half measurements?
4. Is there a more efficient algorithm for the multiply controlled Z which is outside the setting of probability $1/2$ measurements?
5. Studying resource state conversion protocols also informs us about possible values of arbitrary monotones. A related open question is the classification of all possible monotones for non-Clifford states that have certain properties, for example additivity (multiplicativity), faithfulness and strong convexity.

8 Acknowledgements

Circuit diagrams were created using $\langle q|pic\rangle$ [17] and Quantikz [35]. The correctness of many of the circuits was verified using Q# and Microsoft's Quantum Development Kit [41]. We thank Craig Gidney who pointed out a strengthening of our bounds for the adder after the first edition of this paper was released. For completeness, we reproduce his argument, presented in [22], in Appendix A.6.

M.H. is supported by a Royal Society–Science Foundation Ireland University Research Fellowship.

References

[1] Matthew Amy and Michele Mosca. T-count optimization and Reed-Muller codes. *IEEE Transactions on Information Theory*, Mar 2019. arXiv:1601.07363, doi:10.1109/TIT.2019.2906374.

[2] 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. *Physical Review A*, 52:3457–3467, Nov 1995. arXiv:quant-ph/9503016, doi:10.1103/PhysRevA.52.3457.

[3] Michael E. Beverland, Oliver Buerschaper, Robert Koenig, Fernando Pastawski, John Preskill, and Sumit Sijher. Protected gates for topological quantum field theories. *Journal of Mathematical Physics*, 57(2):022201, 2016. URL: https://doi.org/10.1063/1.4939783, arXiv:https://doi.org/10.1063/1.4939783, doi:10.1063/1.4939783.

[4] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *arXiv preprint arXiv:1808.00128*, 2018. arXiv:1808.00128.

[5] Sergey Bravyi and Jeongwan Haah. Magic-state distillation with low overhead. *Physical Review A*, 86:052329, Nov 2012. arXiv:1209.2426, doi:10.1103/PhysRevA.86.052329.

[6] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005. arXiv:quant-ph/0403025, doi:10.1103/PhysRevA.71.022316.

[7] Sergey Bravyi and Robert König. Classification of topologically protected gates for local stabilizer codes. *Physical review letters*, 110:170503, 04 2013. doi:10.1103/PhysRevLett.110.170503.

[8] Earl T. Campbell. Catalysis and activation of magic states in fault-tolerant architectures. *Physical Review A*, 83:032317, Mar 2011. arXiv:1010.0104, doi:10.1103/PhysRevA.83.032317.

[9] Earl T. Campbell. Shorter gate sequences for quantum computing by mixing unitaries. *Physical Review A*, 95:042306, Apr 2017. arXiv:1612.02689, doi:10.1103/PhysRevA.95.042306.

[10] Earl T. Campbell. A random compiler for fast Hamiltonian simulation. *arXiv preprint arXiv:1811.08017*, 2018. arXiv:1811.08017.

[11] Earl T. Campbell and Mark Howard. Unified framework for magic state distillation and multiqubit gate synthesis with reduced resource cost. *Physical Review A*, 95(2):022316, 2017. arXiv:1606.01904, doi:10.1103/PhysRevA.95.022316.

[12] Earl T. Campbell and Mark Howard. Magic state parity-checker with pre-distilled components. *Quantum*, 2:56, March 2018. arXiv:1709.02214, doi:10.22331/q-2018-03-14-56.

[13] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172, 2017. arXiv:1612.07330, doi:10.1038/nature23460.

[14] Andrew M. Childs, Aaron Ostrander, and Yuan Su. Faster quantum simulation by randomization. *arXiv preprint arXiv:1805.08385*, 2018. arXiv:1805.08385.

[15] Richard Cleve and Daniel Gottesman. Efficient computations of encodings for quantum error correction. *Physical Review A*, 56(1):76–82, jul 1997. arXiv:9607030, doi:10.1103/PhysRevA.56.76.

[16] H. Cohen. *Number Theory: Volume I: Tools and Diophantine Equations*. Graduate Texts in Mathematics. Springer New York, 2007. URL: https://books.google.com/books?id=8zC8VPQV8psC.

[17] Thomas G. Draper and Samuel A. Kutin. $\langle q|pic\rangle$: Quantum circuits made easy. , 2019. URL: https://github.com/qpic.

[18] Bryan Eastin and Emanuel Knill. Restrictions on transversal encoded quantum gate sets. *Phys. Rev. Lett.*, 102:110502, Mar 2009. URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.110502, doi:10.1103/PhysRevLett.102.110502.

[19] Simon Forest, David Gosset, Vadym Kliuchnikov, and David McKinnon. Exact synthesis of single-qubit unitaries over Clifford-cyclotomic gate sets. *Journal of Mathematical Physics*, 56(8):082201, aug 2015. arXiv:1501.04944, doi:10.1063/1.4927100.

[20] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86:032324, Sep 2012. arXiv:1208.0928, doi:10.1103/PhysRevA.86.032324.

[21] Craig Gidney. Halving the cost of quantum addition. *Quantum*, 2:74, jun 2018. arXiv:1709.06648, doi:10.22331/q-2018-06-18-74.

[22] Craig Gidney. Producing an N+1 Qubit CCZ State with an N Qubit Adder. , 2019. URL: https://www.https://algassert.com/post/1906.

[23] Craig Gidney and Austin G. Fowler. Efficient magic state factories with a catalyzed |CCZ> to 2|T> transformation. *arXiv preprint arXiv:1812.01238*, 2018. arXiv:1812.01238.

[24] D Gosset, V Kliuchnikov, M Mosca, and V Russo. An algorithm for the T-count. *Quantum Information \& Computation*, 14(15\&16):1261–1276, nov 2014. arXiv:1308.4134.

[25] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, nov 1999. arXiv:quant-ph/9908010, doi:10.1038/46503.

[26] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM. arXiv:quant-ph/9605043, doi:10.1145/237814.237866.

[27] Jeongwan Haah, Matthew B. Hastings, D. Poulin, and D. Wecker. Magic state distillation with low space overhead and optimal asymptotic input count. *Quantum*, 1:31, October 2017. arXiv:1703.07847, doi:10.22331/q-2017-10-03-31.

[28] Matthew B. Hastings. Turning gate synthesis errors into incoherent errors. *arXiv preprint arXiv:1612.01011*, 2016. arXiv:1612.01011.

[29] Luke E. Heyfron and Earl T. Campbell. An efficient quantum compiler that reduces T count. *Quantum Science and Technology*, 4(1):015004, Sep 2018. arXiv:1712.01557, doi:10.1088/2058-9565/aad604.

[30] Mark Howard and Earl T. Campbell. Application of a resource theory for magic states to fault-tolerant quantum computing. *Physical Review Letters*, 118:090501, Mar 2017. arXiv:1609.07488, doi:10.1103/PhysRevLett.118.090501.

[31] Raban Iten, Roger Colbeck, Ivan Kukuljan, Jonathan Home, and Matthias Christandl. Quantum circuits for isometries. *Physical Review A*, 93:032318, Mar 2016. arXiv:1501.06911, doi:10.1103/PhysRevA.93.032318.

[32] Cody Jones. Low-overhead constructions for the fault-tolerant toffoli gate. *Physical Review A*, 87:022328, Feb 2013. arXiv:1212.5069, doi:10.1103/PhysRevA.87.022328.

[33] Cody Jones. Multilevel distillation of magic states for quantum computing. *Physical Review A*, 87(4):042305, 2013. arXiv:1210.3388, doi:10.1103/PhysRevA.87.042305.

[34] Torsten Karzig, Christina Knapp, Roman M. Lutchyn, Parsa Bonderson, Matthew B. Hastings, Chetan Nayak, Jason Alicea, Karsten Flensberg, Stephan Plugge, Yuval Oreg, Charles M. Marcus, and Michael H. Freedman. Scalable designs for quasiparticle-poisoning-protected topological quantum computation with Majorana zero modes. *Phys. Rev. B*, 95(23):235305, June 2017. arXiv:1610.05289, doi:10.1103/PhysRevB.95.235305.

[35] Alastair Kay. Quantikz. 9 2018. URL: https://royalholloway.figshare.com/articles/Quantikz/7000520, doi:10.17637/rh.7000520.v3.

[36] Vadym Kliuchnikov, Alex Bocharov, Martin Roetteler, and John Yard. A Framework for Approximating Qubit Unitaries. *arXiv preprint arXiv:1510.03888*, oct 2015. arXiv:1510.03888.

[37] Emanuel Knill. Approximation by Quantum Circuits. *arXiv preprint arXiv:1812.10145*, pages 1–23, aug 1995. arXiv:quant-ph/9508006.

[38] Emanuel Knill. Quantum computing with realistically noisy devices. *Nature*, 434(7029):39, 2005. arXiv:quant-ph/0410199, doi:10.1038/nature03350.

[39] Yunseong Nam, Yuan Su, and Dmitri Maslov. Approximate quantum fourier transform with o(nlog(n)) T gates. *arXiv preprint arXiv:1803.04933*, 2018. arXiv:1803.04933.

[40] Fernando Pastawski and Beni Yoshida. Fault-tolerant logical gates in quantum error-correcting codes. *Phys. Rev. A*, 91:012305, Jan 2015. URL: https://link.aps.org/doi/10.1103/PhysRevA.91.012305, doi:10.1103/PhysRevA.91.012305.

[41] Microsoft Quantum. Microsoft Quantum Development Kit. , 2019. URL: https://www.microsoft.com/en-us/quantum/development-kit.

[42] Bartosz Regula. Convex geometry of quantum resource quantification. *Journal of Physics A: Mathematical and Theoretical*, 51(4):045303, 2017. arXiv:1707.06298, doi:10.1088/1751-8121/aa9100.

[43] Neil J. Ross and Peter Selinger. Optimal ancilla-free Clifford+T approximation of z-rotations. *Quantum Information and Computation*, 16(11-12):901–953, mar 2016. arXiv:1403.2975.

[44] Peter Selinger. Quantum circuits of t-depth one. *Physical Review A*, 87:042302, Apr 2013. arXiv:1210.0974, doi:10.1103/PhysRevA.87.042302.

[45] V. V. Shende, S. S. Bullock, and I. L. Markov. Synthesis of quantum-logic circuits. *Trans. Comp.-Aided Des. Integ. Cir. Sys.*, 25(6):1000–1010, June 2006. arXiv:quant-ph/0406176, doi:10.1109/TCAD.2005.855930.

[46] Vivek V. Shende, Igor L. Markov, and Stephen S. Bullock. Minimal universal two-qubit controlled-not-based circuits. *Physical Review A*, 69(6):062321, 2004. arXiv:quant-ph/0308033, doi:10.1103/PhysRevA.69.062321.

[47] Vivek V. Shende, Igor L. Markov, and Stephen S. Bullock. Smaller two-qubit circuits for quantum communication and computation. In *Proceedings Design, Automation and Test in Europe Conference and Exhibition*, volume 2, pages 980–985. IEEE, 2004.

[48] Xin Wang, Mark M Wilde, and Yuan Su. Efficiently computable bounds for magic state distillation. *arXiv preprint arXiv:1812.10145*, 2018. arXiv:1812.10145.

A Appendices

A.1 Generic circuits for injecting diagonal gates

In this section we provide an algorithm to implement any n-qubit diagonal unitary U using the corresponding resource state $|U\rangle = U|+\rangle^{\otimes n}$, as mentioned in Section 1.3. It was pointed out in [30] that when U belongs to the third level of the Clifford hierarchy it can be implemented using one resource state $|U\rangle$ via a half-teleportation circuit (see Figure 1(a) in [30]). Here we make this protocol more explicit as well as slightly more general. The algorithm is as follows:

Algorithm A.1 (Apply the diagonal n-qubit unitary U using $|U\rangle$)**.**
Input: *$2n$ qubits, with the first n qubits in the state $|U\rangle$, and the last n qubits in an arbitrary state $|\alpha\rangle$.*

1. *apply* $\mathrm{CNOT}_{n+1,1} \dots \mathrm{CNOT}_{2n,n}$.
2. *measure the first n qubits; the measurement outcomes are $m(1), \dots, m(n)$.*
3. *for each k in $\{1, \dots, n\}$: if $m(k)$ is 1 apply X_{n+k}.*
4. *for each k in $\{1, \dots, n\}$: if $m(k)$ is 1 apply $UX_{n+k}U^\dagger$.*

Output: *The first n qubits are in a known computational basis state, and the last n qubits are in the state $U|\alpha\rangle$.*

Above we use the notation $\mathrm{CNOT}_{a,b}$ for CNOT gate with the control qubit a and target qubit b. Note that step 3 must be completed for all k in $\{\dots n\}$ before proceeding to step 4. Next we prove the correctness of above protocol.

Proposition A.2. *Algorithm A.1 is correct. If the diagonal unitary U belongs to level k of the Clifford hierarchy, then the corrections applied in step (4) are unitaries that belong to at most level $k-1$ of the hierarchy.*

Proof. Let us first show the correctness. We will use the following notation for U and $|\alpha\rangle$:

$$U = \sum_{k\in\{0,1\}^n} e^{i\varphi(k)}|k\rangle\langle k|, \quad |\alpha\rangle = \sum_{k\in\{0,1\}^n} \alpha_k|k\rangle.$$

The initial state can be written as:

$$U|+\rangle^{\otimes n}|\alpha\rangle = \frac{1}{2^{n/2}} \sum_{k,j\in\{0,1\}^n} e^{i\varphi(k)}\alpha_j|k,j\rangle.$$

After applying the CNOT gates in step (1) the state becomes:

$$\frac{1}{2^{n/2}} \sum_{k,j\in\{0,1\}^n} e^{i\varphi(k)}\alpha_j|k\oplus j, j\rangle = \frac{1}{2^{n/2}} \sum_{k,j\in\{0,1\}^n} e^{i\varphi(k\oplus j)}\alpha_j|k,j\rangle.$$

For measurement outcome $m = (m(1), \dots, m(n))$, the state after step (2) will be

$$|m\rangle \otimes \sum_{j\in\{0,1\}^n} e^{i\varphi(j\oplus m)}\alpha_j|j\rangle.$$

After applying the last two steps of the protocol the state of qubits $n+1,\ldots,2n$ will be:

$$U\Big(X^{m(1)}\otimes\ldots\otimes X^{m(n)}\Big)U^\dagger\Big(X^{m(1)}\otimes\ldots\otimes X^{m(n)}\Big)\sum_{j\in\{0,1\}^n} e^{i\varphi(j\oplus m)}\alpha_j|j\rangle.$$

Note that:

$$\begin{aligned}U^\dagger\Big(X^{m(1)}\otimes\ldots\otimes X^{m(n)}\Big)\sum_{j\in\{0,1\}^n} e^{i\varphi(j\oplus m)}\alpha_j|j\rangle &= U^\dagger\sum_{j\in\{0,1\}^n} e^{i\varphi(j\oplus m)}\alpha_j|j\oplus m\rangle,\\ &= U^\dagger\sum_{j\in\{0,1\}^n} e^{i\varphi(j)}\alpha_{j\oplus m}|j\rangle,\\ &= \sum_{j\in\{0,1\}^n}\alpha_{j\oplus m}|j\rangle.\end{aligned}$$

Finally, we see that

$$U\Big(X^{m(1)}\otimes\ldots\otimes X^{m(n)}\Big)\sum_{j\in\{0,1\}^n}\alpha_{j\oplus m}|j\rangle = U|\alpha\rangle.$$

as required.

Finally, we note that if U belong to the ℓ^{th} level of the Clifford hierachy, we have by definition that all $UX_kU^\dagger$ belong to the $(\ell-1)^{\text{th}}$ level.

□

We finish this section with the expression for some explicit corrections $UX_kU^\dagger$ in the following list and in Figure 8 and Figure 9.

- $U=\exp\Big(i\pi|1\rangle\langle 1|/2^k\Big)$, correction: $UXU^\dagger = e^{-i\pi/2^k}\exp\Big(i\pi|1\rangle\langle 1|/2^{k-1}\Big)X$.
- $U=CS=\exp\Big(\frac{\pi i}{2}|11\rangle\langle 11|\Big)$, corrections:
  - $UX_1U^\dagger=\exp\Big(\frac{-\pi i}{2}Z\otimes|1\rangle\langle 1|\Big)X_1=\mathrm{CNOT}_{1,2}S_1S_2^\dagger\mathrm{CNOT}_{1,2}X_1$,
  - $UX_2U^\dagger=\mathrm{SWAP}_{1,2}UX_1U^\dagger\mathrm{SWAP}_{1,2}=\mathrm{CNOT}_{2,1}S_2S_1^\dagger\mathrm{CNOT}_{2,1}X_2$.
- $U=CCZ=\exp(i\pi|111\rangle\langle 111|)$, corrections:
  - $UX_1U^\dagger=\exp(i\pi I\otimes|11\rangle\langle 11|)X_1=CZ_{2,3}X_1$,
  - $UX_2U^\dagger=\mathrm{SWAP}_{1,2}UX_1U^\dagger\mathrm{SWAP}_{1,2}=CZ_{1,3}X_2$,
  - $UX_3U^\dagger=\mathrm{SWAP}_{1,3}UX_1U^\dagger\mathrm{SWAP}_{1,3}=CZ_{1,2}X_3$.

A.2 Reducing the cost of unitary synthesis using $\sqrt{T}$ gates

In this section we describe how to reduce the cost of approximate unitay synthesis using $\sqrt{T}$ states as mentioned in Section 3. We also make use of a trick to reduce the injection cost when applying $\sqrt{T}$ gates sequentially.

Applying a $\sqrt{T}$ gate using magic state injection uses an extra T gate with probability one half. Using the family of conversion protocols $|\sqrt{T}\rangle+(5k+\frac{1}{2})|T\rangle\to(2k+1)|\sqrt{T}\rangle$ to create

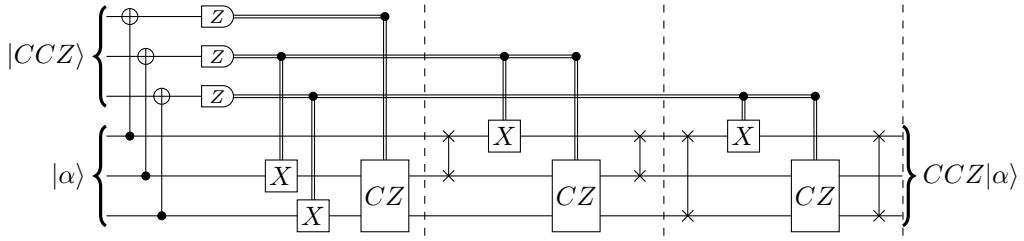
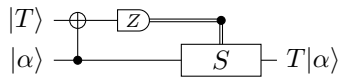
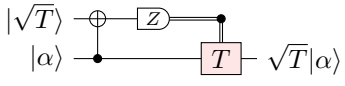


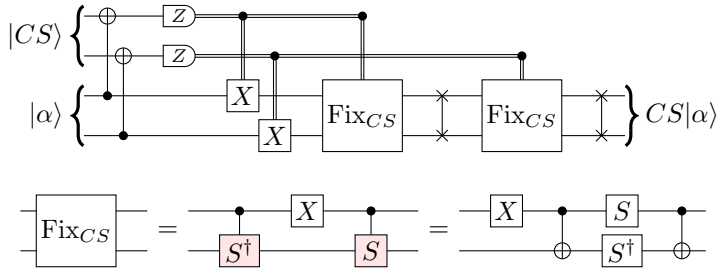
Figure 8: Implementing CCZ using $|CCZ\rangle$.



(a) Implementing T using $|T\rangle$.



(b) Implementing $\sqrt{T}$ using $|\sqrt{T}\rangle$ and $|T\rangle$.



(c) Implementing CS using $|CS\rangle$.

Figure 9: Gate injection circuits to apply some non-Clifford gates using resource ancilla states and Clifford operations. Shaded boxes represent gates in the third level of the Clifford hierarchy, which if necessary could in turn be implemented using a resource state and a Clifford circuit.

$|\sqrt{T}\rangle$ states, applying one $\sqrt{T}$ gate uses on average $3 + 1/(4k)$ T gates. In the worst case, this method will use $3.5 + 1/(2k)$ T gates. We further reduce the number of T gates needed to apply $\sqrt{T}U\sqrt{T}$. This situation is common when $\sqrt{T}$ gates are used for the synthesis of single qubit Z rotations by an arbitrary angle. The circuit shown in Figure 10 uses on average three T gates per $\sqrt{T}$ gate and 3.5 T gates in the worst case. In addition, applying $\sqrt{T}$ gates using the protocol in Figure 10 requires less ancillary qubits in comparison to using conversion protocols $|\sqrt{T}\rangle + (5k + \frac{1}{2})|T\rangle \to (2k + 1)|\sqrt{T}\rangle$ for $k > 1$.

A significant application of the above is to reduce the overhead of circuit synthesis by giving access to a larger gate set. We therefore take an aside here to explain the context and describe how our results imply overhead reduction. Approximating the single qubit rotation $\exp(i\theta|1\rangle\langle 1|)$ to within 1-norm accuracy ε using Clifford and T gates requires less then $3\log_2(1/\varepsilon) + O(\log(\log_2(1/\varepsilon)))$ T gates [43] in the typical case and less then $4\log_2(1/\varepsilon) + O(1)$ in the worst case. Consider now expanding the gate set to Clifford, T, $\sqrt{T}$ and $\sqrt{T}^3$ gates. If N_T, $N_{\sqrt{T}}$ and $N_{\sqrt{T}^3}$ denote the number of T, $\sqrt{T}$ and $\sqrt{T}^3$ gates used to approximate the rotation, then the algorithm described in [36] finds gate sequences with the number of gates satisfying:

$$2N_T + 3\left(N_{\sqrt{T}} + N_{\sqrt{T}^3}\right) < 4\log_2(1/\varepsilon) + O(1).$$

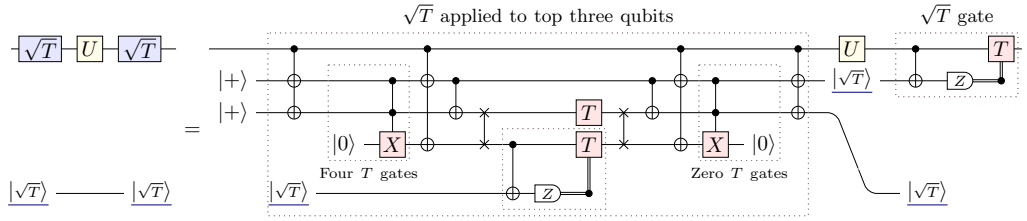


Figure 10: Application of $\sqrt{T}U\sqrt{T}$ catalyzed by a $|\sqrt{T}\rangle$ state. This uses six $|T\rangle$ states on average, and always uses at least five and at most seven $|T\rangle$ states. The $\sqrt{T}$ gate can be injected using the $\left|\sqrt{T}\right\rangle$ state as in Figure 9b.

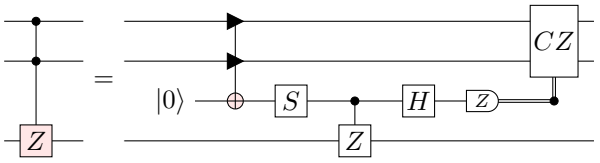
For this algorithm it was also empirically observed that $N_T \approx N_{\sqrt{T}} + N_{\sqrt{T}^3}$. Assuming that applying $\sqrt{T}$ and $\sqrt{T}^3$ gates consumes α $|T\rangle$ states, we see that using Clifford, T, $\sqrt{T}$ and $\sqrt{T}^3$ gates for rotation synthesis will use less than

$$\frac{1+\alpha}{5} \cdot 4\log_2(1/\varepsilon) + O(1)$$

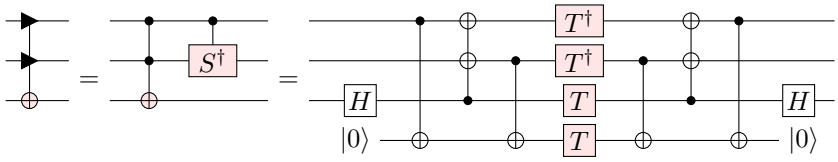
T gates. When the same algorithm uses only Clifford and T gate set it finds sequences with the at most $4\log_2(1/\varepsilon) + O(1)$ T gates. Therefore, we achieve break-even point with Clifford and T synthesis when applying $\sqrt{T}$ gate consumes four T gates. In the best protocol we find so far three T gates are consumed for each $\sqrt{T}$ gate applied on average and 3.5 T gates are consumed in the worst case. This results in an average-case 20% reduction and worst case 10% reduction in the number of T gates used to synthesize single qubit rotation.

A.3 Explicit circuits for some common resource conversions

In this section we include a number of explicit constructions which provide conversion upper bounds that appear in Table 1 and Table 2 in Section 3.

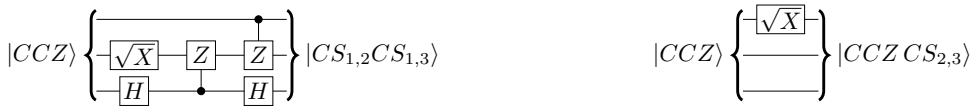


(a) CCZ gate using Toffoli* gate [32].



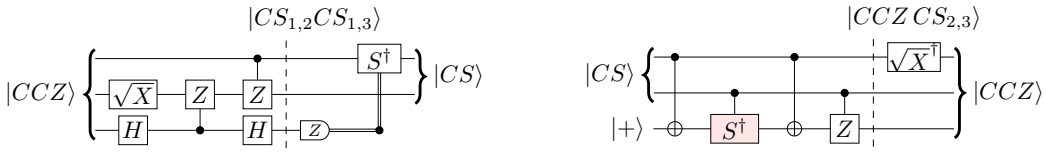
(b) The Toffoli* gate, which differs from the Toffoli gate (as defined in [44]), uses four $|T\rangle$ states.

Figure 11: Known circuits for implementing CCZ gate using four T gates.



(a) Conversion $|CCZ\rangle \leftrightarrow |CS_{1,2}CS_{2,3}\rangle$ (b) Conversion $|CCZ\rangle \leftrightarrow |CCZ\,CS_{2,3}\rangle$

Figure 12: Two way conversion of resource states from [30]. These circuits are useful subroutines for some of our results.



(a) Conversion $|CCZ\rangle \to |CS\rangle$. (b) Conversion $2|CS\rangle \to |CCZ\rangle$.

Figure 13: Conversion between $|CCZ\rangle$ and $|CS\rangle$.

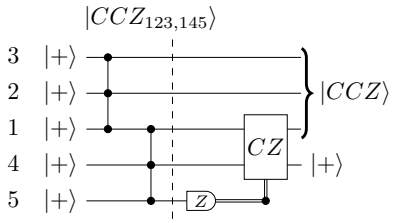


Figure 14: Conversion between $|CCZ\rangle$ and $|CCZ_{123,145}\rangle$.

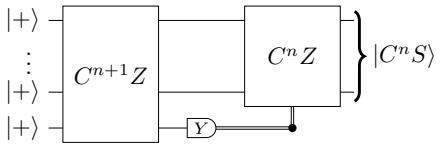


Figure 15: Conversion from $\left|C^{n+1}Z\right\rangle$ to $\left|C^{n-1}S\right\rangle$ or $\left|C^{n-1}S^\dagger\right\rangle$ with probability one half. See Proposition A.4 for the correctness proof.

Proposition A.3. *Let U be a diagonal n-qubit unitary, and let CU be a controlled version of U, then measuring the fist qubit of the state $|CU\rangle = CU|+\rangle^{\otimes(n+1)}$ in Z basis sets the rest of the qubits into the state $|U\rangle$ with probability one half and into the state $|+\rangle^{\otimes n}$ otherwise.*

In particular, this implies the following conversion protocols:

- $|C^nZ\rangle \to \frac{1}{2}|C^{n-1}Z\rangle \to \ldots \to \frac{1}{2^{n-2}}|CCZ\rangle$
- $|C^nS\rangle \to \frac{1}{2}|C^{n-1}S\rangle \to \ldots \to \frac{1}{2^{n-1}}|CS\rangle$

Proof. Note that projectors $(I \pm Z)/2$ commute with CU and therefore applying Z measurement to the first qubit is the same as measuring Z on the first qubit of $|+\rangle^{\otimes(n+1)}$ and then applying CU to $|0\rangle \otimes |+\rangle^{\otimes n}$ or $|1\rangle \otimes |+\rangle^{\otimes n}$ depending on the measurement outcome. We get $|0\rangle$ or $|1\rangle$ on the first qubit with probability one half and therefore we get $|U\rangle$ or $|+\rangle^{\otimes n}$ on the rest of the qubits with probability one half. □

Proposition A.4. *For $n \geq 0$, the probability of measuring the eigenvalue $m = \pm 1$ of Y on the first qubit of $|C^{n+1}Z\rangle$ is $1/2$. After the measurement, the state of the rest of the qubits is $|C^n S^m\rangle$.*

Proof. Let us first show that the probability of measurement outcome is $1/2$. Let us write

$$\left|C^{n+1}Z\right\rangle = |0\rangle \otimes |+\rangle^{\otimes n}/\sqrt{2} + |1\rangle \otimes |C^n Z\rangle^{\otimes n}/\sqrt{2}$$

The probability of measuring $+1$ eigenvalue of Y is:

$$\left\langle C^{n+1}Z\middle|I+Y\middle|C^{n+1}Z\right\rangle/2 = \langle 0|I+Y|0\rangle/4 + \langle 0|I+Y|0\rangle/4 + \alpha\langle 0|I+Y|1\rangle/4 + \alpha^*\langle 1|I+Y|0\rangle/4,$$

where $\alpha = \langle +|^{\otimes n}|C^n Z\rangle$. The probability is half because α is a real number and $\langle 0|I+Y|1\rangle = -\langle 1|I+Y|0\rangle$.

We prove the second part of the proposition by induction on n. When $n = 0$, and the measurement outcome is $+1$, the second qubit will be in the state

$$\frac{I+Y_1}{\sqrt{2}}|CZ\rangle = |i\rangle \otimes \left(|+\rangle/\sqrt{2} - iZ|+\rangle\sqrt{2}\right) = e^{-i\pi/4}|i\rangle \otimes |S\rangle,$$

where $|i\rangle = (1, i)/\sqrt{2}$. Suppose we now we have shown that

$$\frac{I+Y_1}{\sqrt{2}}|C^n Z\rangle = e^{-i\pi/4}|i\rangle \otimes \left|C^{n-2}S\right\rangle$$

Let us now observe that

$$\frac{I+Y_1}{\sqrt{2}}\left|C^{n+1}Z\right\rangle = \frac{I+Y_1}{\sqrt{2}}|+\rangle^{\otimes n} \otimes |0\rangle + \frac{I+Y_1}{\sqrt{2}}|C^n Z\rangle \otimes |1\rangle$$

By induction hypothesis and the fact that $(I+I)/\sqrt{2}|+\rangle = e^{-i\pi/4}|i\rangle$ it follows that:

$$\frac{I+Y_1}{\sqrt{2}}\left|C^{n+1}Z\right\rangle = e^{-i\pi/4}|i\rangle \otimes \left(|+\rangle^{\otimes n-1} \otimes |0\rangle + \left|C^{n-1}S\right\rangle|1\rangle\right) = e^{-i\pi/4}|i\rangle \otimes |C^n S\rangle$$

By applying element-wise complex conjugation to all the equations above we get the proof for the -1 outcome of the measurement, because $Y^* = -Y$. □

A.4 Extent values

Here in Table 4 we list the extent values for some common resource states, which are used to produce some of the bounds in Table 1 and Table 2 in Section 3.3. To rigorously find the exact value of the extent one can perform the following steps:

1. Find approximate numerical solutions to the primal and dual linear programs (that is a decomposition into a linear combination of stabilizer states and a witness state).
2. Guess exact expressions close to the approximate solutions (or use algebraic number reconstruction tools to find them).
3. Plug the (guessed) exact solutions into the linear program and see that min/max for primal/dual problem are equal thereby confirming they are the true solutions.

Note that we did not perform the rigorous extent calculation for some of the multi-qubit states in Table 4. Instead, we computed extent value up to eight digits of precision and reconstructed the exact expression that matches found approximation.

$\|\psi\rangle$	$\xi(\|\psi\rangle)$
$\left\|\sqrt{T}\right\rangle$	$2-\sqrt{2}+1/\sqrt{2+\sqrt{2}}$
$\|T\rangle$	$\frac{4}{2+\sqrt{2}}$
$\|CS\rangle$	$\frac{8}{5}$
$\|CCS\rangle$	$\frac{41}{20}$
$\left\|C^3S\right\rangle$	$\frac{9}{8}+\frac{1}{\sqrt{2}}$
$\|CCZ\rangle$	$\frac{16}{9}$

$\|\psi\rangle$	$\xi(\|\psi\rangle)$
$\left\|C^3Z\right\rangle$	$\frac{9}{4}$
$\left\|C^4Z\right\rangle$	$\frac{9}{8}+\frac{1}{\sqrt{2}}$
$\|CCZ_{123,145}\rangle$	2
$\|W_3\rangle$	$\frac{16}{9}$
$\|W_4\rangle$	$\frac{64}{29}$
$\|W_5\rangle$	$\frac{64}{25}$

Table 4: Exact expressions for the extent of the states used in Table 1 and Table 2. All values are accurate to within eight digits of precision. Note that $\left|C^4Z\right\rangle$ and $\left|C^3S\right\rangle$ have the same extent but they are not Clifford-equivalent; measuring $\left|C^4Z\right\rangle$ on the last qubit in the Y basis produces $\left|C^3S\right\rangle$ or $\left|C^3S^\dagger\right\rangle$ (Proposition A.4), but conversion in the reverse direction is ruled out by the stabilizer nullity.

A.5 Further details on phase polynomial protocols

Here we provide the proofs for Theorem 3.1 and Lemma A.5 in Section 3.1.

Theorem 3.1. *Let $|U\rangle = U|+\rangle^{\otimes n}$ be an n-qubit magic state for a diagonal unitary U from the 3^{rd} level of the Clifford hierarchy, and let $\tau(U)$ be the minimum number of T gates needed to implement U using the gate set $\{CNOT, S, T\}$. The following resource conversion is possible*

$$|U\rangle \xrightarrow{|T\rangle^{\otimes \tau(U)-\nu(|U\rangle)}} |T\rangle^{\otimes 2\nu(|U\rangle)-\tau(U)}. \tag{3}$$

In this theorem, we follow the conversion notation of Definition 2.9 and use ν that was defined earlier as the stabilizer nullity (recall Definition 2.2).

Proof. The proof of the theorem uses the phase polynomial formalism, which we quickly review here and the reader can learn more about in Refs. [1, 11, 29].

For any diagonal unitary in the 3^{rd} level of the Clifford hierarchy we have

$$U_f = \sum_x \exp(if(x)\pi/4)|x\rangle\langle x|, \tag{19}$$

where $f : \mathbb{Z}_2^n \to \mathbb{Z}_8$ is a cubic form. That is, f can be decomposed as the phase polynomial

$$f(x) = \sum_{a_k \neq 0} a_k \lambda_k(x) \pmod 8, \tag{20}$$

where $a_k \in \mathbb{Z}_8$ and each λ_k is a $\mathbb{Z}_2$ linear function. That is, each λ_k has the form

$$\lambda_k(x) = (P_{1,k}x_1) \oplus (P_{2,k}x_2) \dots (P_{n,k}x_n) \pmod 2, \tag{21}$$

where $P_{j,k}$ are binary. Therefore, the function can be described by a binary matrix P and vector a. We only define columns of P for nonzero a_k, so it has a number of columns equal to the number of terms in f.

For a function with a single term $f(x) = a_k\lambda_k(x)$, an easily verified circuit decomposition is

$$U_{\lambda_k} = \sum_x \exp(i\lambda_k(x)\pi/4)|x\rangle\langle x| = V^\dagger_{CNOT(\lambda_k)} T_1^{a_k} V_{CNOT(\lambda_k)} \tag{22}$$

where T_1 is a T gate acting on qubit 1 and $V_{CNOT(\lambda_k)}$ is a cascade of CNOT gates such that

$$V_{CNOT(\lambda_k)}|x\rangle = V_{CNOT(\lambda_k)}|x_1, x_2, \ldots x_n\rangle = |\lambda_k(x), x_2, \ldots x_n\rangle. \tag{23}$$

We note that if a_k is even then $T_1^{a_k} = S_1^{a_k/2}$ is a Clifford and the whole circuit is Clifford. Whereas if a_k is odd then $T_1^{a_k} = T_1 S_1^{(a_k-1)/2}$ and only a single T gate is used. For a phase polynomial f with many terms we have

$$U_f = \prod_k U_{\lambda_k} \tag{24}$$

and so the T-count for the associated circuit is equal to the number of odd valued a_k. If all values are even, then the unitary is Clifford.

We use this insight to split the unitary U_f into a Clifford and non-Clifford part. For each a_k coefficient, we define $b_k \in \mathbb{Z}_4$ and $c_k \in \mathbb{Z}_2$ such that $a_k = 2b_k + c_k$. Notice that $c_k = 1$ if and only if a_k is odd valued. Then we have that $f = g + 2h$ where g and h are the functions

$$g(x) = \sum_{c_k \neq 0} c_k \lambda_k(x) \pmod 8, \tag{25}$$

$$h(x) = \sum_{b_k \neq 0} b_k \lambda_k(x) \pmod 8. \tag{26}$$

We see that $U_f = U_{g+2h} = U_g U_{2h}$ where U_{2h} is a Clifford unitary. The non-Clifford part is U_g and all the terms have odd valued co-coefficients, so the number of terms in g gives an upper bound on $\tau(U_g)$ as discussed earlier. It follows that if the function g has m (odd-valued) terms then the state can be prepared using m many T gates or states. For any given unitary U_g there is an equivalence class of different functions g that all result in the same unitary but with different numbers of terms. Herein we assume that g is the optimal representative with the fewest number of terms, which we denote $\tau(U_g)$. Design of compilers for finding this optimal function is an ongoing research area with several useful heuristics [1, 11, 29]. Furthermore, there is a binary matrix P description of g (as defined above) with a number of columns also equal to $\tau(U_g)$. A trivial, but relevant, example is $U = T^{\otimes n}$ for which $P = \mathbb{1}_n$ and $\tau(T^{\otimes n}) = n$.

The next important step is that given a unitary U_g we may also be able to remove terms from g by applying inverse T gates. More generally, given two such unitaries U_g and $U_{g'}$ with phase polynomials g and g', we have that $U_{g'} = U_g U_\Delta$ where $\Delta = g - g'$. Therefore,

$$|U_{g'}\rangle = U_\Delta |U_g\rangle, \tag{27}$$

and

$$|T\rangle^{\otimes \tau(U_\Delta)} |U_{g'}\rangle \to |U_g\rangle. \tag{28}$$

The number of T states needed is equal to $\tau(U_\Delta)$, which in turn is equal to the number of terms where g and g' differ.

Given any P we can always bring it into row-reduced echelon form using a CNOT circuit, by virtue of the arguments presented in Sec. III of Ref. [30]. Then

$$P = \begin{pmatrix} \mathbb{1}_r & A \\ 0 & 0 \end{pmatrix}, \tag{29}$$

where $\mathbb{1}_r$ is an identity matrix of size equal to $r := \text{rank}(P)$. If P is full rank the additional 0 padding is not present. Note that if P has any 0 rows then the unitary acts trivially on the corresponding qubits leaving them in the $|+\rangle$ state and so $r \geq \mu(U)$. Using our earlier argument, we can always remove from P the columns corresponding to the matrix A using a number of T states equal to the number of columns in A. Since A has $\tau(U_g) - r$ columns, this requires the same quantity of T states. The resulting $U_{g'}$ has $P' = \mathbb{1}_r$ (with possibly some 0 row padding) which corresponds to r copies of T states. Therefore, we can perform

$$|U_g\rangle|T\rangle^{\otimes(\tau(U_g)-r)} \to |T\rangle^{\otimes r}. \tag{30}$$

If $r = \mu(U_g)$ then we have the result of the theorem. If $r > \mu(U_g)$ then we actually have a stronger result and the statement of the theorem still follows. □

The interesting cases of Theorem 3.1 are those where $|U\rangle \to |T\rangle^{\otimes r}$ is forbidden by virtue of the ring argument as presented in Theorem 2.8. We make the following observation

Claim 1. *Let U be a diagonal unitary from the 3rd level of the Clifford hierarchy with phase polynomial matrix P. If all rows of P have even Hamming weight then $U|+\rangle^{\otimes n} \nrightarrow |T\rangle$.*

To see this, note that every diagonal unitary from the 3rd level of the Clifford hierarchy is (up to Cliffords) a product of T, CS and CCZ gates [11]. In the special case that U has phase polynomial matrix with even rows, then the unitary is a product of CS and CCZ gates (see App.D of Ref. [30]). Such a unitary has elements in the ring $\mathbb{Q}(i)$ and so $U|+\rangle^{\otimes n} \nrightarrow |T\rangle$ follows (as discussed in Section 2.3). Though this transform is impossible without a catalyst, Theorem 3.1 gives a recipe for designing catalytic protocols and we next discuss some concrete examples.

For any $n \geq 2$, we define W_n as the unitary with phase polynomial matrix

$$P_n = (\mathbb{1}_n, 1) = \begin{pmatrix} 1 & 0 & & 0 & 1 \\ 0 & 1 & & 0 & 1 \\ & & \ddots & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \tag{31}$$

which is the identity matrix padded with an all-one column. More explicitly, we have W_n

$$W_n = \sum_x \exp(i\pi g(x)/4)|x\rangle\langle x|, \tag{32}$$

with

$$g(x) = (\oplus_{i=1}^n x_i) + \sum_{i=1}^n x_i, \tag{33}$$

where the $\oplus$ sum is performed modulo 2.

With the machinery of phase polynomials and P matrices established, it is now straightforward to prove Lemma A.5,

Lemma A.5. $\tau(W_n) = n + 1$.

Proof. Since P has a width of $n + 1$ columns, we have $\tau(W_n) \leq n + 1$. The only full rank phase polynomial matrices that are square give a unitary that is Clifford equivalent to $T^{\otimes n}$, since this is not the case we conclude $\tau(W_n) = n + 1$. □

Notice that every row of P_n is even weight and so by Claim 1 we know $|W_n\rangle \nrightarrow |T\rangle$. Since P is a full rank matrix, we have $\mu(|W_n\rangle) = n$. Therefore, $2\mu(|U\rangle) - \tau(U) = n - 1$ and by Theorem 3.1 we conclude that

$$|W_n\rangle \implies |T\rangle^{\otimes n-1}. \tag{34}$$

These are the most illuminating examples that one can obtain from Theorem 3.1 because assuming $U \neq T^{\otimes n}$ we know $\tau(U) > \mu(U)$ and then $\tau(U) = \mu(U) + 1$ leads to the best possible catalysis protocols.

At first glance, the W_n unitaries may look unfamiliar. However, W_2 has the same non-Clifford part as CS and so they are equivalent up to Cliffords. The W_2 example is also equivalent to the catalysis protocol first observed by Campbell [8]. For W_3, we have that the state $|W_3\rangle$ is Clifford equivalent to $|CCZ\rangle$ and so $|CCZ\rangle \implies |T\rangle^{\otimes 2}$, which is the catalysis protocol observed by Gidney and Fowler [23]. The Clifford equivalence of $|W_3\rangle$ and $|CCZ\rangle$ may be not obvious and so we comment further on this. We have that $CNOT_{3,2}W_3CNOT_{3,2}$ has the same phase polynomial matrix as $V = CCZ_{1,2,3}CS_{2,3}$. Furthermore, Cody Jones [32] showed that V can be used to synthesize CCZ, which establishes the equivalence.

A.6 Lower bound reduction from the modular adder to the multiply-controlled Z state

Here we reproduce the argument from [21] that the modular adder which acts on a pair of n-qubit registers can be used to produce the n-controlled Z state $|C^nZ\rangle$.

The argument proceeds in two steps. First we show that the controlled modular increment circuit $C\text{Inc}_n$ can be used to produce a $|C^nZ\rangle$ state using Clifford operations. Second, we show that the controlled modular increment circuit can be implemented using the modular adder. The controlled modular increment circuit Inc_n acts as follows on computational basis states

$$C\text{Inc}_n : \ |j\rangle|a + j\rangle \ \mapsto \ |j\rangle|a + j \mod 2^n\rangle, \tag{35}$$

where $j = 0, 1$ and $a = 0, 1, \ldots 2^n - 1$ is stored using binary on an n-qubit state. One can implement the controlled modular increment circuit as show in Figure 16(b).

Consider applying $C\text{Inc}_n$ to the state $|+\rangle^{\otimes n}|0\rangle$. Since the target of all the controlled gates is the X gate, those which have a target qubit in the state $|+\rangle$ (an eigenstate of X) have no action and can be removed from the circuit so that only the last n-controlled not gate remains. The resulting state is clearly Clifford-equivalent to $|C^nZ\rangle$ as shown in Figure 16(c).

We have seen that if one can implement $C\text{Inc}_n$, it is possible to produce the state $|C^nZ\rangle$ by applying it to a stabilizer state and using Clifford gates. Now note that one can implement $C\text{Inc}_n$ with the modular adder in Figure 16(d) by using the last qubit of the first input of the adder as the control and setting the other qubits of the first input to $|0\rangle$.

A.7 Canonical form for post-selected stabilizer computations

The goal of this section is to establish the canonical form for post-selected stabilizer computations described in Theorem 5.3.

First we show we can assume both the input and output states of Theorem 5.3 have trivial stabilizer, i.e. defined on a number of qubits equal to their nullity, due to the following proposition:

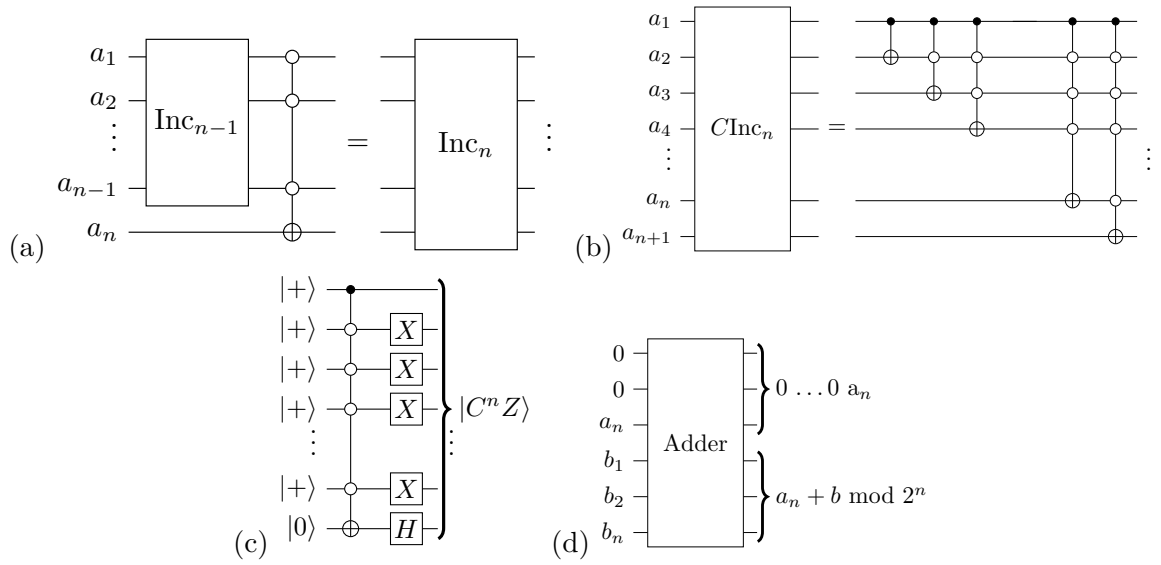


Figure 16: (a) An inductive argument shows that the modular increment circuit Inc_n is built from a sequence of multiply-controlled not gates (where the control is activated on the $|0\rangle$ state of the control qubits rather than the $|1\rangle$ state). (b) The controlled modular increment circuit $C\mathrm{Inc}_n$ is then implemented by including an additional control for each gate on the additional qubit which controls whether or not Inc_n is applied. (c) When applied to the state $|+\rangle^{\otimes n}|0\rangle$, all but the last of the gates in the circuit for $C\mathrm{Inc}_n$ annihilate, and the resulting state is Clifford-equivalent to $|C^n Z\rangle$. (d) One can implement $C\mathrm{Inc}_n$ with the modular adder by using the last qubit of the first input of the adder as the control and setting the other qubits of the first input to $|0\rangle$.

Proposition A.6. *Let $|\phi\rangle$ be an m-qubit state and let $|\mathrm{Stab}(|\phi\rangle)| = 2^r$ for $r > 0$, then there exist a Clifford unitary C such that $|\phi\rangle = C(|0\rangle^r \otimes |\phi'\rangle)$ where $|\psi'\rangle$ has a trivial stabilizer.*

Proof. Recall that for any commutative sub-group $\mathcal{G}$ of the Pauli group that does not contain $-I$ there exist a Clifford C such that $C\mathcal{G}C^\dagger = \langle Z_1, \ldots, Z_m\rangle$ [15]. Choosing $\mathcal{G} = \mathrm{Stab}(|\phi\rangle)$ lets us find the required Clifford, because $\mathrm{Stab}(C|\phi\rangle) = C\ \mathrm{Stab}(|\phi\rangle)C^\dagger$. □

Given this, Theorem 5.3 is inferred from the following theorem (identical to Theorem 5.3 but in which the input and output states have trivial stabilizer) which we prove in the remainder of this section:

Theorem A.7. *Consider a post-selected stabilizer circuit with n-qubit input state $|\psi_{\mathrm{in}}\rangle$ and m-qubit output state $|\psi_{\mathrm{out}}\rangle$, where $m \leq n$ and where $n = \nu(|\psi_{\mathrm{in}}\rangle)$ and $m = \nu(|\psi_{\mathrm{out}}\rangle)$. Then there exists a set of $k = n - m$ independent commuting Pauli operators $P_1, \ldots, P_k$ and a Clifford unitary C such that*

$$|\psi_{\mathrm{out}}\rangle \otimes |S\rangle \propto CM_{P_1} \ldots M_{P_k}|\psi_{\mathrm{in}}\rangle,$$

where $|S\rangle$ is a stabilizer state and where M_P is the projector on the $+1$ eigenspace of P.

Note that if $m = n$, the states $|\psi_{\mathrm{in}}\rangle$ and $|\psi_{\mathrm{out}}\rangle$ can be obtained from one another by applying a Clifford unitary. An interesting feature of Theorem A.7 is that if we wish to enumerate all possible stabilizer circuits that can act on a particular input state, we need only

to consider Pauli measurements that commute with each other. The following proposition gives some intuition for why this is the case.

Proposition A.8. *Let $|\psi\rangle$ be an n-qubit state and let P be a n-qubit Pauli operator such that there exists $Q \in \mathrm{Stab}(|\psi\rangle)$ that anti-commutes with P. Then the measurement of P is equivalent to randomly applying the Clifford unitaries $(I+PQ)/\sqrt{2}$ or $(I-PQ)/\sqrt{2}$ with equal probability.*

Proof. First check that measuring P gives outcome $+1$ or -1 with probability $1/2$. Indeed the probability of measuring $+1$ is $\langle\psi|(I+P)|\psi\rangle/2$ and it is equal to:

$$\langle\psi|Q(I+P)Q|\psi\rangle/2 = \langle\psi|(I+QPQ)|\psi\rangle/2 = \langle\psi|(I-P)|\psi\rangle/2$$

Therefore the probability of measuring $+1$ and -1 is the same and their sum is one. Therefore the probability of each measurement outcome in $1/2$. This means what in case of $+1$ outcome the state becomes $(I+P)|\psi\rangle/\sqrt{2}$ which is equal to $(I+PQ)|\psi\rangle/\sqrt{2}$ which is a Clifford unitary. Similarly in case of -1 outcome we have applied $(I-PQ)|\psi\rangle/\sqrt{2}$. □

The next step towards the proof of Theorem A.7 is to rewrite an arbitrary quantum circuit consisting of Clifford unitaries and post-selected Pauli measurements into a canonical form. This is the subject of the next lemma:

Lemma A.9. *Let $|\psi_{\text{out}}\rangle$ be a non-zero n-qubit state that can be obtained from an n-qubit state $|\psi_{in}\rangle$ using Clifford unitaries and post-selected Pauli measurements. Then there exists a Clifford unitary C and a commutative sub-group $\mathcal{G}$ of the Pauli group that does not contain $-I$ with generators $P_1, \ldots, P_m$ such that:*

- $|\psi_{\text{out}}\rangle \propto C M_{P_m} \ldots M_{P_1} |\psi_{\text{in}}\rangle$,
- *the group generated by $\mathcal{G}$ and* $\mathrm{Stab}(|\psi_{\text{in}}\rangle)$ *is a commutative sub-group of n-qubit Pauli group and does not contain $-I$,*
- *none of the P_k's are in* $\mathrm{Stab}(|\psi_{\text{in}}\rangle)$.

Proof. We write the circuit of Clifford unitaries and post-selected Pauli measurements as:

$$|\psi_{\text{out}}\rangle \propto C_{m'+1} M_{P'_{m'}} C_{m'} M_{P'_{m'-1}} C_{m'-2} \ldots C_2 M_{P'_1} C_1 |\psi_{\text{in}}\rangle,$$

where P'_k are n-qubit hermitian Pauli operators and C_k are n-qubit Clifford unitaries. Next we observe that the projector $M_P = (I+P)/2$ transforms into another Pauli projector under conjugation by a Clifford unitary: $C^\dagger M_P C = M_{C^\dagger P C} = M_{P'}$ where P' is an n-qubit hermitian Pauli operator because Clifford unitaries map Pauli matrices to Pauli matrices. By repeatedly applying this observation we can push each Clifford unitary to the end of computation and therefore:

$$|\psi_{\text{out}}\rangle = C' M_{P''_m} M_{P''_{m-1}} \ldots M_{P''_1} |\psi_{\text{in}}\rangle,$$

where each P''_k is an n-qubit hermitian Pauli operator and C' is an n-qubit Clifford unitary.

Next we describe how to construct $P_1, \ldots, P_m$ out of $P''_1, \ldots, P''_{m'}$. Suppose P''_1 anti-commutes with some Q from $\mathrm{Stab}(|\psi_{\text{in}}\rangle)$. In this case we can replace $M_{P''_1}$ with the Clifford unitary $(I+P''_1 Q)/\sqrt{2}$ as shown in Proposition A.8. Then we pull this Clifford unitary

through the following measurements and absorb it into the Clifford gate applied at the end. If P_1'' commutes with $\mathrm{Stab}(|\psi_{in}\rangle)$, there are several cases we need to consider. If P_1'' is in $\mathrm{Stab}(|\psi_{in}\rangle)$ than $M_{P_1''}$ can be removed from the canonical form, if $-P_1''$ is in $\mathrm{Stab}(|\psi_{in}\rangle)$ then $|\psi_{\text{out}}\rangle$ is the zero state. The remaining case is that P_1'' commutes with $\mathrm{Stab}(|\psi_{in}\rangle)$ but does not belong to it. In this case we set P_1 to be P_1''. We have ensured that P_1 and $\mathrm{Stab}(|\psi_{in}\rangle)$ generate commutative sub-group of a Pauli group that does not contain $-I$ and that P_1 is not in $\mathrm{Stab}(|\psi_{in}\rangle)$. We repeat the described procedure for $P_2'', \ldots, P_{m'}''$ and get the required result. □

Lemma A.9 implies that if the state $|\psi_{\text{in}}\rangle$ can be transformed into the state $|\psi_{\text{out}}\rangle$ by post-selected stabilizer operations, then for some m, m' and n:

$$|0\rangle^{\otimes m'} \otimes |\psi_{out}\rangle \propto C M_{P_1} \ldots M_{P_n} |0\rangle^{\otimes m} \otimes |\psi_{in}\rangle.$$

To prove Theorem A.7 it remains to get rid of the ancillary qubits on the right side of this equation. The following result is a key to this.

Lemma A.10. *Let $|\phi\rangle$ and $|\psi\rangle$ be two states such that for a Clifford unitary C and $n > 0$:*

$$|0\rangle^{\otimes n} \otimes |\psi\rangle = C(|0\rangle^{\otimes n} \otimes |\phi\rangle), \tag{36}$$

then there exists a Clifford unitary C_0 such that $|\psi\rangle = C_0|\phi\rangle$.

We postpone the proof of Lemma A.10 and first complete the proof Theorem A.7 using Lemma A.10.7

Proof of Theorem A.7. Using Lemma A.9 we conclude that there exist a Clifford unitary C and commuting Pauli operators $P_1, \ldots, P_s$ such that:

$$|0\rangle^{\otimes m'} \otimes |\psi_{\text{out}}\rangle \propto C' M_{P_s} \ldots M_{P_1} |0\rangle^{\otimes m} \otimes |\psi_{\text{in}}\rangle. \tag{37}$$

Next we show that operators P_k can be replaced with operators Q_k supported only on the last $n-m$ qubits. Indeed, each of operators P_k must commute with the stabilizer of $|0\rangle^{\otimes m} \otimes |\psi_{in}\rangle$ which consists of all possible operators $Z^{a_1} \otimes \ldots \otimes Z^{a_m} \otimes I_{2^{n-m}}$ for $a_j \in \{0, 1\}$. This implies that each P_k can be written as a tensor product

$$Z^{a_{k,1}} \otimes \ldots \otimes Z^{a_{k,m}} \otimes Q_k, \text{ for some } a_{k,j} \in \{0, 1\}$$

For this reason, applying M_{P_k} to a state $|0\rangle^{\otimes m} \otimes |\psi\rangle$ is equivalent to applying $I_{2^m} \otimes M_{Q_k}$. Let $\mathcal{G}$ be a group generated by $Q_1, \ldots, Q_s$. We rewrite Equation (37) as:

$$|0\rangle^{\otimes m'} \otimes |\psi_{\text{out}}\rangle = C'(|0\rangle^{\otimes m} \otimes (M_{Q_s} \ldots M_{Q_1} |\psi_{\text{in}}\rangle))$$

We remove first m qubits initialized to $|0\rangle$ from the equation above by using Lemma A.10.

Note that after measuring $Q_1, \ldots, Q_s$ on $|\psi_{\text{in}}\rangle$ the stabilizer of the result can be strictly bigger then the group generated by $Q_1, \ldots, Q_s$. We can just add remaining generators to the list of $Q_1, \ldots, Q_s$ to make sure that there are $m - m'$ of them. If $m - m'$ is zero, then s must be zero and input and output states must be Clifford equivalent. This completes the proof. □

A.7.1 Decoupling stabilizer states

Here we prove Lemma A.10. It relies on several simpler results, which we separate into propositions and lemmas after the main proof of Lemma A.10.

Proof of Lemma A.10. Note that it is sufficient to consider the case when $|\phi\rangle$ and $|\psi\rangle$ have a trivial stabilizer. Our proof strategy consists of two steps. First we show that in the Equation (36) we can replace unitary C with a Clifford unitary C_n such that C_n commutes with Pauli matrices Z_k for k from 1 to n. Second we show that the commutation of C_n and Z_1 implies that

$$C_n = |0\rangle\langle 0| \otimes C_{n-1} + |1\rangle\langle 1| C'_{n-1}, \text{ where } C_{n-1} \text{ is a Clifford.} \tag{38}$$

This implies that $|\psi\rangle \otimes |0\rangle^{n-1}$ and $|\phi\rangle \otimes |0\rangle^{n-1}$ are Clifford equivalent. Proceeding by induction completes the proof.

Let us now construct a Clifford C_n with required properties. Consider Pauli matrices $P_a = Z_1^{a(1)} \otimes \ldots \otimes Z_n^{a(n)}$ where each $a(j)$ is either zero or one. These are exactly the matrices that stabilize $|0\rangle^{\otimes n}$. For each a, there exist b such that $CP_aC^\dagger = P_b$ because the stabilizer of $|0\rangle^n \otimes |\psi\rangle$ and $|0\rangle^n \otimes |\phi\rangle$ is exactly the set $\{P_a : a \in \{0,1\}^n\}$. There exist a Clifford D composed only of CNOT gates acting on the first n qubits such that $DCP_aC^\dagger D^\dagger = P_a$. Defining $C_n = DC$ ensures that C_n commutes with Pauli Z_k. Because D is composed only of CNOT gates acting on first n qubits $|0\rangle^n \otimes |\psi\rangle = D|0\rangle^n \otimes |\psi\rangle$. This shows that Equation (36) holds with C replaced by C_n.

Now let us show that C_n is of the form given by Equation (38). Note that C_n commutes with Pauli Z on the first qubit, therefore by Proposition A.11 unitary C_n can be written as $|0\rangle\langle 0| \otimes C_{n-1} + |1\rangle\langle 1| \otimes C'_{n-1}$. To show that C_{n-1} must be a Clifford unitary we rely on Lemma A.13. Indeed, for any positive d, $C_{n-1} \otimes I_d$ maps stabilizer states to stabilizer states because $C_n \otimes I_d$ is a Clifford that maps stabilizer states of the from $|0\rangle \otimes |\alpha\rangle$ to stabilizer states of the form $|0\rangle \otimes |\beta\rangle$. □

The following proposition is a well-known result from the linear algebra and we provide the proof for completeness.

Proposition A.11. *Let U be a unitary that commutes with a Pauli Z matrix on the first qubit then $U = |0\rangle\langle 0| \otimes U_{00} + |1\rangle\langle 1| \otimes U_{11}$.*

Proof. Note that the fact that U commutes with Z_1 implies that U commutes with matrix $M = \lambda_0 |0\rangle\langle 0| \otimes I + \lambda_1 |1\rangle\langle 1| \otimes I$ for arbitrary complex numbers λ_0, λ_1. Let us write $U = \sum_{a,b\in\{0,1\}} |a\rangle\langle b| \otimes U_{ab}$. Next expand MU and UM as:

$$MU = \lambda_0 |0\rangle\langle 0| \otimes U_{00} + \lambda_0 |0\rangle\langle 1| \otimes U_{01} + \lambda_1 |1\rangle\langle 0| \otimes U_{10} + \lambda_1 |1\rangle\langle 1| \otimes U_{11}$$
$$UM = \lambda_0 |0\rangle\langle 0| \otimes U_{00} + \lambda_0 |1\rangle\langle 0| \otimes U_{10} + \lambda_1 |0\rangle\langle 1| \otimes U_{01} + \lambda_1 |1\rangle\langle 1| \otimes U_{11}$$

We see that equality $UM = MU$ is only possible when U_{01} and U_{10} are both zero. □

The next proposition if a convenient characterization of Pauli matrices that we use to establish a necessary condition for a unitary to be a Clifford later in this section.

Proposition A.12. *Let M be an n-qubit matrix such that $\mathrm{Tr}\left(MM^\dagger\right) = 2^n$ and for every Pauli matrix P from $\{I,X,Y,Z\}^{\otimes n}$ the trace $\mathrm{Tr}(MP)$ is either 0 or $\pm 2^n$, then M or $-M$ is a Pauli matrix.*

Proof. Recall that the set $P_n = \{I,X,Y,Z\}^{\otimes n}$ is an orthogonal basis in the vector space of n-qubit matrices with respect to inner product $\langle A,B\rangle = \mathrm{Tr}\left(AB^\dagger\right)$. Matrix M can be represented as a sum $\sum_{P\in P_n} P\langle M,P\rangle/\langle P,P\rangle$. In particular the square norm of M is $2^n = \langle M,M\rangle = \sum_{P\in P_n}|\langle P,M\rangle|^2/\langle P,P\rangle$. The equality is only possible when there is exactly one Pauli matrix P such that $\langle P,M\rangle = \pm 2^n$. □

It is well-known that Clifford unitaries map stabilizer states to stabilizer states. One can show that this is also a necessary condition for unitary to be a Clifford. Here we prove a slightly weaker result.

Lemma A.13. *Let U be an n qubit unitary such that unitary $U\otimes I_n$ maps stabilizer states to stabilizer states then U is a Clifford unitary.*

Proof. We will exploit the fact that the Choi state of U must be a stabilizer state. Recall that the Choi state of unitary U is the result of applying $U\otimes I_n$ to n Bell states. Bell state $(|00\rangle + |11\rangle)/\sqrt{2}$ is stabilized by $X\otimes X$ and $Z\otimes Z$ and its density matrix is proportional to $\sum_{P\in\{I,X,Y,Z\}} P\otimes P$. The density matrix of n Bell states is proportional to $\sum_{P\in\{I,X,Y,Z\}^{\otimes n}} P\otimes P$. The density matrix of the Choi state of U is equal to:

$$\rho = \frac{1}{2^{2n}} \sum_{P\in\{I,X,Y,Z\}^{\otimes n}} UPU^\dagger \otimes P$$

Let us now fix P and show that $M = UPU^\dagger$ is a Pauli matrix by using Proposition A.12. First note that $\mathrm{Tr}\left(MM^\dagger\right) = 2^n$. Next observe that for arbitrary Pauli matrix Q the $\mathrm{Tr}(MQ)$ must be either 0 or 2^n. Note that $\mathrm{Tr}(\rho(Q\otimes P)) = 2^{-n}\mathrm{Tr}(MQ)$. On the other hand, because ρ is a density matrix of a stabilizer state, value $\mathrm{Tr}(\rho(Q\otimes P))$ can only be $0,1,-1$. □

A.8 Conversion protocols for dyadic rational powers of T gate

In this section we look at the creation of many copies of states $|\pi j/2^d\rangle$ which include $|T\rangle$, $|\sqrt{T}\rangle$ when $j=1$, $d=3,4$. We show that in the limit of creating many copies of the same state less than one CCZ gate is required per state. We start with generalizations of some of the results discussed in Section 3.2 in context of producing $|\sqrt{T}\rangle$ states.

Proposition A.14. *Let θ be a real number and let k be a positive integer. The parallel application of $2k+1$ unitaries $\exp(i\theta|1\rangle\langle 1|)$ can be achieved by stabilizer operations with measurements that have probability 50%, one unitary $\exp(i\theta|1\rangle\langle 1|)$, k CCZ gates and the parallel application of k unitaries $\exp(i2\theta|1\rangle\langle 1|)$.*

Proof. We proof the proposition by induction on k. Let us start with the base case $k=1$. Using a circuit similar to Figure 6a we can apply three unitaries $\exp(i\theta|1\rangle\langle 1|)$ in parallel by using one ancilla, one CCZ gate, one unitary $\exp(i2\theta|1\rangle\langle 1|)$ and one unitary $\exp(i\theta|1\rangle\langle 1|)$. Suppose now that we have established the proposition for $k=j$. Let us prove the result for $k=j+1$. We need to apply $2j+3$ unitaries $\exp(i\theta|1\rangle\langle 1|)$ in parallel. We apply first three

of them using a circuit similar to Figure 6a. The circuit will use one $\exp(i2\theta|1\rangle\langle 1|)$ gate, one ancilla, one CCZ gate and one $\exp(i\theta|1\rangle\langle 1|)$ gate. We notice that remaining $2j$ unitaries $\exp(i\theta|1\rangle\langle 1|)$ can be applied in parallel with the newly introduced one. A special case of the induction step is shown on Figure 7. Using the induction hypothesis we see that in total we will need $j+1$ ancillary qubits, $j+1$ CCZ gates, $j+1$ unitaries $\exp(i2\theta|1\rangle\langle 1|)$ and one unitary $\exp(i\theta|1\rangle\langle 1|)$. This completes the proof. □

Next we apply above proposition to obtain a protocol that uses catalysis to apply rotations $R(\theta) = \exp(i\theta|1\rangle\langle 1|)$ by angle $\theta = \pi j/2^d$ for positive integer $d \geq 3$ and odd integer j.

Proposition A.15. *Let k, d be positive integers and let j be an odd integer. The parallel application of $2k$ unitaries $R(\pi j/2^d)$ can be achieved by stabilizer operations with measurements that have probability 50%, using resource state $|\pi j/2^d\rangle$ as a catalyst, k CCZ gates and the parallel application of $k+1$ unitaries $R(\pi j/2^{d-1})$.*

Proof. To apply the required unitary transformation we use the protocol described in Proposition A.14 with the last input set to $|+\rangle$ state. This will ensure that we apply $2k$ unitaries $R(\pi j/2^d)$ in parallel and produce one resource state $|\pi j/2^d\rangle$. To apply one gate $R(\pi j/2^d)$ needed by protocol from Proposition A.14 we use resource state injection protocol. The protocol consumes one state $|\pi j/2^d\rangle$ and with probability 50% requires one application of $R(\pi j/2^{d-1})$. The gate $R(\pi j/2^{d-1})$ used in the injection protocol can be applied in parallel with the rest of $R(\pi j/2^{d-1})$ applied as a part of protocol from Proposition A.14. Therefore in total we will need to apply at most $k+1$ unitaries $R(\pi j/2^{d-1})$ in parallel. We use the same number of CCZ gates as in Proposition A.14 which is equal to k. □

In Section 3.2 we presented a special case of the above proposition for $j = 1$ and $d = 4$. Next we apply above proposition recursively to obtain a family of conversion protocols for resource states $|\pi j/2^d\rangle$ that use states $|\pi j/2^d\rangle, \ldots, |\pi j/2^2\rangle$ as catalysts together with CCZ gates.

Theorem A.16. *Let $k, d \geq 1$ be positive integers and let j be an odd integer and let $a_{d,k} = 2^{d-1}(k-1)+2$. Then $a_{d,k}$ gates $\exp(\pi ij/2^d|1\rangle\langle 1|)$ can be executed in parallel by using stabilizer operations with measurements that have probability 50%, $b_{d,k} = (2^{d-1}-1)(k-1)+d-1$ copies of $|CCZ\rangle$ state and using one copy of each of the states $|\pi j/2^d\rangle, |\pi j/2^{d-1}\rangle, \ldots, |\pi j/2^2\rangle$ as a catalyst. Asymptotically, the state $|\pi j/2^d\rangle$ is produced using $1 - 1/2^{d-1}$ $|CCZ\rangle$ states.*

Proof. We prove the theorem by induction on d. The base case $d = 1$ is true because $R(\pi j/2)$ are Clifford gates and require zero CCZ gates to be applied. Suppose now that we have shown the result for $d = d'$ and let us prove the theorem for $d = d' + 1$. We need to apply $a_{d'+1,k}$ gates $R(j/2^{d'+1})$. According to Proposition A.15 we achieve this using $a_{d'+1,k}/2$ CCZ gates, one resource state $|\pi j/2^{d'+1}\rangle$ used as a catalyst, and the parallel application of $a_{d'+1,k}/2+1$ unitaries $R(\pi j/2^{d'})$. We observe that $a_{d'+1,k}/2+1 = a_{d',k}$. Therefore, by induction hypothesis, the parallel application of unitaries $R(\pi j/2^{d'})$ can be achieved using the resources described in the statement of the theorem. The total number of CCZ gates applied is $b_{d',k} + a_{d'+1,k}/2$ which is equal to $b_{d'+1,k}$ as required. We also added state $|\pi j/2^{d'+1}\rangle$ to the list of the catalysts used in the protocol. Finally we note that $\lim_{k\to\infty} b_{d,k}/a_{d,k} = 1 - 2^{-(d-1)}$. □

A.9 Overview of some definitions and results from Number Theory

The goal of this appendix is to review the results from algebraic number theory needed to define and calculate function v_2 used in Section 6. We aim for a pedagogical and as self-contained as possible exposition of the needed results. The readers with a solid knowledge of algebraic number theory should proceed to Remark A.25.

A.9.1 Definition of v_2 and additivity

Recall that we have defined v_2 in the beginning of Section 6.1 for rational numbers as following. If q is a non-zero rational number then $v_2(q)$ is equal to the power of 2 in the factorization of q into prime numbers. If q is zero, then $v_2(q) = +\infty$. We need to extend v_2 to the real subset of the following family of sets:

$$\mathcal{R}_d = \mathbb{Z}\Big[\exp(i\pi/2^d), 1/2\Big] = \left\{ \frac{1}{2^k} \sum_{j=0}^{2^d-1} a_j \exp(i\pi j/2^d) : \text{ where } a_j, k \text{ are integers} \right\}.$$

and show that v_2 has the following two properties:

- additivity, that is $v_2(x \cdot y) = v_2(x) + v_2(y)$,
- $v_2(x + y) \geq \min(v_2(x), v_2(y))$.

We also need to calculate values $v_2\Big(\cos(\pi k/2^d)\Big)$, $v_2\Big(\sin(\pi k/2^d)\Big)$ for integers k, d.

We will define v_2 on a larger family of sets that includes $\mathcal{R}_d$ and its real subsets

$$\mathbb{Q}(\exp(i\pi/2^d)) = \left\{ \sum_{j=0}^{2^d-1} a_j \exp(i\pi j/2^d) : \text{ where } a_j \text{ are rational numbers} \right\}$$

and shown that it has the required properties. Observe that sets $\mathbb{Q}(\exp(i\pi/2^d))$ are closed under addition and multiplication similarly to sets $\mathcal{R}_d$, so $v_2(x \cdot y)$ and $v_2(x + y)$ are well-defined.

Our strategy for extending v_2 is the following. Later in this section we will define family of functions N_d on sets $\mathbb{Q}(\exp(i\pi/2^d))$ with four properties:

- value of N_d is always rational,
- N_d is multiplicative, that is $N_d(x \cdot y) = N_d(x) \cdot N_d(y)$,
- N_0 is trivial, that is $N_0(x) = x$,
- $N_d(x)^2 = N_{d+1}(x)$.

Using functions N_d and the definition of v_2 on the set of rational numbers we extend v_2 to the family of sets $\mathbb{Q}(\exp(i\pi/2^d))$ as:

$$v_2(x) = v_2(N_d(x))/2^d \tag{39}$$

Above mentioned properties of N_d make sure that v_2 is additive and well-defined. We see that the additive property of v_2 follows immediately from the multiplicative property of N_d.

The definition of v_2 on rational number does not change because $N_0(x) = x$. Finally, the definition of v_2 is consistent. Function v_2 is defined on the family of nested sets:

$$\mathbb{Q} \subset \mathbb{Q}(i) \subset \mathbb{Q}(\exp(i\pi/2^2)) \subset \ldots \subset \mathbb{Q}(\exp(i\pi/2^d)) \subset \mathbb{Q}(\exp(i\pi/2^{d+1})) \subset \ldots.$$

If x belongs to set $\mathbb{Q}(\exp(i\pi/2^d))$, then x also belongs to all the sets $\mathbb{Q}(\exp(i\pi/2^{d+k}))$ for all integer k. For v_2 to be defined consistently, $v_2(N_d(x))/2^d$ must be equal to $v_2(N_{d+k}(x))/2^{d+k}$. This follows, from property $N_d(x)^2 = N_{d+1}(x)$ and the fact that for rational q value $v_2(q^n) = nv_2(q)$.

The rest of this section is dedicated to defining function N_d known as norm functions of $\mathbb{Q}(\exp(i\pi/2^d))$ and establishing their four required properties. Let us start with $\mathbb{Q}(i)$ and N_1. The following three properties of complex conjugation are useful for our purpose:

- if $a + bi$ is in $\mathbb{Q}(i)$, then $(a + bi)^*$ is in $\mathbb{Q}(i)$,
- $(x \cdot y)^* = x^* \cdot y^*$,
- $x = x^*$ if and only if x is in $\mathbb{Q}$,

We define $N_1(x) = x \cdot x^*$. First property of complex conjugation ensures that N_1 is well-defined, the second one ensures multiplicativity of N_1, the third property ensures that N_1 is rational and that $N_1(x) = N_0(x)^2$ when x is rational. To define N_d for $d > 1$ we will need more maps similar to complex conjugation defined on sets $\mathbb{Q}(\exp(i\pi/2^d))$:

$$\sigma_k : \mathbb{Q}(\exp(i\pi/2^d)) \to \mathbb{Q}(\exp(i\pi/2^d)),\ \sigma_k\left(\sum_{j=0}^{2^d-1} a_j \exp(i\pi j/2^d)\right) = \sum_{j=0}^{2^d-1} a_j \exp(i\pi jk/2^d) \quad (40)$$

Note that σ_{-1} is the complex conjugation and σ_1 is the identity map. The next proposition established some well-known properties of maps σ_k. We provide proof for completeness.

Proposition A.17. *For all odd k, maps σ_k have the following properties*

1. *for all x, y, $\sigma_k(x \cdot y) = \sigma_k(x) \cdot \sigma_k(y)$ and $\sigma_k(x + y) = \sigma_k(x) + \sigma_k(y)$*
2. *x from $\mathbb{Q}(\exp(i\pi/2^d))$ is rational if and only if for all odd k $\sigma_k(x) = x$*
3. *for all x from $\mathbb{Q}(\exp(i\pi/2^d))$ and for all odd k, $\sigma_{k+2\cdot 2^d}(x) = \sigma_k(x)$*
4. *or all x, $\sigma_k(\sigma_j(x)) = \sigma_{kj}(x)$*

Proof. Property three follows from 2π periodicity of $\exp(i\phi)$. Property four is a direct consequence of the definition of σ_k. Additivity also follows directly from the definition. The fact that for all rational a, $\sigma(a) = a$ also follows from definition. For rational a, $\sigma_k(a \cdot x) = \sigma_k(a) \cdot \sigma_k(x)$ again by definition of σ_k.

To establish multiplicativity it is sufficient to check that for all j and j' and rational a, b:

$$\sigma_k\Big(a \exp(i\pi j/2^d) \cdot b \exp(i\pi j'/2^d)\Big) = \sigma_k\Big(a \exp(i\pi j/2^d)\Big) \cdot \sigma_k\Big(b \exp(i\pi j'/2^d)\Big),$$

and then use additivity.

It remains to show that $\sigma_k(x) = x$ for all odd k implies that x is rational. Consider

$$x = \sum_{j=0}^{2^d-1} a_j \exp(i\pi jk/2^d) \in \mathbb{Q}(\exp(i\pi/2^d))$$

and let us see what are the implications of the fact $\sigma_{2^d+1}(x) = x$. Observe, that all a_j for odd $j = 2j' + 1$ must be zero. Indeed, $\sigma_{2^d+1}\Big(\exp(i\pi(2j'+1)/2^d)\Big) = -\exp(i\pi(2j'+1)/2^d)$ and therefore $a_j = -a_j$. We have shown, that x belongs to $\mathbb{Q}(\exp(i\pi/2^{d-1}))$. Repeatedely applying above argument we conclude that x must be rational. □

Now we can define N_d as following

$$N_d(x) = \prod_{k=0}^{2^d-1} \sigma_{2k+1}(x) \tag{41}$$

and prove that N_d has required properties:

Proposition A.18. *Maps N_d have the the following properties:*

- *value of N_d is always rational,*
- *N_d is multiplicative, that is $N_d(x \cdot y) = N_d(x) \cdot N_d(y)$,*
- *N_0 is trivial, that is $N_0(x) = x$,*
- *$N_d(x)^2 = N_{d+1}(x)$.*

Proof. Multiplicativity of N_d follows from the multiplicativity of σ_k. Let us check that $\sigma_j(N_d(x)) = N_d(x)$ for all odd j to establish that N_d is rational using the second property of σ_j established in Proposition A.17:

$$\sigma_j(N_d(x)) = \prod_{k\in\{1,3,\ldots,2\cdot 2^d-1\}} \sigma_j(\sigma_k(x)) = \prod_{k\in\{1,3,\ldots,2\cdot 2^d-1\}} \Big(\sigma_{kj \bmod (2\cdot 2^d)}(x)\Big)$$

We used properties three and four from Proposition A.17 to establish the last equality. Above expression is equal to $N_d(x)$ because map $k \mapsto kj \bmod (2 \cdot 2^d)$ maps set

$$\Big\{1, 3, \ldots, 2 \cdot 2^d - 1\Big\}$$

to itself when j is odd.

Map N_0 is equal to σ_1 and therefore trivial. Consider now expression for N_{d+1} for x from $\mathbb{Q}(\exp(i\pi/2^d))$:

$$N_{d+1}(x) = \prod_{k=0}^{2^{d+1}-1} \sigma_{2k+1}(x) = \prod_{k=0}^{2^{d+1}-1} \sigma_{2k+1 \bmod (2\cdot 2^d)}(x)$$

Note that function $k \mapsto 2k + 1 \bmod (2 \cdot 2^d)$ takes the same value for k and $k + 2^d$ and therefore the expression above equals to $N_d(x)^2$. □

A.9.2 Certain values of v_2

To compute many useful values of $v_2(x)$ it is sufficient to know values of $N_d(x)$ given by the following proposition:

Proposition A.19. *For all j, $N_d\Big(\exp\Big(i\pi j/2^d\Big)\Big) = 1$ and $N_d\Big(1-\exp\Big(i\pi(2j+1)/2^d\Big)\Big) = 2$.*

Proof. First note that

$$N_d\Big(\exp\Big(i\pi j/2^d\Big)\Big) = \exp\left(i\pi j\sum_{k=0}^{2^d-1}(2k+1)/2^d\right) = \exp\Big(i\pi j 4^d/2^d\Big) = 1$$

Second recall that polynomial $\Phi_d(x) = x^{2^d}+1$ can be written as

$$\Phi_d(x) = \prod_{k=0}^{2^d-1}\Big(x-\exp\Big(i(2k+1)\pi/2^d\Big)\Big)$$

because each of $\exp\Big(i(2k+1)\pi/2^d\Big)$ for $k=0,\dots,2^{d-1}$ is a root of $\Phi_d(x)$. Expression for $N_d\Big(1-\exp\Big(i\pi(2j+1)/2^d\Big)\Big)$ coincides with the expression for $\Phi_d(1)=2$. □

Using above proposition and properties of v_2 we find that for odd k

$$v_2\Big(\sin(\pi k/2^d)\Big) = v_2\Big(2\sin(\pi k/2^d)\Big) - 1 = v_2\Big(\exp(i\pi k/2^d) - \exp(-i\pi k/2^d)\Big) - 1 = \tag{42}$$

$$= v_2\Big(1-\exp(\pi k/2^{d-1})\Big) - 1 = 1/2^{d-1} - 1 \tag{43}$$

Similar calculation shows that $v_2\Big(\cos(\pi k/2^d)\Big) = 1/2^{d-1}-1$.

In Proposition A.19 we saw that N_d takes integer values for two elements of a ring of cyclotomic integers

$$\mathbb{Z}\Big[\exp(i\pi/2^d)\Big] = \left\{\sum_{j=0}^{2^d-1} a_j\exp(i\pi j/2^d):\ \text{where } a_j \text{ are integers}\right\}$$

This is true more generally

Proposition A.20. *Let x be and element of $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$, then $N_d(x)$ is an integer. If x' is an element of*

$$\mathcal{R}_d = \mathbb{Z}\Big[\exp(i\pi/2^d), 1/2\Big] = \left\{\frac{1}{2^k}\sum_{j=0}^{2^d-1} a_j\exp(i\pi j/2^d):\ \textit{where } a_j, k \textit{ are integers}\right\},$$

then $N_d(x') = a/2^K$ for integers a, K.

Proof. Consider the case when x is from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$. The result follows from a proof technique similar to the proof of rationality of $N_d(x)$, when x is from $\mathbb{Q}\Big(\exp(i\pi/2^d)\Big)$ in Proposition A.18.

The second case follows from representing $x' = x/2^k$ for some x from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ and some integer k. Next we notice that by properties of N_d from Proposition A.18 of $N_d(x') = N_d(x)/2^K$ for $K = 2^d k$. □

The following proposition gives a necessary condition for an element of $\mathcal{R}_d$ to be equal to ± 1, in terms of v_2. This is the key to the proof of the fact that dyadic monotone μ_2 is positive and is zero if and only if the corresponding state is the stabilizer state.

Proposition A.21. *Let x be an element of*

such that for all odd k, $|\sigma_k(x)| \leq 1$, then $v_2(x) \leq 0$ and the equality is achieved if and only if $x = \pm 1$.

Proof. Let us first show that $v_2(x)$ is non-positive. Condition $|\sigma_k(x)| \leq 1$ implies that $N_d(x) \leq 1$. Because x is an element of $\mathcal{R}_d$ it can be written as $z/2^k$ for z from

Following the proof Proposition A.18 of rationality of N_d, one can show that $N_d(z) = n$ is an integer and therefore $N_d(x) = n/2^K$ and has absolute value less or equal to 1. For any number of the form $n/2^K$ with absolute value less or equal to 1 value of v_2 is non-positive and v_2 is zero if an only if $n/2^K = \pm 1$. We see that $v_2(x)$ is non-positive and is zero if and only if $N_d(x) = \pm 1$.

Let us now show that $v_2(x)$ equal zero implies that $x = \pm 1$. We have already shown that $N_d(x) = \pm 1$. We also have condition that $|\sigma_k(x)| \leq 1$ for all k. The only way $N_d(x)$ can be equal to ± 1 is if $\sigma_1(x) = x = 1$ which conclude the proof. □

A.9.3 Inequality $v_2(x+y) \geq \min(v_2(x), v_2(y))$

Recall, that in Section 6.1 the inequality $v_2(x+y) \geq \min(v_2(x), v_2(y))$ for rational x, y was first established for integer x and y and then extended to rationals by using additivity of v_2. We will follow the same strategy in the general case and introduce cyclotomic integers:

$$\mathbb{Z}\Big[\exp(i\pi/2^d)\Big] = \left\{ \sum_{j=0}^{2^d-1} a_j \exp(i\pi j/2^d) : \text{ where } a_j \text{ are integers} \right\}$$

Indeed, for arbitrary x, y from $\mathbb{Q}\Big(\exp(i\pi/2^d)\Big)$ there always exist an integer C, such that $x' = Cx, y' = Cy$ are both cyclotomic integers from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$. The general inequality easily follows from the inequality for cyclotimic integers:

$$v_2(x+y) = v_2(x'+y') + v_2(1/C) \geq \min(v_2(x'), v_2(y')) + v_2(1/C) = \min(v_2(x'/C), v_2(y'/C))$$

To complete the proof of the inequality we need the proposition below. Once this proposition is established, we can follow the same proof idea as for the rational version of the inequality in Section 6.1 with 2 replaced by $1 - \exp(i\pi/2^d)$.

Proposition A.22. *Let x be an element of $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$, then $v_2(x) \geq 0$. Moreover, for $k = 2^d v_2(x)$, x can be written as $x'(1 - \exp(i\pi/2^d))^k$ for x' from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ such that $v_2(x') = 0$.*

Proof. Let us denote $\alpha_d = (1 - \exp(i\pi/2^d))$ and choose k to be the biggest power of α_d that divides [6] x. We can write $x = \alpha_d^k x'$ such that x' is from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ such that α_d does not

[6] For x, y from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$, we say that x divides y if there exist r from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ such that $y = rx$.

divide x'. It remains to show that 2 does not divide $N_d(x')$, because this will establish that $v_2(x') = 0$ and $v_2(x) = kv_a(\alpha_d) = k/2^d$.

Let us show that 2 does not divide $N_d(x')$. Recall that $N_d(x')$ is an integer for any x' from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ according to Proposition A.20. Suppose now that 2 divides $N_d(x')$. This implies that α_d divides $N_d(x')$. Because α_d is prime according to Proposition A.23, α_d must divide $\sigma_{2k+1}(x')$ for some k. There exist j such that $(2j+1)(2k+1) \mathrm{mod}\, 2^{d+1} = 1$ and $\sigma_{2j+1}(\sigma_{2k+1}(x)) = x$ for such j. Therefore $\sigma_{2j+1}(\alpha_d)$ divides x'. However, α_d divides $\sigma_{2j+1}(\alpha_d)$ according to Proposition A.24 and therefore α_d divides x' which is a contradiction. □

Proposition A.23. *$\alpha_d = (1 - \exp(i\pi/2^d))$ is a prime element of $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$. That is, for x or y from $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$, if α_d divides xy then α_d divides x or y.*

Proof. The result follows from the fact that $N_d(\alpha_d) = 2$ is a prime number and the fact that every element of ring of integers of a number field with a prime norm is a prime element of the ring of integers. □

Proposition A.24. *Number $u_j = \Big(1 - \exp\Big(i\pi(2j-1)/2^d\Big)\Big)/\Big(1 - \exp\Big(i\pi/2^d\Big)\Big)$ is a unit in $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$. In other words, u_j and u_j^{-1} are both in $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$.*

Proof. First, note that u_j is an element of $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ because the polynomial $1 - x^{2j-1}$ is divisible by $(1-x)$:

$$(1 - x^{2j-1})/(1-x) = \sum_{k=0}^{2j-2} x^k \implies u_j = \sum_{k=0}^{2j-2} \exp\Big(ik\pi/2^d\Big),$$

To show that the inverse of u_j is an element of $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$ we first find an integer j' such that $j'(2j-1) \equiv 1 \mod 2 \cdot 2^d$ by using the extended Euclidean algorithm and the fact that $(2j-1)$ and 2^{d+1} are coprime, then the inverse is

$$u_j^{-1} = \Big(1 - \exp\Big(i\pi(2j-1)j'\pi/2^d\Big)\Big)/\Big(1 - \exp\Big(i\pi(2j-1)/2^d\Big)\Big).$$

Again using that the polynomial $1 - x^{j'}$ divisible by the polynomial $1 - x$, we conclude that u_j^{-1} is an element of $\mathbb{Z}\Big[\exp(i\pi/2^d)\Big]$. □

Remark A.25. *All the ring of integers of number fields $\mathbb{Q}(\exp(i\pi/2^d))$ have unique ramified prime ideal $\mathfrak{p}_d$ with norm 2. Function v_2 is a $\mathfrak{p}$-adic valuation divided by 2^d. The re-normalization makes sure that v_2 is defined consistently for the whole family of nested fields*

$$\mathbb{Q} \subset \mathbb{Q}(i) \subset \mathbb{Q}(\exp(i\pi/2^2)) \subset \ldots \subset \mathbb{Q}(\exp(i\pi/2^d)) \subset \mathbb{Q}(\exp(i\pi/2^{d+1})) \subset \ldots.$$

All the properties of v_2 follow from the properties of $\mathfrak{p}$-adic valuations and the fact that $\mathfrak{p}_d = \mathfrak{p}_{d+1}^2$.

A.10 Some properties of dyadic monotone μ_2

In this appendix we prove properties of the dyadic monotone that use slightly more advanced techniques from number theory introduced in Appendix A.9.

Proposition A.26. *Let $|\psi\rangle$ be a state with entries in $\mathcal{R}_d$, then $\mu_2|\psi\rangle \geq 0$ and the equality achieved if and only if $|\psi\rangle$ is a stabilizer state. In addition, for every Pauli operator P, if the expectation $\langle\psi|P|\psi\rangle \neq 0$, then $v_2(\langle\psi|P|\psi\rangle) \leq 0$.*

Proof. Consider Pauli P expectation $\alpha = \langle\psi|P|\psi\rangle$. Because P has eigenvalues ± 1, $|\langle\psi|P|\psi\rangle| \leq 1$. Consider now $\sigma_k(\alpha)$. Because σ_k respects addition, multiplication and commutes with complex conjugation according to Proposition A.17, α_k can be written as expectation $\langle\psi_k|P_k|\psi_k\rangle$ where $|\psi_k\rangle$ is the state obtained from $|\psi\rangle$ by applying σ_k element-wise and P_k some other Pauli operator obtained from P by also applying σ_k element-wise. We conclude that $|\alpha_k| \leq 1$. Now using Proposition A.21 we conclude that $v_2(\alpha) \leq 0$ and the equality is achieved if and only if $\alpha = \pm 1$. This implies that μ_2 is always non-negative and equality is achieved if and only if all non-zero Pauli expectations of $|\psi\rangle$ are ± 1. This implies that $|\psi\rangle$ is a stabilizer state, similarly to the proof of Proposition 6.3. □

Next we show that μ_2 is non-increasing for a slightly more general class of measurement than Pauli measurements with outcome probabilities one half.

Proposition A.27. *Let $|\psi\rangle$ be a state with entries in $\mathcal{R}_d$. Let P be a multi-qubit Pauli observable and let $p = \langle\psi|I + P|\psi\rangle/2 > 0$ be a probability of measuring $+1$ eigenvalue of P. Suppose there exist global phase $e^{i\phi}$ such that $|\psi_+\rangle = e^{i\phi}\frac{I+P}{\sqrt{p}}|\psi\rangle$ is the state with entries in $\mathcal{R}_d$, then $\mu_2|\psi_+\rangle \leq \mu_2|\psi\rangle$.*

Proof. We assume that $\langle\psi|P|\psi\rangle \neq 0$, because equality to zero case corresponds to $p = 1/2$ and covered by Proposition 6.5.

Consider Pauli matrix Q and corresponding expectation $\alpha = \langle\psi_+|Q|\psi_+\rangle$. If P and Q anti-commute, the expectation α is zero, because $(I + P)Q(I + P) = (I + P)(I - P)Q = 0$. It remains to consider the case when P and Q commute. In this case the expectation is

$$\alpha = \frac{\langle\psi|QP + P|\psi\rangle}{1 + \langle\psi|P|\psi\rangle}$$

Using multiplicative property of v_2 and inequality $v_2(a + b) \geq \min(v_2(a), v_2(b))$:

$$v_2(\alpha) \geq \min(v_2(\langle\psi|QP|\psi\rangle), v_2(\langle\psi|P|\psi\rangle)) - v_2(1 + \langle\psi|P|\psi\rangle)$$

Recall, that by definition of μ_2, $v_2(\langle\psi|P'|\psi\rangle) \geq -\mu_2|\psi\rangle$ for any Pauli P' including P and PQ. It remains to show that $v_2(1 + \langle\psi|P|\psi\rangle)$ is non-positive.

$$v_2(\langle\psi|P|\psi\rangle) = v_2(\langle\psi|P|\psi\rangle + 1 - 1) \geq \min(v_2(1 + \langle\psi|P|\psi\rangle), 0)$$

We have shown above in Proposition A.26 that $v_2(\langle\psi|P|\psi\rangle)$ is non-positive when the expectation $\langle\psi|P|\psi\rangle$ is non-zero. We see that $v_2(1 + \langle\psi|P|\psi\rangle) = \langle\psi|P|\psi\rangle \leq 0$.

We have shown that for arbitrary Pauli matrix Q, $-v_2(\langle\psi_+|Q|\psi_+\rangle) \leq \mu_2|\psi\rangle$. Inequality $\mu_2|\psi_+\rangle \leq \mu_2|\psi\rangle$ follows from the definition of μ_2. □

One might wonder if above result holds for two or more post-selected Pauli measurements. Below we provide an example showing that post-selecting on two commuting Pauli measurements can increase value of μ_2:

$$|\Psi\rangle = (1, -i, -5i-2, -i, -2i+1, -i-2, -i, -i+2, 1, i, i+2, i, 1, i, i+2, i)/8$$

By direct computation one can check that $\mu_2|\Psi\rangle = 3$. Post-selecting on $+1$ outcome for observables Z_1, Z_2 results in the state $(1, -i, -5i-2)/4\sqrt{2}$ with the value of μ_2 equal to 4.

$\|\psi\rangle$	$\mu_2(\|\psi\rangle)$
$\left\|\sqrt{T}\right\rangle$	3/4
$\|T\rangle$	1/2
$\|CS\rangle$	1
$\|CCS\rangle$	2
$\|C^3S\rangle$	3
$\|CCZ\rangle$	1

$\|\psi\rangle$	$\mu_2(\|\psi\rangle)$
$\|C^3Z\rangle$	2
$\|C^4Z\rangle$	3
$\|CCZ_{123,145}\rangle$	5
$\|W_3\rangle$	1
$\|W_4\rangle$	2
$\|W_5\rangle$	2

Table 5: Exact expressions for the dyadic monotone of the states used in Table 1 and Table 2.

A.10.1 Connections between the dyadic monotone and maximum denominator exponent

Here we show that the maximum denominator exponent of a unitary, which was used in [19] for the exact synthesis and canonical form of Clifford-cyclotomic gate-sets, is proportional to the dyadic monotone of the corresponding Choi state. Recall, that the Choi state's density matrix can be written as

$$\frac{1}{4^n} \sum_{P \in I,X,Y,Z^{\otimes n}} UPU^\dagger \otimes P.$$

The Pauli spectrum of the Choi state consists of values

$$\left| \mathrm{Tr}\left((Q \otimes Q') \sum_{P \in \{I,X,Y,Z\}^{\otimes n}} UPU^\dagger \otimes P \right) \right| / 4^n,$$

for all possible Pauli matrices Q, Q'. Taking into account that $\mathrm{Tr}(PQ) = \delta_{P,Q} \cdot 2^n$, we see that the Pauli spectrum of the Choi state is the set:

$$\left\{ \left| \mathrm{Tr}\left(QUPU^\dagger\right) \right| / 2^n : P, Q \in \{I, X, Y, Z\}^{\otimes n} \right\},$$

which is exactly the set of all entries of U in the channel representation [24] used to compute the maximum denominator exponent of U. Finally, maximum denominator exponent of the entry is proportional to minus the normalized p-adic valuation v_2.

A.11 General lower bounds for approximate unitary synthesis

The goal of this section is to generalize Proposition 5.8, from the tensor power of $|T\rangle$ state to arbitrary resource state for which dyadic monotone is defined. This leads to a lower bound on single qubit unitary approximation which generalizes Lemma 5.6, Lemma 5.7, Lemma 5.9. Finally we show a version of Theorem 5.2 involving T and $\sqrt{T}$ states.

Lemma A.28. *Let $d \geq 1$ and $|\Psi\rangle$ be a state with entries in $\mathcal{R}_d = \mathbb{Z}\left[\exp(i\pi/2^d), 1/2\right]$ when written in computational basis. Let $\{P_1, \ldots, P_m\}$ be independent commuting Pauli operators and let the probability of joint measurement of $+1$ eignevalue of $\{P_1, \ldots, P_m\}$ on $|\Psi\rangle$ be*

$$p = \frac{1}{2^m} \sum_{P \in \langle P_1, \ldots, P_m \rangle} \langle \Psi | P | \Psi \rangle. \tag{44}$$

If the value of p is non-zero, then $\log_2 p \geq -2^{d-1}(m + \mu_2|\Psi\rangle)$.

Proof. The proof consists of two steps. First we lower bound p by $\sqrt{N_d(p)}$. Second, we observe that $N_d(p)$ is given by ratio $a/2^K$ for some odd integer a and non-negative integer K (Proposition A.20) and upper bound K in terms of $\mu_2|\Psi\rangle$.

Let us first recall the expression for N_d:

$$N_d(p) = \prod_{k=0}^{2^d-1} \sigma_{2k+1}(p)$$

Using properties of map σ_k from Proposition A.17, expression for N_d can be rewritten as:

$$N_d(p) = \prod_{k=0}^{2^{d-1}-1} \sigma_{2k+1}(p)\sigma_{-(2k+1)}(p) = \left(\prod_{k=0}^{2^{d-1}-1} \sigma_{2k+1}(p) \right)^2$$

Above we used qualities $\sigma_{-1}(x) = x^*$, $\sigma_{kj}(x) = \sigma_k(\sigma_j(x))$ with $j = -1$, and took into account that p is a real number. To lower bound p in terms of $N_d(p)$ it remains to notice that $\sigma_k(p)$ is equal to:

$$\frac{1}{2^m} \sum_{P \in \langle P'_1, \ldots, P'_m \rangle} \langle \Psi_k | P | \Psi_k \rangle.$$

Where P'_j are some Pauli matrices obtained from P_j by element-wise application of σ_k and $|\Psi_k\rangle$ is a state obtained from $|\Psi\rangle$ by element-wise application of σ_k. Because $\sigma_k(p)$ is probability of some measurement it must be less than 1. We see that $p \geq \sqrt{N_d(p)}$.

Recall that we have defined $v_2(p)$ as

$$v_2(N_d(p))/2^d.$$

Therefore, $K = -2^d v_2(p)$. Now using inequality $v_2(a+b) \geq \min(v_2(a), v_2(b))$, we have:

$$v_2(p) \geq -m + \min_{P \in \langle P'_1, \ldots, P'_m \rangle} v_2(\langle \Psi_k | P | \Psi_k \rangle) \geq -m - \mu_2|\Psi\rangle.$$

We conclude that $K \leq 2^d \mu_2(|\Psi\rangle)$ and therefore $p \geq \sqrt{2}^{-2^d(m+\mu_2|\Psi\rangle)}$. □

Next we follow the proof technique of Lemma 5.6 and use the proposition above and establish the following result:

Lemma A.29. *Let $d \geq 1$ and $|\Psi\rangle$ be a state with entries in $\mathcal{R}_d = \mathbb{Z}\left[\exp(i\pi/2^d), 1/2\right]$ when written in computational basis. Suppose that qubit state $|\psi\rangle$ is approximated to within trace distance ε using stabilizer operations and has post-selection with input $|\Psi\rangle$. Inequalities $\varepsilon < 1/8$ and $\varepsilon < |\langle\psi|0\rangle|^2 < 3\varepsilon$ imply $\mu_2|\Psi\rangle + \nu|\Psi\rangle \geq \frac{1}{2^{d-1}} \log_2(1/\varepsilon) - \frac{1}{2^{d-2}}$.*

Proof. First note that we can assume without loss of generality that $\nu|\Psi\rangle$ is equal to the number of of qubits on which $|\Psi\rangle$ is defined. The rest of the proof is similar to Lemma 5.6. □

For example, above result implies a lower bound when approximating using N_T copies of $|T\rangle$ state, and $N_{\sqrt{T}}$ copies of $|\sqrt{T}\rangle$ and $|\sqrt{T}^3\rangle$ states:

$$N_{\sqrt{T}} + \frac{6}{7} N_T \geq \frac{1}{7} \log_2(1/\varepsilon) - \frac{1}{14}.$$

Above inequality leads to the following generalization of Theorem 5.2.

Theorem A.30. *Consider a protocol that uses $\mathcal{N}_{\sqrt{T}}(U, \varepsilon)$ copies of $|\sqrt{T}\rangle$ and $|\sqrt{T}^3\rangle$ states, $\mathcal{N}_T(U, \varepsilon)$ copies of $|T\rangle$ state and stabilizer operations to approximate a one-qubit unitary U to within precision ε (measured by the diamond norm). For any positive $C > 1$ and $\varepsilon < 1/(2^8 C)$ there exists a unitary U such that the following inequality must hold*

$$\mathcal{N}_{\sqrt{T}}(U, \varepsilon) + \frac{6}{7} \mathcal{N}_T(U, \varepsilon) \geq \frac{1}{14} \log_2(1/\varepsilon) - \frac{1}{14} \log_2(C) - \frac{3}{14}$$

with probability at least $(C-1)/C$. In particular, this is the case for all unitaries U such that $2\sqrt{C\varepsilon} \leq |\langle 0|U|1\rangle|^2 \leq 6\sqrt{C\varepsilon}$.