Universal, fault-tolerant gate sets

Recall: and efficient ☺

① The Clifford group "is" FT but classically simulable

② Transversal gates can't get us to universality

How can we proceed?

We proceed by relaxing our definition of universality and then hoping we can figure out how to scrape by with a non-transversal gate...

Explicitly, we will show that

$$C_n \cup \{T\}$$

suffices with this relaxed notion, and then give fault-tolerant constructions of the T gate

Approximating unitaries

Given a finite set of gates G, $\langle G \rangle$ (circuits over G) is countable, but $U(n)$ ($n \times n$ unitary matrices) is uncountable

Since we can't hope to implement every unitary exactly, we need to have a notion of approximation error.

(Approximation error)
Given a unitary U, the error when U is approximated by V is (operator norm)

$$\|U - V\| = \max_{|\psi\rangle} \|(U - V)|\psi\rangle\|$$

Note: $\|\cdot\|$ is a norm, so $\|A + B\| \leq \|A\| + \|B\|$
also $\|UA\| = \|A\|$, U unitary

We use the operator norm because the difference in measurement probabilities after apply U or V is bounded by $2\|U - V\|$

Def'n

A finite set of gates G, G is universal for quantum computation if for any unitary U and error rate $\epsilon > 0$, there exists a circuit V over G such that

$$\|U - V\| \leq \epsilon$$

Approximation of single-qubit unitaries suffices

We first show that it suffices to only approximate single-qubit unitaries assuming CNOT is given. For this it is helpful to note an important fact about how errors in a circuit combine

(Approximation errors add) — important!

Observe that if U_1 and U_2 are approximated to error ϵ by V_1 and V_2, resp., then

$$\begin{aligned} \|U_1U_2 - V_1V_2\| &= \|U_1U_2 - U_1V_2 + U_1V_2 - V_1V_2\| \\ &= \|U_1(U_2-V_2) + (U_1-V_1)V_2\| \\ &\le \|U_2-V_2\| + \|U_1-V_1\| \\ &= 2\epsilon \end{aligned}$$

The preceding point is useful in compilation because it tells us that if we want to approximate a circuit with d gates to error ϵ, we need to approximate each gate to error ϵ/d

Prop.

Given a set of single-qubit gates G which is universal for 1-qubit unitaries, $G \cup \{CNOT\}$ is universal for n-qubit unitaries.

Pf.

Write a unitary over CNOT + single qubit rotations, then approximate each to error ϵ/d

Universal sets of quantum gates (approximately) (single-qubit)

Thm.

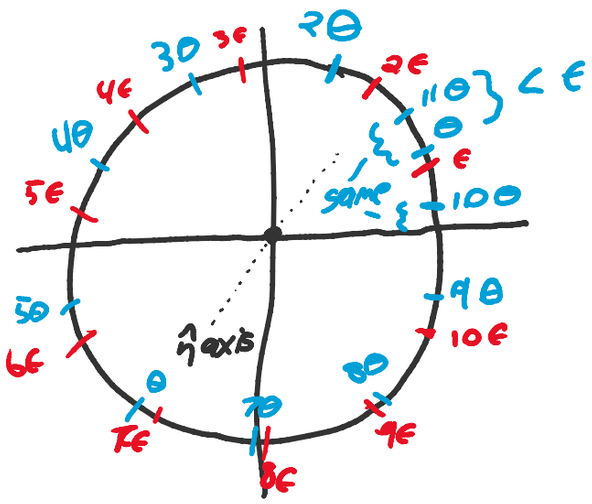
$\{H, T\}$ is universal for single-qubit unitaries

Pf. Sketch

Let $R_1 = THTH$, $R_2 = HR_1H = HTHT$.

It can be shown that R_1 & R_2 are rotations of an angle θ which is an irrational multiple of 2π (i.e. $k\theta \neq 2\pi \ \forall k \in \mathbb{Z}$) around non-parallel axes $\hat{n}, \hat{m}$ resp.

Since θ is irrational, we can rotate by infinitely many angles around either axis. Intuitively, say we want to approximate $R_{\hat{n}}(\alpha)$ to accuracy ϵ.



$\therefore \|R_{\hat{n}}(10\theta) - I\| < \epsilon$

so multiples of $R_{\hat{n}}(10\theta)$ fill the space with points $< \epsilon$ distance apart (called an ϵ-net)

So write any single-qubit unitary as

$$R_{\hat{n}}(\beta) R_{\hat{m}}(\delta) R_{\hat{n}}(\gamma)$$

and approximate each ☺

Speed of approximation

Naively, this method generates an ϵ-net with circuits of length $O(1/\epsilon^2)$ - roughly $2\pi/\epsilon$ rotations to get within ϵ of I, then $2\pi/\epsilon$ multiples to cover the circle.

This sucks!

Thm. (Solovay-Kitaev) / proven unnecessary just last year!!!

Given a gate set G closed under inverses that is universal for single-qubit unitaries, a single-qubit unitary U can be approximated over G to ϵ error with at most $O(\log^c(1/\epsilon))$ gates from G

The Solovay-Kitaev theorem is fundamental to the possibility of FTQEC, as it states that we can use fault-tolerant gates with sub-polynomial overhead. In practice it is important for FTQEC implementations of

- The Quantum Fourier transform (basis of Shor + others)
- Hamiltonian simulation

both of which require many single-qubit gate approximations.

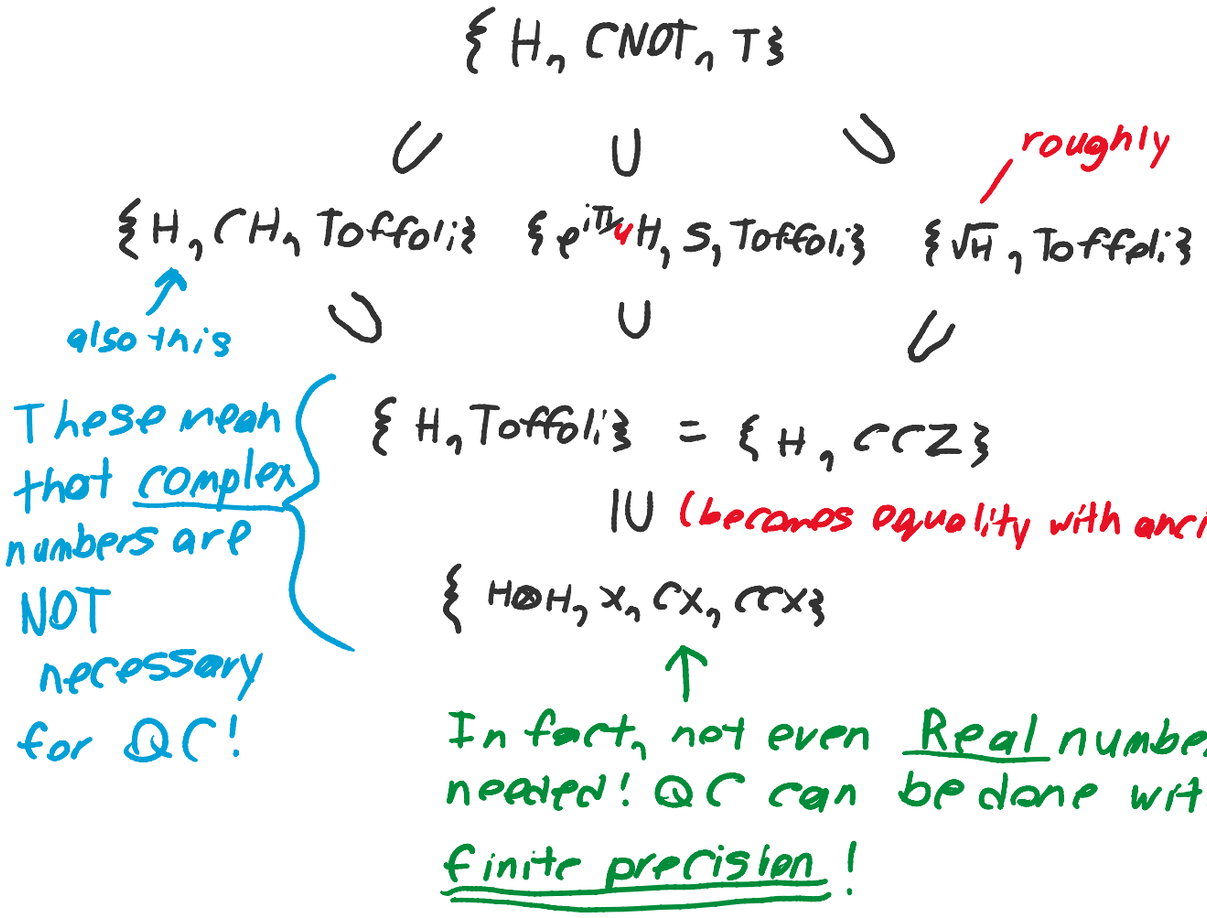
Other universal gate sets

Thm.

The Clifford group plus any non-Clifford unitary is approximately universal

(The proof of this theorem brings up interesting connections to lattice theory. Certain gate sets correspond to automorphisms of well-known lattices. In the Clifford case, this is the Barnes-Wall lattice, and the most commonly cited proof for this fact falls out of this connection)

The following are some approximately universal gate sets and their inclusions:



These mean that complex numbers are NOT necessary for QC!

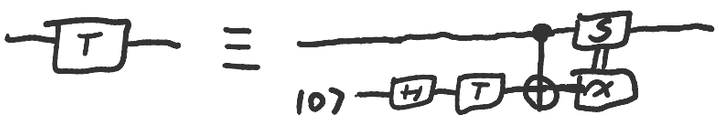
In fact, not even Real numbers are needed! QC can be done with finite precision!

Non-transversal T gates

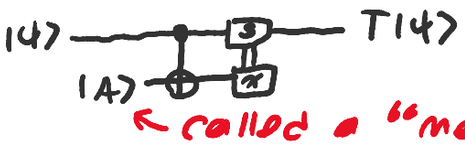
We know we need at least one gate (T) which is NOT transversal to get a universal gate set. Fortunately there exist other more costly methods for certain gates, namely Gate teleportation + magic states.

(Teleporting the T gate)

The following circuits are equivalent up to global phase



If we write $|A\rangle = TH|0\rangle$, then up to global phase



↖ called a "magic state"

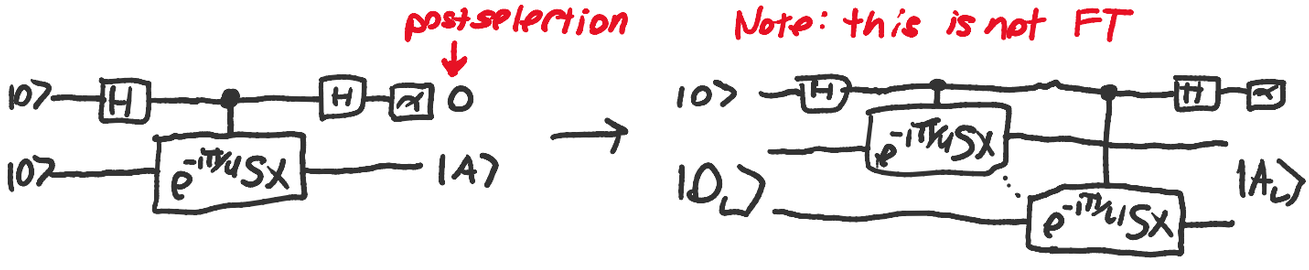
Upshot: if we can prepare the state $|A\rangle$, perform Clifford gates, and measure fault-tolerantly, we can perform the T gate fault-tolerantly!

($|A\rangle$ states circa early 2000's)

The $|A\rangle$ state can be prepared fault-tolerantly by using transversal Cliffords. In particular,

$$e^{-i\pi/4 SX}|A\rangle = |A\rangle$$

So we can project a state onto the $|A\rangle$ state by measuring $e^{-i\pi/4 SX}$, which is Clifford. Explicitly,



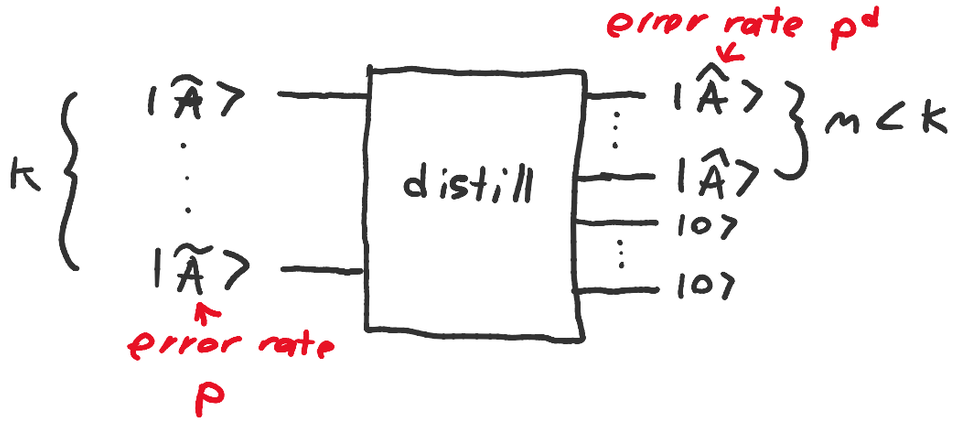
State distillation

In many modern codes, the old-style $|A\rangle$ state preparation is very inefficient.

Bravyi & Kitaev (2005) gave an alternative method which scales better called **magic state distillation**

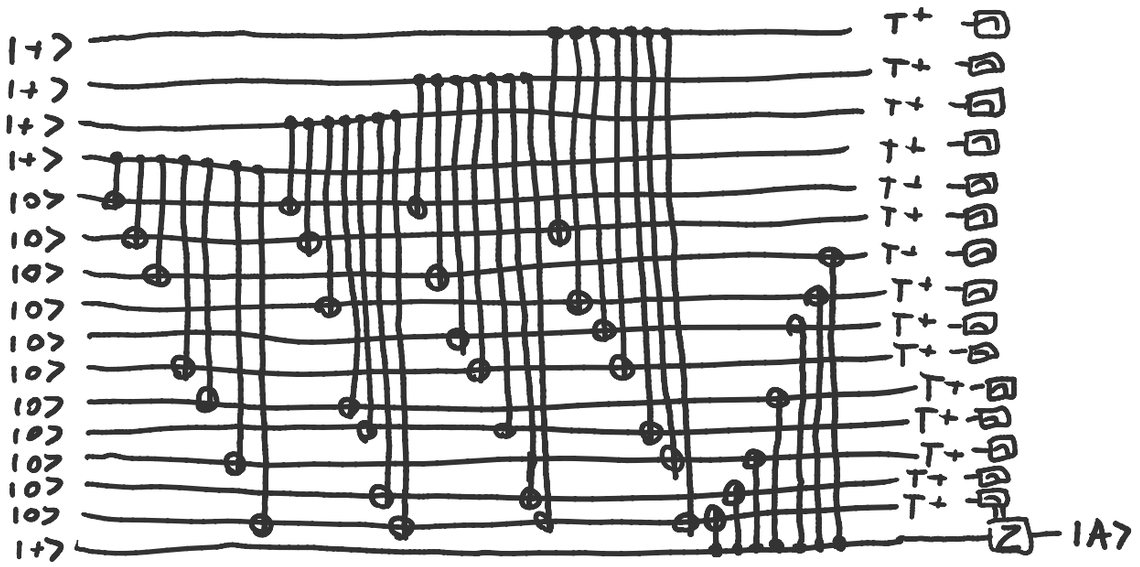
(magic state distillation)

The basic idea of magic state distillation is to take many **faulty** copies of the magic state, and distill them into **one or a few** better copies



(15-1 protocol)

The 15-1 protocol uses a (punctured) Reed-Muller code for which the T gate is transversal to get **cubic** error reduction (i.e. $p \xrightarrow{\text{distill}} O(p^3)$). The details are unimportant, but the **time and space** it uses is!



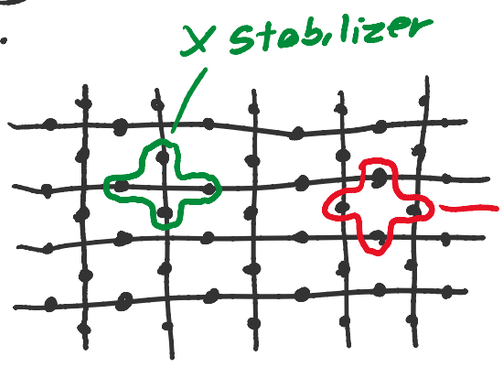
Topological codes

Since the 2010's, Kitaev's **surface** and more generally **topological codes** have, with the help of state distillation, supplanted the leading CSS code (**Steane code**).

(Surface code)

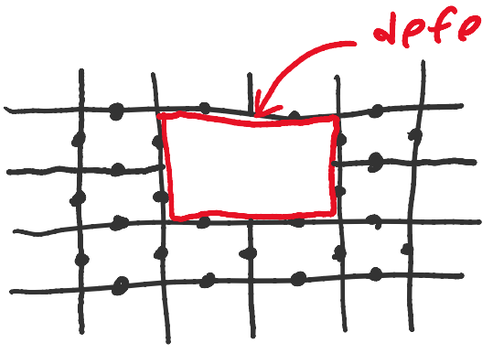
The surface code is a stabilizer code laid out on a 2-D lattice. The generators are importantly **local**, corresponding to X's around each vertex and Z's around each face.

(i.e. only neighbouring interactions needed)



actually depends on boundaries..

On a square lattice, this gives **no logical qubits** so logical qubits are made by cutting **defects** from the stabilizer generators



Defects can then be **moved** with only measurement of the stabilizers

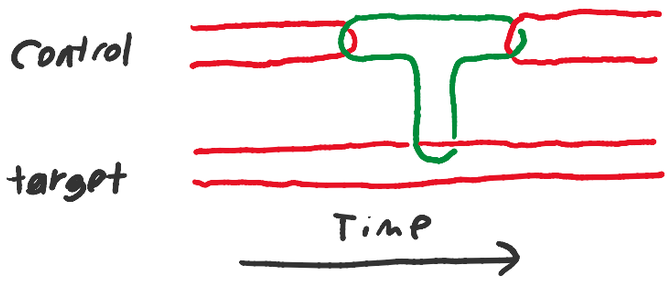
Fault-tolerant topological gates

Fault-tolerant gates in a topological QECC have different flavours, but generally **Clifford gates are CHEAP, T gates are EXPENSIVE.** In particular, Clifford gates are generally performed by moving defects, while T gates use magic state distillation.

(Braided surface code – until 2020's)

The first main FTQEC proposal based on surface codes. Defects (primal or dual) are in pairs, implement gates by braiding the defects

e.g. CNOT:



"only the topology matters"

↓

time could be vertical...

spawned compilation work looking at topological compaction

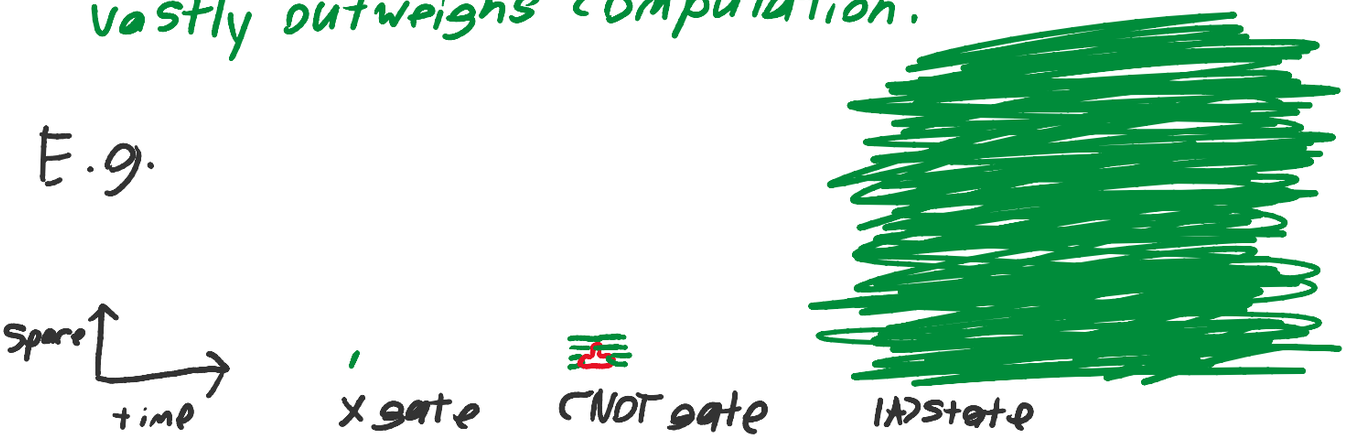
(Lattice surgery – the new kid)

Leading proposals now use lattice surgery. Main idea is to compute (Cliffords) by merging and splitting defects.

Relative volumes

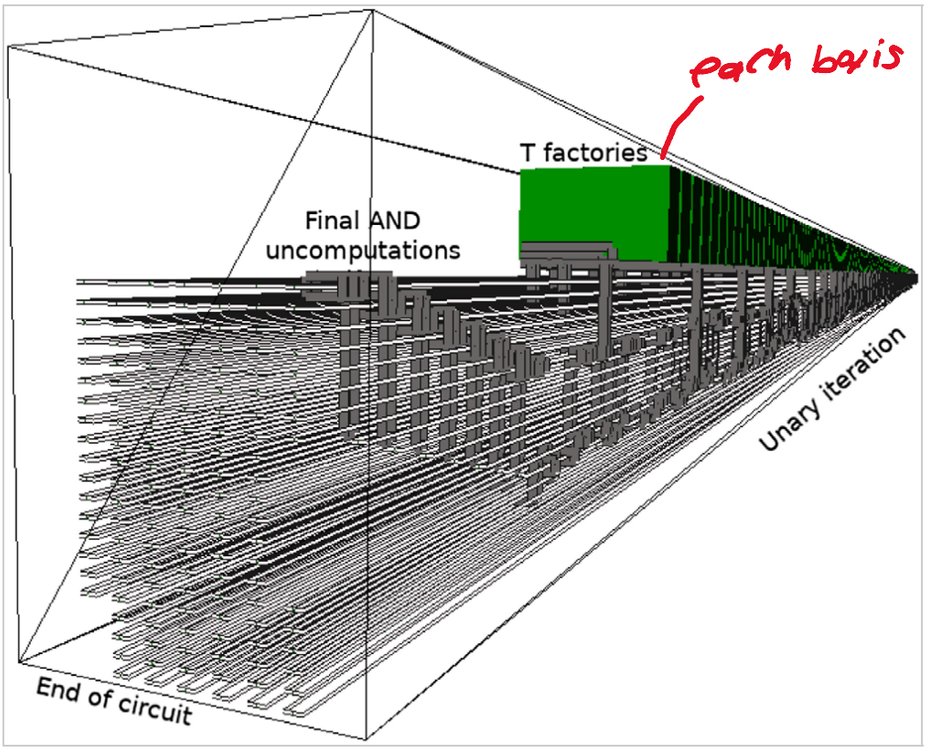
Computation in topological codes is typically measured in volume. By volume, state distillation vastly outweighs computation!

E.g.



Result: The T gate is ORDERS of MAGNITUDE more expensive in surface codes than Clifford gates

E.g. from arXiv:1805.03662, Hamiltonian simulation with linear T-complexity



Further notes about resources

In FTQEC, space and time usage are linked and increases in either of the logical layer tend to compound and affect both physically.

Thm. (Threshold theorem)

A threshold theorem for a FTQEC method states that if physical error rates are below p_{th}, then a computation can be performed to arbitrary precision by increasing the amount of error correction.

Ex.

p_{th} for the Steane code is proven to be 10^{-5}

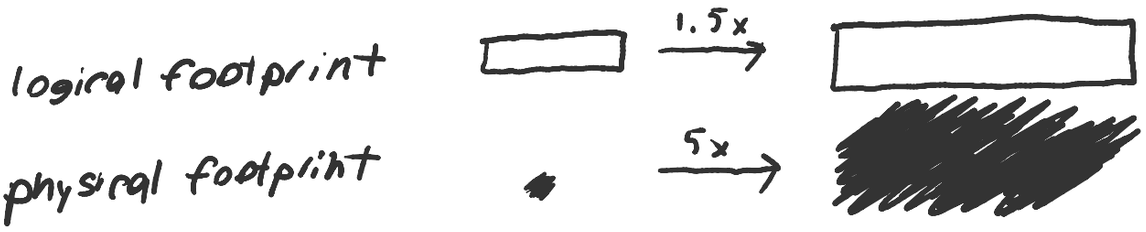
p_{th} for surface code is estimated at 10^{-2}

(code distance)

The distance of a QECC is a measure of how resilient to errors it is. Increasing distance requires (typically) increasing physical qubits per logical qubit.

To understand how resource usage compounds in FTQEC, imagine that one circuit C_1 takes twice as long (i.e. depth) as another circuit C_2. When applying error correction, C_1 has more chance for errors to accumulate, and so may take more physical space as well. If C_1 has twice as many T gates as well, each needs a magic state which is 50% more precise, leading to larger & longer state preparation circuits.

e.g. 15-1 becomes 15^2-1...



Beyond Solovay-Kitaev

Since gate approximation produces long circuits with many T gates, optimizing single qubit gate approximations in Clifford+T is a fundamental problem in circuit compilation

For many years, the best-known constructive (i.e. produces a circuit) method (Dawson & Nielsen 2005) gave $C > 3$ for $\{H, T\}$. An information-theoretic lower bound of $3\log(1/\epsilon)$ suggested better was possible but wasn't met until the Number theoretic method was developed in 2013-2014. To this day, for $\{H, T\}$ the best known (pure circuit) algorithm is Ross & Selinger (2014)'s Grid-Synth which produces approximations of length

$$3\log(1/\epsilon) + O(\log\log 1/\epsilon)$$

If measurement is allowed, the PQF method of Bocharov, Roetteler & Svore (2015) reduces this to an expected length

$$\log(1/\epsilon) + O(\log\log 1/\epsilon)$$