Distillation of Non-Stabilizer States for Universal Quantum Computation

Guillaume Duclos-Cianci[1] and Krysta M. Svore[2]

1 **Départment de Physique, Université de Sherbrooke**
Sherbrooke, Québec, J1K 2R1 (Canada)
`Guillaume.Duclos-Cianci@USherbrooke.ca`
2 **Quantum Architectures and Computation Group, Microsoft Research**
Redmond, WA 98052 (USA)
`ksvore@microsoft.com`

─── **Abstract** ───

Magic state distillation is a fundamental technique for realizing fault-tolerant universal quantum computing, and produces high-fidelity Clifford eigenstates, called magic states, which can be used to implement the non-Clifford $\pi/8$ gate. We propose an efficient protocol for distilling other *non-stabilizer* states that requires only Clifford operations, measurement, and magic states. One critical application of our protocol is efficiently and fault tolerantly implementing arbitrary, non-Clifford, single-qubit rotations in average *constant* online circuit depth and polylogarithmic (in precision) offline resource cost, resulting in significant improvements over state-of-the-art decomposition techniques. Finally, we show that our protocol is robust to noise in the resource states.

1 Introduction

Given recent progress in quantum algorithms, quantum error correction, and quantum hardware, a *scalable* quantum computer is becoming closer and closer to reality. For many proposed quantum computer architectures, e.g., the surface-code model based on code deformation [1], Clifford operations, stabilizer-state preparations, and measurements can be implemented efficiently. However, these operations alone are not sufficient for quantum universality and can be simulated classically [2]. Magic state distillation [3, 4, 5, 6] produces Clifford eigenstates, which in turn can be used to realize a non-Clifford operation, e.g., the single-qubit $\pi/8$ gate, T.

In this paper, we present an efficient protocol for distilling other *non-stabilizer* states. Our protocol uses only $|H\rangle$-type magic resource states, Clifford operations, and measurements, and is robust to noise in the resource states. One notable application of our protocol is producing an arbitrary single-qubit, fault-tolerant unitary operation. Most commonly, a single-qubit unitary U is decomposed into a discrete set of gates, typically $\{H, T\}$, using Solovay-Kitaev decomposition [7, 8], which efficiently produces an approximate fault-tolerant implementation of U with circuit depth $\Theta(\log^c(1/\epsilon))$, where ϵ is the precision and c is around 3.97 [9, 8]. Remarkably, efficient decomposition algorithms have recently been proposed which lower c to 1 [10, 11]. Each T gate in the decomposed sequence requires a number of copies of a quantum magic state $|H\rangle$, dependent on the specific state distillation protocol and

$$|Z(\theta)\rangle \;-\boxed{X(Z)}-\boxed{\measuredangle}= \;|m\rangle$$
$$|\psi\rangle \;-\!\!\bullet\!\!-\qquad Z(X)((-1)^m\theta)|\psi\rangle$$
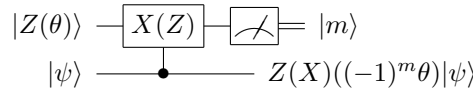
Figure 1 Circuit to rotate by angle $\pm\theta$ around the $Z(X)$-axis.

purity of the state [3, 4, 5, 6]. We show that our protocol requires only *constant* online circuit depth and fewer resources than state-of-the-art decomposition techniques. Our protocols may be useful for other applications as well.

2 Distilling Magic States and Implementing Rotations

We first review how to perform an arbitrary rotation about the Z-axis using a resource state. A state $|\psi\rangle$ is *magic* if we can "distill" a purer $|\psi\rangle$ state from a Clifford circuit applied to n noisy copies of $|\psi\rangle$. We focus on the $+1$ eigenstate of the Hadamard operation H, $|H\rangle = \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle$. We assume throughout that Clifford operations are perfect and resource states are arbitrarily pure. We can arbitrarily purify these states by applying a distillation protocol recursively [3, 4, 5, 6]. We concentrate on single-qubit states found in either the XZ- or XY-plane of the Bloch sphere; note that a state can be rotated from one plane to the other through application of the Clifford $HSHX$ operation.

Suppose we have states $|Z(\theta)\rangle = |0\rangle + \mathrm{e}^{i\theta}|1\rangle$ and $|\psi\rangle = a|0\rangle + b|1\rangle$. The circuit to implement a rotation around the Z-axis using $|Z(\theta)\rangle$ as a resource state is presented in Fig. 1. Upon measurement of the first qubit in the computational basis, we obtain either

$$\xrightarrow{m=0}\quad a|0\rangle + b\mathrm{e}^{i\theta}|1\rangle,\ \text{or}$$
$$\xrightarrow{m=1}\quad a\mathrm{e}^{i\theta}|0\rangle + b|1\rangle = a|0\rangle + b\mathrm{e}^{-i\theta}|1\rangle,$$

each with probability $1/2$. Thus, the rotation angle is randomly either θ or $-\theta$, up to global phase. An analogous circuit performs a rotation about the X-axis [1].

As an example, consider the XY-plane version of $|H\rangle$:

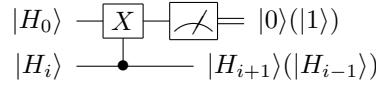$$|Z(\pi/4)\rangle = HSHX|H\rangle = |0\rangle + \mathrm{e}^{i\pi/4}|1\rangle.$$

Using the circuit in Fig. 1, we can implement a Z-rotation of angle $\pm\pi/4$, producing at random either the T gate or its adjoint, $T^\dagger$. We can deterministically *correct* the angle by applying the phase gate S: $ST^\dagger|\psi\rangle = T|\psi\rangle$. For general rotations, deterministic correction is not possible.

3 Distilling Other Non-Stabilizer States

We now present our protocol for producing other non-stabilizer states using a very simple two-qubit Clifford circuit and $|H\rangle$ states as an initial resource.

Consider the circuit of Fig. 2. One can easily verify that it measures the parity of the two input qubits and decodes the resulting state into the second qubit. Consider the two inputs to be $|H\rangle$ states and define $\theta_0 = \frac{\pi}{8}$ and $|H\rangle = |H_0\rangle = \cos\theta_0|0\rangle + \sin\theta_0|1\rangle$. Then upon application of the controlled-NOT gate $\Lambda(X)$,

$$|H_0\rangle|H_0\rangle \xrightarrow{\Lambda(X)} \cos^2\theta_0|00\rangle + \sin^2\theta_0|01\rangle + \cos\theta_0\sin\theta_0(|11\rangle + |10\rangle).$$

$$|H_0\rangle \quad \boxed{X} \quad \boxed{\measuredangle} \mathrel{=}\!\!\!= |0\rangle(|1\rangle)$$
$$|H_i\rangle \quad\longrightarrow\!\!\bullet\!\!\longrightarrow \quad |H_{i+1}\rangle(|H_{i-1}\rangle)$$

■ **Figure 2** Two-qubit circuit used to obtain new $|H_i\rangle$ states from initial resource states $|H_0\rangle$. Upon measuring the 0 (1) outcome, the output state is $|H_{i+1}\rangle$ ($|H_{i-1}\rangle$).

Upon measurement m of the first qubit, we have

$$\xrightarrow{m=0} \frac{\cos^2\theta_0\,|0\rangle + \sin^2\theta_0\,|1\rangle}{\cos^4\theta_0 + \sin^4\theta_0}, \text{ or } \xrightarrow{m=1} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

We define θ_1 such that

$$\cos\theta_1\,|0\rangle + \sin\theta_1\,|1\rangle = \frac{\cos\theta_0\,|0\rangle + \sin\theta_0\,|1\rangle}{\cos^4\theta_0 + \sin^4\theta_0},$$

from which we deduce $\cot\theta_1 = \cot^2\theta_0$. We define $|H_1\rangle = \cos\theta_1\,|0\rangle + \sin\theta_1\,|1\rangle$, a non-stabilizer state obtained from $|H\rangle$ states, Clifford operations, and measurements. If the measurement outcome is 1, then we obtain a stabilizer state and discard the output (see Fig. 2). The measurement outcomes occur with respective probabilities $p_0 = \cos^4\theta_0 + \sin^4\theta_0 = \frac{3}{4}$ and $p_1 = 1 - p_0 = \frac{1}{4}$.

We now recurse on this protocol using the non-stabilizer states produced by the previous round of the protocol as input to the circuit in Fig. 2. We define $|H_i\rangle = \cos\theta_i\,|0\rangle + \sin\theta_i\,|1\rangle$, where $\cot\theta_i = \cot^{i+1}\theta_0$. Using as input the previously produced $|H_i\rangle$ state and a new $|H_0\rangle$ state, we have

$$|H_0\rangle\,|H_i\rangle \xrightarrow{\Lambda(X)} \cos\theta_0\cos\theta_i\,|00\rangle + \sin\theta_0\sin\theta_i\,|01\rangle + \sin\theta_0\cos\theta_i\,|10\rangle + \cos\theta_0\sin\theta_i\,|11\rangle.$$

Upon measurement of the first qubit, we have

$$\xrightarrow{m=0} (\cos\theta'\,|0\rangle + \sin\theta'\,|1\rangle),$$
$$\xrightarrow{m=1} (\cos\theta''\,|0\rangle + \sin\theta''\,|1\rangle), \text{ where}$$
$$\cot\theta' = \cot\theta_i\cot\theta_0 = \cot^{i+2}\theta_0 = \cot\theta_{i+1},$$
$$\cot\theta'' = \cot\theta_i\tan\theta_0 = \cot^i\theta_0 = \cot\theta_{i-1}.$$

Thus, if we measure $m = 0$, we obtain the state $|H_{i+1}\rangle$ and if we measure $m = 1$, we obtain $|H_{i-1}\rangle$. The probability of measuring 0 is given by $p_{0,i} = \cos^2\theta_i\cos^2\theta_0 + \sin^2\theta_i\sin^2\theta_0$. Note that $\frac{3}{4} \le p_{0,i} < \cos^2\frac{\pi}{8} = 0.853\ldots$

We can view this recursive process as a semi-infinite random walk with biased non-homogeneous probabilities, as Fig. 3 illustrates. Every time a step is taken along this "ladder" of states, one $|H\rangle \equiv |H_0\rangle$ is consumed, except at the first step of the ladder when we require two $|H\rangle$ states; if $m = 1$ at the first node, we discard the output and start with two new $|H\rangle$ states.

4 A Denser Ladder

We can produce a denser ladder of states by using additional resource states $\left|\psi_0^{0,1,2}\right\rangle$. Consider the Clifford circuit of Fig. 4(a) that takes as input four $|H\rangle$ states. With probability $3(2 + \sqrt{2})/32 \approx 0.320$, the measurement outcome is 000 and the resulting state is $\left|\psi_0^0\right\rangle =$
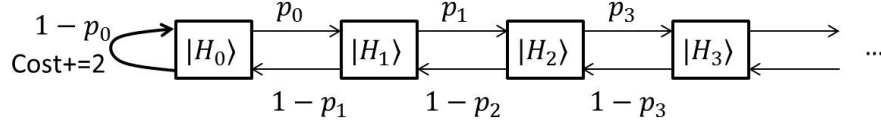
Figure 3 Obtaining non-stabilizer states from initial $|H\rangle$ states. Using $|H_i\rangle$ and $|H_0\rangle$ states probabilistically yields a $|H_{i-1}\rangle$ or $|H_{i+1}\rangle$ using the circuit of Fig. 2. Each ladder step costs one $|H_0\rangle$ state, except the first one which costs two.
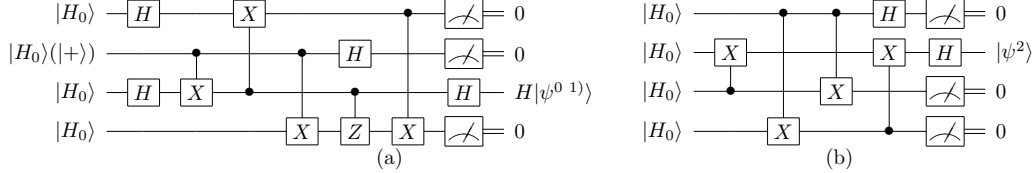
Figure 4 (a) Circuit to produce $|\psi_0^0\rangle$ ($|\psi_0^1\rangle$) states. (b) Circuit to produce $|\psi_0^2\rangle$ states.

$\cos\phi_0^0 |0\rangle + \sin\phi_0^0 |1\rangle$ with $\phi_0^0 = \frac{\pi}{2} - \cot^{-1}\left(\frac{2+3\sqrt{2}}{6+5\sqrt{2}}\right) \approx 0.446$. Otherwise the output is discarded. Since the probability of success is 0.320 and every trial consumes four copies of $|H_0\rangle$, the average cost to produce $|\psi_0^0\rangle$ is 12.50 $|H_0\rangle$ states.

Another interesting state is obtained using the same circuit with one input state replaced with a $|+\rangle$ state. Measurement 000 is obtained with probability $(6+\sqrt{2})/32 \approx 0.232$, resulting in the state $|\psi_0^1\rangle = \cos\phi_0^1 |0\rangle + \sin\phi_0^1 |1\rangle$ with $\phi_0^1 = \frac{\pi}{2} - \cot^{-1}\left(\frac{2\sqrt{2}}{3+\sqrt{2}}\right) \approx 0.570$. Since the probability of success is 0.232 and every trial consumes three $|H_0\rangle$ states, the average cost to produce $|\psi_0^1\rangle$ is 12.95 $|H_0\rangle$ states. Fig. 4(b) shows a circuit which produces the output state $|\psi_0^2\rangle = \cos\phi_0^2 |0\rangle + \sin\phi_0^2 |1\rangle$ with $\phi_0^2 = \frac{\pi}{2} - \cot^{-1}\left(\frac{7}{6\sqrt{2}}\right) \approx 0.690$, when measurement 000 is obtained (with probability $11/32 \approx 0.344$). The probability of success is 0.344 and the average cost to produce $|\psi_0^2\rangle$ is 11.64 $|H_0\rangle$ states.

Now we can use one of these non-stabilizer states as input to the circuit in Fig. 2 in place of the top $|H_0\rangle$ state. Begin with states $|\psi_0^i\rangle$ and $|H_0\rangle$. If $m = 1$, the state is discarded. Otherwise, we obtain $|\psi_1^i\rangle = \cos\phi_1^i |0\rangle + \sin\phi_1^i |1\rangle$, where $\cot\phi_1^i = \cot\phi_0^i \cot\theta_0$. As before, we define $\left|\psi_i^j\right\rangle = \cos\phi_i^j |0\rangle + \sin\phi_i^j |1\rangle$, where $\cot\phi_i^j = \cot\phi_0^j \cot^i\theta_0$. If we input states $\left|\psi_i^j\right\rangle$ and $|H_0\rangle$, we obtain

$$|H_0\rangle \left|\psi_i^j\right\rangle \xrightarrow{\Lambda(X)} \cos\theta_0\cos\phi_i^j |00\rangle + \sin\theta_0\sin\phi_i^j |01\rangle + \sin\theta_0\cos\phi_i^j |10\rangle + \cos\theta_0\sin\phi_i^j |11\rangle,$$

such that the output state is, depending on the measurement outcome,

$$\xrightarrow{m=0} \left|\psi_{i+1}^j\right\rangle, \qquad \text{or} \xrightarrow{m=1} \left|\psi_{i-1}^j\right\rangle.$$

Denser "ladders" of states can be obtained using $\left|\psi_0^{0,1,2}\right\rangle$ as inputs in place of the top $|H_0\rangle$ state.

5   Noisy states

A priori, noise in the $|H_0\rangle$ resource states could be amplified by the circuit in Fig. 2 and affect the purity of the $|H_i\rangle$ states. However, we show this is not the case. We measure the accuracy of the imperfect $|H_i\rangle$ states using the trace distance on states ρ and
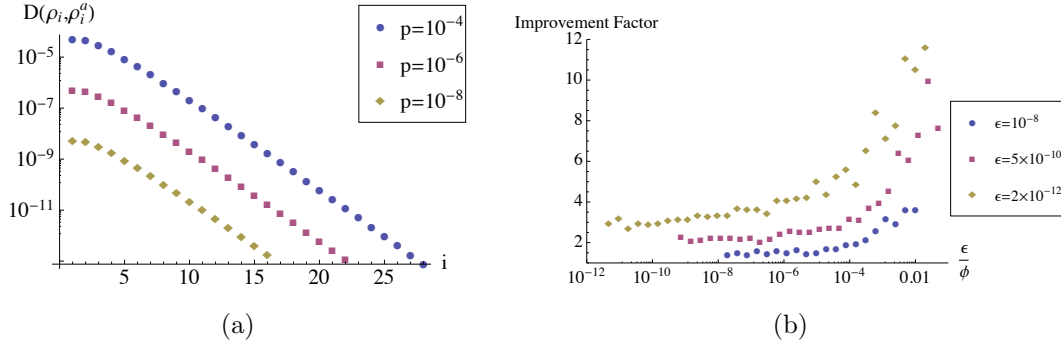
Figure 5 (a) Evolution of the trace distance between imperfect ρ_i^a and perfect $|H_i\rangle$ states with noise p. Exponential decay fits give $(2.08 * 10^{-3}) \times 2.31^{-i}$, $(1.63 * 10^{-5}) \times 2.28^{-i}$ and $(1.26 * 10^{-7}) \times 2.24^{-i}$ for the circle, square and diamond data set, respectively. (b) Improvement factor of the total offline cost using the noisiest $|H_0\rangle$ states to distill $|H_i\rangle$ states of precision ϵ as a function of the relative precision of the rotation ϵ/ϕ.

σ: $D(\rho, \sigma) = \text{tr}(\sqrt{(\rho - \sigma)^\dagger (\rho - \sigma)})/2$. We assume errors only occur on the $|H_0\rangle$ states. We numerically study three types of errors. For the first error, we assume that the mixed state, ρ_0^a, is on the line joining the center of the Bloch sphere and the the perfect state, i.e.,

$$\rho_0^a(p) \quad = \quad (1-p)|H_0\rangle\langle H_0| + p|-H_0\rangle\langle -H_0|,$$

where $|-H_0\rangle = \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle$ is the state orthogonal to $|H_0\rangle$. We denote the imperfect version of $|H_i\rangle$ obtained from ρ_0^a states as ρ_i^a. We can always bring any mixed state into this form using twirling [4]. For the protocol to be practical, we require it to remain stable under the two following types of errors, where we assume the state is pure and the rotation is off of the desired axis by δ:

$$\rho_0^b(\delta) \quad = \quad \frac{1}{2}\left(I + \sin\left(\frac{\pi}{4} + \delta\right)X + \cos\left(\frac{\pi}{4} + \delta\right)Z\right),$$
$$\rho_0^c(\delta) \quad = \quad \frac{1}{2}\left(I + \sin\frac{\pi}{4}\cos\delta X + \sin\frac{\pi}{4}\sin\delta Y + \cos\frac{\pi}{4}Z\right).$$

We numerically generated 1000 pseudo-random instances of the protocol to produce $|H_i\rangle$ states for different values of i for each error type and for noise strengths $10^{-4}, 10^{-6}$, and 10^{-8}. Figure 5(a) shows an exponential decay of the distance between erroneous and ideal states; if we start with a $|H_0\rangle$ state distilled to our target accuracy, all subsequent derived $|H_i\rangle$ states will also be distilled to at least that accuracy. This further suggests that for larger values of i, noisier $|H_0\rangle$ states could be used to still achieve the desired accuracy, and in turn decrease the number of distillation recursions (and resources) necessary to prepare the $|H_0\rangle$ states.

Extrapolating from Fig.5(a), one could for example prepare ρ_{12} states with accuracy 10^{-9} using only input $|H_0\rangle$ states of accuracy 10^{-6}, saving at least one round of distillation prior to our protocol, reducing the total offline cost (including magic state distillation). Using states as noisy as possible and using the costs and accuracies presented in Table I of [4], we were able to estimate, via numerical simulations, the improvement factor to be gained in offline cost for different rotations and precisions. The results are presented in Fig. 5(b). Two important behaviors are noted. First, for any given realtive precision ϵ/ϕ, the improvement factor increases as the absolute precision ϵ goes down. Second, and more importantly, there is as much as an order of magnitude to be gained for rotation angles that are comparable to

the desired accuracy ϵ, e.g., for $\epsilon = 5 \times 10^{-10}$ and $\phi \sim 100\epsilon$, there is a factor ~ 11 reduction in resource offline cost.

6   Application to Single-qubit Rotations

We now show how to use the ladders of states to enable the fault-tolerant approximation of any single-qubit rotation. Results do not include the improvements in offline cost discussed in previous section, so an additional gain factor between 2 and 10, depending on ϵ, is expected. Recall the circuit given in Fig. 1. If we input either $HSHX |H_i\rangle$ or $HSHX \left|\psi_i^j\right\rangle$ in place of the top qubit, we obtain rotation $Z(\pm 2\theta_i)$ on $|\psi\rangle$. Note that there is a factor of two difference between the angle θ_i involved in the description of the state and the rotation applied, e.g., the $|H_0\rangle$ state is over $\theta_0 = \frac{\pi}{8}$, and can be used to implement a $\frac{\pi}{4}$ rotation. Also, since $0 < \theta_i < \frac{\pi}{4}$ ($\forall i$), the discontinuity of cotangent is not a problem.

Although the circuit in Fig. 1 randomly applies $\pm\theta$, our protocols still result in efficient application of the desired Z-rotation.

We propose the following protocol to approximate a Z-rotation $Z(\phi)$:

1. Set desired accuracy ϵ.
2. Pick a target rotation angle $0 < \phi < 2\pi$.
3. Find the state $|H_i\rangle$ (or denser state $\left|\psi_i^j\right\rangle$) such that $2\theta_i$ is close to ϕ.
4. Simulate an instance of the ladder to obtain that state and add its cost to the offline cost.
5. Apply a rotation using $|H_i\rangle$ (or denser state $\left|\psi_i^j\right\rangle$) as input to the circuit of Fig. 1 and add one to the online cost.
6. Recurse on steps 3 through 5 until the desired accuracy is reached.

Thus, one has to implement a sequence of j rotations $\{Z(2\theta_{i_j})\}$ on $|\psi\rangle$ using the sequence of states $\{\left|H_{i_j}\right\rangle\}$, such that $Z(\phi) \approx \prod_j Z(2\theta_{i_j})$. The online cost is also given by $\left|\{\left|H_{i_j}\right\rangle\}\right|$.

We define the accuracy of the applied rotation V compared to the target rotation $U = Z(\phi)$ as

$$\max_{|\psi\rangle} D(U |\psi\rangle \langle\psi| U^\dagger, V |\psi\rangle \langle\psi| V^\dagger),$$

where $D(\rho, \sigma)$ is the trace distance between states ρ and σ. If U and V are rotations about the same axis, one can show that in our case, for small angles of rotation, this reduces to the difference of rotation angles: $\epsilon = \Delta\phi$. In [9], the distance measure used is $D(U, V) = \sqrt{(2 - |\text{tr}(UV^\dagger)|)/2}$. In the case of rotations about the same axis, it can be reduced to $\sqrt{1 - |\cos(\Delta\phi)|} \approx \Delta\phi/\sqrt{2}$ for small $\Delta\phi$.

We define an online and offline cost to apply a unitary gate. The *online cost*, C_{on}, is the expected number of $|H_i\rangle$ states required to implement the unitary. The *offline cost*, C_{off}, is the total number of distilled $|H_0\rangle$ states required to obtain all of the intermediate $|H_i\rangle$ states used to perform the given unitary. For Solovay-Kitaev decomposition, the offline cost equals the online cost and is given by the total number of T and $T^\dagger$ gates in the decomposition. In both cases, we do not count the cost of initially distilling $|H_0\rangle$ states.

We simulated $\sim 1.8 \times 10^4$ instances of our $|H\rangle$ protocol, each for a random angle ϕ and target accuracy between $10^{-12} < \epsilon < 10^{-4}$. We assume that $C_{\text{on}} \sim \ln_{\text{on}}^c(\frac{1}{\epsilon})$, and $C_{\text{off}} \sim \ln^{c_{\text{off}}}(\frac{1}{\epsilon})$, where C_{on} and C_{off} are the online and offline costs, respectively, such that $\ln C_{\text{on}} \sim c_{\text{on}} \ln \ln(\frac{1}{\epsilon})$, and $\ln C_{\text{off}} \sim c_{\text{off}} \ln \ln(\frac{1}{\epsilon})$. From linear fits to the data, we find $\ln(C_{\text{on}}) = -0.21 + 1.23 \ln(\ln(1/\epsilon))$ with a standard deviation around the mean of $\ln(\Delta C_{\text{on}}) = -0.30 + 0.83 \ln(\ln(1/\epsilon))$, and $\ln(C_{\text{on}}) = -0.44 + 2.22 \ln(\ln(1/\epsilon))$ with a standard
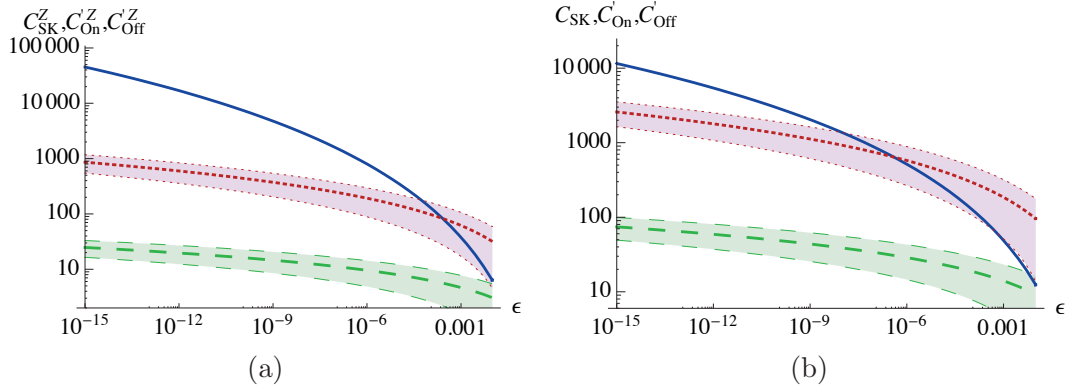
Figure 6 Cost of (a) random Z-rotations and (b) random unitaries as a function of precision ϵ. Solid line: SK decomposition [9]. Dotted line: Offline cost using $|H\rangle$, or $\{|\psi^0\rangle, |\psi^1\rangle, |\psi^2\rangle\}$ as initial resources. Dashed line: Online cost using $|H\rangle$, or $|\psi^0\rangle, |\psi^1\rangle, |\psi^2\rangle\}$ as initial resources. The shaded regions around the dashed and dotted lines represent the standard deviation around the mean. (a) $\ln(C'^Z_{\text{SK}}) = -4.88 + 4.41 \ln(\ln(1/\epsilon))$; $\ln(C'^Z_{\text{On}}) = -0.46 + 1.04 \ln(\ln(1/\epsilon))$; $\ln(C'^Z_{\text{Off}}) = 0.96 + 1.64 \ln(\ln(1/\epsilon))$. (b) $\ln(C'_{\text{SK}}) = -2.67 + 3.40 \ln(\ln(1/\epsilon))$; $\ln(C'_{\text{On}}) = -0.46 + 1.04 \ln(\ln(1/\epsilon)) + \ln 3$; $\ln(C'_{\text{Off}}) = 0.96 + 1.64 \ln(\ln(1/\epsilon)) + \ln 3$.

deviation around the mean of $\ln(\Delta C_{\text{on}}) = 0.02 + 1.87 \ln(\ln(1/\epsilon))$. We deduce that $c_{\text{on}} \sim 1.23$ and $c_{\text{off}} \sim 2.22$ for our protocol.

For the denser protocol, the offline costs are 12.50, 12.95, and 11.64 for $|\psi_0^0\rangle$, $|\psi_0^1\rangle$, and $|\psi_0^2\rangle$, respectively. The denser set of states results in improved scalings for both the online and offline costs: $c'_{\text{on}} \sim 1.04$ and $c'_{\text{off}} \sim 1.64$, where $'$ denotes the denser protocol. However, the offline costs of our new states $|\psi_0^i\rangle$ are improved only when precisions are smaller then $\epsilon \approx 1.28 \times 10^{-5}$.

Figure. 6 shows the behavior of the protocols on Z rotations and arbitrary rotations. For an arbitrary rotation, recall that a single-qubit unitary U is composed of three rotations around the X- and Z-axes [12]: $U \propto X(\alpha)Z(\beta)X(\gamma)$, for some angles α, β, γ. We can use our protocol to implement both Z and X rotations as previously outlined. Fig. 6(a) plots the fit for Solovay-Kitaev decomposition [9] (solid line), the online cost (dashed), and offline cost (dotted). For all practical precisions, the online cost of our proposed scheme is consistently smallest. The offline cost is advantageous when $\epsilon \leq 4.41 \times 10^{-4}$ for Z-rotations and $\epsilon < 1.03 \times 10^{-6}$ for random unitaries.

7   Minimizing Online Cost

We can further minimize the online cost by considering instead the following protocol to implement a Z rotation by angle ϕ: Prepare *offline* the state $|Z(\phi)\rangle$ using the protocol described to apply $|Z(\phi)\rangle$ to a $|0\rangle$ ancilla. Then, use $|Z(\phi)\rangle$ *online* to apply the rotation to the desired qubit. With probability $\frac{1}{2}$, the rotation $Z(\phi)$ is applied and the online cost is 1. If it fails, prepare *offline* $|Z(2\phi)\rangle$; with probability $\frac{1}{2}$, $Z(\phi)$ is applied online and the online cost is 2. If it fails, prepare *offline* $|Z(4\phi)\rangle$, and so on. The probability of success after n iterations decreases exponentially with n; the process is a negative binomial of parameter $p = \frac{1}{2}$ and the expected number of online rotations for success is $\sim \frac{1}{p} = 2$. We simulated this process for random angles $0 < \phi < 2\pi$ and accuracies $10^{-12} < \epsilon < 10^{-4}$ and found the expected number of online rotations is $\langle C''_{\text{on}} \rangle = 1.99$ and the offline cost is $c''_{\text{off}} \sim 1.75$. Note
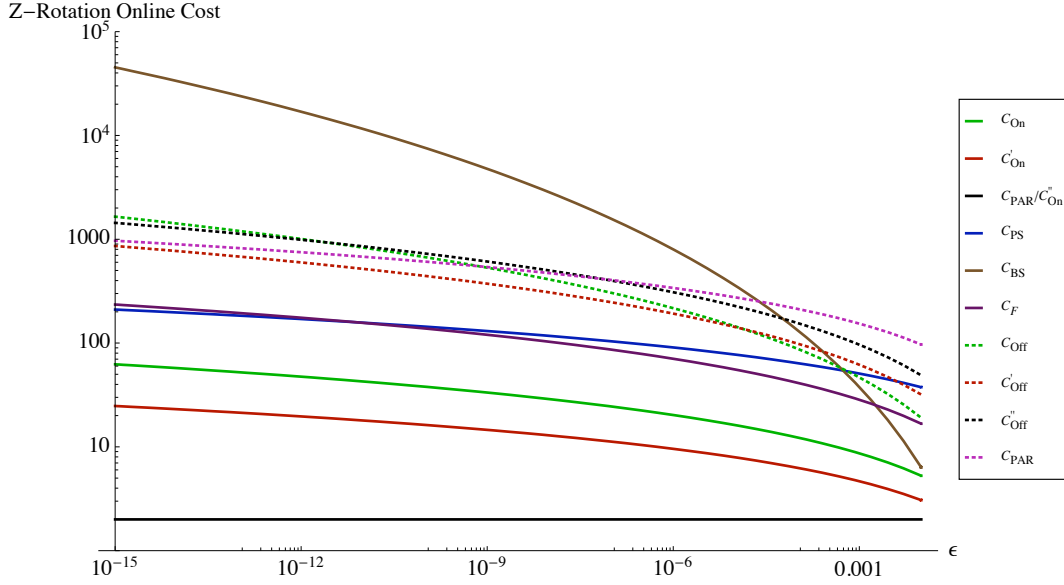
■ **Figure 7** Comparison of online (solid) and offline (dashed) costs to decompose Z rotations vs. accuracy ϵ. Methods plotted include C_{PAR} [13], C_{PS} [10], C_{BS} [9], C_{F} [14]. C_{On}, C_{Off}, C'_{On}, C'_{Off}, C''_{On}, C''_{Off} represent our $|H\rangle$ ladder, dense ladder, and minimal online cost with the dense ladder, respectively. The offline costs for $C_{\{\text{BS,PS,F}\}}$ are equal to their online costs.

that any method can be used to prepare the ancilla state offline, and here we use our protocol for preparation. We discovered after writing that a similar technique was described in [13].

Figure 7 compares the cost of state-of-the-art decomposition techniques with our protocols. The plot highlights the tradeoffs between the various methods. Note that we only plotted two methods, our protocol C'' and C_{PAR} (which uses C_{F} to prepare the state), using the minimal online framework, but the other techniques could also be used to prepare the state offline, yielding an expected online cost of 2 and a roughly doubled offline cost. Our protocols C, C', and C'' (red, green, black) exhibit a very clear tradeoff between online circuit depth and offline cost. For example, if operations on logical qubits must be minimized (due to noise), then trading offline resources for low online circuit depth is desirable, making C, C', and C'' advantageous compared to $C_{\{\text{BS,F,PS}\}}$. C'' is competitive with the minimal-online versions of C_{F} (plotted as C_{PAR}) and C_{PS} (not plotted). In practice, several decomposition techniques will be used throughout the compilation of a quantum algorithm.

Finally, our protocol can be used to fault-tolerantly implement elements of the V basis, which consists of $V_{\{1,2,3\}} = (I + 2i\{X,Y,Z\})/\sqrt{5}$ and their inverses. The V basis was shown to be *efficiently universal*, guaranteeing decompositions to be of depth $O(\log(1/\epsilon))$ [15]. It was previously dismissed as a candidate basis for decomposition due to the inability to implement the gates fault-tolerantly. However, our protocol enables fault-tolerant implementation: $V = Z(\pi/4)Z(2\theta_2)$, which is a T gate followed by a rotation using the $|H_2\rangle$ resource state. On average, it requires an offline cost of 10 $|H_0\rangle$ states. This has prompted the development of decomposition algorithms targeted to the V basis that may outperform those for the $\{H,T\}$ basis [16].

8 Conclusions

We have proposed a protocol to distill non-stabilizer states efficiently using magic states, Clifford operations, and measurements. One application of our protocol is implementing arbitrary single-qubit rotations with lower resource cost than state-of-the-art decomposition methods and *constant* online circuit depth. An extension of our work is to study other stabilizer circuits as "ladders" of states, or to use SH eigenstates distilled using the protocols of [3, 5]. Finally optimizing the sequence of angles required to implement the desired rotation, or determining when to use a given decomposition technique, will be a necessary component of any quantum compiler. We thank Alex Bocharov and Cody Jones for many useful discussions.

References

1 A. G. Fowler, A. M. Stephens, and P. Groszkowski, Phys. Rev. A **80**, 052312 (2009).

2 S. Aaronson and D. Gottesman, Physical Review A **70**, 052328 (2004), http://arxiv.org/abs/arXiv:quant-ph/0406196 .

3 S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).

4 A. M. Meier, B. Eastin, and E. Knill, "Magic state distillation with the four-qubit code," (2012).

5 S. Bravyi and J. Haah, "Magic state distillation with low overhead," (2012), 1209.2426 .

6 N. C. Jones, "Multilevel distillation of magic states for quantum computing," (2012), 1210.3388 .

7 A. Kitaev et al., *Classical and Quantum Computation* (American Mathematical Society, Providence, RI, 2002).

8 C. M. Dawson and M. A. Nielsen, "The Solovay-Kitaev algorithm," (2005), arxiv:quant-ph/0505030 .

9 A. Bocharov and K. M. Svore, Phys. Rev. Lett. **109**, 190501 (2012).

10 P. Selinger, "Efficient clifford+T approximation of single-qubit operators," (2012), 1212.6253 .

11 V. Kliuchnikov, D. Maslov, and M. Mosca, "Practical approximation of single-qubit unitaries by single-qubit quantum clifford and T circuits," (2012), 1212.6964 .

12 M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

13 N. C. Jones, J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. V. Meter, A. Aspuru-Guzik, and Y. Yamamoto, "Simulating chemistry efficiently on fault-tolerant quantum computers," (2012), 1204.0567 .

14 A. Fowler, Quantum Information and Computation **11**, 867 (2011), quant-ph/0411206 .

15 A. W. Harrow, B. Recht, and I. L. Chuang, J. Math. Phys. **43** (2002).

16 A. Bocharov, Y. Gurevich, and K. M. Svore, in *Quantum Information Processing (QIP) 2012, Poster* (2012).