
Article:
Campbell, E. and Howard, M. (2018) Magic state parity-checker with pre-distilled components. Quantum, 2. 56. ISSN: 2521-327X

https://doi.org/10.22331/q-2018-03-14-56

Magic state parity-checker with pre-distilled components

Earl T. Campbell and Mark Howard

Department of Physics & Astronomy, University of Sheffield, Sheffield, S3 7RH, United Kingdom.

February 23, 2018

Magic states are eigenstates of non-Pauli operators. One way of suppressing errors present in magic states is to perform parity measurements in their non-Pauli eigenbasis and postselect on even parity. Here we develop new protocols based on non-Pauli parity checking, where the measurements are implemented with the aid of pre-distilled multiqubit resource states. This leads to a two step process: pre-distillation of multiqubit resource states, followed by implementation of the parity check. These protocols can prepare single-qubit magic states that enable direct injection of single-qubit axial rotations without subsequent gate-synthesis and its associated overhead. We show our protocols are more efficient than all previous comparable protocols with quadratic error reduction, including the protocols of Bravyi and Haah.

Error corrected quantum computers require additional gadgets and tricks to enable fully universal, fault-tolerant quantum computation (see Ref. [1] for a review). Of the various competing approaches, magic state distillation is especially efficient. At first, magic state distillation was conceived of as a way to implement the T gate (also called the $\pi/8$ phase gate), with successive generations of protocols making ever greater improvements [2–6]. Any desired unitaries could be approximated by efficient synthesis of T gates and Clifford gates, with recent years also bringing new, optimised synthesis methods [7–11]. However, we can circumvent the need for synthesis if we instead prepare magic states tailored for injecting specific gates. Preparing tailored single qubit magic states can directly provide single qubit rotations other than the T-gate and distillation routines for these states have been proposed [12–16]. Complex multi-qubit circuits can also be directly injected once multi-qubit magic states have been distilled [6, 17–20].

Here we propose a family of two-step protocols for distilling magic states that are tailored for injecting a desired single-qubit Z-axis rotation. A distinctive feature is that the first step creates a multi-qubit magic state using the synthillation protocol [19, 20]. In the second step, this multi-qubit resource is then used to fault-tolerantly perform a parity check in a non-Pauli basis. Combined, the protocol takes in single qubit states and outputs single qubit states, with multi-qubit magic states appearing only fleetingly.

A single round of our protocol will quadratically reduce errors using fewer inputs per output than any previous protocol with quadratic error reduction. Higher order error reductions can be achieved by concatenation of our protocols. Higher order reductions in noise are possible without concatenation by using sophisticated protocols that give a large

Earl T. Campbell: earltcampbell@gmail.com, https://earltcampbell.com/

Mark Howard: m.howard@sheffield.ac.uk, https://markhoward.info

error reduction in a single round [5, 6, 21, 22], but there are important practical considerations for why one might favour a concatenated approach (see Sec. 6.2 for a discussion).

We begin by covering some basic notation. Sec. 2 gives an overview of our approach. Note that the first step is our previously proposed synthillation method [19, 20], so the details will not be repeated here. Rather we focus on how non-Pauli parity checking is possible given these pre-distilled resources, giving a detailed explanation in Sec. 3. The protocol's performance is analysed in Sec. 4. In Sec. 5, we present a small bonus result that proves equivalent performance is unlikely to be possible using codes with conventional transversal gate constructions.

1 Notation

We denote axial rotations about the Pauli-Z axis as

$$R(\theta) = \exp(i\theta Z) = \cos(\theta)\mathbb{1} + i\sin(\theta)Z. \tag{1}$$

If the angle is $\theta = \pi/2^\ell$ for integer ℓ then $R(\theta)$ belongs in the ℓ^{th} level of the Clifford hierarchy [23]. Therefore, $R(\pi/8)$ is the $\pi/8$-phase gate, also known at the T gate. Unitaries inside the Clifford hierarchy are special because they can be realised using state-injection and a bounded number of appropriate magic states. However, all the analysis in this paper holds for any θ, even values corresponding to unitaries and magic states not connected to the Clifford hierarchy.

We use $W(\theta)$ for the Hermitian operator $W(\theta) := R(\theta)XR(\theta)^\dagger = R(2\theta)X$. Note that $W(\pi/8)$ is a Clifford and plays a similar role to the Hadamard; it interchanges X and Y whereas the Hadamard interchanges X and Z. The relevant magic states are eigenstates of $W(\theta)$ and sit on the equator of the Bloch sphere

$$|R(\theta)\rangle = W(\theta)|R(\theta)\rangle = R(\theta)|+\rangle. \tag{2}$$

More generally, when U is a diagonal gate (acting on n qubits) we use $|U\rangle := U(|+\rangle^{\otimes n})$. In this notation, the familiar T state is $|R(\pi/8)\rangle$. We use CZ for control-Z and CCZ for control-control-Z.

2 Overview of new protocols

2.1 Protocols for $\pi/8$ phase gates

The protocols presented here can be used both for distillation of T-states that can implement $\pi/8$ phase gates, or more generally for distillation of $|R(\theta)\rangle$ magic states for smaller angle rotations. We begin by sketching the simple case of T-state distillation. We say a protocol is an $n \to k$ protocol if it takes n inputs and outputs k magic states with some success probability (typically this probability approaches unity in the low noise limit). Our two-step protocols for T-state distillation are $3k + 4 \to k$ protocols for even k with quadratic reduction of noise. The resource overhead of the protocol is roughly captured by $n/k = 3 + \frac{4}{k}$, which approaches 3 for large k. For $k = 2$ we have a $10 \to 2$ protocol and so the protocol is very similar to the MEK proposal proposed by Meier, Eastin and Knill [3]. Therefore, compared to MEK, we reduce the n/k overhead from 5 to 3 by going to larger blocks sizes (larger k). Another class of protocols was proposed by Bravyi and Haah [4], which are $3k + 8 \to k$ protocols for even k with quadratic reduction of noise. The Bravyi-Haah protocols also have $n/k \to 3$ in the large protocol limit. Both our protocols and the

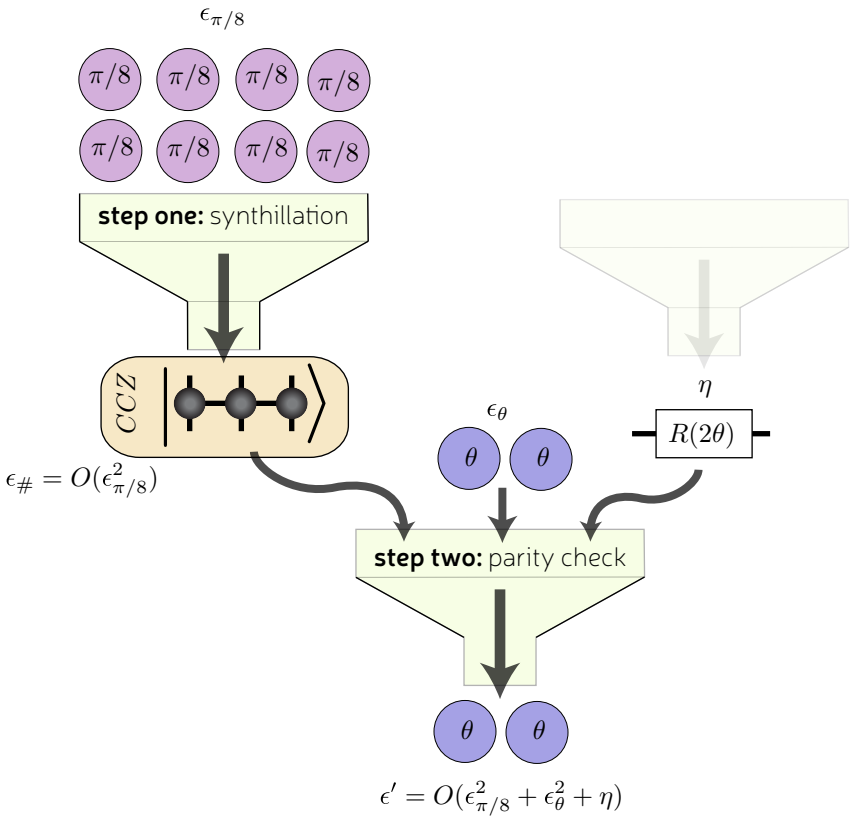

Figure 1: An illustration of how the two steps of our protocol are chained together for the special case $N = 1$. The combined process is described in Eq. (9) for general N.

Bravyi-Haah protocols have the same asymptotic limit, but our protocols approach this limit faster. For example, to achieve $n/k = 4$ we can use modest size $16 \to 4$ protocols, whereas the comparable Bravyi-Haah protocol is $32 \to 8$ with double the block size. This effect becomes more pronounced as the protocol is concatenated.

Furthermore, our work presents a different approach to distillation as we break the process up into two steps, making use of a multi-qubit magic state resource. Let us describe how this two-step process works, first considering the $10 \to 2$ case. In the first step we prepare a "pre-distilled" magic state that can inject a CCZ (or Toffoli) gate. It has been known for several years that a single CCZ magic state can be prepared from 8 T-states with quadratic error reduction [17, 18]. It is no longer appropriate to describe this process in simple $n \to k$ notation and so we introduce the more detailed description of these protocols as being

$$\left\{ \begin{array}{c} 8 \\ |R(\pi/8)\rangle \\ \epsilon_{\pi/8} \end{array} \right\} \to \left\{ \begin{array}{c} 1 \\ |CCZ\rangle \\ O(\epsilon_{\pi/8}^2) \end{array} \right\}. \tag{3}$$

In this notation, the left hand side gives the input resources and the right hand size gives the output resources. The top line gives the quantity of resources, the middle describes the species of magic state and the bottom line gives the infidelity. Our second step, which was not previously known, is to notice that a single $|CCZ\rangle$ resource can be used to check

the parity in a pair of $|R(\pi/8)\rangle$ states, which implements the transform

$$\left\{\begin{array}{cc} 2 & 1 \\ |R(\pi/8)\rangle & |CCZ\rangle \\ \epsilon_{\pi/8} & \epsilon_{\#} \end{array}\right\} \rightarrow \left\{\begin{array}{c} 2 \\ |R(\pi/8)\rangle \\ O(\epsilon_{\pi/8}^2 + \epsilon_{\#}) \end{array}\right\}. \tag{4}$$

Chaining these steps together so that $\epsilon_{\#} = O(\epsilon_{\pi/8}^2)$ yields

$$\left\{\begin{array}{c} 10 \\ |R(\pi/8)\rangle \\ \epsilon_{\pi/8} \end{array}\right\} \rightarrow \left\{\begin{array}{cc} 2 & 1 \\ |R(\pi/8)\rangle & |CCZ\rangle \\ \epsilon_{\pi/8} & O(\epsilon_{\pi/8}^2) \end{array}\right\} \rightarrow \left\{\begin{array}{c} 2 \\ |R(\pi/8)\rangle \\ O(\epsilon_{\pi/8}^2) \end{array}\right\}, \tag{5}$$

or simply $10 \rightarrow 2$ for short.

Next we outline the two-step process for larger block protocols. Instead of using $|CCZ\rangle$ resources, larger block protocols will use a resource that we denote as $|CCZ_{\#N}\rangle$. This resource can inject a gate $CCZ_{\#N}$, which is N copies of the CCZ gate all sharing one control qubit in common, and the relevant magic state is simply

$$|CCZ_{\#N}\rangle := CCZ_{\#N}(|+\rangle^{\otimes 2N+1}). \tag{6}$$

In the first step, we borrow results on synthillation [19, 20] that provide protocols implementing

$$\left\{\begin{array}{c} 4N+4 \\ |R(\pi/8)\rangle \\ \epsilon_{\pi/8} \end{array}\right\} \rightarrow \left\{\begin{array}{c} 1 \\ |CCZ_{\#N}\rangle \\ O(\epsilon_{\pi/8}^2) \end{array}\right\}. \tag{7}$$

This protocol is described in the section "$U_{N\#}$ family" of Ref. [20]. To avoid repetition, here we simply treat this synthillation routine as a black box with known properties. Rather, here we focus on the second step — the main technical contribution of this work — by showing that a pre-distilled $|CCZ_{\#N}\rangle$ can be used to parity check on $2N$ magic states

$$\left\{\begin{array}{cc} 2N & 1 \\ |R(\pi/8)\rangle & |CCZ_{\#N}\rangle \\ \epsilon_{\pi/8} & \epsilon_{\#} \end{array}\right\} \rightarrow \left\{\begin{array}{c} 2N \\ |R(\pi/8)\rangle \\ O(\epsilon_{\pi/8}^2 + \epsilon_{\#}) \end{array}\right\}. \tag{8}$$

Chaining steps one and two, so that $\epsilon_{\#} = O(\epsilon_{\pi/8}^2)$, we obtain a family of protocols for all integer N

$$\left\{\begin{array}{c} 6N+4 \\ |R(\pi/8)\rangle \\ \epsilon_{\pi/8} \end{array}\right\} \rightarrow \left\{\begin{array}{cc} 2N & 1 \\ |R(\pi/8)\rangle & |CCZ_{\#N}\rangle \\ \epsilon_{\pi/8} & O(\epsilon_{\pi/8}^2) \end{array}\right\} \rightarrow \left\{\begin{array}{c} 2N \\ |R(\pi/8)\rangle \\ O(\epsilon_{\pi/8}^2) \end{array}\right\}, \tag{9}$$

or simply $6N + 4 \rightarrow 2N$ for short. Alternatively, using $k = 2N$ we have a family of $3k + 4 \rightarrow k$ protocols with even k. So we have returned to the choice of symbols used by Bravyi and Haah who found $3k + 8 \rightarrow k$ protocols with even k. Fig. 1 illustrates this for the $N = 1$ ($k = 2$) case. This completes the outline of our protocols for T-state distillation.

2.2 Protocols for general phase gates

Next, we describe a family of protocols for other equatorial magic states $|R(\theta)\rangle$. Step one will remain the same, again making use of synthillation of $|CCZ_{\#N}\rangle$ resource states. Step two generalises to

$$\left\{\begin{array}{ccc} 2N & 1 & N \\ |R(\theta)\rangle & |CCZ_{\#N}\rangle & R(2\theta) \\ \epsilon_\theta & \epsilon_\# & \eta \end{array}\right\} \to \left\{\begin{array}{c} 2N \\ |R(\theta)\rangle \\ O(\epsilon_\theta^2 + \epsilon_\# + \eta) \end{array}\right\}. \tag{10}$$

When $R(2\theta)$ appears uncluttered by a ket, as it does above, it refers to inputting a phase gate $R(2\theta)$ rather than a magic state. Most interesting is the case when $R(2\theta)$ sits in the Clifford hierarchy and so can be injected using appropriate magic states. Proving the validity of the above mapping is at the core of this paper. Chaining this with synthillation yields an overall protocol

$$\left\{\begin{array}{ccc} 2N & 4N+4 & N \\ |R(\theta)\rangle & |R(\pi/8)\rangle & R(2\theta) \\ \epsilon_\theta & \epsilon_{\pi/8} & \eta \end{array}\right\} \to \left\{\begin{array}{ccc} 2N & 1 & N \\ |R(\theta)\rangle & |CCZ_{\#N}\rangle & R(2\theta) \\ \epsilon_\theta & O(\epsilon_{\pi/8}^2) & \eta \end{array}\right\} \to \left\{\begin{array}{c} 2N \\ |R(\theta)\rangle \\ O(\epsilon_\theta^2 + \epsilon_{\pi/8}^2 + \eta) \end{array}\right\}. \tag{11}$$

It is important to notice that there is no noise reduction in η. Therefore, it is crucially important they are predistilled (e.g. $\eta \sim \epsilon^2$) and so we refer to the rotations $R(2\theta)$ as pivotal. Despite the need for high fidelity pivotal rotations, other protocols have used pivotal rotations and found significant reductions of resource costs compared against using gate-synthesis. For instance, the $N = 1$ two-step protocol has an identical resource cost to the protocols introduced by Campbell and O'Gorman [15]. Pivotal rotations (though not under this name) played a similar role in the protocols proposed by Duclos-Cianci and Poulin [14], which in our notation can be described as

$$\left\{\begin{array}{ccc} 2N & 8N+4 & N \\ |R(\theta)\rangle & |R(\pi/8)\rangle & R(2\theta) \\ \epsilon_\theta & \epsilon_{\pi/8} & \eta \end{array}\right\} \to \left\{\begin{array}{c} 2N \\ |R(\theta)\rangle \\ O(\epsilon_\theta^2 + \epsilon_{\pi/8}^2 + \eta) \end{array}\right\}. \tag{12}$$

Note that for the majority of their paper, Duclos-Cianci and Poulin only discuss the $N = 1$ case, but they do sketch the higher N case later in the paper.

One could say our protocols are compressed as they essentially give a slight compression in T-cost of the Duclos-Cianci and Poulin protocols [14]. Our protocols also have very different inner workings. This provides a new perspective on magic state distillation, but also the two-step feature has a potentially significant practical advantage. The exotic resources are higher value since the $R(2\theta)$ pivotal rotation and $|R(\theta)\rangle$ resources are more difficult to prepare than standard $|R(\pi/8)\rangle$ magic states. However, in the two-step protocols, one does not risk using the exotic resources until the first step has succeeded. This contrasts, with both the Duclos-Cianci-Poulin protocols and Campbell-O'Gorman protocols for which all the resources are committed at the same time, with a single error anywhere leading to loss of all resources.

3 Implementing step two

Here we show how to use CCZ circuits, implemented using synthillation, to perform a parity check. We will construct circuits that measure the parity of $2N$ qubits in the basis

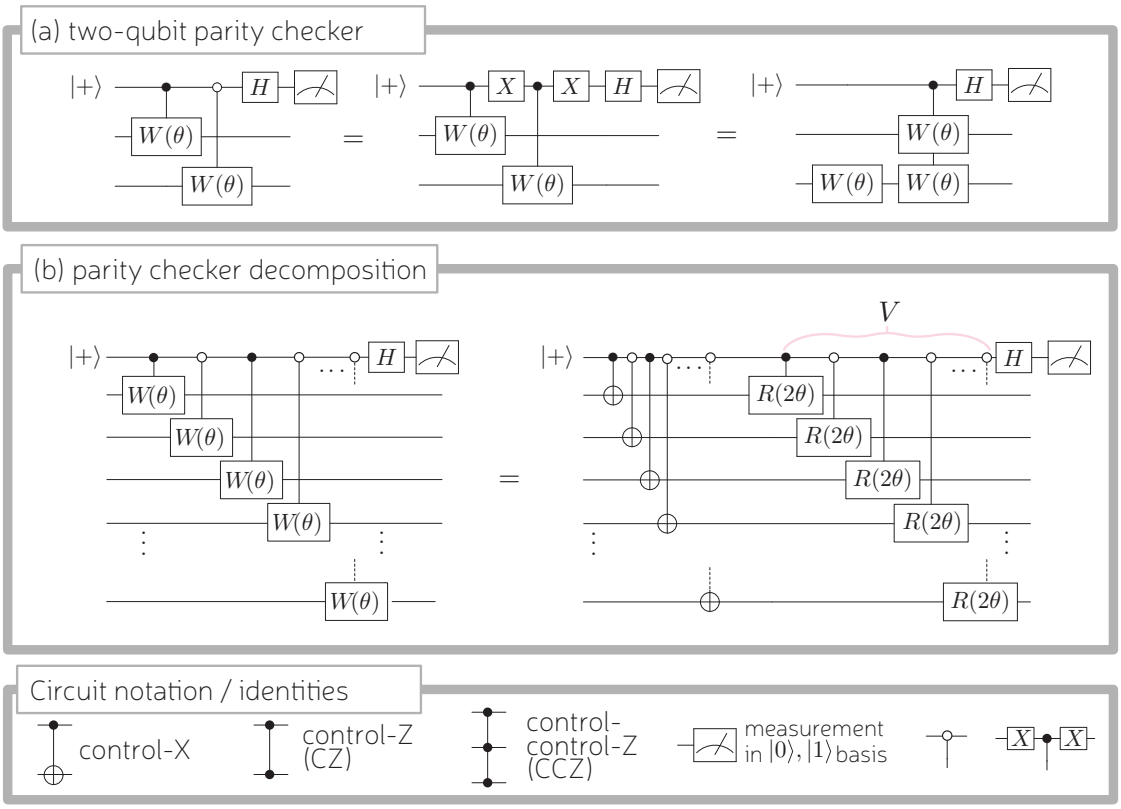

Figure 2: Gadgets for measuring parity in the $W(\theta)$ basis. In (a) we illustrate how control-$W(\theta)$ gates are used to measure parity. In (b) we show how the control-$W(\theta)$ gates can be decomposed into CNOT and control-phase gates. In the text we describe this phase-gate circuit by V (see Eq. (14)), which can be broken up into a product of U_j gates (see Eq. (17)).

of $W(\theta)$. Using an ancilla and a control-$W(\theta)^{\otimes 2N}$ gate would achieve this, but a smaller resource overhead is needed if we instead use the parity check gadgets in Fig. 2. We use a combination of control-$W(\theta)$ that triggers when the control bit is in the $|1\rangle$ state with an unconventional control-$W(\theta)$ (shown with open circles) that triggers when the control bit is in the $|0\rangle$ state. This parity check gadget measures in the $W(\theta)$ basis, but with an additional $W(\theta)$ rotation on half the qubits. This resulting Kraus operator for the desired even parity outcome is

$$\begin{aligned} K &= \frac{1}{2}(\mathbb{1}^{\otimes N} \otimes W(\theta)^{\otimes N} + W(\theta)^{\otimes N} \otimes \mathbb{1}^{\otimes N}) \\ &= \frac{1}{2}(\mathbb{1}^{\otimes N} \otimes W(\theta)^{\otimes N})(\mathbb{1}^{\otimes 2N} + W(\theta)^{\otimes 2N}). \end{aligned} \tag{13}$$

We assume the noisy $|R(\theta)\rangle$ magic states are diagonal in the $W(\theta)$ basis, and so the additional $W(\theta)$ rotations have no effect. The assumption of diagonal noise is mild and can be removed at the expense of hideously complicating the noise analysis (see e.g. App D of Ref. [15]).

Using $W(\theta) = R(2\theta)X$ it follows that this parity measurement circuit can be split into a sequence of control-X gates followed by a phase gate circuit (see Fig. 2b). Algebraically, this phase gate circuit is

$$V = \prod_{j=1,\ldots,N} \left[|0\rangle\langle 0|_0 R_{2j}(2\theta) + |1\rangle\langle 1|_0 R_{2j-1}(2\theta)\right], \tag{14}$$

where the subscripts denote which qubits the rotations act on, with qubit labels running from 0 to $2N$. Using the shorthand

$$U_j := |0\rangle\langle 0|_0 R_{2j}(2\theta) + |1\rangle\langle 1|_0 R_{2j-1}(2\theta), \tag{15}$$

we have

$$V = \prod_{j=1,\ldots,N} U_j. \tag{16}$$

To recap, given $|R(\theta)\rangle$ of error rate ϵ and the ability to implement V we can parity check in the $W(\theta)$ basis, outputting $|R(\theta)\rangle$ of error rate $O(\epsilon^2)$ with some probability $p = 1 - O(\epsilon)$.

Next, we show how to implement V with some pre-distilled resources. First, we use the decomposition $R(2\theta) = \cos(2\theta)\mathbb{1} + i\sin(2\theta)Z$ to expand out U_j as

$$U_j = |0\rangle\langle 0|_0(\cos(2\theta)\mathbb{1} + i\sin(2\theta)Z_{2j}) + |1\rangle\langle 1|_0(\cos(2\theta)\mathbb{1} + i\sin(2\theta)Z_{2j-1}). \tag{17}$$

Collecting the cos and sin terms, we have

$$\begin{aligned} U_j &= \cos(2\theta)\mathbb{1} + i\sin(2\theta)[|0\rangle\langle 0|_0 Z_{2j} + |1\rangle\langle 1|_0 Z_{2j-1}] \\ &= \cos(2\theta)\mathbb{1} + i\sin(2\theta)M_j, \end{aligned} \tag{18}$$

where we have introduced further shorthand

$$\begin{aligned} M_j &:= |0\rangle\langle 0|_0 Z_{2j} + |1\rangle\langle 1|_0 Z_{2j-1} \\ &= Z_{2j}(|0\rangle\langle 0|_0 + |1\rangle\langle 1|_0 Z_{2j} Z_{2j-1}) \\ &= Z_{2j} C Z_{0,2j} C Z_{0,2j-1}, \end{aligned} \tag{19}$$

which is unitary, Hermitian and Clifford.

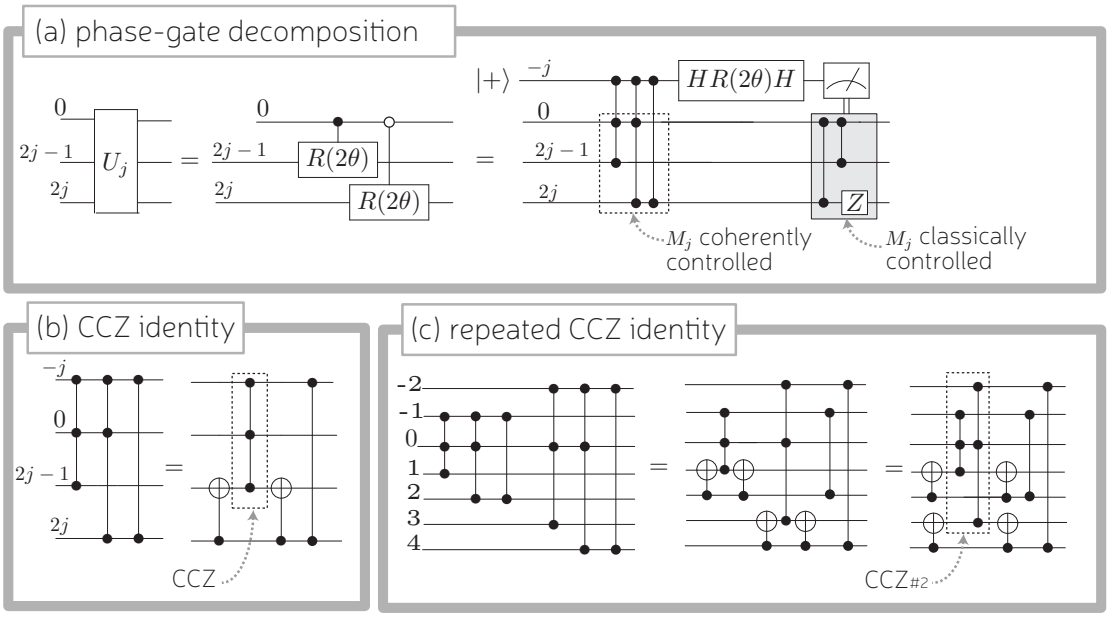

Figure 3: In (a) we show how U_j (a pair of control phase gates) can be implemented using an ancilla, a control-M_j gate and a pivotal rotation. An algebraic proof of this equivalence starts at Eq. (17) and ends after Eq. (23). In (b) we show how a pair of CCZ gates, which share a pair of controls, is Clifford equivalent to only a single CCZ gate. In (c) we uses this identity twice to simplify a more complex circuit down to a $CCZ_{\#2}$. In general, one finds that a sequence of N control-M_j can be simplified to a $CCZ_{\#N}$ circuit (see Eq. (24) and Eq. (26)).

Next we show that each U_j can be implemented with access to a single $|+\rangle$ ancilla, a control-M_j unitary and a $R(2\theta)$ rotation (see Fig. 3a). We prepare the ancilla in the $|+\rangle$ state and use it as the control qubit for the control-M_j unitary, which gives the state $|0\rangle|\psi\rangle + |1\rangle(M_j|\psi\rangle)$. Next we rotate the ancilla by $HR(2\theta)H$ and measure in the computational basis. This is equivalent to measuring with projections

$$\langle 0|HR(2\theta)H = \cos(2\theta)\langle 0| + i\sin(2\theta)\langle 1| \tag{20}$$
$$\langle 1|HR(2\theta)H = \cos(2\theta)\langle 1| + i\sin(2\theta)\langle 0|. \tag{21}$$

In the eventuality of a "+1" outcome, we find

$$|0\rangle|\psi\rangle + |1\rangle(M_j|\psi\rangle) \to (\cos(2\theta)\mathbb{1} + i\sin(2\theta)M_j)|\psi\rangle = U_j|\psi\rangle, \tag{22}$$

as desired. However, when a "1" outcome is measured we have

$$|0\rangle|\psi\rangle + |1\rangle(M_j|\psi\rangle) \to (i\sin(2\theta)\mathbb{1} + \cos(2\theta)M_j)|\psi\rangle = M_jU_j|\psi\rangle. \tag{23}$$

We see that an M_j gate will correct for the different measurements outcomes, and since M_j is Clifford this does not contribute to the resource cost. This completes the proof of the identity in Fig. 3a.

Similarly we chain together N such circuits, assuming we have access to N copies of $R(2\theta)$ and the circuit

$$\tilde{V} = \prod_{j=1,\ldots,N} (|0\rangle\langle 0|_{-j} + |1\rangle\langle 1|_{-j}M_j), \tag{24}$$
$$= \prod_{j=1,\ldots,N} CZ_{-j,2j}CCZ_{-j,0,2j}CCZ_{-j,0,2j-1}. \tag{25}$$

This is composed of a sequence of control-M_j gates, where each gate is controlled from a different ancillary qubit. We have labelled these new ancilla with negative integers from -1 to $-N$. Each control-M_j consists of a CZ gate and two CCZ gates. On first inspection this seems to imply $2N$ CCZ gates are needed. However, using the identity in Fig. 3b we see a pair of CZZ gates can sometimes be realised using a single CCZ. Note this identity only works because the pair of CCZ gates share two control bits in common. Applying this identity repeatedly (see e.g. Fig. 3c) can reduce the circuit to one using only N CCZ gates. Algebraically, the identity is

$$\tilde{V} = \left(\prod_{j=1}^{N} CX_{2j,2j-1}\right)\left(\prod_{j=1}^{N} CCZ_{-j,0,2j-1}\right)\left(\prod_{j=1}^{N} CX_{2j,2j-1}\right)\left(\prod_{j=1}^{N} CZ_{-j,2j}\right), \tag{26}$$

where $CX_{c,t}$ is a control-X gate with control qubit c and target qubit t. The resource intensive part is the non-Clifford component of N CCZ gates all sharing one single control qubit in common (qubit 0). Here we denote such a circuit as $\mathrm{CCZ}_{\#N}$, which has been elsewhere called $\mathrm{Tof}_{\#N}$. For these gates, the problem of optimal synthesis into CNOT + T gates has been solved and the circuit requires $4N + 3$ T gates (see Example IV.2. of Ref. [20]). Recall that we are requiring that $\mathrm{CCZ}_{\#N}$ is predistilled to a higher fidelity. The most efficient known method to achieve this is to use the synthillation protocol that can perpare $\mathrm{CCZ}_{\#N}$ using only $(4N+4)$ T-states of $\epsilon_{\pi/8}$ error rate, and so this is the first-step of our two-step protocol.

We have demonstrated how step two works using a series of circuit identities (for any integer N). For completeness, we show in Fig. 4 how these circuit identities plug together for $N = 1$ and $N = 2$. The circuits could be further expanded by replacing the non-Clifford gate $CCZ_{\#N}$ with the magic state $|CCZ_{\#N}\rangle$ and the appropriate Clifford injection circuit.

4 Noise analysis

This section presents a performance analysis for one round of our protocols. Subsection 4.1 reports some results of numerical simulations for smaller size protocols with $N = 1$ and $N = 2$. Subsection 4.2 focuses on providing a simple, yet rigorous derivation, of analytic upper bounds on output noise. The analytic results hold for all N, but are loose and actual performance will be much better than analytically bounded.

4.1 Numerical analysis

We performed full state vector numerical simulations using IBM's QISKit (code available as ancillary file). We simulated the effect of leading order errors for circuits with $N = 1$ and $N = 2$. The output error probabilities are for a single qubit with other output qubits traced out. Numerical results were independent of which output qubit is chosen and independent of θ (up-to numerical accuracy of 2 significant figures).

For $N = 1$, we found that

$$\epsilon' = 8\epsilon_{\pi/8}^2 + \epsilon_\theta^2 + \frac{1}{4}\eta + O(\epsilon_{\pi/8}^3, \epsilon_\theta^2, \eta^2, \ldots). \tag{27}$$

The leading order coefficients for output error are identical to those for the MEK_k protocols proposed by Campbell and O'Gorman (themselves a modified form of MEK) and so it seems that our new protocols (with $N = 1$) perform identically in this regard. We have a slight

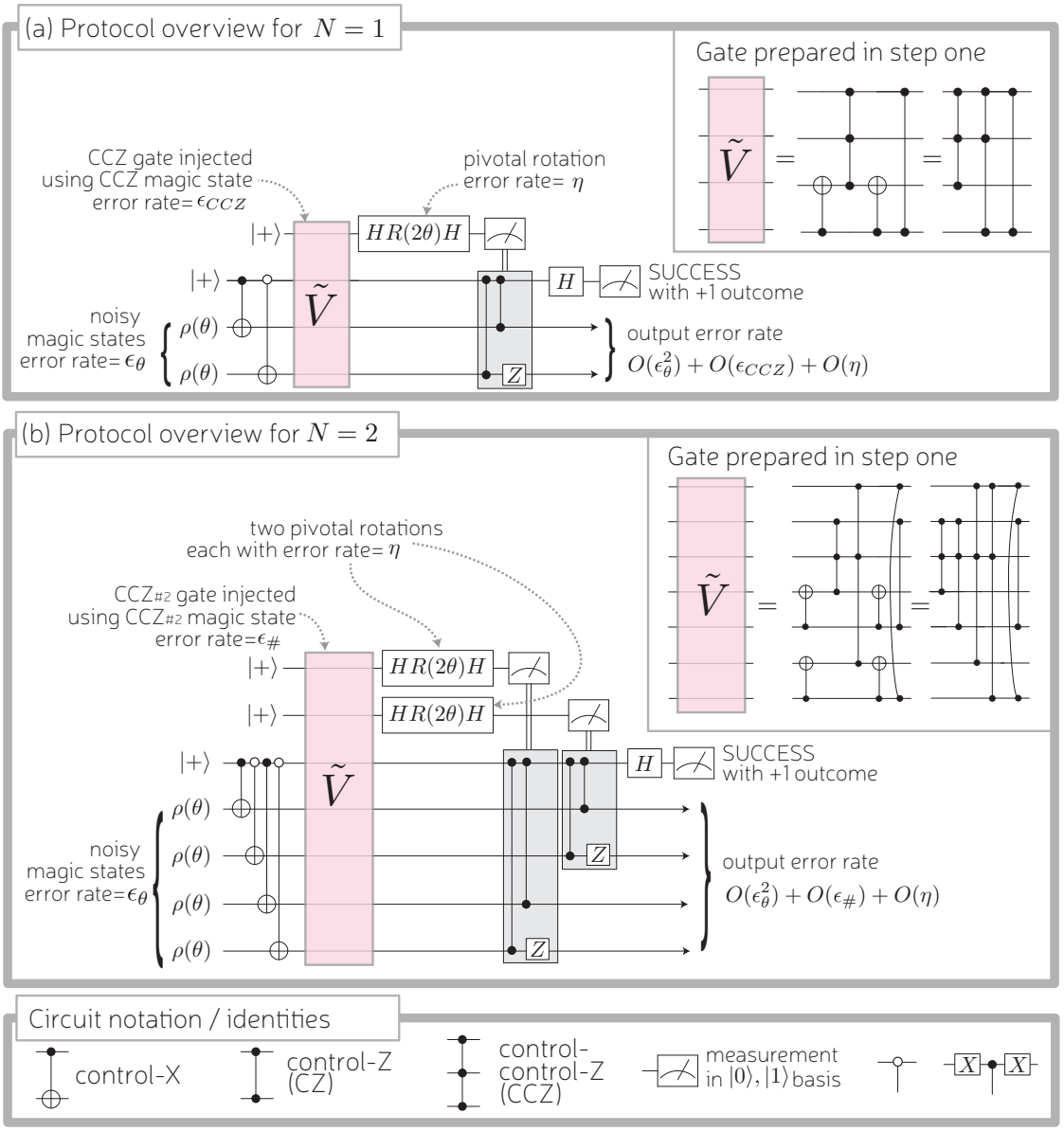

Figure 4: An overview of the second-step of our protocol for $N = 1$ and $N = 2$, which can be extended to any integer N. The first-step is to use synthillation to inject the non-Clifford gates shown in the pink section of the circuit. The only other non-Clifford elements of the circuits are the pivotal rotations $R(\theta)$. Input to the circuit are $2N$ mixed states $\rho(\theta)$, which are $|R(\theta)\rangle$ magic states with ϵ phase noise. In the absence of noise, the circuit measures the parity in the $W(\theta)$ bases. Therefore, when we see a SUCCESS event, the $\rho(\theta)$ states are output with quadratically reduced noise. In the event of a FAILURE, we discard the qubits and attempt again.

performance advantage in terms of success probability due to the two step nature. We found

$$p_{\text{synth}} = 1 - 8\epsilon_{\pi/8} + O(\epsilon_{\pi/8}^2), \tag{28}$$
$$p_{\text{parity}} = 1 - 2\epsilon_\theta - \frac{1}{2}\eta + O(\epsilon_{\pi/8}^2, \epsilon_\theta^2, \eta^2, \ldots),$$

where p_{synth} is the probability of step one succeeding and p_{parity} is the probability of step two succeeding. To leading order, the previous MEK$_k$ protocols had a success probability equal to $p_{\text{mek}} = p_{\text{synth}}p_{\text{parity}}$. Here, we don't commit to the second step until the first step is successful, which will lead to superior rates of generating magic states. For the setting $\theta = \pi/8$, the protocol simplifies to a $10 \to 2$ protocol with $\epsilon' = 9\epsilon_{\pi/8}^2 + O(\epsilon_{\pi/8}^3)$ and overhead $n/k = 10/2 = 5$, very similar to the original MEK protocol.

For $N = 2$, we found that

$$\epsilon' = 16\epsilon_{\pi/8}^2 + 3\epsilon_\theta^2 + \frac{1}{4}\eta + O(\epsilon_{\pi/8}^3, \epsilon_\theta^2, \eta^2, \ldots), \tag{29}$$
$$p_{\text{synth}} = 1 - 12\epsilon_{\pi/8} + O(\epsilon_{\pi/8}^2), \tag{30}$$
$$p_{\text{parity}} = 1 - 6\epsilon_\theta - \eta + O(\epsilon_{\pi/8}^2, \epsilon_\theta^2, \eta^2, \ldots), \tag{31}$$

By going to $N = 2$, we incur only mildly worse constant prefactors, but gain a significant efficiency improvement in terms of magic states output per input. For the setting $\theta = \pi/8$, the protocol simplifies to a $16 \to 4$ protocol with $\epsilon' = 19\epsilon_{\pi/8}^2 + O(\epsilon_{\pi/8}^3)$ and overhead $n/k = 16/4 = 4$. To obtain the same n/k overhead using Bravyi-Haah protocols (which are limited to $\theta = \pi/8$), we need to go to a larger size $32 \to 8$ protocol with $\epsilon' = 25\epsilon_{\pi/8}^2 + O(\epsilon_{\pi/8}^3)$. This confirms that our protocols can obtain similar resource overheads with a smaller scale quantum computer, and without any sacrifice in terms of error suppression or success probability.

4.2 Algebraic analysis

Here we take an analytic approach. We do not know of an analytic method of determining the exact expressions for ϵ', but can prove a rigourous upper bound using standard norm inequalities. Actual performance will be much better than proven here. We begin by considering the effect of noise on the input states

$$\rho(\theta) := (1 - \epsilon_\theta)|R(\theta)\rangle\langle R(\theta)| + \epsilon_\theta Z|R(\theta)\rangle\langle R(\theta)|Z. \tag{32}$$

We extend later to account for CCZ$_{\#N}$ noise and pivotal rotation noise, but when these components work perfectly the circuit implements

$$\rho(\theta)^{\otimes 2N} \to \mathcal{E}(\rho(\theta)^{\otimes 2N}) = K\rho(\theta)^{\otimes 2N}K^\dagger \tag{33}$$

where K is the parity projecting Kraus operator introduced in Eq. (13). Rather than the whole multi-qubit output, we are interested in the fidelity of a single output qubit, and so introduce the channel

$$\mathcal{E}_i(\rho(\theta)^{\otimes 2N}) = \text{tr}_i[K\rho(\theta)^{\otimes 2N}K^\dagger], \tag{34}$$

where $\text{tr}_i[\cdots]$ is the partial trace over all but the i^{th} qubit. The output of this channel is the unnormalised state

$$\mathcal{E}_i(\rho(\theta)^{\otimes 2N}) = p_{\text{good}}|R(\theta)\rangle\langle R(\theta)| + p_{\text{bad}}Z|R(\theta)\rangle\langle R(\theta)|Z. \tag{35}$$

The renormalisation constant $p_{\text{good}} + p_{\text{bad}}$ is the probability of the parity check yielding a "+1" outcome. When the parity check process is error-free, this occurs whenever the input states contains no errors or an even number of errors, and so

$$p_{\text{good}} + p_{\text{bad}} = \frac{1}{2}(1 + (1 - 2\epsilon_\theta)^{2N}) \tag{36}$$

The term p_{bad} is the probability of an error on i^{th} qubit and an odd number of errors on the remaining $2N - 1$ qubits

$$\begin{aligned} p_{\text{bad}} &= \epsilon_\theta \sum_{j=1}^{N} \tbinom{2N-1}{2j-1} \epsilon_\theta^{2j-1} (1 - \epsilon_\theta)^{2N-2j}, \\ &= \frac{\epsilon_\theta}{2}(1 - (1 - 2\epsilon_\theta)^{2N-1}), \\ &< (2N-1)\epsilon_\theta^2, \end{aligned} \tag{37}$$

where in the first line we have a binomial coefficient. The inequality follows from Bernoulli's inequality. This shows quadratic noise suppression in ϵ_θ.

Next, we account for $\epsilon_\#$ noise in the $\text{CCZ}_{\#N}$ gate. We can write the corresponding magic state as

$$\rho_{\#N} = (1 - \epsilon_\#)\Psi_{\#N} + \epsilon_\# \sigma_{\#N}, \tag{38}$$

where

$$\Psi_{\#N} = |CCZ_{\#N}\rangle\langle CCZ_{\#N}|, \tag{39}$$

and $\sigma_{\#N}$ carries some Z noise. We define $\mathcal{F}$ as the channel describing the action of the whole circuit (including implicit injection gadget for $CCZ_{\#N}$), assuming ideal pivotal rotations, acting on $\rho_{\#N}$ and $\rho(\theta)^{\otimes 2N}$. We will use that $\Psi_{\#N}$ leads to a parity

$$\mathcal{F}(\Psi_{\#N} \otimes \rho(\theta)^{\otimes 2N}) = \mathcal{E}(\rho(\theta)^{\otimes 2N}) \tag{40}$$

and by linearity of $\mathcal{F}$ we deduce

$$\mathcal{F}(\rho_{\#N} \otimes \rho(\theta)^{\otimes 2N}) = (1 - \epsilon_\#)\mathcal{E}(\rho(\theta)^{\otimes 2N}) + \epsilon_\# \mathcal{F}(\sigma_{\#N} \otimes \rho(\theta)^{\otimes 2N}). \tag{41}$$

Again, we are interested in only the single output qubit, and so introduce $\mathcal{F}_i = \text{tr}_i \mathcal{E}_i$, which straightforwardly yields

$$\mathcal{F}_i(\rho_{\#N} \otimes \rho(\theta)^{\otimes 2N}) = (1 - \epsilon_\#)\mathcal{E}_i(\rho(\theta)^{\otimes 2N}) + \epsilon_\# \mathcal{F}_i(\sigma_{\#N} \otimes \rho(\theta)^{\otimes 2N}). \tag{42}$$

This yields a single qubit state of the form in Eq. (35) with new parameters p'_{good} and p'_{bad}, which are tricky to exactly calculate but can again be bounded. The joint probability $p'_{\text{good}} + p'_{\text{bad}}$ can be lower bounded by assuming $\mathcal{F}_i(\sigma_{\#N} \otimes \rho(\theta)^{\otimes 2N}) = 0$ and so

$$p'_{\text{good}} + p'_{\text{bad}} \geq (1 - \epsilon_\#)(p_{\text{good}} + p_{\text{bad}}) \tag{43}$$

The p'_{bad} term can be upper bounded by considering the worst-case scenario that $\mathcal{F}_i(\sigma_{\#N} \otimes \rho(\theta)^{\otimes 2N})$ leads to a logical error with unit probability, and so

$$p'_{\text{bad}} \leq (1 - \epsilon_\#)p_{\text{bad}} + \epsilon_\# < (2N-1)(1 - \epsilon_\#)\epsilon_\theta^2 + \epsilon_\#, \tag{44}$$

where the second inequality follows from Eq. (37). These bounds are very loose and overestimate p'_{bad} by quite a lot. Nevertheless, they are simple to obtain and rigorous.

Next, we further consider phase noise on the pivotal rotation, each failing with probability η. In other words, all pivotal rotations act perfectly with probability $(1-\eta)^N$. Therefore, the channel implemented is not $\mathcal{F}_i$ but something of the form

$$\mathcal{G}_i = (1-\eta)^N \mathcal{F}_i + (1-(1-\eta)^N)\mathcal{F}'_i, \tag{45}$$

where $\mathcal{F}'_i$ is the noisy part of the channel with diamond norm not exceeding unity. Therefore,

$$\begin{aligned}\mathcal{G}_i(\rho_{\#N} \otimes \rho(\theta)^{\otimes 2N}) =& (1-\eta)^N (p'_{\text{good}} |R(\theta)\rangle\langle R(\theta)| + p'_{\text{bad}} Z|R(\theta)\rangle\langle R(\theta)|Z) \\ &+ (1-(1-\eta)^N)\mathcal{F}'_i(\rho_{\#N} \otimes \rho(\theta)^{\otimes 2N}).\end{aligned} \tag{46}$$

The worst case scenario is that $\mathcal{F}'_i$ always generates an error, adding a $(1-(1-\eta)^N)$ contribution to the error term. Therefore, after renormalising the error probability is bounded by

$$\begin{aligned}\epsilon' &\leq \frac{(1-\eta)^N p'_{\text{bad}} + (1-(1-\eta)^N)}{(1-\eta)^N (p'_{\text{good}} + p'_{\text{bad}}) + (1-(1-\eta)^N)} \\ &\leq \frac{(1-\eta)^N ((2N-1)(1-\epsilon_{\#})\epsilon_\theta^2 + \epsilon_{\#}) + (1-(1-\eta)^N)}{(1-\eta)^N (1-\epsilon_{\#})(p_{\text{good}} + p_{\text{bad}}) + (1-(1-\eta)^N)}\end{aligned}$$

The result scales as $O(\epsilon_{\#})$, but this error rate is itself the output of performing the synthillation protocol using noisy $|R(\pi/8)\rangle$-states of error rate $\epsilon_{\pi/8}$. In particular, Eq. (128) of Ref. [20] shows that

$$\epsilon_{\#} \leq 1 - \frac{2(1-\epsilon_{\pi/8})^{4N+4}}{1+(1-2\epsilon_{\pi/8})^{4N+4}} = (6+14N+8N^2)\epsilon_{\pi/8}^2 + O(\epsilon_{\pi/8}^3) \tag{47}$$

This suffices to conclude that

$$\epsilon' \leq (2N-1)\epsilon_\theta^2 + (6+14N+8N^2)\epsilon_{\pi/8}^2 + N\eta + O(\epsilon_\theta^3, \epsilon_{\pi/8}^3, \eta^2, \ldots) \tag{48}$$

where $O(\epsilon^3)$ collects all higher order terms. For instance, for $N=1$ and $N=2$ this yields

$$\epsilon' \leq \begin{cases} \epsilon_\theta^2 + 28\epsilon_{\pi/8}^2 + \eta + O(\epsilon_\theta^3, \epsilon_{\pi/8}^3, \eta^2, \ldots) & \text{for } N=1 \\ 3\epsilon_\theta^2 + 66\epsilon_{\pi/8}^2 + 2\eta + O(\epsilon_\theta^3, \epsilon_{\pi/8}^3, \eta^2, \ldots) & \text{for } N=2 \end{cases} \tag{49}$$

Comparing this with the numerical expressions Eq. (27) and Eq. (29), we see the analytic upper bound is very loose and grossly overestimates the prefactors.

5 No small triorthogonal codes

Many other distillation protocols are based upon projections into codespaces with a transversal non-Clifford, with transversality proofs typically using some notion of triorthogonal matrices [4]. While the protocols proposed here do not manifestly have this form, it is natural to ask whether there is some codespace projection with equivalent performance. Indeed, Jones' first-level distiller protocol [5] is effectively equivalent to projecting onto the codespace of Bravyi-Haah triorthogonal codes [4], and Haah has recently introduced level-lifting as a general methodology for finding such equivalences [16]. Furthermore, it has long been known that for any distillation protocol there exists a codespace projection

that achieves the same error suppression [24], though it may not achieve the same success probability or admit a transversal non-Clifford gate.

In this section, we show that there exist no triorthogonal codes with fewer than 14 qubits. This bound is tight since the smallest Bravyi-Haah code is a 14 qubit triorthogonal construction. It follows that the $10 \to 2$ MEK protocol is not equivalent to a projection onto a triorthogonal code. What distinguishes MEK is that it is a highly compressed circuit that is obtained from taking a larger circuit and cancelling some T-gates. This suggests that something happens during the compression process of eliminating extraneous T-gates that breaks the equivalence to triorthogonal codes. Since our protocols can be understood as a generalisation of MEK protocols, it seems unlikely similar performance parameters will be achievable using projections onto codes with exotic transversality properties.

We present the definition of triorthogonality

Definition 1 *(Def 1. of Ref. [4])A binary matrix G of size $m \times n$ is called triorthogonal iff the supports of any pair and any triple of its rows have even overlap, that is,*

$$\sum_{j=1}^{n} G_{a,j} G_{b,j} = 0 \pmod 2 \tag{50}$$

for all pairs of rows $1 \leq a < b \leq m$ and

$$\sum_{j=1}^{n} G_{a,j} G_{b,j} G_{c,j} = 0 \pmod 2 \tag{51}$$

for all triples of rows $1 \leq a < b < c \leq m$.

The definition of triorthogonality allows a matrix to have either odd or even rows, and it is standard to use a horizontal line to demarcate the split

$$G = \left(\frac{G_1}{G_0}\right), \tag{52}$$

so G_1 contains odd weight rows and G_0 contains even weight rows. Assuming G is row-wise linearly independent, it describes an $[[n, k, d]]$ quantum code where: n is the number of columns in G; k is the number of rows in G_1; and $d \geq 2$ if and only if G_0 is non-trivially supported on every column. We also use the notion of a biorthogonal matrix, which obeys the constraint for pairs of rows but not for triples of rows.

Let G be a triorthogonal matrix with block matrix form

$$G = \left(\frac{G_1}{G_0}\right) = \left(\begin{array}{cc} A & B \\ \hline C & D \\ \mathbf{1} & \mathbf{0} \end{array}\right), \tag{53}$$

where $\mathbf{1}$ and $\mathbf{0}$ are the all-1 and all-0 row vectors of appropriate width. Using column permutation the matrix can always be brought into this form. Without loss of generality, we assume that the last row has weight w and is the highest weight row in the span of G_0. Let B and D be width u, so that the total matrix width is $n = w + u$.

From triorthogonality of G, it follows that the submatrix

$$L = \left(\begin{array}{c} A \\ C \end{array}\right), \tag{54}$$

is biorthogonal with all rows being even weight and that

$$R = \left(\frac{B}{D} \right), \tag{55}$$

is biorthogonal with B containing odd weight rows. Since B contains odd weight rows, the matrix D cannot contain the all-1 vector as this would violate biorthogonality. However, since the code is distance 2, the matrix D must be supported on every column. Therefore, there must exist at least 2 non-trivial rows in D. The smallest possible width for D is then achieved by

$$D = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}, \tag{56}$$

which has width $u = 6$ and contains weight 4 vectors, and so we can infer that $w \geq 4$. Other D are possible, but one cannot obtain smaller parameters: There are at least two rows and for any pair of rows they must overlap on at least 2 columns and also have support on at least 2 other non-overlapping columns. From this we see that $n = w+u \geq 4+6 = 10$. So this is already enough to prove there are no $[[n, k, 2]]$ triorthogonal codes with $k \leq 1$ and n less than 10.

However, the bound $w \geq 4$ was obtained based on the rows of D only, but w is the max weight across all rows in the span of G_0. If C contains a row of weight w_c, then we know that $w \geq w_c + 4$. For any row of C we can add the last row of G_0, which generates a row of weight $w_c' = w - w_c$, entailing that $w \geq w_c' + 4 = w - w_c + 4$ and so $w_c \geq 4$. Putting this together yields $w \geq 8$ and $u \geq 6$ so that $n \geq 14$. There are no triorthogonal codes with fewer than 14 qubits.

6 Discussion

6.1 Variation of the two-step protocol

This subsection discusses one possible variant of the two-step protocol. Consider a quantum algorithm that needs many magic states of the form $|R(\theta)\rangle$, but with different values of θ. As presented, our main protocol cannot be used to full effect as the very large N limit assumes that we need many magic states with the same θ. However, it is straightforward to check that one can distill pairs of states with the same θ_j. That is, we may input states of the form $\bigotimes_{j=1}^{N} \rho(\theta_j)^{\otimes 2}$ and use N pivotal rotations with corresponding angles $2\theta_j$.

6.2 Quadratic vs higher order error suppression

In our introduction, we remarked on the existence of distillation routines that offer much larger reductions in errors without using concatenation [5, 6, 21, 22]. The appeal of these protocols is better asymptotic performance in the limit of large quantum computers and large error reduction. The analysis underpinning these results assumes that it is appropriate to quantify resource costs by the ratio of input to output states. But a more realistic picture is given by an involved full space time analysis; also accounting for the cost of Clifford gates and quantum error correction. In such an analysis, it is possible to scale the size of error correction codes between rounds of magic state distillation [25–27]. This scaling trick is extremely effective, and is arguably the most important tool in the arsenal of magic state distillation techniques. Although the idea has been known for some time, it has gone without a name. In an effort to popularise this trick, O'Gorman and Campbell recently proposed the phrase "balanced investment" [27].

Balanced investment relies on distinct rounds of magic state distillation with successive error reduction. Therefore, balanced investment is more compatible with protocols giving quadratic error reduction, such as presented here, than with the protocols of Refs. [5, 6]. This argument is qualitative, and we need detailed resource analyses to make concrete quantitative statements. However, such full resource investigations are difficult and time-consuming and have only been undertaken for the Reed-Muller and Bravyi-Haah protocols. Naturally, such a numerical investigation also falls outside the scope of this paper.

7 Conclusions

We presented a new two-step method of magic state distillation that is very competitive at preparing single-qubit magic states, offering a way to circumvent the need for costly gate-synthesis of single-qubit rotations. An important aspect of these new protocols are the preparation of multi-qubit magic states using synthillation. Pressing open questions include how these competing approach fare when all resource costs are considered, though such an analysis will depend heavily on the architecture considered.

We would also like to explore whether the synthillation driven techniques proposed here could be extended to protocols with larger than quadratic error reduction [5, 6]. After completing this work, Hastings and Haah proposed some new approaches to synthillation that may provide a starting point for attacking this problem [28].

8 Acknowledgements

This research was supported by the EPSRC (grant ref EP/M024261/1). We thank IBM and developers of the QISKit, which was used for numerical simulations.

References

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

[2] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Phys. Rev. A*, 71:022316, 2005. DOI: 10.1103/PhysRevA.71.022316.

[3] Adam M. Meier, Bryan Eastin, and Emanuel Knill. Magic-state distillation with the four-qubit code. *Quant. Inf. and Comp.*, 13:195, 2013.

[4] Sergey Bravyi and Jeongwan Haah. Magic-state distillation with low overhead. *Phys. Rev. A*, 86:052329, Nov 2012. DOI: 10.1103/PhysRevA.86.052329.

[5] Cody Jones. Multilevel distillation of magic states for quantum computing. *Phys. Rev. A*, 87:042305, Apr 2013. DOI: 10.1103/PhysRevA.87.042305.

[6] 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. ISSN 2521-327X. DOI: 10.22331/q-2017-10-03-31.

[7] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. Asymptotically optimal approximation of single qubit unitaries by clifford and t circuits using a constant number of ancillary qubits. *Phys. Rev. Lett.*, 110:190502, May 2013. DOI: 10.1103/PhysRevLett.110.190502.

[8] David Gosset, Vadym Kliuchnikov, Michele Mosca, and Vincent Russo. An algorithm for the t-count. *Quant. Inf. & Comp.*, 14(15-16):1261–1276, 2014.

[9] Neil J Ross and Peter Selinger. Optimal ancilla-free clifford+ t approximation of z-rotations. *Quant. Inf. and Comp.*, 16:901, 2016.

[10] Adam Paetznick and Krysta M Svore. Repeat-until-success: Non-deterministic decomposition of single-qubit unitaries. *Quant. Inf. & Comp.*, 14(15-16):1277–1301, 2014.

[11] Alex Bocharov, Martin Roetteler, and Krysta M. Svore. Efficient synthesis of probabilistic quantum circuits with fallback. *Phys. Rev. A*, 91:052317, May 2015. DOI: 10.1103/PhysRevA.91.052317.

[12] Andrew J Landahl and Chris Cesare. Complex instruction set computing architecture for performing accurate quantum z rotations with less magic. *arXiv preprint arXiv:1302.3240*, 2013. URL `https://arxiv.org/pdf/1302.3240.pdf`.

[13] Cody Jones. Distillation protocols for fourier states in quantum computing. *arXiv preprint arXiv:1303.3066*, 2013. URL `https://arxiv.org/pdf/1303.3066.pdf`.

[14] Guillaume Duclos-Cianci and David Poulin. Reducing the quantum-computing overhead with complex gate distillation. *Phys. Rev. A*, 91:042315, Apr 2015. DOI: 10.1103/PhysRevA.91.042315.

[15] Earl T Campbell and Joe O'Gorman. An efficient magic state approach to small angle rotations. *Quantum Science and Technology*, 1(1):015007, 2016. DOI: doi:10.1088/2058-9565/1/1/015007.

[16] Jeongwan Haah. Towers of generalized divisible quantum codes. *arXiv preprint arXiv:1709.08658*, 2017. URL `https://arxiv.org/pdf/1709.08658.pdf`.

[17] Cody Jones. Low-overhead constructions for the fault-tolerant toffoli gate. *Phys. Rev. A*, 87:022328, Feb 2013. DOI: 10.1103/PhysRevA.87.022328.

[18] Bryan Eastin. Distilling one-qubit magic states into toffoli states. *Phys. Rev. A*, 87: 032321, Mar 2013. DOI: 10.1103/PhysRevA.87.032321.

[19] Earl T. Campbell and Mark Howard. Unifying gate synthesis and magic state distillation. *Phys. Rev. Lett.*, 118:060501, Feb 2017. DOI: 10.1103/PhysRevLett.118.060501.

[20] Earl T. Campbell and Mark Howard. Unified framework for magic state distillation and multiqubit gate synthesis with reduced resource cost. *Phys. Rev. A*, 95:022316, Feb 2017. DOI: 10.1103/PhysRevA.95.022316.

[21] Jeongwan Haah, Matthew B Hastings, D Poulin, and D Wecker. Magic state distillation at intermediate size. *Quant. Inf. and Comp.*, 18:0114, 2018.

[22] Matthew B. Hastings and Jeongwan Haah. Distillation with sublogarithmic overhead. *Phys. Rev. Lett.*, 120:050504, Jan 2018. DOI: 10.1103/PhysRevLett.120.050504.

[23] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402: 390, 1999. DOI: 10.1038/46503.

[24] Earl T. Campbell and Dan E. Browne. On the structure of protocols for magic state distillation. *Lecture Notes in Computer Science*, 5906:20, 2009. DOI: 10.1007/978-3-642-10698-9_3. arXiv:0908.0838.

[25] R. Raussendorf, J. Harrington, and K. Goyal. A fault-tolerant one-way quantum computer. *Annals of Physics*, 321(9):2242 – 2270, 2006. ISSN 0003-4916. DOI: 10.1016/j.aop.2006.01.012.

[26] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A*, 86: 032324, Sep 2012. DOI: 10.1103/PhysRevA.86.032324.

[27] Joe O'Gorman and Earl T. Campbell. Quantum computation with realistic magic-state factories. *Phys. Rev. A*, 95:032338, Mar 2017. DOI: 10.1103/PhysRevA.95.032338.

[28] Jeongwan Haah and Matthew B Hastings. Codes and protocols for distilling t, controlled-s, and toffoli gates. *arXiv preprint arXiv:1709.02832*, 2017. URL `https://arxiv.org/pdf/1709.02832.pdf`.