"Magic" ancillas and their distillation

Main principles

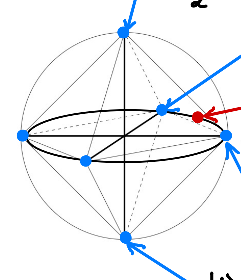
-- Certain states, e.g. $|A\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle\right)$ or $|H\rangle = \left(\cos\frac{\pi}{8}\right)|0\rangle + \left(\sin\frac{\pi}{8}\right)|1\rangle = \underbrace{e^{-i\frac{\pi}{8}} K H\, |A\rangle}_{\text{Clifford-equivalent to } |A\rangle}$
together with Clifford+classical operations provide computational universality.

-- An imperfect version of such a state can be prepared and then encoded with some stabilizer code.

-- It is intrinsically easier to deal with errors in a specific state than in qubits carrying some information. If an error is detected, we may discard the state and start from scratch.

Non-universal and potentially universal states: the Bloch sphere picture

A general single-qubit state: $\rho_{\vec{n}} = \frac{1 + \vec{n}\cdot\vec{\sigma}}{2}$



The vertices of the octahedron are stabilizer states. The states inside the octahedron are mixtures of stabilizer states. Applying Clifford operations to them gives mixtures of multi-qubit stabilizer states, and this sort of computation can be efficiently simulated classically.

But this state lies outside the octahedron:

$$\boxed{|A\rangle\langle A| = \frac{1+A}{2}} = \rho_{\vec{n}}, \quad \text{where} \quad \vec{n} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}$$

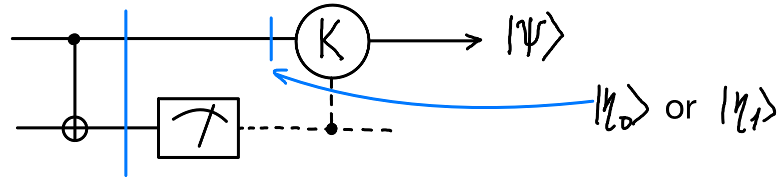
$$\boxed{\begin{aligned} A &= \frac{1}{\sqrt{2}}\left(\sigma^x + \sigma^y\right) = T\sigma^x T^{-1} \\ &= e^{-i\frac{\pi}{4}} K \sigma^x \quad \text{(Clifford but not Pauli)}\end{aligned}}$$

Therefore, it *may* be universal.

Implementation of T uzing $|A\rangle$ and Clifford operations

$$|\xi\rangle = c_0|0\rangle + |c_1\rangle|1\rangle$$

$$|A\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle\right)$$



$$|\eta\rangle = \frac{1}{\sqrt{2}}\left(c_0|00\rangle + c_0 e^{i\frac{\pi}{4}}|01\rangle + c_1|11\rangle + c_1 e^{i\frac{\pi}{4}}|10\rangle\right)$$

Depending on the measurement outcome, $|\eta\rangle$ collapses to $|\eta_0\rangle$ or $|\eta_1\rangle$

$$\sqrt{p_0}\,|\eta_0\rangle = \left(I \otimes \langle 0|\right)|\eta\rangle = \frac{1}{\sqrt{2}}\left(c_0|0\rangle + c_1 e^{i\frac{\pi}{4}}|1\rangle\right) = \frac{1}{\sqrt{2}}\,T|\xi\rangle$$

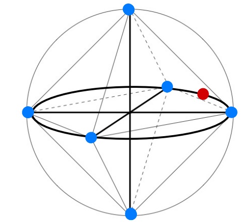
post-measurement states $|\eta_0\rangle, |\eta_1\rangle$ up to phase

$$\sqrt{p_1}\,|\eta_1\rangle = \left(I \otimes \langle 1|\right)|\eta\rangle = \frac{1}{\sqrt{2}}\left(c_0 e^{i\frac{\pi}{4}}|0\rangle + c_1|1\rangle\right) = \frac{e^{i\frac{\pi}{4}}}{\sqrt{2}}\,T^{-1}|\xi\rangle$$

$$\left(p_0 = p_1 = \frac{1}{2}\right)$$

Output state: $|\psi\rangle = K^a|\eta_a\rangle = T|\xi\rangle$ (up to phase)

What other states provide computational universality?

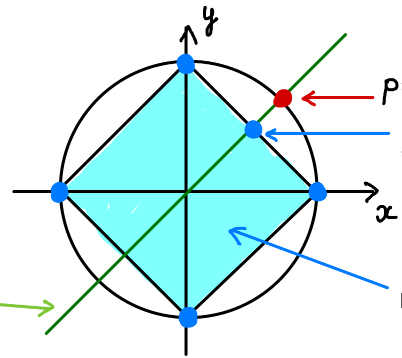


-- All pure states, except the 6 stabilizer states, are known to be universal

-- The general answer seems to be unknown, but we will study this question for special states of the form

$$\underline{\rho(p) = (1-p)\,|A\rangle\langle A| + p\,|-A\rangle\langle -A|}\,, \qquad |-A\rangle = \frac{|0\rangle - e^{i\frac{\pi}{4}}|1\rangle}{\sqrt{2}} = \sigma^z |A\rangle$$

Equatorial cross section of the Bloch sphere:



$p=0$

$p = p_* = \dfrac{\sqrt{2}-1}{2\sqrt{2}} \approx 0.146$

not universal

$\rho(p)$ lies on the green line

We can hope that the states $\rho(p)$ for $0 \le p < p_*$ are universal, and this is indeed true.

Any state can be converted to $\rho(p)$ **for some** p **by applying the superoperator** $\frac{1}{2} I \cdot I + \frac{1}{2} A \cdot A^{\dagger}$

$$\text{general state} = \rho_{++}|A\rangle\langle A| + \underbrace{\rho_{+-}|A\rangle\langle -A| + \rho_{-+}|-A\rangle\langle A|}_{\text{these terms are eliminated}} + \rho_{--}|-A\rangle\langle A|$$

Distillation of "magic ancillas"

$$\boxed{|A\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle\right) = T|+\rangle, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}}$$

<u>Available operations</u>: classically controlled Clifford unitaries, $|0\rangle$ ancillas and $\{|0\rangle, |1\rangle\}$ measurements.

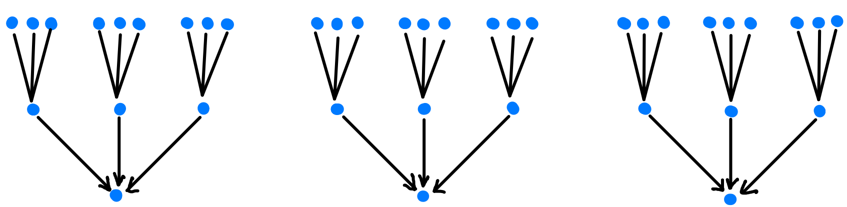
<u>Goal</u>: convert a number of imperfect copies of $|A\rangle$,

$$\boxed{\rho(p) = (1-p)\,|A\rangle\langle A| + p\,|-A\rangle\langle -A| = T\left((1-p)\,|+\rangle\langle +| + p\,|-\rangle\langle -|\right)T^{-1}}$$

into fewer copies of $\rho_{out} = \rho(p_{out})$ with an arbitrarily small error parameter p_{out}

Distillation scheme based on the 15-qubit code $\mathcal{M} = CSS(\underbrace{RM''(4,2)}_{D_z}, \underbrace{RM''(4,1)}_{D_x})$ (which is invariant under $T^{\otimes 15}$)

<u>Idea</u>: Prepare the input state $\rho(p)^{\otimes 15}$ and push it to the code subspace as if we were correcting or detecting errors. We will get something very close to the logical state $|A\rangle\langle A|$.



N copies of $\rho(p)$

$\approx \frac{N}{15}$ copies of $\rho(f(p))$, where $f(p) \ll p$

$\approx \frac{N}{15^2}$ copies of $\rho(f(f(p)))$, ...

The 15-to-1 distillation procedure (can be iterated)

$$A=\begin{pmatrix}0 & e^{-i\pi/4}\\ e^{i\pi/4} & 0\end{pmatrix}=T\sigma^x T^{-1}$$

Note that T commutes with σ^z

1) Take 15 copies of $\rho(p)$ and measure the Z syndrome μ of the code $\mathcal{M}$.

2) Find the corresponding "error" $g=g(\mu)$; apply $A(g):=A^{g_1}\otimes\cdots\otimes A^{g_{15}}$ (similar to $\sigma^x(g)$)

3) Measure the X syndrome ν. If $\nu\neq 0$, the distillation procedure has failed.

4) If $\nu=0$, decode the logical state. Apply $K=\begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}$.

Analysis of steps 1 and 2

$\underbrace{\mu=(\mu_1,\dots,\mu_{10})}_{\sigma^z(f)|\psi\rangle=(-1)^{\mu_j}|\psi\rangle \text{ for basis vectors } f_j\in D_z}$, $g=g(\mu)$ satisfies the condition $(g,f_j)=\mu_j$ for all j (g is defined up to an arbitrary element of $D_z^\perp$)

$$I=\sum_\mu \Pi_\mu^{(z)},\qquad \Pi_\mu^{(z)}=\prod_j\frac{1+(-1)^{\mu_j}\sigma^z(f_j)}{2}\qquad \Pi_\mu^{(z)}=\sigma^x(g)\,\Pi_0^{(z)}\sigma^x(g)=\boxed{A(g)\,\Pi_0^{(z)}A(g)\quad\text{for}\quad g=g(\mu)}$$

$$\rho_{in}=\rho(p)^{\otimes 15}\longmapsto\sum_\mu A(g(\mu))\left(\Pi_\mu^{(z)}\rho_{in}\Pi_\mu^{(z)}\right)A(g(\mu))^\dagger=\Pi_0^{(z)}\Big(\sum_\mu \underbrace{A(g(\mu))\,\rho_{in}\,A(g(\mu))^\dagger}_{=\rho_{in}}\Big)\Pi_0^{(z)}$$

Analysis of the distillation procedure (cont.)

$$\rho(p) = \underbrace{\big((1-p)\, I\cdot I + p\, Z\cdot Z\big)}_{\text{error superoperator}}\ \underbrace{|A\rangle\langle A|}_{T|+\rangle\langle+|T^{-1}}$$

Steps 1,2 yield $\rho_2 = |D_z|\cdot \Pi_0^{(z)}\rho_{in}\Pi_0^{(z)}$, $\qquad |D_z| = 2^{10}$

where $\rho_{in} = \rho(p)^{\otimes 15} = T^{\otimes n}\underbrace{\left(\sum_{h\in\mathbb{F}_2^n}(1-p)^{n-|h|}p^{|h|}\, Z(h)\,\big(|+\rangle\langle+|\big)^{\otimes n} Z(h)\right)}_{\gamma}\big(T^{\otimes n}\big)^{-1}$, $\qquad n = 15$

Step 3

$$\rho_3 = \Pi_0^{(x)}\rho_2\Pi_0^{(x)} = |D_z|\,\underbrace{\Pi_0^{(x)}\Pi_0^{(z)}}_{\Pi}\,T^{\otimes n}\gamma\big(T^{\otimes n}\big)^{-1}\underbrace{\Pi_0^{(z)}\Pi_0^{(x)}}_{\Pi} = T^{\otimes n}\underbrace{\big(|D_z|\,\Pi\gamma\Pi\big)}_{\tilde\gamma}\big(T^{\otimes n}\big)^{-1}$$

Key fact: Π (projector onto $\mathcal{M}$) commutes with $T^{\otimes n}$

$\tilde\gamma$ is a sum of $|D_z|\cdot\Pi_0^{(x)}\Pi_0^{(z)}Z(h)\,|+\rangle^{\otimes n}\langle+|^{\otimes n}Z(h)\,\Pi_0^{(z)}\Pi_0^{(x)} = \boxed{\Pi_0^{(x)}Z(h)\,|+\rangle_L\langle+|_L\,Z(h)\,\Pi_0^{(x)}}$

We have used the fact that $\boxed{\sqrt{|D_z|}\,\Pi_0^{(z)}|+\rangle^{\otimes n} = |+\rangle_L}$

because $\Pi_0^{(z)} = \sum_{g\in D_z^\perp}|g\rangle\langle g|$, $\qquad \Pi_0^{(z)}|+\rangle^{\otimes n} = 2^{-n/2}\sum_{g\in D_z^\perp}|g\rangle = \frac{1}{\sqrt{|D_z|}}|+\rangle_L$

The outcome depends on the type of Z-error h present in the input state:

$$\Pi_0^{(x)}Z(h)\,|+\rangle_L = \begin{cases} |+\rangle_L & \text{if } h\in D_z \text{ (trivial error)}\\ 0 & \text{if } h\notin D_x^\perp \text{ (detectable error)}\\ |-\rangle_L & \text{if } h\in D_z+[1] \text{ (bad error)}\end{cases}$$

Completing the analysis

Step 3 yields the unnormalized state

$$\rho_3 = \sum_{h\in\mathbb{F}_2^n} (1-p)^{n-|h|}\, p^{|h|}\, |\Psi_h\rangle\langle\Psi_h|$$

$$(n=15)$$

$$|\Psi_h\rangle = \begin{cases} T^{\otimes 15}|+\rangle_L & \text{if } h\in D_Z \text{ (trivial error)}\\ 0 & \text{if } h\notin D_X^{\perp} \text{ (detectable error)}\\ T^{\otimes 15}|-\rangle_L & \text{if } h\in D_Z+[1] \text{ (bad error)}\end{cases}$$

$T^{\otimes 15}$ realizes T_L^{-1}

Step 4 (decoding and the application of $K=T^2$):

$$T^{\otimes 15}|+\rangle_L \xrightarrow{\text{decode}} T^{-1}|+\rangle \overset{K}{\mapsto} T|+\rangle = |A\rangle$$

Final state

$$T^{\otimes 15}|-\rangle_L \xrightarrow{\text{decode}} T^{-1}|-\rangle \overset{K}{\mapsto} T|-\rangle = |{-A}\rangle$$

$$\rho_{out} = \underbrace{\left(\sum_{h\in D_Z}(1-p)^{n-|h|}p^{|h|}\right)}_{f_+(p)=W_{D_Z}(1-p,\,p)\ \text{(weight enumerator)}} |A\rangle\langle A| + \underbrace{\left(\sum_{h\in D_Z+[1]}(1-p)^{n-|h|}p^{|h|}\right)}_{f_-(p)=f_+(1-p)} |{-A}\rangle\langle {-A}|$$

MacWilliams identity

$$W_C(x,y) = \frac{1}{|C|}\, W_{C^\perp}(x+y,\, x-y)$$

Counting errors

$$D_Z = \text{lin. span}\,\{[x_1],[x_2],[x_3],[x_4],[x_1x_2],[x_1x_3],[x_1x_4],[x_2x_3],[x_2x_4],[x_3x_4]\}$$

$$D_Z^{\perp} = \text{lin. span}\,\{[1],[x_1],[x_2],[x_3],[x_4]\}$$

$$W_{D_Z^\perp}(u,v) = \sum_{g\in D_Z^\perp} u^{n-|g|}v^{|g|} = \underbrace{u^{15}}_{|g|=0} + 15\,\underbrace{u^{7}v^{8}}_{|g|=8} + \underbrace{v^{15}}_{|g|=15} + 15\,\underbrace{u^{8}v^{7}}_{|g|=7} \Rightarrow$$

$$f_\pm(p) = 2^{-5}\left(1 \pm 15(1-2p)^7 + 15(1-2p)^8 \pm (1-2p)^{15}\right)$$

Conclusions

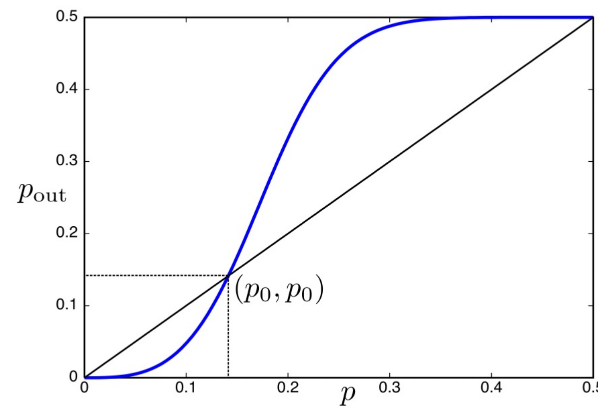
The distillation procedure transforms 15 copies of $\rho(p)$ into 1 copy of

$$\boxed{\rho_{out} = w(p)\cdot\rho(f(p))}$$

where $w(p)$ is the success probability = Pr [undetectable *Z*-error] ,

$$w(p) = \underbrace{f_+(p)}_{\text{Pr [trivial error]}} + \underbrace{f_-(p)}_{\text{Pr [bad error]}} = \frac{1}{16}\left(1+15(1-2p)^8\right) = 1 - O(p)$$

New error parameter: $f(p) = \dfrac{f_-(p)}{w(p)} = 35\underbrace{p^3}_{\text{because } d(D_z)=3} + O(p^4)$



Threshold

The error parameter *p* is below threshold if the sequence $p,\ f(p),\ f(f(p)),\ \ldots$ converges to 0.

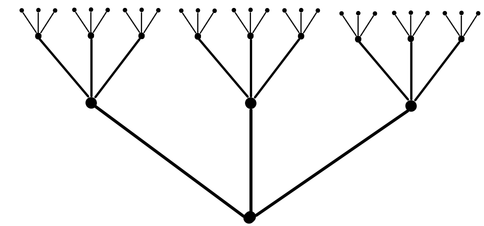
This condition is satisfied if $p < p_0$,

$$\boxed{p_0 \approx 0.141 < p_* = \frac{\sqrt{2}-1}{2\sqrt{2}} \approx 0.146}$$

Multi-level scheme and asymptotic overhead

At each distillation level, the number of ancillas N and their error parameter p are transformed as follows:

$$N \mapsto \underbrace{\frac{w(p)}{15}}_{\text{yield } \approx \frac{1}{15}} N, \qquad p \to c\, p^3 \qquad \text{(and hence, } \sqrt{c}\, p \mapsto (\sqrt{c}\, p)^3 \text{)}$$



k levels: $$p \to p_{out} \sim \frac{1}{\sqrt{c}} (\sqrt{c}\, p)^{3^k}, \qquad \text{yield } \frac{1}{n}$$

$$n \sim 15^k \sim \left(\frac{\ln(\sqrt{c}\, p_{out})}{\ln \sqrt{c}\, p} \right)^{\log_3 15} \sim \boxed{\left(\ln \frac{1}{p_{out}} \right)^{\gamma}, \quad \gamma = \log_3 15}$$

Different distillation schemes

1) Based on Steane's code (Reichardt): 7 to 1 (when successful); optimal threshold but low yield

2) The scheme just described (Knill; Bravyi and Kitaev): 15 to 1; slightly lower threshold but less overhead

3) Meier, Eastin, and Knill: 10 to 2; even less overhead, though the threshold is low

4) Bravyi and Haah: A family of protocols with high asymptotic efficiency