Supplementary material for "Quantifying magic for multi-qubit operations", published in Proceedings of the Royal Society A, 20190251. http://dx.doi.org/10.1098/rspa.2019.0251

James R. Seddon[1] and Earl T. Campbell[2]

[1] *Department of Physics and Astronomy, University College London, London, UK** and

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

(Dated: June 27, 2019)

I. PROPERTIES OF ROBUSTNESS OF THE CHOI STATE

Here we confirm that the robustness of the Choi state, $\mathcal{R}(\Phi_{\mathcal{E}})$ has the properties convexity and submultiplicativity under tensor product. We then give an example to show that it is not submultiplicative under composition.

Convexity: This follows immediately from convexity of robustness of magic. Consider a real linear combination of n-qubit channels: $\mathcal{E} = \sum_k q_k \mathcal{E}_k$. The Choi state for $\mathcal{E}$ is:

$$\Phi_{\mathcal{E}} = (\mathcal{E} \otimes \mathbb{1}_n) |\Omega_n\rangle\langle\Omega_n| \tag{S1}$$

$$= \sum_k q_k (\mathcal{E}_k \otimes \mathbb{1}_n) |\Omega_n\rangle\langle\Omega_n| = \sum_k q_k \Phi_{\mathcal{E}_k}, \tag{S2}$$

where in the last line we identified $(\mathcal{E}_k \otimes \mathbb{1}_n) |\Omega\rangle\langle\Omega|$ as the Choi state for $\mathcal{E}_k$. Then by convexity of robustness of magic:

$$\mathcal{R}(\Phi_{\mathcal{E}}) \leq \sum_k |q_k| \mathcal{R}(\Phi_{\mathcal{E}_k}), \tag{S3}$$

which shows $\mathcal{R}(\Phi_{\mathcal{E}})$ is convex in $\mathcal{E}$.

Submultiplicativity under tensor product: The maximally entangled state $|\Omega_{n+m}\rangle^{AA'|BB'}$ as defined by equation (4.7) in the main text can be factored as $|\Omega_{n+m}\rangle^{AA'BB'} = |\Omega_n\rangle^{A|B} |\Omega_m\rangle^{A'|B'}$. So the Choi state for a channel $\mathcal{E}^{AA'} = \mathcal{E}^A \otimes \mathcal{E}'^{A'}$, where $\mathcal{E}^A$ and $\mathcal{E}'^{A'}$ are respectively n-qubit and m-qubit channels, can be written:

$$\Phi_{\mathcal{E}} = \left(\mathcal{E}^A \otimes \mathcal{E}'^{A'} \otimes \mathbb{1}_{n+m}\right) |\Omega_{n+m}\rangle\langle\Omega_{n+m}|^{AA'|BB'} \tag{S4}$$

$$= \left(\mathcal{E}^A \otimes \mathbb{1}_n\right) |\Omega_n\rangle\langle\Omega_n|^{A|B} \otimes \left(\mathcal{E}'^{A'} \otimes \mathbb{1}_m\right) |\Omega_m\rangle\langle\Omega_m|^{A'|B'} = \Phi_{\mathcal{E}^A} \otimes \Phi_{\mathcal{E}'^{A'}}. \tag{S5}$$

Then by submultiplicativity of robustness of magic for states, we have:

$$\mathcal{R}\left(\Phi_{\mathcal{E}^A \otimes \mathcal{E}'^{A'}}\right) \leq \mathcal{R}(\Phi_{\mathcal{E}^A}) \mathcal{R}\left(\Phi_{\mathcal{E}'^{A'}}\right),$$

which is the desired property.

Failure of submultiplicativity under composition: Let $\mathcal{E}_1$ be the single-qubit Z-reset channel defined by Kraus operators $\{|0\rangle\langle 0|, |0\rangle\langle 1|\}$, and let $\mathcal{E}_2$ be the conditional channel defined by $\{|T\rangle\langle 0|, |1\rangle\langle 1|\}$, where $|T\rangle = T|+\rangle$. These channels respectively have Choi states $\Phi_{\mathcal{E}_1} = |0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2}$, and $\Phi_{\mathcal{E}_2} = \frac{1}{2}(|T0\rangle\langle T0| + |11\rangle\langle 11|)$, with robustness of magic $\mathcal{R}(\Phi_{\mathcal{E}_1}) = 1$ and $\mathcal{R}(\Phi_{\mathcal{E}_2}) \approx 1.207$.

* james.seddon.15@ucl.ac.uk

The composed channel $\mathcal{E}_2 \circ \mathcal{E}_1$ has a Kraus representation $\{|T\rangle\langle 0|, |T\rangle\langle 1|\}$, and so has a Choi state $\Phi_{\mathcal{E}_2\circ\mathcal{E}_1} = |T\rangle\langle T| \otimes \frac{\mathbb{1}}{2}$, with $\mathcal{R}(\Phi_{\mathcal{E}_2\circ\mathcal{E}_1}) \approx 1.414 > \mathcal{R}(\Phi_{\mathcal{E}_2})\mathcal{R}(\Phi_{\mathcal{E}_1})$. So it is not the case that the robustness of the Choi state is submultiplicative under composition.

More intuitively, such counterexamples arise for channels $\mathcal{E}$ where the stabilizer state $|\phi_*\rangle$ that results in maximal final robustness $\mathcal{R}[(\mathcal{E} \otimes \mathbb{1}_n) |\phi_*\rangle\langle\phi_*|]$ is not the maximally entangled state $|\Omega_n\rangle$, as then we can always boost the output robustness by using a stabilizer-preserving operation to prepare $|\phi_*\rangle$ before applying $\mathcal{E}$.

II. PROPERTIES OF CHANNEL ROBUSTNESS

Faithfulness: Suppose $\mathcal{E}$ is an n-qubit CPTP map. There are two cases:

(i) $\mathcal{E} \in \mathrm{SP}_{n,n}$. In this case, $\Phi_{\mathcal{E}}$ is itself a mixed stabilizer state, and since $\mathcal{E}$ is trace-preserving it satisfies $\mathrm{Tr}(\Phi_{\mathcal{E}}) = \mathbb{1}_n/2^n$. So $\Phi_{\mathcal{E}}$ is already trivially a decomposition of the correct form, with $p = 0$, so that $R_*(\mathcal{E}) = 1 + 2p = 1$.

(ii) $\mathcal{E} \notin \mathrm{SP}_{n,n}$. Then by faithfulness of robustness of the Choi state (Lemma 4.2 in the main text), $\Phi_{\mathcal{E}}$ has $\mathcal{R}(\Phi_{\mathcal{E}}) > 1$. Since the definition of $\mathcal{R}_*$ is a restriction of $\mathcal{R}(\Phi)$, it must be the case that $\mathcal{R}(\Phi_{\mathcal{E}}) \leq \mathcal{R}_*(\mathcal{E})$. Therefore $\mathcal{R}_*(\mathcal{E}) > 1$.

Convexity: Suppose we have a set of Choi states $\Phi_{\mathcal{E}_j}$ corresponding to channels $\mathcal{E}_j$, with optimal decompositions:

$$\Phi_{\mathcal{E}_j} = (1 + p_j)\rho_{j_+} - p_j\rho_{j_-}, \tag{S6}$$

where each $\rho_{j_\pm}$ separately satisfies the condition $\mathrm{Tr}_A(\rho_\pm) = \frac{\mathbb{1}_n}{2^n}$, so that $\mathcal{R}_*(\mathcal{E}_j) = 1 + 2p_j$. Now take a real linear combination of such channels:

$$\mathcal{E} = \sum_i q_i\mathcal{E}_i = \sum_{j\in P} q_j\mathcal{E}_j + \sum_{k\in N} q_k\mathcal{E}_k, \tag{S7}$$

where P is the set of indices such that $q_j \geq 0$, and N is the set such that $q_k < 0$. We assume that $\sum_i q_i = 1$ so that $\mathrm{Tr}(\Phi_{\mathcal{E}}) = 1$. Then the corresponding Choi state for channel $\mathcal{E}$ is:

$$\Phi_{\mathcal{E}} = \sum_{j\in P} q_j\Phi_{\mathcal{E}_j} - \sum_{k\in N} |q_k|\Phi_{\mathcal{E}_k} \tag{S8}$$

$$= \sum_{j\in P} q_j\left[(1+p_j)\rho_{j_+} - p_j\rho_{j_-}\right] - \sum_{k\in N} |q_k|\left[(1+p_k)\rho_{k_+} - p_k\rho_{k_-}\right] \tag{S9}$$

$$= \left(\sum_{j\in P} q_j(1+p_j)\rho_{j_+} + \sum_{k\in N} |q_k|p_k\rho_{k_-}\right) - \left(\sum_{j\in P} q_jp_j\rho_{j_-} + \sum_{k\in N} |q_k|(1+p_k)\rho_{k_+}\right). \tag{S10}$$

Note that the terms inside the brackets all have positive coefficients, hence we can interpret as non-normalized mixtures over stabilizer states. To normalize them we can define:

$$\widetilde{\rho}_+ = \frac{\sum_{j\in P} q_j(1+p_j)\rho_{j_+} + \sum_{k\in N} |q_k|p_k\rho_{k_-}}{\sum_{j\in P} q_j(1+p_j) + \sum_{k\in N} |q_k|p_k}, \tag{S11}$$

and

$$\widetilde{\rho}_- = \frac{\sum_{j\in P} q_jp_j\rho_{j_-} + \sum_{k\in N} |q_k|(1+p_k)\rho_{k_+}}{\sum_{j\in P} q_jp_j + \sum_{k\in N} |q_k|(1+p_k)}. \tag{S12}$$

Then writing:

$$\widetilde{p} = \sum_{j\in P} q_j p_j + \sum_{k\in N} |q_k|(1+p_k), \tag{S13}$$

one can check that:

$$1+\widetilde{p} = \sum_{j\in P} q_j(1+p_j) + \sum_{k\in N} |q_k| p_k. \tag{S14}$$

This allows us to rewrite the Choi state as: $\Phi_{\mathcal{E}} = (1+\widetilde{p})\widetilde{\rho}_+ - \widetilde{p}\widetilde{\rho}_-$. Since $\widetilde{\rho}_\pm$ are convex mixtures over stabilizer states satisfying $\mathrm{Tr}_A(\rho_{j\pm}) = \frac{\mathbb{1}_n}{2^n}$, they must satisfy the same condition. We also know that $\widetilde{p} \geq 0$, so it is clear that the decomposition is in the form required for the definition of $\mathcal{R}_*$, except that it is not necessarily optimised to minimise $1+2\widetilde{p}$. So we have:

$$\mathcal{R}_*\left(\sum_j q_j \mathcal{E}_j\right) \leq 1+2\widetilde{p} = \sum_{j\in P} q_j(1+p_j) + \sum_{k\in N}|q_k| + \sum_{j\in P} q_j p_j + \sum_{k\in N}|q_k|(1+p_k) \tag{S15}$$

$$= \sum_{j\in P}|q_j|(1+2p_j) + \sum_{k\in N}|q_k|(1+2p_k) \tag{S16}$$

$$= \sum_i |q_i| \mathcal{R}_*(\mathcal{E}_j), \tag{S17}$$

which gives us the required result.

Invariance under tensor with identity: In Section 5 of the main text we saw that $\mathcal{R}_*\big(\mathcal{E}^A \otimes \mathbb{1}\big) \leq \mathcal{R}_*\big(\mathcal{E}^A\big)$. We now complete the proof that $\mathcal{R}_*\big(\mathcal{E}^A \otimes \mathbb{1}\big) = \mathcal{R}_*\big(\mathcal{E}^A\big)$ by showing that $\mathcal{R}_*\big(\mathcal{E}^A\big) \leq \mathcal{R}_*\big(\mathcal{E}^A \otimes \mathbb{1}\big)$.

Consider an optimal decomposition for $\Phi_{\mathcal{E}^A \otimes \mathbb{1}_m} = (1+p')\rho'_+ - p'\rho'_-$, such that $\mathcal{R}_*(\mathcal{E}^A \otimes \mathbb{1}_m) = 1+2p'$, where $\mathrm{Tr}_{AA'}(\rho_\pm) = \mathbb{1}_{n+m}/2^{n+m}$. Here we do not assume that $\rho'_\pm$ are products across the partition $AB|A'B'$, as was the case in equation (5.12) in the main text. However, we have seen from equation (5.11) in the main text that $\Phi_{\mathcal{E}^A \otimes \mathbb{1}_m}$ can be written as a product, so that by tracing out systems $A'B'$ we obtain:

$$\Phi_{\mathcal{E}^A} = (1+p')\,\mathrm{Tr}_{A'B'}(\rho'_+) - p'\,\mathrm{Tr}_{A'B'}(\rho'_-). \tag{S18}$$

Partial trace of a stabilizer state remains a stabilizer state, so this is a stabilizer decomposition. We just need to check that the partial trace condition holds, so we want to show:

$$\mathrm{Tr}_A(\mathrm{Tr}_{A'B'}(\rho'_\pm)) = \mathrm{Tr}_{AA'B'}(\rho'_\pm) = \frac{\mathbb{1}_n}{2^n}, \tag{S19}$$

but this is clearly the case from the fact that $\rho'_\pm$ were constrained such that $\mathrm{Tr}_{AA'}(\rho_\pm) = \mathbb{1}_{n+m}/2^{n+m}$. Hence again we have a valid, not necessarily optimal decomposition and:

$$\mathcal{R}_*(\mathcal{E}^A) \leq 1+2p' = \mathcal{R}_*(\mathcal{E}^A \otimes \mathbb{1}^B). \tag{S20}$$

Combining with the inequality $\mathcal{R}_*(\mathcal{E}^A \otimes \mathbb{1}^B) \leq \mathcal{R}_*(\mathcal{E}^A)$, shown in the main text, we obtain the equality:

$$\mathcal{R}_*(\mathcal{E} \otimes \mathbb{1}) = \mathcal{R}_*(\mathcal{E}) = \mathcal{R}_*(\mathbb{1} \otimes \mathcal{E}). \tag{S21}$$

III. OPTIMISATION PROBLEM FOR CHANNEL ROBUSTNESS

In Howard and Campbell [1], the optimisation problem for calculating robustness of magic for states was cast as follows:

$$\begin{aligned} \textbf{minimise} \quad & \|\vec{q}\|_1 \\ \textbf{subject to} \quad & A\vec{q} = \vec{b}, \end{aligned}$$

where $\vec{q}$ is a vector of coefficients, $\vec{b}$ is the vector of Pauli expectation values for the target state $\Phi_{\mathcal{E}}$, and A is a matrix whose columns are the Pauli vectors for the stabilizer states. For n-qubit channels, we have $2n$-qubit Choi states, so the number of generalised Paulis is $N_P = 4^{2n}$, and the number of stabilizer states is $N_S = 2^{2n}\prod_{j=1}^{2n}(2^j+1)$ [1]. Then $\vec{b}$ has N_P entries, $\vec{q}$ has N_S entries, and the dimension of A is $(N_P \times N_S)$. From this construction we can recover optimal decompositions of the form: $\Phi_{\mathcal{E}} = \sum_j q_j |\phi_j\rangle\langle\phi_j|$, where $\sum_j q_j = 1$ and $|\phi_j\rangle$ are the pure stabilizer states.

We want to restrict the problem to decompositions of the form:

$$\Phi_{\mathcal{E}} = (1+p)\rho_+ - p\rho_-, \tag{S22}$$

where $p \geq 0$ and $\rho_\pm$ correspond to trace-preserving channels, and can in general be mixed. Rather than enumerating all the extreme points of the set of stabilizer states corresponding to maps in $\mathrm{SP}_{n,n}$, it is more convenient to retain the same A matrix and modify the constraints. We still need to start from a finite set of extreme points, i.e. pure stabilizer states, so first rewrite as:

$$\Phi_{\mathcal{E}} = \sum_j q_{j+}\rho_j + \sum_j q_{j-}\rho_j = \sum_j p_{j+}\rho_j - \sum_j p_{j-}\rho_j, \tag{S23}$$

where q_{j+} are the positive quasiprobabilities, q_{j-} are the negative quasiprobabilities, and $p_{j\pm} = |q_{j\pm}|$. In the Pauli vector picture we can write this as $\vec{b} = A\vec{p}_+ - A\vec{p}_-$, where all the entries of $\vec{p}_\pm$ are non-negative. We define a new variable vector $\vec{p}$ which will have twice the length of the previous $\vec{q}$, i.e. $2N_S$ entries:

$$\vec{p} = \begin{pmatrix} \vec{p}_+ \\ \vec{p}_-, \end{pmatrix} \tag{S24}$$

and define a new $(N_P \times 2N_S)$ matrix A' in block form, $A' = \begin{pmatrix} A & -A \end{pmatrix}$. Then we have:

$$A'\vec{p} = \begin{pmatrix} A & -A \end{pmatrix} \begin{pmatrix} \vec{p}_+ \\ \vec{p}_- \end{pmatrix} = A\vec{p}_+ - A\vec{p}_- = \vec{b}. \tag{S25}$$

So now we need to minimise $\|\vec{p}\|_1 = \sum_j p_j$ subject to $A'\vec{p} = \vec{b}$ and $\vec{p} \geq 0$.

Next, we need the trace-preserving condition. Provided $\mathcal{E}$ is CPTP, if one part of the decomposition is trace-preserving, then the other will be as well, so we only need enforce the constraint on one of ρ_+ or ρ_-. Assume that we check ρ_+. The condition for a Choi state $\Phi^{AB} = \mathcal{E}^A \otimes \mathbb{1}^B(|\Omega\rangle\langle\Omega|^{AB})$ to be trace-preserving is:

$$\mathrm{Tr}_A(\Phi^{AB}) = \frac{\mathbb{1}}{d}, \tag{S26}$$

where d is the dimension of the subsystem. We need to convert this to a constraint on the vector $\vec{b}_+$ corresponding to ϕ_+, which is given by $\vec{b}_+ = A\vec{p}_+$. First, note that all Paulis are traceless except for the identity $P_0 = \mathbb{1}$, so for the maximally mixed state:

$$\langle P_j \rangle = \mathrm{Tr}\left(P_j \frac{\mathbb{1}}{d}\right) = \frac{\mathrm{Tr}(P_j)}{d} = \delta_{j,0}, \tag{S27}$$

so if the first entry in a Pauli vector is always $\langle \mathbb{1} \rangle$, the maximally mixed state has Pauli vector:

$$\vec{b}_B = \begin{pmatrix} 1 \\ \vec{0} \end{pmatrix}. \tag{S28}$$

where $\vec{0}$ is the zero vector. However, we need this to hold just for the reduced state on B rather than the full Pauli vector. Consider that if the whole state is written $\Phi^{AB} = \sum_{j,k} r_{j,k} P_j \otimes P_k$ for some set of coefficients $r_{j,k}$, then the expectation values are given by:

$$\langle P_l \otimes P_m \rangle = \sum_{j,k} r_{j,k} \,\mathrm{Tr}(P_l P_j \otimes P_m P_k) = \sum_{j,k} r_{j,k} d^2 \delta_{j,l} \delta_{m,k} = d^2 r_{l,m}. \tag{S29}$$

The reduced state is:

$$\mathrm{Tr}_A(\Phi^{AB}) = \sum_{j,k} r_{j,k} \,\mathrm{Tr}_A(P_j \otimes P_k) = \sum_{j,k} r_{j,k} d \delta_{j,0} P_k = d \sum_k r_{0,k} P_k. \tag{S30}$$

and the entries of the reduced Pauli vector will be:

$$\langle P_m \rangle = d \sum_k r_{0,k} \,\mathrm{Tr}(P_m P_k) = d^2 r_{0,m} = \langle P_0 \otimes P_m \rangle. \tag{S31}$$

So for condition (S26) to hold for the reduced state on B, we combine equations (S27) and (S31) to get:

$$\langle P_m \rangle = \langle P_0 \otimes P_m \rangle = \delta_{m,0}. \tag{S32}$$

That is, we just need to look at the entries of $\vec{b}_+$ corresponding to Paulis of the form $\mathbb{1} \otimes P_j$. These should all be zero except the first entry, which corresponds to $\langle \mathbb{1} \otimes \mathbb{1} \rangle$. Note that $\vec{b}_+ = A\vec{p}_+$ will in general not be normalised, but this does not matter, since we are only interested in whether or not entries are zero. We can use a binary matrix M to pick out the values of interest. As an example we consider the two-qubit case, and assume that the entries are ordered as:

$$\vec{b}_+ = \begin{pmatrix} \langle \mathbb{1} \otimes \mathbb{1} \rangle \\ \langle \mathbb{1} \otimes X \rangle \\ \langle \mathbb{1} \otimes Y \rangle \\ \langle \mathbb{1} \otimes Z \rangle \\ \langle X \otimes \mathbb{1} \rangle \\ \vdots \\ \langle Z \otimes Z \rangle \end{pmatrix}. \tag{S33}$$

Here, we are only interested in the 2nd, 3rd and 4th entries. We form a new vector $\vec{c}$ by left multiplying with M:

$$\vec{c} = M\vec{b}_+ = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \end{pmatrix} \vec{b}_+ = \begin{pmatrix} \langle \mathbb{1} \otimes X \rangle \\ \langle \mathbb{1} \otimes Y \rangle \\ \langle \mathbb{1} \otimes Z \rangle \end{pmatrix}. \tag{S34}$$

Then the condition we need is just $\vec{c} = 0$. To convert this to a condition on the $2N_S$-entry variable

$\vec{p} = \begin{pmatrix} \vec{p}_+ \\ \vec{p}_- \end{pmatrix}$, we first pad A with zeroes: $A_+ = \begin{pmatrix} A & \bar{0} \end{pmatrix}$, where $\bar{0}$ is the $(N_P \times N_S)$ zero matrix. We then have:

$$\vec{b}_+ = A\vec{p}_+ = A\vec{p}_+ + \bar{0}\vec{p}_- = \begin{pmatrix} A & \bar{0} \end{pmatrix} \begin{pmatrix} \vec{p}_+ \\ \vec{p}_- \end{pmatrix} = A_+\vec{p}, \tag{S35}$$

so that $\vec{c} = M\vec{b}_+ = MA_+\vec{p}$. Therefore, our condition for trace-preserving ρ_+ is $MA_+\vec{p} = 0$. We can therefore specify the new optimisation problem as:

$$\begin{aligned} \textbf{minimise} \quad & \|\vec{p}\|_1 = \sum_j p_j \\ \textbf{subject to} \quad & A'\vec{p} = \vec{b}, \\ & \vec{p} \geq 0, \\ & MA_+\vec{p} = 0 \end{aligned}$$

where $A' = \begin{pmatrix} A & -A \end{pmatrix}$, and $A_+ = \begin{pmatrix} A & \bar{0} \end{pmatrix}$, with A and $\vec{b}$ having the same definitions as previously, $\bar{0}$ is the zero matrix with dimension the same as A, and with M being the binary matrix that picks out the $\langle \mathbb{1} \otimes P_j \rangle$ entries from the vector $A_+\vec{p}$. Most of this is straightforward to implement. The step that requires some care is in correctly constructing the matrix M, as it will depend on the choice of ordering of Pauli operators in the construction of A and $\vec{b}$. If the B subsystem has n qubits, then we will need to constrain $4^n - 1$ non-trivial $\langle \mathbb{1} \otimes P_j \rangle$ expectation values to zero, so M should have dimension $((4^n - 1) \times N_P)$. If the Paulis are ordered as in the example given above for 2-qubit Choi states, then the construction is just $M = \begin{pmatrix} \vec{0} & \mathbb{1}' & \vec{0} & \cdots & \vec{0} \end{pmatrix}$, where $\mathbb{1}'$ is the $((4^n - 1) \times (4^n - 1))$ identity, and $\vec{0}$ denotes a column of zeroes. We have implemented this linear program in MATLAB, using the convex optimisation package CVX [2], and have made the code available from the repository Ref. [3].

IV. PROPERTIES OF MAGIC CAPACITY

Faithfulness: For any n-qubit stabilizer-preserving CPTP channel Λ, if $\rho \in \mathrm{STAB}_{2n}$ is a stabilizer state, then $(\Lambda \otimes \mathbb{1}_n)\rho$ is also a stabilizer state. So by the faithfulness of robustness of magic, $\mathcal{R}((\Lambda \otimes \mathbb{1}_n)\rho) = 1$ for any input stabilizer state $\rho \in \mathrm{STAB}_{2n}$, and $\mathcal{C}(\Lambda) = 1$.

Suppose instead that $\mathcal{E}$ is non-stabilizer-preserving, but still CPTP. Then there exists at least one stabilizer state $\rho \in \mathrm{STAB}_{2n}$ such that $(\mathcal{E} \otimes \mathbb{1})\rho$ is a normalised state, but not a stabilizer state. Then by faithfulness of $\mathcal{R}$ when applied to states, $\mathcal{R}((\mathcal{E} \otimes \mathbb{1})\rho) > 1$, and so $\mathcal{C}(\mathcal{E}) > 1$.

Convexity: Suppose we have a real linear combination of n-qubit CPTP maps $\mathcal{E}_k$:

$$\mathcal{E} = \sum_k q_k \mathcal{E}_k. \tag{S36}$$

There exists some optimal stabilizer state ρ_* that achieves $\mathcal{C}(\mathcal{E}) = \mathcal{R}(\mathcal{E} \otimes \mathbb{1}(\rho_*))$. Then

$$\mathcal{R}((\mathcal{E} \otimes \mathbb{1}_n)\rho_*) = \mathcal{R}\left(\sum_k q_k [(\mathcal{E}_k \otimes \mathbb{1}_n)\rho_*] \right) \tag{S37}$$

$$\leq \sum_k |q_k| \mathcal{R}((\mathcal{E}_k \otimes \mathbb{1}_n)\rho_*), \tag{S38}$$

where the last line follows by convexity of the robustness of magic. But each robustness

$\mathcal{R}((\mathcal{E}_k \otimes \mathbb{1}_n)\rho_*)$ can be no larger than $\mathcal{C}(\mathcal{E}_k)$. So we have:

$$\mathcal{C}\left(\sum_k q_k \mathcal{E}_k\right) \leq \sum_k |q_k| \mathcal{C}(\mathcal{E}_k). \tag{S39}$$

V. CALCULATING MONOTONES FOR DIAGONAL CHANNELS

A. Reducing the problem size

As mentioned earlier, the size of the optimisation problem for calculating our monotones (as well as $\mathcal{R}(\Phi_{\mathcal{E}})$) quickly becomes prohibitively large for n-qubit states, since the number of stabilizer states increases super-exponentially with n (Table I). The issue is even worse than it first

n	N_S
1	6
2	60
3	1,080
4	36,720
5	2,423,520
6	315,057,600

TABLE I. Number of pure stabilizer states N_S for number of qubits n.

appears, since for an n-qubit channel we must in general consider $2n$-qubit stabilizer states. Direct calculation of either monotone is impractical for n-qubit channels with $n > 2$. This difficulty is aggravated when calculating the capacity as in principle we have to repeat the optimisation for every $(\mathcal{E} \otimes \mathbb{1}_n) |\phi\rangle\langle\phi|$ such that $|\phi\rangle \in \mathrm{STAB}_{2n}$. In some cases we can ameliorate these problems by looking for Clifford gates that commute with the channel of interest. Here we consider the case where $\mathcal{E}$ is a diagonal channel, meaning it has a Kraus representation where each Kraus operator is diagonal in the computational basis. This of course includes diagonal unitaries as a special case. One could likely reduce the problem size further by exploiting symmetries of channels using techniques similar to those used in Ref. [4], but we will not consider this strategy here.

It is straightforward to see how the problem can be simplified for calculating $\mathcal{R}(\Phi_{\mathcal{E}})$ and $\mathcal{R}_*(\mathcal{E})$. If $\mathcal{E}$ is diagonal, the operation $\mathcal{E} \otimes \mathbb{1}_n$ commutes with any sequence of CNOTs targeted on the last n qubits. But the maximally entangled state $|\Omega_n\rangle$ can be written:

$$|\Omega_n\rangle = U_C(|+\rangle^{\otimes n} \otimes |0\rangle^{\otimes n}). \tag{S40}$$

Here $U_C = \otimes_{j=1}^n U_j$, where U_j is the CNOT controlled on qubit j and targeted on qubit $n+j$. By the monotonicity of robustness of magic, we immediately see that:

$$\mathcal{R}(\Phi_{\mathcal{E}}) = \mathcal{R}[(\mathcal{E} \otimes \mathbb{1}_n) |\Omega_n\rangle\langle\Omega_n|] = \mathcal{R}\big[\mathcal{E}\big(|+\rangle\langle+|^{\otimes n}\big) \otimes |0\rangle\langle 0|^{\otimes n}\big] = \mathcal{R}\big[\mathcal{E}\big(|+\rangle\langle+|^{\otimes n}\big)\big]. \tag{S41}$$

For the channel robustness we would like to decompose $\mathcal{E}\big(|+\rangle\langle+|^{\otimes n}\big)$ in terms of states $\rho_\pm \in \mathrm{STAB}_n$, but need to take care that the trace condition $\mathrm{Tr}_A(\rho'_\pm) = \mathbb{1}_n/2^n$ is satisfied for the equivalent $2n$-qubit Choi states $\rho'_\pm$. In Section V B of this Supplementary Material we show that the criterion is satisfied provided all diagonal elements of $\rho_\pm$ are equal to $1/2^n$. So for diagonal channels we can write:

$$\mathcal{R}_*(\mathcal{E}) = \min_{\rho_\pm \in \mathrm{STAB}_n} \left\{1 + 2p : (1+p)\rho_+ - p\rho_- = \mathcal{E}(|+\rangle\langle+|^{\otimes n}),\, p \geq 0,\, \langle x|\rho_\pm|x\rangle = \frac{1}{2^n}, \forall x\right\}. \tag{S42}$$

So calculation of $R_*(\mathcal{E})$ and $\mathcal{R}(\Phi_\mathcal{E})$ is tractable up to five qubits provided $\mathcal{E}$ is diagonal. We will see below in Section V C that this is also true for the magic capacity.

B. Trace condition for diagonal channels

Consider that the Choi state for a diagonal channel has a decomposition

$$\Phi_\mathcal{E} = U_C\big(\mathcal{E}(|+\rangle\langle+|^{\otimes n}) \otimes |0\rangle\langle 0|^{\otimes n}\big)U_C^\dagger = (1+p)\rho_+ - p\rho_-, \tag{S43}$$

where $U_C = \otimes_{j=1}^n U_j$ is the tensor product of CNOTs U_j that are controlled on the jth qubit and targeted on the $n+j$th. Then

$$\mathcal{E}(|+\rangle\langle+|^{\otimes n}) \otimes |0\rangle\langle 0|^{\otimes n} = (1+p)\rho'_+ - p\rho'_-, \tag{S44}$$

where $\rho'_\pm$ are still stabilizer states since U_C is Clifford. Now consider the stabilizer-preserving channel $\mathbb{1}_n \otimes \Lambda$ that resets the last n qubits to $|0\rangle\langle 0|^{\otimes n}$. Applying this to both sides of equation (S44) we get a new decomposition

$$\mathcal{E}(|+\rangle\langle+|^{\otimes n}) \otimes |0\rangle\langle 0|^{\otimes n} = (1+p)\rho''_+ \otimes |0\rangle\langle 0|^{\otimes n} - p\rho''_- \otimes |0\rangle\langle 0|^{\otimes n}. \tag{S45}$$

Then referring back to equation (S43), we obtain $\rho_\pm = U_C\big(\rho''_+ \otimes |0\rangle\langle 0|^{\otimes n}\big)U_C^\dagger$. So, the trace-preserving condition becomes:

$$\frac{\mathbb{1}_n}{2^n} = \mathrm{Tr}_A\left(\rho_\pm\right) = \mathrm{Tr}_A\left(U_C\big(\rho''_+ \otimes |0\rangle\langle 0|^{\otimes n}\big)U_C^\dagger\right) \tag{S46}$$

$$= \sum_x \langle x|^A\, U_C\big(\rho''_+ \otimes |0\rangle\langle 0|^{\otimes n}\big)U_C^\dagger\, |x\rangle^A, \tag{S47}$$

where $|x\rangle$ are the computational basis states on subsystem A. Recalling that U_C can be written as a tensor product of CNOTs $U_C = \otimes_{j=1}^n U_j$ one can check that this equation can be written:

$$\frac{\mathbb{1}_n}{2^n} = \sum_x \langle x|\, \rho''_\pm\, |x\rangle\, |x\rangle\langle x|. \tag{S48}$$

Therefore, the decomposition corresponds to a pair of trace-preserving channels provided that all diagonal elements of $\rho''_\pm$ are equal to $1/2^n$.

For a given diagonal channel, there always exists a decomposition that satisfies these conditions and has ℓ_1-norm equal to the channel robustness as defined for the full Choi state. We do not give the full proof here, but sketch the argument. Given any diagonal channel $\mathcal{E}$ decomposition of the full Choi state $\Phi_\mathcal{E} = (1+p)\rho_+ - p\rho_-$ satisfying the trace condition, one can always find a new decomposition $\Phi_\mathcal{E} = (1+p)\Lambda(\rho_+) - p\Lambda(\rho_-)$ where $\Lambda(\rho_\pm)$ still satisfy $\mathrm{Tr}_A(\Lambda(\rho_\pm)) = \mathbb{1}_n/2^n$, but are now the Choi states for diagonal channels. The map Λ used to obtain this decomposition is in effect an error correction circuit that takes general stabilizer Choi states to the subspace corresponding to the diagonal channels. Specifically, we note that the Choi states for diagonal maps $\mathcal{T}$ have the general form:

$$\Phi_\mathcal{T} = \frac{1}{2^n}\sum_{j,k} c_{j,k}\, |j\rangle^A\, |j\rangle^B\, \langle k|^A\, \langle k|^B. \tag{S49}$$

In general $c_{j,k}$ can be complex or zero, but terms on the diagonal are constrained. In particular, trace-preserving diagonal channels cannot change the weight of particular computational basis states, so the probability distribution for computational basis states will be the same as for $|\Omega\rangle$:

$$\langle p,q|\,\Phi_{\mathcal{T}}\,|p,q\rangle = \frac{1}{2^n}\delta_{p,q}. \tag{S50}$$

The circuit Λ is defined by the following steps. For each j from 1 to n:

1. Perform a parity measurement ($Z\otimes Z$) between qubits j and $n+j$.
2. If even parity (+1 outcome), do nothing. If odd parity (-1 outcome), perform an X gate on qubit j.

This stabilizer-preserving channel leaves Choi states for diagonal maps (and crucially, the target Choi state $\Phi_{\mathcal{E}}$) invariant, but updates general Choi states to have the form (S49). One can check that the circuit preserves the property $\mathrm{Tr}_A(\rho_\pm) = \mathbb{1}_n/2^n$. We then obtain a decomposition in the desired form:

$$\Lambda(\Phi_{\mathcal{E}}) = \Phi_{\mathcal{E}} = (1+p)\Lambda(\rho_+) - p\Lambda(\rho_-) \tag{S51}$$

Where $\Lambda(\rho_\pm) = (\mathcal{E}_\pm\otimes\mathbb{1})\,|\Omega\rangle\langle\Omega|$ are Choi states for n-qubit diagonal channels $\mathcal{E}_\pm$. But as described above, the CNOT sequence U_C commutes with diagonal channels acting on the first n qubits, so we can obtain n-qubit representatives of these channels:

$$\mathcal{E}_\pm(|+\rangle\langle+|^{\otimes n})\otimes|0\rangle\langle 0|^{\otimes n} = U_C((\mathcal{E}_\pm\otimes\mathbb{1})\,|\Omega\rangle\langle\Omega|)U_C^\dagger. \tag{S52}$$

Discarding the last n qubits we obtain the desired n-qubit decomposition:

$$\mathcal{E}(|+\rangle\langle+|^{\otimes n}) = (1+p)\mathcal{E}_+(|+\rangle\langle+|^{\otimes n}) - p\mathcal{E}_-(|+\rangle\langle+|^{\otimes n}). \tag{S53}$$

C. Magic capacity in the affine space picture

In this section we will make use of the formalism due to Dehaene and De Moor, in which stabilizer states are cast in terms of affine spaces and quadratic forms over binary vectors [5, 6], to prove the following theorem:

Theorem S1 (Capacity for diagonal operations)**.** *Suppose the n-qubit channel $\mathcal{E}_D$ is diagonal. Let:*

$$|\mathcal{K}\rangle = \frac{1}{|\mathcal{K}|^{1/2}}\sum_{x\in\mathcal{K}}|x\rangle\,, \tag{S54}$$

where $x\in\mathbb{F}_2^n$ are binary vectors and $\mathcal{K}\subseteq\mathbb{F}_2^n$ is an affine space. Then:

$$\mathcal{C}(\mathcal{E}_D) = \max_{\mathcal{K}}\mathcal{R}(\mathcal{E}_D(|\mathcal{K}\rangle\langle\mathcal{K}|)). \tag{S55}$$

That is, given an n-qubit channel $\mathcal{E}$, provided the channel is diagonal, the capacity $\mathcal{C}(\mathcal{E})$ may be calculated by optimisation over only the n-qubit states $|\mathcal{K}\rangle$ as defined in equation (S54), rather than over all $2n$-qubit stabilizer states.

We first review the formalism of Ref. [5]. Computational basis states $|x\rangle$ can be labelled by binary column vectors $x = (x_1,\ldots,x_n)^T\in\mathbb{F}_2^n$, so that $x_j\in\{0,1\}$ relates to the jth qubit. Any pure n-qubit stabilizer state may be written:

$$|\mathcal{K},q,d\rangle = \frac{1}{|\mathcal{K}|^{1/2}}\sum_{x\in\mathcal{K}} i^{d^Tx}(-1)^{q(x)}\,|x\rangle\,, \tag{S56}$$

where $\mathcal{K} \subseteq \mathbb{F}_2^n$ is an affine space, d is some fixed binary vector, and $q(x)$ has the form:

$$q(x) = x^T Q x + \lambda^T x. \tag{S57}$$

Here Q is a binary, strictly upper triangular matrix, λ is a vector, and addition is modulo 2. Conversely, any state that can be written in this way is a stabilizer state.

An affine space $\mathcal{K}$ is a linear subspace $\mathcal{L}$ shifted by some constant binary vector h, modulo 2: $\mathcal{K} = \mathcal{L} + h$. Every affine space is related in this way to exactly one linear subspace, and the dimension $k = \dim(\mathcal{K})$ of an affine space means the dimension of the corresponding subspace. Instead of enumerating all elements of an affine space, we can specify it by a shift vector h and an $n \times k$ matrix where each column is one of the generators of the corresponding linear space:

$$G = \begin{pmatrix} \vec{g}_1 & \vec{g}_2 & \cdots & \vec{g}_k \end{pmatrix} = \begin{pmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,k} \\ \vdots & \vdots & & \vdots \\ g_{j,1} & g_{j,2} & \cdots & g_{j,k} \\ \vdots & \vdots & & \vdots \\ g_{n,1} & g_{n,2} & \cdots & g_{n,k} \end{pmatrix}. \tag{S58}$$

We have freedom in our choice of k independent generators, and we can transform between equivalent generating sets by adding any two columns of G. We are also free to swap any two columns. A general transform between generating sets can therefore be represented by an invertible matrix S of dimension $k \times k$, multiplying on the right $G \longrightarrow GS$.

Any non-trivial linear transformation of the affine space can be fully specified by the transformation of the generators and the shift vector. In particular, we can represent the action of a single CNOT by multiplication on the left by a matrix C. If the CNOT has control qubit j and target qubit k, then C has 1s on the diagonal, a 1 in the jth element of the kth row, and zeroes everywhere else. A sequence for a $2n$-qubit system, in which CNOTs are always controlled on the first n qubits, and targeted on the last n qubits can be represented in block form:

$$C = \begin{pmatrix} \mathbb{1} & 0 \\ M & \mathbb{1} \end{pmatrix}, \tag{S59}$$

where each block has dimension $n \times n$, and M can be any binary matrix. We use this formalism to prove the following lemma, which leads directly to Theorem S1:

Lemma S1 (Equivalences for diagonal channels)**.** *Suppose $\mathcal{E}_D$ is a diagonal CPTP channel. Then:*

1. *All input stabilizer states with the same affine space $\mathcal{K}$ result in the same final robustness:*

$$\mathcal{R}((\mathcal{E}_D \otimes \mathbb{1}) |\mathcal{K}, q, d\rangle\langle\mathcal{K}, q, d|) = \mathcal{R}\big((\mathcal{E}_D \otimes \mathbb{1}) \big|\mathcal{K}, q', d'\big\rangle\big\langle\mathcal{K}, q', d'\big|\big), \quad \forall q, q', d, d'. \tag{S60}$$

2. *Given a $2n$-qubit state $|\phi\rangle \in \mathrm{STAB}_{2n}$, there exists some n-qubit $|\phi'\rangle \in \mathrm{STAB}_n$ such that:*

$$\mathcal{R}((\mathcal{E}_D \otimes \mathbb{1}_n) |\phi\rangle\langle\phi|) = \mathcal{R}\big(\mathcal{E}_D\big(\big|\phi'\big\rangle\big\langle\phi'\big|\big)\big). \tag{S61}$$

Proof. We first prove statement 1. Since robustness of magic is invariant under Clifford unitaries, we need to show that there exists a Clifford unitary U that converts $(\mathcal{E}_D \otimes \mathbb{1}) |\mathcal{K}, q, d\rangle\langle\mathcal{K}, q, d|$ to $(\mathcal{E}_D \otimes \mathbb{1}) |\mathcal{K}, q', d'\rangle\langle\mathcal{K}, q', d'|$. A suitable choice for U is one such that $U |\phi_{\mathcal{K},q,d}\rangle = \big|\phi_{\mathcal{K},q',d'}\big\rangle$, and, crucially, that commutes with the channel $\mathcal{E}_D$. Since $\mathcal{E}_D$ is given to be diagonal, any diagonal Clifford U will suffice. The affine space $\mathcal{K}$ remains unchanged, so we only need show there is always a diagonal Clifford that maps $q \to q'$ and $d \to d'$ for any q, q', d and d'. That this is always possible is perhaps already evident from Ref. [5], but for completeness we give the argument here.

We can convert d to d' using appropriately chosen S_j gates, meaning the gate $\mathrm{diag}(1,i)$ acting on the jth qubit. Consider the action of S_j on a basis vector:

$$S_j |x\rangle = \begin{cases} |x\rangle \text{ if } x_j = 0 \\ i|x\rangle \text{ if } x_j = 1 \end{cases} . \tag{S62}$$

If we define basis vector e_j so that it has 1 in the jth position and zeroes elsewhere, we can write the action of S_j as:

$$S_j |x\rangle = i^{e_j^T x} |x\rangle . \tag{S63}$$

Note that the form of this equation is independent of the value of x, so we can write:

$$S_j |\phi_{\mathcal{K},q,d}\rangle = \frac{1}{|\mathcal{K}|^{1/2}} \sum_{x\in\mathcal{K}} i^{d^T x} (-1)^{q(x)} S_j |x\rangle \tag{S64}$$

$$= \sum_{x\in\mathcal{K}} i^{(d^T + e_j^T)x} (-1)^{q(x)} S_j |x\rangle . \tag{S65}$$

So, we can flip any bit of d by applying the correct S gate. The quadratic form $q(x)$ is left unchanged.

Now consider $q(x) = x^T Q x + \lambda^T x$, which we must convert to some other $q'(x) = x^T Q' x + \lambda'^T x$. We can use the same trick as above to convert any λ to any other λ', by replacing S_j with the Z_j gate, i.e. $\mathrm{diag}(1,-1)$ acting on the jth qubit. For Q we can use the controlled-Z gate between the jth and kth qubit, which we denote CZ_{jk}. This has the following effect on a basis state:

$$CZ_{jk} |x\rangle = (-1)^{x^T M_{jk} x} |x\rangle , \tag{S66}$$

where M_{jk} is the $n\times n$ matrix with a 1 in position (j,k) and zeroes everywhere else. The set of all $\{M_{jk}\}$ form a basis for $n\times n$ binary matrices, hence we can convert any Q to any other Q' by an appropriately chosen sequence of CZ gates, leaving d and λ untouched. This completes the proof of statement 1.

Now to prove statement 2. From statement 1 any stabilizer state $|\phi\rangle$ is equivalent to:

$$|\mathcal{K}\rangle = \frac{1}{|\mathcal{K}|^{1/2}} \sum_{x\in\mathcal{K}} |x\rangle , \tag{S67}$$

up to some diagonal Clifford, for some $\mathcal{K}$. The strategy is to find a Clifford unitary U that commutes with $\mathcal{E}_D$, and converts the $2n$-qubit stabilizer state $|\mathcal{K}\rangle$ to some product of two n-qubit states $|\mathcal{K}'\rangle = |\mathcal{K}'_A\rangle \otimes |\mathcal{K}'_B\rangle$. Then we have:

$$\mathcal{R}[(\mathcal{E}_D \otimes \mathbb{1}_n) |\mathcal{K}\rangle\langle\mathcal{K}|] = \mathcal{R}\left[(\mathcal{E}_D \otimes \mathbb{1}_n)(|\mathcal{K}'_A\rangle\langle\mathcal{K}'_A| \otimes |\mathcal{K}'_B\rangle\langle\mathcal{K}'_B|)\right] \tag{S68}$$

$$= \mathcal{R}\left[\mathcal{E}_D(|\mathcal{K}'_A\rangle\langle\mathcal{K}'_A|) \otimes |\mathcal{K}'_B\rangle\langle\mathcal{K}'_B|\right] = \mathcal{R}\left[\mathcal{E}_D(|\mathcal{K}'_A\rangle\langle\mathcal{K}'_A|)\right], \tag{S69}$$

where the last step follows as $|\mathcal{K}'_B\rangle$ is a stabilizer state so makes no contribution to the robustness. The final state $|\mathcal{K}'\rangle$ can be factored as $|\mathcal{K}'_A\rangle \otimes |\mathcal{K}'_B\rangle$ provided its generator G' can be written in block matrix form as:

$$G' = \begin{pmatrix} G'_A & 0 \\ 0 & G'_B \end{pmatrix} , \tag{S70}$$

where G'_A and G'_B have n rows, and represent the generators for affine spaces $\mathcal{K}'_A$ and $\mathcal{K}'_B$.

We now show that we can always reach this form by a Clifford U_C comprised of a sequence of CNOTs targeted on the last n qubits. Such a sequence always commutes with $\mathcal{E}_D \otimes \mathbb{1}_n$. Suppose we have some $2n \times k$ generator G for an affine space $\mathcal{K}$ with $k = \dim(\mathcal{K})$:

$$G = \begin{pmatrix} G_A \\ G_B \end{pmatrix}, \tag{S71}$$

where G_A and G_B are each $n \times k$ submatrices. The full matrix G will have rank k, and G_A will have some rank $m \leq k$. Either G_A is already full rank ($m = k$), or it can be reduced to the following form by elementary column operations, which is equivalent to multiplication on the right by a $k \times k$ matrix S:

$$G_A \longrightarrow G_A S = \begin{pmatrix} G'_A & 0 \end{pmatrix}, \tag{S72}$$

where G'_A is $n \times m$ (and hence full column rank), and 0 is $n \times (k - m)$. Multiplying G on the right by S, we interpret as a change in the choice of generating set:

$$G \longrightarrow GS = \begin{pmatrix} G_A S \\ G_B S \end{pmatrix} = \begin{pmatrix} G'_A & 0 \\ G''_B & G'_B \end{pmatrix}. \tag{S73}$$

Now, apply the Clifford U_C described by the matrix C in equation (S59). This transforms the generator to:

$$G' = CGS = \begin{pmatrix} \mathbb{1} & 0 \\ M & \mathbb{1} \end{pmatrix} \begin{pmatrix} G'_A & 0 \\ G''_B & G'_B \end{pmatrix} = \begin{pmatrix} G'_A & 0 \\ MG'_A + G''_B & G'_B \end{pmatrix}. \tag{S74}$$

Note that if G_A was already full rank, the change of generating set is not necessary. If we can set the bottom-left submatrix to zero, then $U_C \left|\mathcal{K}\right\rangle$ can be factored as described above. This is possible if there exists a binary matrix M such that $MG'_A = G''_B$. But G'_A has full column rank m, so there exists an $m \times n$ left-inverse $G'^{-1}_{A,\text{left}}$ such that $G'^{-1}_{A,\text{left}} G'_A = \mathbb{1}$, where $\mathbb{1}$ is $m \times m$. Then we can set $M = G''_B G'^{-1}_{A,\text{left}}$, so that:

$$MG'_A = G''_B G'^{-1}_{A,\text{left}} G'_A = G'_B \mathbb{1} = G''_B. \tag{S75}$$

Then $G' = CGS$ is in the form (S70), so $U_C \left|\mathcal{K}\right\rangle = \left|\mathcal{K}'_A\right\rangle \otimes \left|\mathcal{K}'_B\right\rangle$, as required. □

Lemma S1 shows that if $\mathcal{E}_D$ is diagonal then for any $2n$-qubit stabilizer state $\left|\phi\right\rangle$ we have that $\mathcal{R}((\mathcal{E}_D \otimes \mathbb{1}_n) \left|\phi\right\rangle\!\left\langle\phi\right|) = \mathcal{R}(\mathcal{E}_D(\left|\mathcal{K}\right\rangle\!\left\langle\mathcal{K}\right|))$ for some n-qubit affine space $\mathcal{K}$. This shows that the capacity can be calculated by maximising over just the representative states $\left|\mathcal{K}\right\rangle$, proving Theorem S1. Table II illustrates the reduction in problem size. For example, whereas naively for a two-qubit channel we would need to calculate robustness for all 36,720 four-qubit stabilizer states, using the result above we only need check one stabilizer state for each of the 7 non-trivial affine spaces. Cases up to five qubits are now tractable using this method.

n	Stabilizer states	Total affine spaces	Non-trivial affine spaces
2	60	11	7
3	1,080	51	43
4	36,720	307	291
5	2,423,520	2451	2419

TABLE II. Number of n-qubit stabilizer states compared with number of affine spaces. By trivial affine spaces we mean those comprised of a single element, which correspond to computational basis states. Diagonal CPTP channels act as the identity on such states.

D. Dimension of affine space

Here we make further observations that will help interpret numerical results from Section 8 of the main text.

Observation S1 (Dimension of affine space limits achievable robustness)**.** *Suppose U is a diagonal unitary acting on n qubits, and suppose $|\mathcal{K}\rangle$ is a stabilizer state associated with some affine space $\mathcal{K}$, $k = \dim(\mathcal{K})$. Then $\mathcal{R}(U|\mathcal{K}\rangle) = \mathcal{R}(U'|\phi'\rangle)$ where $U'|\phi'\rangle$ is a state on only k qubits, and U' is some k-qubit unitary. Therefore $\mathcal{R}(U|\mathcal{K}\rangle)$ is upper-bounded by the maximum robustness achievable for a k-qubit state.*

Proof. We prove the result by showing that there is a sequence of Clifford gates that takes $U|\mathcal{K}\rangle$ to the product of a k-qubit state and an $(n-k)$-qubit stabilizer state. We know from Lemma S1 that for diagonal unitaries, all states with same affine space result in the same robustness, so it is enough to consider the state:

$$|\mathcal{K}\rangle = \frac{1}{\sqrt{|\mathcal{K}|}} \sum_{x\in\mathcal{K}} |x\rangle . \tag{S76}$$

A diagonal unitary will map this to:

$$U|\mathcal{K}\rangle = \frac{1}{\sqrt{|\mathcal{K}|}} \sum_{x\in\mathcal{K}} e^{i\theta_x} |x\rangle , \tag{S77}$$

where $\{e^{i\theta_x}\}$ will be some subset of the diagonal elements of U. The affine space $\mathcal{K}$ will have a generator matrix of rank k. As we saw in Lemma S1, a sequence of elementary row operations on the generator matrix can be realised by a sequence of CNOT gates. So we can use Clifford gates to transform any rank k generator matrix as:

$$G \longrightarrow G' = AG = \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} , \tag{S78}$$

where $\mathbb{1}$ is the $k\times k$ identity. Each element of $\mathcal{K}$ can be written $x = \sum_j g_j + h$, where $\sum_j g_j$ is some combination of columns of G, and h is a fixed shift vector. The transformation A corresponds to a sequence of CNOTs that we collect in a single Clifford unitary U_A, that acts on n-qubit computational basis states $|x\rangle$, where $x \in \mathcal{K}$, as follows:

$$U_A|x\rangle = |y(x)\rangle \otimes |h'\rangle , \tag{S79}$$

where h' is an $(n-k)$-length vector, and $y(x)$ is a k-length vector given by:

$$\begin{pmatrix} y(x) \\ h' \end{pmatrix} = Ax = \sum_j Ag_j + Ah. \tag{S80}$$

Note that $y(x)$ is only defined for $x \in \mathcal{K}$, and that h' is independent of x. Elements $x \in \mathbb{F}_2^n$ that are not in $\mathcal{K}$ could be mapped to a vector where the last $n-k$ bits are not h', but these never appear as terms of $U|\mathcal{K}\rangle$. Since U_A must preserve orthogonality, each $|x\rangle$, where $x \in \mathcal{K}$, maps to a distinct element of the k-qubit basis set $\{|y\rangle\}$. In fact, since y are length k and there are 2^k distinct elements, they must form the k-bit linear space $\mathcal{L}' = \mathbb{F}_2^k$. So we can write:

$$U_A U|\mathcal{K}\rangle = \frac{1}{\sqrt{|\mathcal{L}'|}} \sum_{y\in\mathcal{L}'} e^{i\theta'_y} |y\rangle \otimes |h'\rangle \tag{S81}$$

$$= (U'|\mathcal{L}'\rangle) \otimes |h'\rangle , \tag{S82}$$

where $|\mathcal{L}'\rangle$ is a k-qubit stabilizer state, and U' is the k-qubit diagonal unitary with $e^{i\theta'_{y(x)}} = e^{i\theta_x}$ as the non-zero elements. The state $|h'\rangle$ is a stabilizer state, so cannot contribute to the robustness of $U_A U |\mathcal{K}\rangle$, and therefore $\mathcal{R}(U|\mathcal{K}\rangle) = \mathcal{R}(U_A U |\mathcal{K}\rangle) = \mathcal{R}(U'|\mathcal{L}'\rangle)$, where $U'|\mathcal{L}'\rangle$ is a k-qubit state. $\square$

Recall that in Section 8 of the main text, we found that highly structured examples of diagonal unitaries U exist where $\mathcal{C}(U)$ is strictly larger than $\mathcal{R}(\Phi_U)$, whereas for all the random diagonal unitaries sampled, we found them to be exactly equal. We can now explain this by a concentration effect, in conjunction with Observation S1. The n-qubit random diagonal gates concentrate (with high probability) within a narrow range of values for the magic capacity, close to the maximum possible magic capacity for an n-qubit diagonal gate. If $\mathcal{R}(\Phi_U) < C(U)$ then by Theorem S1 we must have that $C(U) = \mathcal{R}(U|\mathcal{K}\rangle\langle\mathcal{K}|U^\dagger)$ for some affine space $\mathcal{K}$ of non-maximal dimension. However, $U|\mathcal{K}\rangle\langle\mathcal{K}|U^\dagger$ is Clifford equivalent to an $(n-1)$-qubit stabilizer state acted on by a diagonal unitary. Then $\mathcal{R}(U|\mathcal{K}\rangle\langle\mathcal{K}|U^\dagger)$ would be upper bounded by the maximum $C(\mathcal{E})$ for $(n-1)$-qubit diagonal unitaries. But if $C(\mathcal{E})$ is close to the maximum possible for n-qubit diagonal unitaries, then it is impossible for $U|\mathcal{K}\rangle\langle\mathcal{K}|U^\dagger$ to achieve the magic capacity.

Finally, we consider the special case of multi-control phase gates $M_{t,n}$, which we defined in the main text as:

$$M_{t,n} = \mathrm{diag}(\exp\left(i\pi/2^t\right), 1, 1, \ldots, 1), \quad t \in \mathbb{Z}. \tag{S83}$$

Note that the gate $M_{t,n}$ acts as the identity on states $|\mathcal{K}\rangle$ unless $\mathcal{K}$ contains the zero vector $0^n = (0, \ldots, 0)^T$, so if $0^n \notin \mathcal{K}$, we get $\mathcal{R}(M_{t,n}|\mathcal{K}\rangle) = 1$. But if $0^n \in \mathcal{K}$, then $\mathcal{K}$ is a linear subspace. So for this type of gate, to find all possible values of $\mathcal{R}(M_{t,n}|\mathcal{K}\rangle) > 1$ we need only consider linear subspaces. The following theorem implies that we actually only need solve one optimisation for each possible *dimension* of linear subspace rather than one for every linear subspace.

Theorem S2. *Consider the n-qubit gate $M_{t,n}$ defined by equation* (S83)*, and let $\mathcal{L}_A$ and $\mathcal{L}_B$ be linear subspaces such that* $\dim(\mathcal{L}_A) = \dim(\mathcal{L}_B) = k$*. Then:*

$$\mathcal{R}(M_{t,n}|\mathcal{L}_A\rangle) = \mathcal{R}(M_{t,n}|\mathcal{L}_B\rangle). \tag{S84}$$

Proof. We largely repeat the arguments of Observation S1, for the special case where the phases are given by:

$$\theta_x = \begin{cases} \pi/2^t & \text{if}\, x = \vec{0} \\ 0 & \text{otherwise} \end{cases} \tag{S85}$$

Since $\dim(\mathcal{L}_A) = \dim(\mathcal{L}_B)$, their generator matrices G_A and G_B have the same rank. It follows from the arguments of Observation S1 that there exists an invertible C, corresponding to a sequence of CNOT gates, such that $G_B = CG_A$, and $|\mathcal{L}_A\rangle = U_C|\mathcal{L}_A\rangle$, where U_C is a unitary Clifford operation.

If we consider instead the state $M_{t,n}|\mathcal{L}_A\rangle$, which involves terms in the same basis vectors as $|\mathcal{L}_A\rangle$, we just need to track what happens to the phase $\exp(i\theta_0)$. Clearly, since any CNOT acts as the identity on $|0^n\rangle$, we obtain:

$$U_C M_{t,n}|\mathcal{L}_A\rangle = \frac{1}{2^{k/2}} \sum_{x\in\mathcal{L}_B} \exp(i\theta_x)|x\rangle = M_{t,n}|\mathcal{L}_B\rangle \tag{S86}$$

Since U_C is a reversible Clifford operation, by monotonicity of robustness of magic, equation (S84) follows. $\square$

From Theorem S2, then, to find $\mathcal{C}(M_{t,n})$, we only need calculate $\mathcal{R}(M_{t,n}|\mathcal{L}\rangle)$ for a single representative subspace for each possible value of $\dim(\mathcal{L})$. Recall that for n-qubit stabilizer states $|\mathcal{L}\rangle$, $k = \dim\mathcal{L}$ can take integer values from 0 to n. The states with $k = 0$ correspond to single computational basis states without superposition, so are unaffected by phase gates. That is, for n-qubit multicontrol phase gates we only have to calculate n robustnesses. Compare this to the number of optimisation problems we would need to solve without using the above observations (Table II).

We can go further. From Observation S1 we know that for a subspace with $\dim(\mathcal{L}) = k < n$, it must be the case that $M_{t_n} |\mathcal{L}\rangle$ is Clifford-equivalent to $(U' |\mathcal{L}'\rangle) \otimes |h'\rangle$ for the k-qubit state $|\mathcal{L}'\rangle$ and $(n-k)$-qubit computational basis state $|h'\rangle$, and some diagonal k-qubit unitary U'. By inspection of the phases given by equation (S85), U' can only be the k-qubit multicontrol gate $M_{t,k}$. This leads to the following statement:

Observation S2 (n-qubit multicontrol gates)**.** *For any fixed t and n-qubit state $|\mathcal{L}\rangle$ where $\dim(\mathcal{L}) = k < n$, we have:*

$$\mathcal{R}(M_{t,n} |\mathcal{L}\rangle) = \mathcal{R}\big(M_{t,k} \left|\mathcal{L}'\right\rangle\big) \tag{S87}$$

where $|\mathcal{L}'\rangle$ is the k-qubit state with $\mathcal{L}' = \mathbb{F}_2^k$.

Linear subspace dimension, k	Number of qubits, n			
	2	3	4	5
1	1.414	1.414	1.414	1.414
2	**1.849**	1.849	1.849	1.849
3	-	**2.195**	2.195	2.195
4	-	-	**2.264**	**2.264**
5	-	-	-	2.195

TABLE III. Final robustness after multicontrol-T gate applied to input stabilizer states $|\mathcal{L}\rangle$ with $k = \dim(\mathcal{L})$. In each column, the maximum robustness (i.e. the capacity) is highlighted red.

Observation S2 partially justifies our Conjecture 8.1 in Section 8 of the main text, that for fixed t, the maximum increase in robustness achievable for $M_{t,n}$, over any n, is given by $\mathcal{R}\left(M_{t,K} |+\rangle^{\otimes K}\right)$, for some finite number of qubits K. To unpack this claim further, let us consider the maximisation over input stabilizer states performed to calculate the capacity $\mathcal{C}$. In this Supplementary Material, we have seen that for the family of gates $M_{t,n}$, we only need to calculate robustness for one representative input stabilizer state for each possible *dimension* of linear subspace; that is, for $M_{t,n}$ there are only n robustnesses to calculate. In Table III we present the relevant values for the family of multicontrol-T gates ($t = 2$) and make two observations. First, looking across the rows of Table III, notice that the values for fixed k are constant with n, assuming $k \leq n$. Indeed, this is a generic feature of the $M_{t,n}$ gates as formalised by Observation S2. Second, looking down the last column of Table III, we see that up until $k = 4$, $\mathcal{R}(M_{t,n} |\mathcal{L}\rangle)$ increases with $\dim(\mathcal{L})$, but at $k = 5$ the value drops. With a little thought we can see that this is necessarily the case if $\mathcal{R}\big(\Phi_{M_{t,5}}\big) < \mathcal{C}(M_{t,5})$; we saw earlier that for diagonal gates U the Choi state robustness is equal to $\mathcal{R}\big(U |+\rangle^{\otimes n}\big)$, and $|+\rangle^{\otimes n}$ is a representative state for the $k = n$ case.

Our current techniques limit us to five-qubit operations, so we are unable to confirm whether $\mathcal{R}(M_{t,n} |\mathcal{L}\rangle)$ continues to decrease with increasing $\dim(\mathcal{L})$. An intuition for why a decrease is plausible goes as follows. A stabilizer state $|\mathcal{L}\rangle$ with $\dim(\mathcal{L}) = k$ will have 2^k equally weighted terms when written in the computational basis, so will have a normalisation factor of $2^{-k/2}$. The non-stabilizer state $M_{t,n} |\mathcal{L}\rangle$ is identical to $|\mathcal{L}\rangle$ apart from the phase on the all-zero term $|0\ldots0\rangle$. As k becomes large, the amplitude of the term $\frac{e^{i\pi/2^t}}{2^{k/2}} |0\ldots0\rangle$ becomes very small, so that $M_{t,n} |\mathcal{L}\rangle$ has high fidelity with the stabilizer state $|\mathcal{L}\rangle$. We would therefore expect $M_{t,n} |\mathcal{L}\rangle$ to have a small robustness if k is large.

[1] M. Howard and E. Campbell, Physical Review Letters **118**, 090501 (2017), arXiv:1609.07488.
[2] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," `http://cvxr.com/cvx` (2014).
[3] J. R. Seddon, "Channel magic - tools for calculating measures of magic for channels using MATLAB and CVX," `https://github.com/jamesrseddon/channel_magic` (2019).
[4] M. Heinrich and D. Gross, Quantum **3**, 132 (2019), arXiv:1807.10296.
[5] J. Dehaene and B. De Moor, Physical Review A **68**, 042318 (2003), arXiv:quant-ph/0304125.
[6] D. Gross and M. Van den Nest, Quantum Info. Comput. **8**, 263 (2008), arXiv:0707.4000.