A. Improving our abilities to include arbitrary output stabilizer states.

The existing analysis works fine when the state τ' has stabilizer eigendecomposition (i.e. it's orthonormal basis is along the stabilizer directions...e.g. along $|0\rangle\langle 0|, |1\rangle\langle 1|$ in the Bloch sphere for a qubit). However it can lie inside the stabilizer polytope but not have such an eigenbasis. Therefore we'd like to include this fully general possibility.

To do this we just need to analyse how the properties of the output state relate to the magnitudes of its Wigner components. Here let τ be a single full-rank stabilizer qutrit state – we are really interested in the output state τ' but let's ignore primes for now....they just clutter expressions unnecessarily. For simplicity denote the Wigner components of the magic state as

$$\boldsymbol{w} = (-v, u, u, u, u, u, u, u, u), \tag{1}$$

and we index from $i = 0$ to $i = 7$, so that e.g. $w_0 = -v$ and $w_i = u$ for $i \neq 0$. Let $\boldsymbol{t}$ denote the Wigner distribution for τ, so that

$$\boldsymbol{t} = (t_0, t_1, t_2, t_3, t_4, t_5, t_6, t_7), \tag{2}$$

and we assume that $t_i > 0$ for all i. We also note that using Clifford unitaries we can always choose to shift either the smallest entry of $\boldsymbol{t}$ or the largest entry of $\boldsymbol{t}$ to the $i = 0$ slot. We might need to exploit this in a moment.

The rescaled distribution is then $\tilde{\boldsymbol{w}} \coloneqq (w_i/t_i)$, and we define $v_0 \coloneqq v/t_0$ and $u_i \coloneqq u/t_i$ for $i \neq 0$. We now need to consider the n–copy version of this, denoted $\tilde{\boldsymbol{w}}^{\otimes n}$, and determine how the largest component arises. A general component of $\tilde{\boldsymbol{w}}^{\otimes n}$ is given by,

$$\tilde{w}_{n,\alpha,i_1,\ldots,i_8} = (-v_0)^{i_0} u_1^{i_1} u_2^{i_2} \cdots u_8^{i_8}, \tag{3}$$

with the assumption that firstly n is always even, and the indices obeying the following conditions:

$$\begin{aligned} 0 \leq i_j \leq n \text{ for each } j = 0, 1, \ldots, 8 \\ i_0 + i_1 + i_2 + \ldots i_8 = n. \end{aligned} \tag{4}$$

Now the largest entry must be positive, and so we require that α is an even number, and so we consider

$$\tilde{w}_{n,i_0,i_1,\ldots,i_8} = v_0^{i_0} u_1^{i_1} u_2^{i_2} \cdots u_8^{i_8}, \tag{5}$$

over the above integers with $i_0 \in \{0, 2, 4, \ldots, n\}$. Let $a = \max\{v_0, u_1, \ldots, u_8\}$, then clearly the maximum value attained ($\tilde{w}_{\max} \coloneqq \max_{i_0,i_1,\ldots,i_8}(\tilde{w}_{n,i_0,i_1,\ldots,i_8})$) is given by

$$\tilde{w}_{\max} = a^n. \tag{6}$$

If the numbers $\{v_0, u_1, u_2, \ldots, u_8\}$ are distinct then this maximum value occurs at a unique $i_k = n$ for some $k \in \{0, 1, \ldots, 8\}$. Let's consider this case first.

(No multiplicities) We expect this is the generic case, but I think it always provides a sufficient condition even if multiplicities exist. We can assume that noise level is below 3/7 and so we have that $v > u$, so the 0 component in $\boldsymbol{w}$, and so also $\boldsymbol{w}^{\otimes n}$, is the largest. But this needs rescaling by $\boldsymbol{t}$.

We now make use of the freedom to do a Clifford unitary transformation that sends the smallest entry in $\boldsymbol{t}$ to the 0 component. Note that in doing so we must also account for this Clifford unitary on the output magic state too! To get around this we can simply say that we consider $\rho_S(\epsilon)^{\otimes n} \to \rho_S(\epsilon')^{\otimes m}$ modulo local Clifford unitaries on the output system. This changes the rules of the game to suit our needs. Instead of distilling out the precise noisy Strange state, we allow the protocol to have arbitrary post-processing on the individual qutrits via reversible Cliffords. I.e. the magic content is the same in each case, but when we consider the problem of how $\tau \to \tau'$ this can be used to our advantage to simplify our analysis. Need to reflect on this variant and how it affects the content of our result.

So given the ability to assume that the smallest entry of $\boldsymbol{t}$ is the t_0 component this means that for noise levels below 3/7 the largest component of $\tilde{w}_{n,i_0,i_1,\ldots,i_8}$ occurs at $(i_0, i_1, \ldots, i_8) = (n, 0, 0, \ldots, 0)$. We therefore have, for the no-multiplicity case, that the first elbow occurs at the point (t_0^n, v^n). This will lead to a bound expression in terms of ϵ for the output state, together with whatever value t_0 takes on. [give expression]. While this is fine, we need to link t_0 to the physics of the state $\tau = e^{-\beta H}/\mathcal{Z}$. Well, just from the definition we have that

$$t_0 = \frac{1}{d\mathcal{Z}} \mathrm{tr}[A_{0,0} e^{-\beta H}], \tag{7}$$

where, of course, $d = 3$. So in t_0^n we immediately get a term $\mathcal{Z}^{-n} = e^{-n\beta F}$, which is nice and general. Using the expansion of the thermal state into its energy eigenstates we obtain

$$\begin{aligned} t_0^n &= \left(\frac{1}{d\mathcal{Z}} \mathrm{tr}[A_{0,0} e^{-\beta H}] \right)^n \\ &= \frac{e^{-n\beta F}}{3^n} (\mathrm{tr}[A_{0,0} \sum_k e^{-\beta E_k} |\varphi_k\rangle\langle\varphi_k|])^n \\ &= \frac{e^{-n\beta F}}{3^n} (\sum_k \alpha_k e^{-\beta E_k})^n, \end{aligned} \tag{8}$$

where we define

$$\alpha_k \coloneqq \langle\varphi_k| A_{0,0} |\varphi_k\rangle, \text{ for } k = 0, 1, 2. \tag{9}$$

However $A_{0,0} = \sum_{r \in \mathbb{Z}_3} |-r\rangle\langle r|$ and so

$$\alpha_k \coloneqq \sum_{r \in \mathbb{Z}_3} \langle\varphi_k| - r\rangle\langle r|\varphi_k\rangle. \tag{10}$$

This just encodes how the energy eigenbasis of the Hamiltonian H relates to the *stabilizer* basis. Anyhow, the

nice thing is that we always get the Free energy term emerging in the t_0^n and so we can include this into the distillation bound, with the α_k terms simply handling the messy issue of how the two important bases relate. We can package the messy details into some definition in some way. E.g. the form of the sum involving α_k looks a bit like a partition function, so we can define something like

$$\zeta := \sum_k \alpha_k e^{-\beta E_k}, \tag{11}$$

so that

$$t_0^n = \frac{e^{-n\beta F}}{3^n}\zeta^n, \tag{12}$$

which is tidy....and we could also then define ϕ as a kind of free energy term via $\zeta =: e^{-\beta\phi}$. This then simplifies things to

$$t_0^n = \frac{e^{-n\beta(F+\phi)}}{3^n}, \tag{13}$$

where F corresponds to the physical free energy of τ while ϕ is some weird term that captures the degree to which the energy basis differs from the stabilizer basis. [Details need tidying/checking and we need the final bounds to be computed.]

(Multiplicity case) I don't think we really need to delve into this. Firstly no multiplicities will be generic. Secondly, if multiplicities appear this simply means the above Lorenz curve point is not an elbow but instead as point on the first line segment. Imposing the conditions for this point is *also a necessary condition* and therefore suffices to get a bound.

I. LOWER BOUNDS IN σ–FRAGMENTS

[Am throwing down some rough material, to be polished later.]

In order to obtain lower bounds one must now make precise what the free operations actually are in the theory, beyond the condition of preserving Wigner-positivity.

We first consider the unital fragment, and consider the transformations possible using Clifford unitaries, and convex mixtures of Clifford unitaries. We denote the Clifford group as $\mathcal{C}$, and given some quantum state ρ the accessible states in the unital fragment are given by $\mathcal{E}(\rho) = \sum_{g\in\mathcal{C}} p(g)U(g)\rho U(g)^\dagger$.

A. Symplectic majorisation of magic state Wigner distributions.

We now exploit the the group structure, which leads to a more general notation majorisation that has been extensively studied in the classical context, but to our knowledge has not yet been used in quantum information theory.

Definition 1. *Given a group G that acts on a vector space V, we say that $\boldsymbol{x}$* G-majorises *$\boldsymbol{y}$ precisely when $\boldsymbol{y} \in H(\boldsymbol{x})$, where $H(\boldsymbol{x})$ is the convex hull of $\{g\boldsymbol{x} : g \in G\}$. We denote this as $\boldsymbol{y} \prec_G \boldsymbol{x}$.*

Now any quantum state ρ has a Wigner distribution $W_\rho(\boldsymbol{x})$. Given a Clifford unitary $\mathcal{U} \in \mathcal{C}$ we have that its representation in the Heisenberg-Weyl frame is given by

$$W_{\mathcal{U}(\rho)}(\boldsymbol{x}) = \text{[fill in]}. \tag{14}$$

This group action corresponds to the action of the affine symplectic group on the discrete phase space $\mathcal{P}_d$. To proceed we can study the discrete translational action, and the symplectic action separately.

B. Cyclic majorisation conditions.

We can consider for G, the cyclic group of order N, which is described by $G = \langle g|g^N = e\rangle$. in Terms of the Wigner distributions, this arises for the discrete lattice translations arising from displacement operators.

A set of necessary and sufficient conditions for cyclic majorisation has been obtained and is given as follows. Suppose $\boldsymbol{x}, \boldsymbol{y}$ are two real vectors in $\mathbb{R}^n$. Let Δ be the elementary shift operator, defined by

$$\Delta\boldsymbol{x} = \Delta(x_1, x_2, \ldots, x_n) := (x_n, x_1, x_2, \ldots, x_{n-1}), \tag{15}$$

and from this it is clear that Δ generates a representation of the abelian group $(\mathbb{Z}_n, +)$. We now say that $\boldsymbol{x}$ *cyclically majorises* $\boldsymbol{y}$, written $\boldsymbol{x} \succ_C \boldsymbol{y}$, if and only if

$$\boldsymbol{y} = \sum_{k=0}^{n-1} p_k \Delta^k \boldsymbol{x}, \tag{16}$$

for some probability distribution $p = (p_k)$. This means that $\boldsymbol{y}$ lies in the convex hull of the orbit of $\boldsymbol{x}$ under cyclic shifts. Thus, we may write this condition as $\boldsymbol{y} = L(\boldsymbol{p})\boldsymbol{x}$, where

$$L(\boldsymbol{p}) := \sum_{k=0}^{n-1} p_k \Delta^k, \tag{17}$$

is a linear operator, as a function of an unknown distribution $\boldsymbol{p}$. We now have the following key identity for linear operators of this form:

$$L(\boldsymbol{p})\boldsymbol{x} = QL(\boldsymbol{x})\boldsymbol{p}, \tag{18}$$

where Q is a permutation matrix sending $e_0 := (1, 0, 0, \ldots, 0)$ to itself, and otherwise sending e_k to e_{n-1-k}, where $e_k = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ is the k'th basis vector. We thus have

$$\begin{aligned}
\boldsymbol{x} \succ_C \boldsymbol{y} &\Leftrightarrow \boldsymbol{y} = L(\boldsymbol{p})\boldsymbol{x} \text{ for some dist. } \boldsymbol{p}. \\
&\Leftrightarrow \boldsymbol{y} = QL(\boldsymbol{x})\boldsymbol{p} \text{ for some dist. } \boldsymbol{p}. \\
&\Leftrightarrow [QL(\boldsymbol{x})]^{-1}\boldsymbol{y} = \boldsymbol{p} \text{ for some dist. } \boldsymbol{p}. \\
&\Leftrightarrow [QL(\boldsymbol{x})]^{-1}\boldsymbol{y} = \boldsymbol{p} \geq \boldsymbol{0} \text{ and normalized.}
\end{aligned} \tag{19}$$

This implies that to check whether $\boldsymbol{x} \succ_C \boldsymbol{y}$ it suffices to compute the components of the left-hand side and ensure non-negativity (I suspect it is normalised if $\boldsymbol{x}$ and $\boldsymbol{y}$ are normalised).

Note though that by using Fourier analysis we can simplify this in terms of computational demands to checking if the vector

$$\boldsymbol{p} = C(\boldsymbol{x}')Q\boldsymbol{y}, \tag{20}$$

has non-negative components, where

$$\boldsymbol{x}' := n\mathcal{F}_n[\mathcal{F}_n\boldsymbol{x}]^{-1}, \tag{21}$$

with $\mathcal{F}_n$ being the discrete Fourier transform, and the bracket term $[\mathcal{F}_n\boldsymbol{x}]^{-1}$ denotes the vector obtained by inverting the components of $\mathcal{F}\boldsymbol{x}$ individually. So the recipe for cyclic majorisation is:

1. Given inputs $\boldsymbol{x}$ and $\boldsymbol{y}$.
2. Compute the vector $\boldsymbol{x}'$ via two Fourier transforms and n inversions.
3. Compute the circulant matrix $C(\boldsymbol{x}')$.
4. Compute $\mathcal{C}(\boldsymbol{x}')Q\boldsymbol{y}$ and check if all components are non-negative.

•

C. Symplectic majorisation for qutrit magic

The qutrit system provides a good illustration of the techniques. In this case the symplectic group $SL(2,\mathbb{Z}_3)$ is isomorphic (up to $\pm\mathbb{1}$) the symmetry group of the tetrahedron, which in turn is isomorphic to S_4, the permutation group on 4 symbols. Therefore we expect in this case that symplectic majorisation on $\mathcal{P}_3$ corresponds to a restricted form of majorisation.

D. Fundamental Regions of a group G

[This is theory for the appendices, and also to help flesh out the theory of G majorisation. It looks a bit technical, but the core idea is simple enough once you get it.] Given a group G that acts on a vector space V, we now have the following concept.

Definition 2. *A* fundamental region F *of* G *in* V *is any open set* F *such that* $F \cap gF = \varnothing$ *for* $g \neq e$ *and moreover*

$$V = \bigcup_g \overline{gF}. \tag{22}$$

Here, the overline denotes the closure of a set. Loosely speaking, we can view F as obtained by quotienting the vector space V via the group action. We then have the following key theorem, which is proved in [CITE].

Theorem 3. *If* G *has a fundamental region* F *that is unique, modulo actions of the group* $F \to gF$*, then* $\bar{F}$ *is a closed, convex cone and for any* $\boldsymbol{x}, \boldsymbol{y} \in \bar{F}$ *we have*

$$\boldsymbol{y} \prec_G \boldsymbol{x} \Leftrightarrow \mathbf{a}\cdot\boldsymbol{y} \leq \mathbf{a}\cdot\boldsymbol{x}, \text{ for all } \mathbf{a} \in \bar{F}. \tag{23}$$

If the cone $\bar{F}$ is finitely generated, namely

$$\bar{F} = \text{cone}(\mathbf{c}_1, \ldots, \mathbf{c}_N), \tag{24}$$

for some finite set of vectors, then the majorisation condition reduces to checking a *finite* set of inequalities, namely checking $\mathbf{c}_k \cdot \boldsymbol{y} \leq \mathbf{c}_k \cdot \boldsymbol{x}$ for $k = 1, \ldots, N$.

Note now that any group action G on a vector space V always has a fundamental region. Consider any $\boldsymbol{x} \in V$ such that $g\boldsymbol{x} \neq \boldsymbol{x}$ unless $g = e$. For this we define

$$K = \{\mathbf{a} \in V : \sup_g [(g\mathbf{a})\cdot\boldsymbol{x}] = \mathbf{a}\cdot\boldsymbol{x}\}, \tag{25}$$

then $F = K_{\text{int}}$ is a fundamental region of G in V, where K_{int} is the interior of the set K. In simple terms, this fundamental region corresponds to the set of vectors that are 'close' to $\boldsymbol{x}$, in the sense that any non-trivial group action $\mathbf{a} \to g\mathbf{a}$ on them moves them away from $\boldsymbol{x}$ with respect the inner product. See [CITE] for a proof of this statement.

E. Computing the fundamental region F for the qutrit

Firstly, note that we actually have that $SL(2,\mathbb{Z}_3)$ is represented by $U(g)$ on $\mathcal{P}_3$, and so we should use that notation for precision.

Since the symplectic group for $d = 3$ is a reflection group, it turns out that this guarantees that an essentially unique fundamental region exists, and so we can reduce to a *finite* set of majorisation conditions (I think 5 in this case). Note that the K defined above is always a closed, convex cone. This means it should be 'easy' to determine the fundamental region. Reflection groups are Coexter groups, and I believe this set K is essentially a Weyl chamber in that language. Anyhow, let's not get distracted by abstraction.

The concrete recipe going forward:

1. Pick your favorite $\boldsymbol{x} \in P_3$ that is *not* stabilized by $g \neq e$. I.e. a vector that moves under all non-trivial group actions.
2. For each g_k look for the extremal cases of $\boldsymbol{y}$ such that have constant inner product with $\boldsymbol{x}$. This corresponds to

   $$[(U(g_k) - \mathbb{1})\boldsymbol{y}]\cdot\boldsymbol{x} = 0 \text{ for } k = 1, \ldots, 24. \tag{26}$$

   Or equivalently,

   $$\boldsymbol{y}\cdot A^T\boldsymbol{x} = 0, \tag{27}$$

   where $A := U(g_k) - \mathbb{1}$.

3. Write down the matrix equation for each g_k.

4. Each of these conditions defines a hyperplane H_k in $\mathcal{P}_3$ corresponding to the boundary of K.

5. We possibly don't need to range over all 24 group elements....but am not sure on this.

6. From this, we should be able to extract a generating set of vectors.

F. Toy example to see how the algorithm works

Okay, let's see how it works in a simple case. Let's consider the case of $V = \mathbb{R}^2$ and $G = S_2 = \langle g | g^2 = e \rangle$ with the group action $g.(x, y) = (y, x)$ that swaps the components of the vector. We should get standard majorisation out of this G–majorisation.

First we find an $\boldsymbol{x}$ that transforms non-trivially under non-trivial group actions. The vector $\boldsymbol{x} = (1, 0)$ does this. To construct a fundamental region we now look at the equation

$$[g.(x, y) - (x, y)] \cdot (1, 0) = 0 \tag{28}$$

and solve for x, y. This becomes,

$$(y - x, x - y) \cdot (1, 0) = 0, \tag{29}$$

which implies the hyperplane (line!) $y = x$ is the *boundary* of the fundamental region. Since we decided to start with $(1, 0)$ we can take the region to the right of this line, and so:

$$F = \{(x, y) : x > y\} \tag{30}$$

where we note that we use the strict inequality to get the interior. The region $\bar{F}$ is a half-space, but this is actually a cone, and can be written as

$$F = \text{cone}((1, 1), (-1, -1), (5, 0)). \tag{31}$$

The first two vectors give the boundary, and we just need one other vector inside F to generate it fully, chosen arbitrarily to be $(5, 0)$. This last one is needed since $(1, 1)$ and $(-1, -1)$ are linearly dependent! Therefore

$$\begin{aligned} \mathbf{c}_1 &= (1, 1) \\ \mathbf{c}_2 &= (-1, -1) \\ \mathbf{c}_3 &= (5, 0). \end{aligned} \tag{32}$$

The S_2–majorisation ordering is then given by

$$\boldsymbol{y} \prec \boldsymbol{x} \Leftrightarrow \mathbf{c}_k \cdot \boldsymbol{y} \leq \mathbf{c}_k \cdot \boldsymbol{x}, \tag{33}$$

for $k = 1, 2, 3$ whenever $\boldsymbol{x}, \boldsymbol{y} \in F$. Note that is gives the ordering for all vectors in the space, since any general vector can be transformed into F via a group action, since F is a fundamental region. Indeed mapping an $\boldsymbol{x}$ into F is simply $\boldsymbol{x} \to \boldsymbol{x}^{\downarrow}$, the sorting maneuver! The generating vectors $(1, 1)$ and $(-1, -1)$ imply that both

$$y_1 + y_2 \leq x_1 + x_2, \tag{34}$$

and

$$y_1 + y_2 \geq x_1 + x_2, \tag{35}$$

which becomes the majorisation condition

$$y_1 + y_2 = x_1 + x_2. \tag{36}$$

Easy! The final condition from $\mathbf{c}_3 = (1, 0)$ is then the condition that

$$5y_1 \leq 5x_1 \Rightarrow y_1 \leq x_1, \tag{37}$$

which is the final majorisation condition. Thus we have shown how the familiar majorisation structure corresponds to the theory of G–majorisation, and the cone ordering structure that comes from fundamental regions.

G. Another toy example – non-standard majorisation this time

Okay, let's look at another option for $V = \mathbb{R}^2$. Now consider the group $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ that can be represented as $G = \langle g_1, g_2 | g^1 = e, g_2 = e, g_1 g_2 = g_2 g_1 \rangle$, and acts on vectors as $g_1.(x, y) = (-x, y), g_2.(x, y) = (x, -y)$. In other words the group just flips vectors about X/Y axes.

A vector that moves under the non-trivial group actions is $(1, 1)$, so we use this one and solve for $[g.(x, y) - (x, y)] \cdot (1, 1) = 0$. We have for $g = g_1$ the equation

$$[(-x, y) - (x, y)] \cdot (1, 1) = 0, \tag{38}$$

which implies the line $x = 0$.

For $g = g_2$ we have

$$[(x, -y) - (x, y)] \cdot (1, 1) = 0, \tag{39}$$

which implies the line $y = 0$.

For $g = g_1 g_2$ we have

$$[(-x, -y) - (x, y)] \cdot (1, 1) = 0, \tag{40}$$

which implies the line $x + y = 0$. The open region bounded by these three conditions is the set

$$F := \{(x, y) : x > 0, y > 0\}. \tag{41}$$

So the positive quadrant of $\mathbb{R}^2$ is a fundamental region for G, which makes sense since via sign flips we can map this onto the whole plane, once we close the set. This is again a convex cone and $\bar{F} = \text{cone}((1, 0), (0, 1))$. Thus given any two vectors $\boldsymbol{x}, \boldsymbol{y}$ we first act with G to flip their signs so that their components are all positive, $\boldsymbol{x} \to \boldsymbol{x}^{\star} \geq \mathbf{0}$, then $\boldsymbol{y} \prec_G \boldsymbol{x}$ if and only if $y_1 \leq x_1$ and $y_2 \leq x_2$. This implies that

$$\boldsymbol{y} \prec_G \boldsymbol{x} \Leftrightarrow \boldsymbol{y}^{\star} \leq \boldsymbol{x}^{\star}. \tag{42}$$

So the $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ majorisation gives the componentwise ordering on vectors once you sort them. It is clear that this extends to any $G = \mathbb{Z}_2^{\times n}$.

H. A more complex example S_3 on $\mathbb{R}^3$.

So the symmetric group S_n always gives majorisation. But the symmetric group has $|S_n| = n!$ elements to it, which is super-exponentially big in n. We clearly don't have to check $n!$ equations, since the majorisation conditions involve just $n+1$ inequalities, and so it must be that the basic generating conditions of the group suffice to determine the fundamental region of G.

Let's look at $V = \mathbb{R}^3$ and $G = S_3$ acting on $\boldsymbol{x} = (x_i)$ as $g.\boldsymbol{x} = (x_{g^{-1}(i)})$. A vector that transforms non-trivially under all non-trivial actions is $(2,1,0)$. Let's look at the transpositions, and the boundary planes they give first.

For $g = (1\ 2)$ we have the equation

$$[(y,x,z) - (x,y,z)] \cdot (2,1,0) = 0, \tag{43}$$

which implies the plane $2(y-x) + x - y = 0$, namely,

$$H_1 = \{(x,y,z) : y = x\}. \tag{44}$$

The other choices of transpositions work the same and give the planes

$$H_2 = \{(x,y,z) : y = z\} \tag{45}$$
$$H_3 = \{(x,y,z) : x = z\}. \tag{46}$$

Note that their intersection gives the line $\{(x,y,z) : x = y = z\}$, which is the line of uniform vectors at the bottom of the majorisation pre-order.

What information do the 3-cycles give? Okay, let's now look at $g = (1\ 2\ 3)$. The core equation for this becomes

$$[(z,x,y) - (x,y,z)] \cdot (2,1,0) = 0, \tag{47}$$

and so gives

$$\begin{aligned} 2(z-x) + (x-y) &= 0 \\ \Rightarrow x + y &= 2z. \end{aligned} \tag{48}$$

Namely the plane

$$H_4 = \{(x,y,z) : x + y = 2z\}. \tag{49}$$

The remaining 3–cycle $(3\ 2\ 1)$ gives

$$H_5 = \{(x,y,z) : y + z = 2x\}. \tag{50}$$

Note that $|S_3| = 6$, but the identity element $g = e$ gives no constraint, and so there are exactly 5 bounding planes. Note also that the line $\{(t,t,t)\}$ is again the intersection of all the planes.

These 5 planes bound a cone containing $(2,1,0)$, namely the fundamental region F. What's a good algorithm to obtain a generating set of vectors for this cone?

Firstly, since the line $\{(t,t,t)\}$ being the intersection of all the planes, is definitely on the boundary of F, this means we must have $\mathbf{c}_1 = (1,1,1)$ and $\mathbf{c}_2 = (-1,-1,-1)$ in our generating set. Perhaps a good algorithm is to compute intersections, obtain generating vectors for these intersections of the hyperplanes, H_{ij} and then let K_i and K_j be the half-spaces in which $(2,1,0)$ resides. Now the intersection of two cones, is itself a cone so we can just keep intersecting the half-space/cones. Thus, we compute

$$\begin{aligned} H_{12} &:= H_1 \cap H_2 && (51)\\ &= \{(t,t,t) : t \in \mathbb{R}\} =: L. && (52) \end{aligned}$$

Similarly, $H_{13} = H_{23} = L$. What about the other intersections? Well $H_{14} = H_{15} = L$. In fact all intersections are the same!

$$H_{ij} = H_i \cap H_j = L \text{ for all } i,j. \tag{53}$$

The fundamental region is given by the intersection of K_1 and K_2, or any other two half-spaces. This is generated by $\mathbf{c}_1$ and $\mathbf{c}_2$ above, and any two independent vectors in the intersection of these half-spaces. We can choose

$$\mathbf{c}_3 = (1,1,0) \tag{54}$$
$$\mathbf{c}_4 = (1,0,0), \tag{55}$$

note that $\mathbf{c}_3 \in H_1$ and $\mathbf{c}_4 \in H_2$, while $\mathbf{c}_3 \in K_2$ and $\mathbf{c}_4 \in K_1$ and so this works correctly. We thus have

$$\bar{F} = \text{cone}(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4). \tag{56}$$

The first two inequalities give the condition that

$$x_1 + x_2 + x_3 = y_1 + y_2 + y_3, \tag{57}$$

while the remaining two conditions give the inequalities

$$y_1 \leq x_1 \tag{58}$$
$$y_1 + y_2 \leq x_1 + x_2. \tag{59}$$

This is the standard majorisation relation. Geometrically, the cone $\bar{F}$ is actually a funny one to visualise: it consists of a half-line radiating out from $(0,0,0)$ to $(-1,-1,-1)$, together with the triangular "infinite pyramid" generated by positive combinations of $(1,1,1),(1,1,0),(1,0,0)$. But we can also take positive linear combinations of $(-1,-1,-1)$ and e.g. $(1,1,0)$ and so it is still a bit tricky to visualise.

I. Symplectic majorisation for $d = 5$ and beyond.

For the case $d = 5$ we also have a particularly simple group structure, where now the symplectic group $SL(2,\mathbb{Z}_5)$ is isomorphic to the symmetry group of the icosahedron. This is another reflection group and therefore we can again obtain the finite set of majorisation conditions to describe it. Here the order of the group is 60 and so we probably shouldn't make Nick compute this by hand.

II. EXTENSION TO GENERAL QUANTUM RESOURCE THEORIES

So far we have introduced the notion of σ–fragments for any resource theory of magic. In this section we briefly generalise this concept to arbitrary resource theories and explain precisely how it connects with resource monotones. The busy reader more focussed on magic may skip this section.

State convertibility within a given resource theory is often a hard question to address due to the intricate structure of the theory. In general, the structure of a theory $\mathcal{R}$ is described by a pre-order $\prec_{\mathcal{R}}$ and usually resource monotones are employed to reduce this structure into a simple real number ordering [CITE]. The subdivisions of magic theories into σ–fragments suggests a new approach towards investigating state convertibility which retains more structure of the originaltheory than a measure can.

Monotones reduce the structure of the resource theory $\mathcal{R}$ to a *total* order on the real numbers. Therefore, two states, even if incomparable in $\mathcal{R}$, are always mapped onto ordered real numbers. We now generalise this idea of a theory projection that preserves comparability between states.

Definition 4 (**Covariant projection**). *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a resource theory with pre-order $\prec_{\mathcal{R}}$. Then a* covariant resource projection *of $\mathcal{R}$ to a resource theory $\mathcal{R}'$ with pre-order $\prec_{\mathcal{R}'}$, is a pair of mappings (Π_s, Π_o), where Π_s maps quantum states in $\mathcal{R}$ to quantum states in $\mathcal{R}'$, and Π_o maps free operations in $\mathcal{R}$ to free operations in $\mathcal{R}'$. Moreover, these obey*

1. $\Pi_{\mathrm{s}}(\rho_1) \prec_{\mathcal{R}'} \Pi_{\mathrm{s}}(\rho_2)$ *whenever* $\rho_1 \prec_{\mathcal{R}} \rho_2$;
2. $\Pi_{\mathrm{o}}(\mathcal{E}) = \Pi_{\mathrm{o}}(\mathcal{E}_1) \circ \Pi_{\mathrm{o}}(\mathcal{E}_2)$ *whenever* $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2$.

We call $\mathcal{R}'$ a covariant fragment *of $\mathcal{R}$.*

Resource monotones can now be clearly seen as a special case of covariant resource projections.

Proposition 5 (**Totally ordered covariant theories**). *Any resource monotone $\mathcal{M}$ of a resource theory $\mathcal{R}$ is a covariant projection for which $\prec_{\mathcal{R}'}$ is a total order. Conversely, any such covariant projection corresponds to a resource monotone $\mathcal{M}$.*

Proof. Consider a monotone $\mathcal{M}$ in the context of a general resource theory $\mathcal{R} = (\mathcal{F}, \mathcal{O})$. State order is covariantly preserved due to the defining property of a monotone, stated in **??**, where the pre-order $\prec_{\mathcal{R}'}$ is simply the total order $\leq$ on $\mathbb{R}$.

Operational composition is covariantly preserved when we simply choose $\Pi_{\mathrm{o}}(\mathcal{E}) = 1_{\times}$, namely the 'multiplication by 1' operation on real numbers. The definition of a resource monotone then automatically implies covariance.

Conversely, given any covariant projection of $\mathcal{R}$ for which $\prec_{\mathcal{R}'}$ is a total order, we may map the totally ordered set of elements $\Pi_{\mathrm{s}}(\rho)$ via an injective, non-decreasing function f into $\mathbb{R}$. Then, $\mathcal{M}(\rho) := f(\Pi_s(\rho))$ provides a numerical value for each ρ that obeys the definition of a monotone.

□

We can also view σ–fragments as an example of reducing the structure of a magic theory $\mathcal{R}$ to a subtheory with a tractable pre-order. However, states which are incomparable in $\mathcal{R}$ remain incomparable and conversions between states which are comparable in $\mathcal{R}$ may no longer be possible.

Definition 6 (**Contravariant projection**). *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a resource theory with pre-order $\prec_{\mathcal{R}}$. Then a* contravariant resource projection *of $\mathcal{R}$ onto a resource theory $\mathcal{R}'$ with pre-order $\prec_{\mathcal{R}'}$, is a pair of mappings (Π_s, Π_o), where Π_s maps quantum states in $\mathcal{R}$ onto quantum states in $\mathcal{R}'$, and Π_o maps free operations in $\mathcal{R}$ onto free operations in $\mathcal{R}'$. Moreover, these obey*

1. $\rho_1 \prec_{\mathcal{R}} \rho_2$ *whenever* $\Pi_{\mathrm{s}}(\rho_1) \prec_{\mathcal{R}'} \Pi_{\mathrm{s}}(\rho_2)$;
2. $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2$ *whenever* $\Pi_{\mathrm{o}}(\mathcal{E}) = \Pi_{\mathrm{o}}(\mathcal{E}_1) \circ \Pi_{\mathrm{o}}(\mathcal{E}_2)$.

We call $\mathcal{R}'$ a contravariant fragment *of $\mathcal{R}$.*

The use of covariant and contravariant in Definitions 4 and 6 refers to the direction of implication between the two pre-orders and operation compositions[1]

Proposition 7. *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a resource theory, and let $\mathcal{O}' \subseteq \mathcal{O}$ be a non-empty subset of the free operations that is closed under composition, and moreover $\mathcal{O}'$ is the largest such subset, in the sense that for any $\mathcal{E}_1 \notin \mathcal{O}'$ and any $\mathcal{E}_2 \in \mathcal{O}'$ we have that both $\mathcal{E}_1 \circ \mathcal{E}_2$ and $\mathcal{E}_2 \circ \mathcal{E}_1$ are not in $\mathcal{O}'$. Then $\mathcal{R}' = (\mathcal{F}, \mathcal{O}')$ of $\mathcal{R}$ defines a contravariant fragment of $\mathcal{R}$.*

Proof. We first define $\Pi_{\mathrm{s}}(\rho) = \rho$ for all ρ. It is clear that since $\mathcal{O}'$ is a subset of $\mathcal{O}$ any operation in $\mathcal{O}'$ will map the set of free states into itself. Moreover the identity channel id is necessarily in $\mathcal{O}'$, due to the maximality assumption. For Π_{o} we let $\Pi_{\mathrm{s}}(\mathcal{E}) = \mathcal{E}$ if $\mathcal{E} \in \mathcal{O}'$ and otherwise $\Pi_{\mathrm{s}}(\mathcal{E}) = id$. Now consider $\Pi_{\mathrm{o}}(\mathcal{E}_1 \circ \mathcal{E}_2)$. Either the triple $\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_1 \circ \mathcal{E}_2\}$ are all in $\mathcal{O}'$ or they are all outside of $\mathcal{O}'$. For the former case $\Pi_{\mathrm{o}}(\mathcal{E}_1 \circ \mathcal{E}_2) = \mathcal{E}_1 \circ \mathcal{E}_2 = \Pi_{\mathrm{o}}(\mathcal{E}_1) \circ \Pi_{\mathrm{o}}(\mathcal{E}_2)$, while for the latter $\Pi_{\mathrm{o}}(\mathcal{E}_1 \circ \mathcal{E}_2) = id = \Pi_{\mathrm{o}}(\mathcal{E}_1) \circ \Pi_{\mathrm{o}}(\mathcal{E}_2)$, which proves that compositions are respected under the map. Finally, $\rho \prec_{R'} \sigma$ implies there exists $\mathcal{E} \in \mathcal{O}' \subseteq \mathcal{O}$ such that $\mathcal{E}(\rho) = \sigma$, and since $\mathcal{E} \in \mathcal{O}$ this implies $\rho \prec_{\mathcal{R}} \sigma$, as required, which completes the proof.

□

[1] Note that strictly these are not projections in the sense of $\Pi^2 = \Pi$, but are instead morphisms. Here our use of the term projection is motivated by the idea that one one generally loses information about $\mathcal{R}$ under the mapping.

[Suppose $\mathcal{E}_2$ is the replacement channel $\mathcal{E}_2(\rho) = \sigma$. This is a stabiliser operation. Then $\mathcal{E}_2 \circ \mathcal{E}_1 \in \mathcal{O}_\sigma$ even if $\mathcal{E}_1 \notin \mathcal{O}_\sigma$. Now suppose $\mathcal{E}$ is a Hadamard unitary, then $id = \mathcal{E} \circ \mathcal{E}_{reverse}$ but $id \in \mathcal{O}_{|0\rangle\langle 0|}$ while $\mathcal{E}, \mathcal{E}_{reverse} \notin \mathcal{O}_{|0\rangle\langle 0|}$.] [Ah ok, agreed. This is fine for now. Let's not spend ages trying to generalise this so as to include the majorisation fragments. It's just good to explore the possibilities a bit, to illustrate the non-trivial aspects.] As an immediate corollary of Proposition 7, a σ–fragment of any magic theory $\mathcal{R}$ is a contravariant fragment of $\mathcal{R}$.

Proposition 8. *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a resource theory, and let $\mathcal{D} \in \mathcal{O}$ be a free operation, which is reversible by $\mathcal{D}_{\text{rev}} \in \mathcal{O}$, so that $\mathcal{D}_{\text{rev}} \circ \mathcal{D} = 1_{\text{C}}$.*

Then, we can define a contravariant projection of $\mathcal{R}$, by acting on all quantum states with $\mathcal{D}$.

Proof. We show that the theory $\mathcal{R}' = (\mathcal{F}', \mathcal{O})$, with $\mathcal{F}' = \{\mathcal{D}(\rho) : \rho \in \mathcal{F}\}$, is a contravariant fragment of $\mathcal{R}$.

Let Π_{s} map every state ρ to $\mathcal{D}(\rho)$ and suppose $\mathcal{D}(\rho_1) \prec \mathcal{D}(\rho_2)$. Then, there exists $\mathcal{E} \in \mathcal{O}$ such that $\rho_1 = (\mathcal{D}_{\text{rev}} \circ \mathcal{E} \circ \mathcal{D})(\rho_2)$, so $\rho_1 \prec \rho_2$.

Finally, let Π_{o} map every free operation to itself, so that composition of operations is trivially preserved.

[If $\mathcal{D}$ is a recovery map, so that $\mathcal{D} \circ \mathcal{D}_{\text{rev}} = 1_{\text{C}}$, then this is a covariant projection instead.

If $\mathcal{D}$ is not reversible, this mapping is in general NOT contravariant (consider the replacement map $\mathcal{D}(\rho) = \frac{1}{d}\mathbb{1}$ for a strange state and stabiliser state - surely there is such a counterexample in thermodynamics theory if we consider a highly coherent state and one with the same energy population but no coherences.] [Ok. Again, let's not spend time worrying about this now. It's clear there is various fine-print to these cases...but they're not essential to our work so let's put this on hold.] □

Important examples of resource fragments appear in several established resource theories. [Need to check if the thermodynamics example works, include magic theories as fragments of $\mathcal{R}_{\max}$, include Nielsen's bipartite entanglement.] [Don't worry about these things now – let's get the computations section improved]

III. COMPARISON WITH EXISTING DISTILLATION RATES

We have derived a *distillation bound* $R(\epsilon, \epsilon', \beta)$ for the process

$$\rho_{\text{S}}(\epsilon)^{\otimes n} \longrightarrow \rho_{\text{S}}(\epsilon')^{\otimes n'} \otimes \gamma_\beta^{\otimes(n-n')}, \tag{60}$$

with even n, n' and $0 \leq \epsilon' < \epsilon \leq 3/7$, where the n-copy, ϵ-noisy Strange state is

$$\rho_{\text{S}}(\epsilon)^{\otimes n} := \left[(1-\epsilon)\,|\text{S}\rangle\langle\text{S}| + \epsilon\frac{1}{3}\mathbb{1}\right]^{\otimes n}. \tag{61}$$

The bound informs us that any distillation protocol in the γ_β-fragment with an ϵ-noisy input and ϵ'-noisy output has a *distillation rate*

$$\frac{n'}{n} \leq R(\epsilon, \epsilon', \beta). \tag{62}$$

We apply our results on the ternary Golay code distillation protocol suggested in [**?**]. In this protocol, a single Strange state is distilled with noise level

$$\epsilon'(n, \epsilon) \approx \frac{1}{\alpha}(\alpha\epsilon)^{n^\xi}, \tag{63}$$

where $\alpha \approx 1.75$ and $\xi \approx 0.112$ are constants specific to the protocol. This can be rewritten in terms of the protocol distillation rate,

$$R_{\text{Golay}}(\epsilon, \epsilon') = \frac{1}{n} \approx \left(\frac{\log(\alpha\epsilon)}{\log(\alpha\epsilon')}\right)^{\frac{1}{\xi}} \tag{64}$$

We compare the distillation rate of the protocol with our distillation bound,

$$R_{\text{Golay}}(\epsilon, \epsilon') \leq R(\epsilon, \epsilon', \beta) \tag{65}$$

in figure Fig. (1)

We can also rewrite the expression as

$$\epsilon_{\text{Golay}}(\epsilon', n) \approx \frac{1}{\alpha}(\alpha\epsilon')^{n^{-\xi}}, \tag{66}$$

and compare with the numerically optimal error bounds provided by majorisation as shown in Fig. (2)

A recent distillation bound was introduced in [**?**], where they show that for the deterministic conversion of n copies of a qubit state ρ to a target pure magic state $|\psi\rangle$, the following bound on the distillation rate holds,

$$\frac{n'}{n} \leq \frac{\log \Lambda^+(\rho)}{\log F(\psi)^{-1}}, \tag{67}$$

where $\Lambda^+(\rho)$ is the generalised robustness of state ρ, calculable via a costly SDP and $F(\psi)$ denotes the stabiliser fidelity of state $|\psi\rangle$. They demonstrate the performance of their bound by considering the purification process of the qubit magic state $|H\rangle$, with their figure replicated

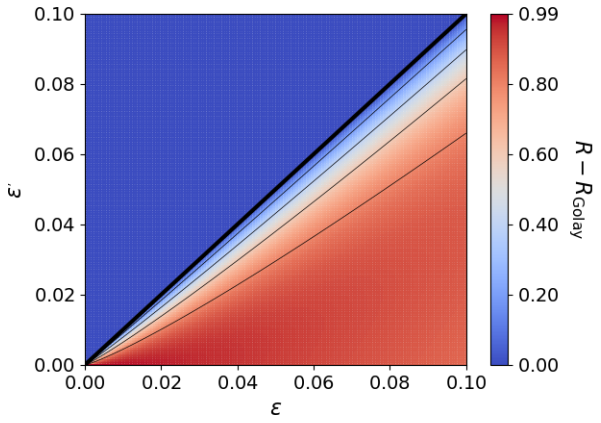


FIG. 1. **Comparison with ternary Golay distillation rate.** Our bound is not violated but gets close to saturated for some noise levels. The thick diagonal line indicates where $\epsilon' = \epsilon$.

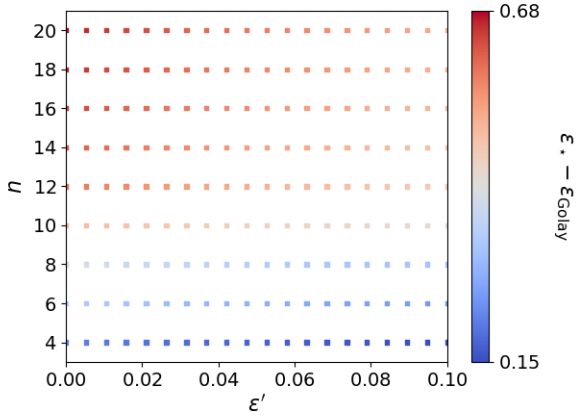


FIG. 2. **Comparison with ternary Golay output error rate.** Our bound is not violated but achieves a maximum saturation of $\epsilon_\star - \epsilon_{\text{Golay}} = 0.147$.

in Fig. (3). For high initial noise level ($\epsilon = 0.25$), their bound dictates that $R = 1/150$, which is pretty good.

Similar concrete rates have been studied for qudits in two occasions: First, the mana bound, which we have demonstrated in **??**. Secondly, Wang *et al.*'s max-thauma bound [**?**], which as far as I can tell, is less tight than mana (?). It is defined via an SDP, but we can use its properties (super-additivity and zero at free states) to readily get a bound on noisy Strange states,

$$\theta_{\max}(\rho_{\mathrm{S}}(\epsilon)) \geq (1-\epsilon)\,|\mathrm{S}\rangle\langle\mathrm{S}| + \epsilon\frac{1}{3}\mathbb{1} = (1-\epsilon)\theta_{\max}(\rho_{\mathrm{S}}(0)). \tag{68}$$

Their bound simply states that any deterministic single-copy distillation process $\rho^{\otimes n} \longrightarrow \tau$ requires an initial number of copies $n \geq \theta_{\max}(\tau)/\theta_{\max}(\rho)$. Therefore, the rate of our pure Strange state distillation process **??** can

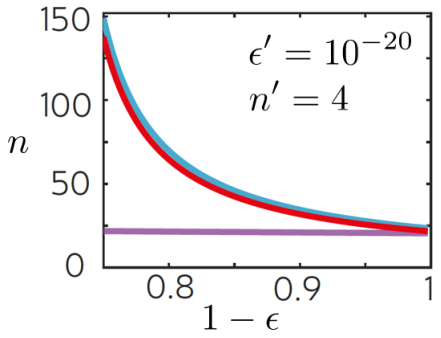


FIG. 3. **Comparison of Seddon *et al.*'s bound [**? **] in Eq. (67) with a generic resource bound in [**? **].** The process they consider is purifying n copies of $(1-\epsilon)\,|H\rangle\langle H| + \epsilon/3\mathbb{1}$ to n' copies of $(1-\epsilon')\,|H\rangle\langle H| + \epsilon'/3\mathbb{1}$. Their bound (red and blue) is pretty good compared with the generic (violet) bound. Not sure why it looks like their bound gives $n \to 25$ as $1-\epsilon \to 1$.

only get as tight as

$$R = 1/n = \frac{\theta_{\max}(\rho_{\mathrm{S}}(\epsilon))}{\theta_{\max}(\rho_{\mathrm{S}}(0))} \leq 1-\epsilon, \tag{69}$$

which is less tight than mana as could be readily observed if plotted in **??**.

An n-copy state ρ^n has *error rate* δ iff ρ has an overlap of at least $1-\delta$ with magic state $|\mathrm{S}\rangle$,

$$\langle\mathrm{S}|\,\rho\,|\mathrm{S}\rangle \geq 1-\delta. \tag{70}$$

The ϵ-noisy Strange state $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ has error rate $\delta = \frac{2}{3}\epsilon$. A general state of n qutrits with error rate δ can be converted to the form $\rho_{\mathrm{S}}(\epsilon_0)^{\otimes n}$ for some noise level ϵ not necessarily equal to δ.

Given an $[[m,k,d]]$ error correcting code, a distillation protocol that achieves the process in Eq. (61) requires that the following bound holds,

$$\frac{n}{n'} = O(\log^\gamma(1/\epsilon')),\ \gamma = \log_d\left(\frac{m}{k}\right) \tag{71}$$

In [**?** **?**], the asymptotic rate $O(\log^\gamma(1/\epsilon'))$ is called the *distillation cost*, and in [**?** **?**] the parameter γ is called the *overhead*.

There are many distillation protocols in the literature, e.g. [**?** **?**], that involve a specific value of γ and convert a specific number of copies n of input states to $n' = 1$ output state [what about ϵ?]. They should all satisfy

$$\frac{1}{R(\epsilon,\epsilon',\beta)} \leq \frac{n}{n'} = O(\log^\gamma(1/\epsilon')). \tag{72}$$

Therefore, given process parameters $n, n', \epsilon, \epsilon'$, we can scan through the stabiliser states to find a range of β for which inequality Eq. (72) is valid. In principle, the higher the value of parameter γ, the lower the output error rate ϵ' can be while Eq. (72) is still true. In [**?**] [CITE

MORE], they report a value of $\gamma = 0.6779\ldots$, while in [**?**], they report a scaling $\gamma = O(1/\log d)$ with respect to the system's odd prime dimension d. The latter tends to 0 asymptotically.

We can rephrase this result, saying that any distillation protocol in the γ_β-fragment with an n-copy input and n'-copy, ϵ'-noisy desired output, must have input noise level

$$\epsilon \leq E\left(\frac{n'}{n}, \epsilon', \beta\right) \tag{73}$$

for some error bound E related to R.

Furthermore, in [**? ?**], two different distillation protocols are proposed with output error rate scalings of the form

$$\epsilon' \propto (\alpha\epsilon)^{n^\xi}, \tag{74}$$

where α is some constant, $\xi = 1/\log_3 15 \approx 0.4$ in [**?**] for qubit codes and $\xi = 1/\log_3 19008 \approx 0.1$ for qutrit codes. They also report a certain input threshold error rate $\epsilon_\star$ above which distillation is not possible. A first requirement is that $\epsilon_\star \leq E(n'/n, \epsilon', \beta)$, but we can also again find a range of β for which the inequality

$$\epsilon \leq E\left(\frac{n'}{n}, \epsilon'(\epsilon), \beta\right) \tag{75}$$

$$\epsilon' \sim O\left((n\epsilon)^d\right) \tag{76}$$

holds, given some distillation rate n'/n and input error rate ϵ.

IV. RANDOM INFO

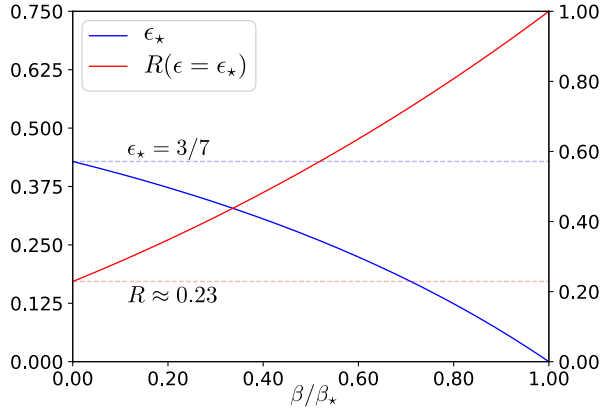


FIG. 4. **Noise threshold $\epsilon_\star$ and rate $R(\beta, \epsilon_\star)$ versus β.** At $\beta = 0$ and threshold noise level $\epsilon = \epsilon_\star = 3/7$, the rate is $R(0, \epsilon_\star) = \frac{\ln(9/7)}{\ln 3} \approx 0.23$.

We generalise the analysis of unital fragments into any circuit with thermalisation noise parametrised by the temperature β^{-1}. In Fig. (5), we examine the majorisation bound for the same purification process of **??** with $\epsilon' = 0.05$. The curves plotted suggest that there exists a fragment $\mathcal{O}_{\gamma_{\beta_{\max}}}$ which allows for a highest noise threshold at any given number of copies of the initial state. Adding more copies results in a lower optimal temperature β^{-1}.

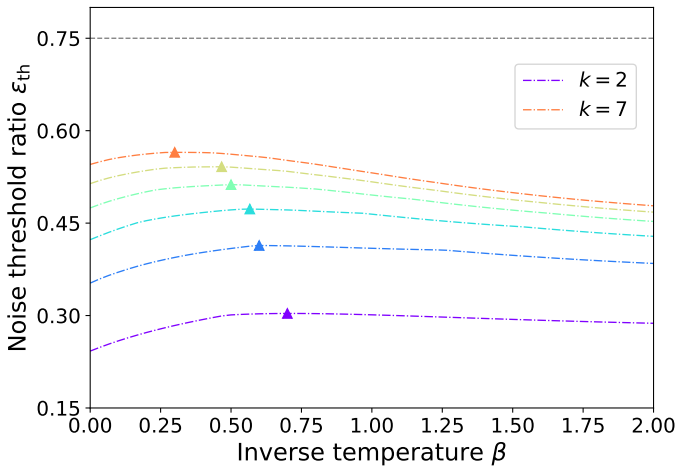


FIG. 5. **Threshold dependence on temperature in thermal fragments.** Lorenz curve ratios for the Strange state purifying process in **??** with $\epsilon' = 0.05$. The peaks of each curve indicate the optimal temperature $\beta_{\max}^{-1}$ that allows for the highest noise threshold at every given number of initial state copies k. The line $\epsilon = \frac{3}{4}$ indicates the threshold noise beyond which the Strange state no longer contains negativities.

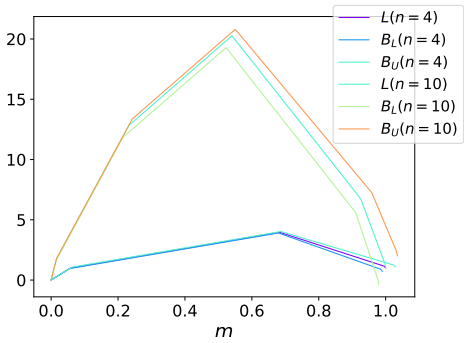


FIG. 6. **Lorenz curves and bounding curves.** The figure explores the distillation process of a 10-copy 0.25–noisy Strange state to a 4-copy 0.05–noisy Strange state in the unital fragment. Clearly, by looking at the Lorenz curves, the distillation process is possible. However the bounding curves, derived in **??**, intersect at the end, suggesting that the process is not necessarily possible.

V. PHI BOUNDS

The experiment is described by binomial statistics and we define the left tail of the cumulative binomial distribution,

$$\Phi_\ell(m;n,p) \coloneqq \sum_{j=0}^{m} \binom{n}{j} p^j (1-p)^{n-j}, \tag{77}$$

where $m \in [0,n]$.

The symmetry between the left tail of a p–coin distribution and the right tail of a $(1-p)$–coin distribution dictates that

$$\Phi_\ell(m;n,p) + \Phi_\ell(n-m-1;n,1-p) = 1, \ \ m \in [0,n-1], \tag{78}$$

Entropic bounds on Φ_ℓ [**?**].

Lemma 9. *Given fixed $n > 0$ and p, Φ_ℓ satisfies the following bounds:*

1. $\Phi_\ell(m;n,p) \geq \left[8m\left(1-\frac{m}{n}\right)\right]^{-\frac{1}{2}} 2^{-nS(p||q)}, \ \ m \in [1,n-1]$
2. $\Phi_\ell(m;n,p) \leq 2^{-nS(p||q)}, \ \ m \in [0,np]$
3. $\Phi_\ell(m;n,p) \leq 1 - \left[8(m+1)\left(1-\frac{m+1}{n}\right)\right]^{-\frac{1}{2}} 2^{-nS\left(\frac{m+1}{n}||p\right)}, \ \ m \in [0,n-2]$
4. $\Phi_\ell(m;n,p) \geq 1 - 2^{-nS\left(\frac{m+1}{n}||p\right)}, \ \ m \in [np+1,n-2]$

Proof. [cite proof] □

Loose bounds follow

Lemma 10. *Given fixed $n > 0$ and p, Φ satisfies the following bounds:*

1. $\Phi_{\pm}(m;n,p) \geq \left[8m\left(1-\frac{m}{n}\right)\right]^{-\frac{1}{2}} 2^{-nS(p||q)}$, $m \in [1, n-1]$
2. $\Phi_{\pm}(m;n,p) \leq 2^{-nS(p||q)}$, $m \in [0, np]$
3. $\Phi_{\pm}(m;n,p) \leq 1 - \left[8(m+1)\left(1-\frac{m+1}{n}\right)\right]^{-\frac{1}{2}} 2^{-nS\left(\frac{m+1}{n}||p\right)}$, $m \in [0, n-2]$

Proof. [paste proof] □

We can therefore use Lemma 10 and **??** to bound the standard Lorenz curves.

A. Strange state MSD in the unital fragment

Consider the Strange state MSD process in the unital fragment,

$$\rho_{\mathrm{S}}(n', \epsilon, 0) \xrightarrow{\mathcal{O}_{1/3}} \rho_{\mathrm{S}}(n', \epsilon', n-n'). \tag{79}$$

We denote input state indices without a prime and target state indices with a prime,

$$I(i,j,k=1) = j + \sum_{\ell=0}^{i-1} m_\ell(n,\epsilon), \tag{80}$$

$$I'(i',j',k') = k' + \left[(j'-1) + \sum_{\ell=0}^{i'-1} m_\ell(n',\epsilon')\right] 9^{n-n'}. \tag{81}$$

Pointwise Lorenz curve comparison requires $x_I = x_{I'}$, so the question is:

Given a triplet (i',j',k'), what is the tuple (i,j) such that $I(i,j) = I'(i',j',k')$?

According to **??** which is proved in **??**, for standard Lorenz curves, we need to match indices at the target state elbows, so the requirement on the indices is finding a tuple (i,j), such that for a given $i' = 0, \ldots, n'$,

$$j + \sum_{\ell=0}^{i-1} m_\ell(n,\epsilon) = \sum_{\ell=0}^{i'} m_\ell(n',\epsilon'). \tag{82}$$

As a basic example, consider the process

$$\rho_{\mathrm{S}}(\epsilon)^{\otimes 4} \xrightarrow{\mathcal{O}_{1/3}} \rho_{\mathrm{S}}(\epsilon')^{\otimes 2} \otimes \left(\frac{1}{3}\mathbb{1}\right)^{\otimes 2}. \tag{83}$$

We want to check which Lorenz curve is higher at the first elbows of the target state, i.e. we want to verify or reject the first inequality below:

$$L_{I(i,j)} \geq L'_{I'(0,1,81)} \tag{84}$$

where the first multiplicity of the target state is $m_0 = 1$. The multiplicities of the initial state are $(1, 384, 4096)$.

The challenge is to find i,j such that $i(i,j) = i'(0,1,81)$. [By trial and error], we find that $(i,j) = (1,80)$. Now we can use **??** to directly calculate

$$\begin{aligned} L'_{i'(0,1,81)} = L'_0 &= \left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^2 \Phi_+\left(0;2,4\frac{3-\epsilon}{15-8\epsilon}\right), \\ L_{i(1,80)} &= \left(1-\frac{80}{384}\right) L_0 + \frac{80}{384} L_1 \\ &= \left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^4 \left[\frac{19}{24}\Phi_+\left(0;4,4\frac{3-\epsilon}{15-8\epsilon}\right)\right. \\ &\quad \left. + \frac{5}{24}\Phi_+\left(2;4,4\frac{3-\epsilon}{15-8\epsilon}\right)\right], \end{aligned}$$

and then compare them.

With the help of this lemma, we directly arrive at

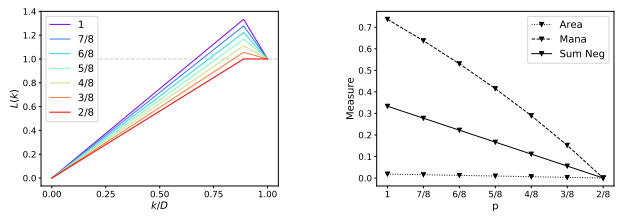
Theorem 11. *Given fixed positive integer n and probability p, Φ_+, Φ_- satisfy the following bounds:*

1. $\Phi_+(m;n,p) \geq \sum_{\ell=0}^{m/2} \left[16\ell\left(1-\frac{2\ell}{n}\right)\right]^{-\frac{1}{2}} 2^{-nS\left(\frac{2\ell}{n}||p\right)}$, *for all even $m \in [2,n]$*
2. $\Phi_+(m;n,p) \leq \sum_{\ell=0}^{m/2} \left[4\pi\ell\left(1-\frac{2\ell}{n}\right)\right]^{-\frac{1}{2}} 2^{-nS\left(\frac{2\ell}{n}||p\right)}$, *for all even $a \in [2,n]$*
3. $\Phi_-(m;n,p) \geq \sum_{\ell=1}^{(m-1)/2} \left[16(\ell+1)\left(1-\frac{2\ell+1}{n}\right)\right]^{-\frac{1}{2}} \times$ $\times\, 2^{-nS\left(\frac{2\ell+1}{n}||p\right)}$, *for all odd $a \in [1,n]$*
4. $\Phi_-(m;n,p) \leq \sum_{\ell=1}^{(m-1)/2} \left[4\pi(\ell+1)\left(1-\frac{2\ell+1}{n}\right)\right]^{-\frac{1}{2}} \times$ $\times\, 2^{-nS\left(\frac{2\ell+1}{n}||p\right)}$, *for all odd $a \in [1,n]$*

Proof. All four statements follow from application of **??** on the combinatorial coefficient and the defintion of relative entropy given in **??** □

VI. AREA MONOTONE

[I don't believe in the area monotone - should we keep?]



(a) (b)

FIG. 7. (a) Lorenz curve of $p\,|\mathrm{S}\rangle\langle\mathrm{S}| + (1-p)\frac{1}{d}\mathbb{1}$ for p given in the legend.(b) Different measures for the states on the left. [can replace ?? with a more informative version of this figure]

Let $L_{>1}$ be the set of points on the Lorenz curve $L_{\rho|\sigma}(k)$ that lie above 1 of state ρ in the σ–fragment. If ρ is a free state, $L_{>1}$ is empty and $\mathcal{A}_\sigma(\rho) = 0$. Otherwise, $\mathcal{A}_\sigma(\rho) > 0$ and it can be calculated exactly using the trapezium rule or the shoelace formula. Let k be the index of the first point $\big(x_k, L_{\rho|\sigma}(k)\big)$ lying above 1. Then $L_\rho(k)$ crosses 1 at

$$x_{\text{int}} = x_k - \frac{x_k - x_{k-1}}{L_{\rho|\sigma}(k) - L_{\rho|\sigma}(k-1)} L_{\rho|\sigma}(k), \tag{85}$$

as well as at $\big(x_d, L_{\rho|\sigma}(d)\big) = (1,1)$

Now we can define

$$L^+_{>1} = \{(x_i, y_i)\}_{0\le i\le n} \tag{86}$$

such that it contains the initial point of intersection $(x_0, y_0) \equiv (x_{n+1}, y_{n+1}) \coloneqq (x_{\text{int}}, 1)$, all points (x_i, y_i), labelled by $i = 1, \ldots, n-1$ that lie above $y = 1$ contained in $L_{>1}$ and finally the second point of intersection $(x_n, y_n) = (1,1)$. Then,

$$\mathcal{A}_\sigma(\rho) = \frac{1}{2}\sum_{i=0}^{n}(x_{i+1}y_i - x_i y_{i+1}). \tag{87}$$

We note that single-copy Lorenz curves are additive in noise,

$$L_{(1-\epsilon)\rho+\epsilon\sigma|\sigma} = (1-\epsilon)L_{\rho|\sigma} + \epsilon L_{\sigma|\sigma}, \tag{88}$$

for any σ, ρ. This is not true for higher number of copies.

[In particular], it can be used in the discussion of standard Lorenz curves to set the necessary constraints,

$$\begin{aligned} L\left(x_{I'(0,m'_0,9^{n-n'})}\right) &\ge L\left(x_{I'(0,m'_0,9^{n-n'})}\right), \\ L\left(x_{I'(1,m'_1,9^{n-n'})}\right) &\ge L\left(x_{I'(1,m'_1,9^{n-n'})}\right), \\ &\vdots \\ L\left(x_{I'(n',m'_r,9^{n-n'})}\right) &\ge L\left(x_{I'(n',m'_{n'},9^{n-n'})}\right), \end{aligned}$$

where $x_{I'(0,m'_0,9^{n-n'})}, x_{I'(1,m'_1,9^{n-n'})}, \ldots, x_{I'(n',m'_{n'},9^{n-n'})} \in T$ are the locations of the elbows of $L_{\rho_{\mathrm{S}}(\epsilon')^{\otimes n'}|\frac{1}{d}\mathbb{1}}$ (excluding 0 and 1).

VII. STABILIZER PROOF DETAILS

Whether we are dealing with (C1) or (C2) depends on the physical parameters β and ϵ as follows:

- $\beta = 0$. Then, $\epsilon_\star = 3/7$ and (C1) holds for all noise levels ϵ. This is the unital fragment.
- $0 < \beta \leq \beta_\star$. Then, $0 \leq \epsilon_\star < 3/7$ and (C1) holds for $\epsilon \leq \epsilon_\star$, while (C2) holds for $\epsilon \geq \epsilon_\star$.
- $\beta > \beta_\star$. Then, $\epsilon_\star = 0$ and (C2) holds for all ϵ.

Consider pure Strange state distillation at high temperatures, $\beta \leq \beta_\star$. Depending on the noise level of the initial state, we have either a (C1) $\to$ (C1) or a (C2) $\to$ (C1) scenario.

At every $\beta > 0$, the rates indicate a transition at $\epsilon_\star$. As β increases, the rate becomes stricter. At the high temperature limit, $\beta \to 0$, we get $\beta F_\beta \to -\ln 3$, and we retrieve the unital fragment bound,

$$R(\beta) \xrightarrow{\beta\to 0} \frac{\ln(3-4\epsilon)}{\ln 3}. \tag{89}$$

Consider pure Strange state distillation at low temperatures, $\beta \to \infty$. Independently of the noise level of the initial state, we have a (C2) $\to$ (C2) scenario.

As β increases, the rate becomes looser. At the zero temperature limit, F_β tends to zero, so $R(\epsilon,\beta) \xrightarrow{\beta\to\infty} 1$.

A. First elbow location

Consider a magic state interconversion in a stabilizer σ–fragment, as in Eq. (60), where we remind that $n \geq n'$ and we denote by (x_0, L_0) and (x_0', L_0') the first elbow coordinates of the initial and target states respectively.

Here, we show that $x_0 \leq x_0'$ for any of the three scenarios outlined in the proof of our main theorem in **??**.

(C1) $\to$ (C1). We know from statistical physics that $e^{-\beta E_0}/\mathcal{Z}_\beta \leq 1$, so

$$\begin{aligned} x_0 &= \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^n = \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n-n'} \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n'} \\ &< \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n'} = x_0' \end{aligned}$$

(C2) $\to$ (C1). We now use the slightly altered inequality $e^{-\beta E_{\max}}/\mathcal{Z}_\beta \leq 1$ to proceed,

$$\begin{aligned} x_0 &= \left(\frac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^n = \frac{e^{-\beta(nE_{\max}-n'E_0)}}{\mathcal{Z}_\beta^{n-n'}} \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n'} \\ &\leq \frac{e^{-\beta(nE_{\max}-n'E_0)}}{e^{-\beta(n-n')E_{\max}}} \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n'} \\ &= e^{-\beta n'(E_{\max}-E_0)} \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n'} \\ &\leq \left(\frac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^{n'} = x_0'. \end{aligned}$$

(C2) $\to$ (C2). Similarly in this scenario,

$$\begin{aligned} x_0 &= \left(\frac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^n = \left(\frac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^{n-n'} \left(\frac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^{n'} \\ &\leq \left(\frac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^{n'} = x_0' \end{aligned}$$

B. Deriving distillation bounds from the last elbow

[RESTRUCTURE]

To find the coordinates of the **last** elbow (x_E, L_E), where E is the number of elbows, we need to evaluate the minimum rescaled component,

$$\begin{aligned} &\boldsymbol{w}(\rho_S|\sigma)_{\min} \coloneqq (3\mathcal{Z})^n \times \\ &\min_{i,j,k} \left\{ (-v)^{n-\alpha} u^\alpha e^{\beta(n-\alpha)E_0} e^{\beta(iE_0+jE_1+kE_2)} \right\}, \end{aligned} \tag{90}$$

where $0 \leq i,j,k \leq n$ and $\alpha \coloneqq i+j+k \leq n$. Notice that for $0 \leq \epsilon \leq 3/7$, we have $v \geq u$. We need the sum $\alpha = i+j+k$ to be odd for the expression to be negative.

Given an odd value for the sum α, the term $v^{n-\alpha}u^\alpha e^{-\beta(n-\alpha)E_0}$ is fixed (and negative), so the expression is minimised by setting the coefficient of the highest energy $E_{\max}$ equal to α. Hence, we have

$$\begin{aligned} &\boldsymbol{w}(\rho_S|\sigma)_{\min} = \\ &-(3\mathcal{Z})^n v^n e^{n\beta E_0} \max_{\substack{\alpha=1,3,\\ \ldots,n-1}} \left\{ \left(\frac{u}{v} e^{\beta(E_{\max}-E_0)}\right)^\alpha \right\}. \end{aligned} \tag{91}$$

If the expression $\frac{u}{v}e^{\beta(E_{\max}-E_0)}$ is less than 1 then the minimum occurs at $\alpha = 1$, otherwise it occurs at $\alpha = n-1$.

The minimum rescaled component can then be expressed as

$$\boldsymbol{w}(\rho_S|\sigma)_{\min} = \begin{cases} -(3\mathcal{Z})^n v^{n-1} u e^{\beta[(n-1)E_0+E_{\max}]}, & \epsilon \leq \epsilon_\star, \quad \text{(C1)} \\ -(3\mathcal{Z})^n v u^{n-1} e^{\beta[E_0+(n-1)E_{\max}]}, & \epsilon > \epsilon_\star. \quad \text{(C2)} \end{cases} \tag{92}$$

Case (C1) can correspond to $(i,j,k) = (1,0,0)$ if $E_{\max} = E_0$, when the multiplicity is $m_{100} = 2n$ and the corresponding Wigner components are $\boldsymbol{w}(\rho_S)_{100}, \boldsymbol{w}(\sigma)_{100}$ or it

can correspond to $(i,j,k)=(0,1,0)$ ($(i,j,k)=(0,0,1)$), if $E_{\max}=E_1$ ($E_{\max}=E_2$), when the multiplicity is $m_{010}=3n$ ($m_{001}=3n$) and the corresponding Wigner components are $\boldsymbol{w}(\rho_{\mathrm{S}})_{010},\boldsymbol{w}(\sigma)_{010}$ ($\boldsymbol{w}(\rho_{\mathrm{S}})_{001},\boldsymbol{w}(\sigma)_{001}$). Case (C2) corresponds to $(i,j,k)=(0,n-1,0)$ ($(i,j,k)=(0,0,n-1)$) if we have $E_{\max}=E_1$ ($E_{\max}=E_2$), when the multiplicity is $m_{0,n-1,0}=3^{n-1}n$ ($m_{0,0,n-1}=3^{n-1}n$) and the corresponding Wigner components are $\boldsymbol{w}(\rho_{\mathrm{S}})_{0,n-1,0},\boldsymbol{w}(\sigma)_{0,n-1,0}$ ($\boldsymbol{w}(\rho_{\mathrm{S}})_{0,0,n-1},\boldsymbol{w}(\sigma)_{0,0,n-1}$).

The last elbow coordinates can finally be derived by sutracting the appropriate rescaled component from 1, $(x_E,L_E)=$

$$\begin{cases}\left(1-\dfrac{2n}{(3\mathcal{Z}_\beta)^n}e^{-\beta[(n-1)E_0+E_{\max}]},\ 1+2nv^{n-1}u\right), & \\ \text{if } E_{\max}=E_0, & \text{(C1a)}\\ \left(1-\dfrac{3n}{(3\mathcal{Z}_\beta)^n}e^{-\beta[(n-1)E_0+E_{\max}]},\ 1+3nv^{n-1}u\right), & \\ \text{if } E_{\max}>E_0,\epsilon\le\epsilon_\star, & \text{(C1b)}\\ \left(1-\dfrac{n}{3\mathcal{Z}_\beta^n}e^{-\beta[E_0+(n-1)E_{\max}]},\ 1+3^{n-1}nvu^{n-1}\right), & \\ \text{if } E_{\max}>E_0,\epsilon\ge\epsilon_\star. & \text{(C2)}\end{cases} \tag{93}$$

Assuming that $E_{\max}>E_0$ for clarity, we have the same three scenarios as in the case of the first elbow bound:

1. (C1) $\to$ (C1) if $\beta<\beta_\star$ and $\epsilon'<\epsilon\le\epsilon_\star$.
2. (C2) $\to$ (C1) if $\beta<\beta_\star$ and $\epsilon'\le\epsilon_\star<\epsilon$.
3. (C2) $\to$ (C2) if $\beta<\beta_\star$ and $\epsilon_\star\le\epsilon'<\epsilon$ or $\beta\ge\beta_\star$.

Now using last elbow constraint, derived in **??**, we can calculate new distillation bounds. The last elbow bounds can be rephrased in terms of the first elbow bound, leading to $R_{\mathrm{last-elb}}-R_{\mathrm{first-elb}}=$

$$\begin{cases}\dfrac{1}{n}\dfrac{\log\frac{u(\epsilon)v(\epsilon')}{u(\epsilon')v(\epsilon)}}{\ln\left(1-\frac{4}{3}\epsilon'\right)+\beta(E_0-F(\beta))}, & \text{(C1)}\to\text{(C1)},\\ \dfrac{1}{n}\dfrac{\log\frac{v(\epsilon)v(\epsilon')}{u(\epsilon)u(\epsilon')}-2\beta(E_{\max}-E_0)}{\ln\left(1-\frac{4}{3}\epsilon'\right)+\beta(E_0-F(\beta))}, & \text{(C2)}\to\text{(C1)},\\ \dfrac{1}{n}\dfrac{\log\frac{v(\epsilon')u(\epsilon')}{v(\epsilon')u(\epsilon)}}{\ln\left(\frac{1}{2}-\frac{1}{6}\epsilon'\right)+\beta(E_{\max}-F(\beta))}, & \text{(C2)}\to\text{(C2)},\end{cases} \tag{94}$$

Note that the last elbow bounds have a dependence on n, which interestingly is due to n being even. Had we considered **odd** number of copies n, then the first elbow bound would depend on n and the last elbow bound would not. This, along with the unital fragment analysis, makes me think that calculating the bounds again for odd n or $\epsilon>3/7$ would reveal nice symmetries of the analysis, but it's probably not worth the time, so including last elbow analysis properly feels incomplete.

Determining the sign of the difference between the last and first elbow bounds tells us which bound is better. The term $\log\frac{u(\epsilon)v(\epsilon')}{u(\epsilon')v(\epsilon)}$ is always positive for $\epsilon>\epsilon'$, therefore the last elbow bound, compared to the first elbow bound, is always **worse** in the first scenario and always **better** in the third scenario (remember, this is the yellow-ish right part of **??**).

In the second scenario ((C2) $\to$ (C1)), the expression looks very weird to me, seems like the difference tends to $-\infty$ as $E_{\max}\to\infty$. I am a bit stuck with this expression and I am not sure it's worth putting more time in it.

Given a state ρ and a complete set of component-multiplicity pairs describing its Wigner distribution W_ρ, we now provide a method of computing the components (and multiplicities) of the n–copy distribution $W_\rho{}^{\otimes n}$.

Lemma 12. *Let W be a distribution defined by a complete set of component-multiplicity pairs $\{(w_i,m_i)\}_{i=1,\dots,D}$ with $D\le\dim W$ and consider the distribution $W^{\otimes n}$ obtained by the n-fold (Kronecker) product $W\otimes\cdots\otimes W$ between n copies of W.*

Denote by $C_D^n:=\{\boldsymbol{k}\}$ the set of all vectors $\boldsymbol{k}:=(k_1,\dots,k_D)$ with non-negative integer components that sum to n, i.e.

$$0\le k_1,\dots,k_D\le n \text{ and } k_1+\cdots+k_D=n.$$

Then, $W^{\otimes n}$ admits a complete set of component-multiplicity pairs $\{(W_{\boldsymbol{k}},M_{\boldsymbol{k}})\}_{\boldsymbol{k}\in C_D^n}$, where

$$M_{\boldsymbol{k}}=\frac{n!}{k_1!\dots k_D!}\prod_{i=1}^{D}{m_i}^{k_i}, \tag{95}$$

$$W_{\boldsymbol{k}}=\prod_{i=1}^{D}{w_i}^{k_i}. \tag{96}$$

Here we prove two simple majorization constraints, one arising by considering only the ascending part of the Lorenz curves between the origin $(0,0)$ and the first elbow and the other by considering only the descending part of the curves between the last elbow and the endpoint $(1,1)$.

C. Deriving distillation bounds from the last elbow

[I would prefer to remove this last elbow constraint analysis, and just mention it in the main text. If we include it then I need to go through the details to ensure they're correct – which is very time-consuming at this stage. We *can't* put glaringly sloppy/incorrect things into a paper that will be public to the world. Happy to remove and just mention in passing?]

We can run a similar analysis, in order to utilise the descending part of the Lorenz curves to extract bounds. These bounds do not offer any additional physical insight

and are expressed by , but can be tighter in some scenarios. Here we derive an equivalent bound to our main theorem using the last elbow constraint, which we prove first.

Proposition 13. *Consider a magic state process $\rho \longrightarrow \rho'$ with input and output Lorenz curves $L_{\rho|\sigma}(x), L_{\rho'|\sigma'}(x)$ and denote the coordinates of the last elbow of $L_{\rho|\sigma}(x)$ by (X_E, L_E) and the coordinates of the first elbow of $L_{\rho'|\sigma'}(x)$ by (X'_E, L'_E).*

Then, given any coordinates (x, L) and (x', L') of the input and output Lorenz curves respectively, where $1 > x \geq X_E$ and $1 > x' \geq X'_E$, the process is possible only if

$$\frac{L-1}{1-x} \geq \frac{L'-1}{1-x'}. \tag{97}$$

Proof. Since both pairs of coordinates are located between the endpoint $(1,1)$ and the last elbow of their respective curves, we can derive the bound via a simple interpolation on the line segment connecting the endpoint and the appropriate pair of coordinates.

First assume that $x > x'$. Then, point $(x, L_{\rho'|\sigma'}(x))$ lies on the lines segment connecting $(1,1)$ and (x', L'), so by interpolating between points $(1,1)$ and (x', L'), we find that

$$L_{\rho'|\sigma'}(x) = 1 + \frac{1-x}{1-x'}(L'-1). \tag{98}$$

Considering the Lorenz curve constraint at x, the process is possible only if

$$L = L_{\rho|\sigma}(x) \geq L_{\rho'|\sigma'}(x) = 1 + \frac{1-x}{1-x'}(L'-1), \tag{99}$$

which is a rearrangement of Eq. (97).

If instead, $x \leq x'$, the point $(x', L_{\rho|\sigma}(x'))$ lies on the lines segment connecting $(1,1)$ and (x, L), so by interpolating between points $(1,1)$ and (x, L), we find that

$$L_{\rho|\sigma}(x') = 1 + \frac{1-x'}{1-x}(L-1). \tag{100}$$

Considering the Lorenz curve constraint at x', the process is possible only if

$$1 + \frac{1-x'}{1-x}(L-1) = L_{\rho|\sigma}(x') \geq L_{\rho'|\sigma'}(x) = L', \tag{101}$$

which is again a rearrangement of Eq. (97). □

In order to derive a bound based on the last elbow constraint, we assume that the smallest component of the equilibrium state distribution is $W_\tau(\boldsymbol{x}_\star)$ and the second smallest is $W_\tau(\boldsymbol{y}_\star)$. We now perform a Clifford operation on ρ_S so that its Wigner component is $-v(\epsilon)$ at $\boldsymbol{x}_\star$ and $u(\epsilon)$ at $\boldsymbol{y}_\star$. [Is this always possible? I think not] Since we have that $u \leq v$ because $\epsilon \leq 3/7$, and $W_\tau(\boldsymbol{x}_\star) \geq W_\tau(\boldsymbol{y}_\star) \geq W_\tau(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{Z}_3^2$, the largest rescaled component occurs when $i_{\boldsymbol{x}_\star} = n-1$, $i_{\boldsymbol{y}_\star} = 1$ and $i_{\boldsymbol{x}} = 0$ for $\boldsymbol{x} \notin \{\boldsymbol{x}_\star, \boldsymbol{y}_\star\}$ and it is equal to $(-v^{n-1}u/(W_\tau(\boldsymbol{x}_\star)^{n-1}W_\tau(\boldsymbol{y}_\star))$. Accordingly, the coordinates of the last Lorenz curve point before $(1,1)$ are given by

$$(x_E, L_E) = \big(1 - W_\tau(\boldsymbol{x}_\star)^{n-1}W_\tau(\boldsymbol{y}_\star), 1 + v(\epsilon)^{n-1}u(\epsilon)\big) \tag{102}$$

for the input state and we express the same coordinates with primed quantities for the output state.

We now define the quantities ϕ_α defined in terms of $\alpha_k = \sum_{r\in\mathbb{Z}_3}\langle E_k|A_{\boldsymbol{x}_\star}|E_k\rangle$. and ϕ_λ defined in terms of $\lambda_k = \sum_{r\in\mathbb{Z}_3}\langle E_k|A_{\boldsymbol{y}_\star}|E_k\rangle$. This allows us to rewrite $W_\tau(\boldsymbol{x}_\star) = e^{\beta(F-\phi_\alpha)}/3$ and $W_\tau(\boldsymbol{y}_\star) = e^{\beta(F-\phi_\lambda)}/3$.

Using the last elbow constraint derived in Proposition 13, we can get the new bound expression,

$$\begin{aligned} R \leq & \frac{\ln\left(1-\frac{4}{3}\epsilon\right) + \beta(\phi_\alpha - F)}{\ln\left(1-\frac{4}{3}\epsilon'\right) + \beta(\phi'_\alpha - F')} + \\ & \frac{1}{n}\frac{\ln\frac{u(\epsilon')v(\epsilon)}{u(\epsilon)v(\epsilon')} + \beta(\phi_\alpha - \phi'_\alpha + \phi'_\lambda - \phi_\lambda)}{\ln\left(1-\frac{4}{3}\epsilon'\right) + \beta(\phi'_\alpha - F')} \end{aligned} \tag{103}$$

We note that it is possible to perform various Clifford transformations and get a different location $(W_\tau(\boldsymbol{x}))^n$, with $\boldsymbol{x} \neq \boldsymbol{0}$, since components $W_\rho(\boldsymbol{x}) = u(\epsilon)$ are all the same as long as $\boldsymbol{x} \neq \boldsymbol{0}$. The effect of this is to alter the expressions for the quantities α_k, but not the resulting bound expression.

Proof. Since both pairs of coordinates are located between $(0,0)$ and the first elbow of their respective curves, we can derive the bound via a simple interpolation on the line segment connecting the origin and the appropriate pair of coordinates.

First assume that $x < x'$. Then, point $(x, L_{\rho'|\sigma'}(x))$ lies on the lines segment connecting $(0,0)$ and (x', L'), so by interpolating between points $(0,0)$ and (x', L') at location x, we find that

$$L_{\rho'|\sigma'}(x) = \frac{x}{x'}L'. \tag{104}$$

Considering the Lorenz curve constraint at x, the process is possible only if

$$L = L_{\rho|\sigma}(x) \geq L_{\rho'|\sigma'}(x) = \frac{x}{x'}L', \tag{105}$$

which is a rearrangement of **??**.

If instead, $x \geq x'$, the point $(x', L_{\rho|\sigma}(x'))$ lies on the lines segment connecting $(0,0)$ and (x, L), so by interpolating between points $(0,0)$ and (x, L) at location x', we find that

$$L_{\rho|\sigma}(x') = \frac{x'}{x}L. \tag{106}$$

Considering the Lorenz curve constraint at x', the process is possible only if

$$\frac{x'}{x}L = L_{\rho|\sigma}(x') \geq L_{\rho'|\sigma'}(x) = L', \tag{107}$$

which is again a rearrangement of **??**. □

Conversely, assume that $L_{\rho|\sigma}(x_i) \geq L_{\rho'|\sigma'}(x_i)$ for all $i = 1, \ldots, r$. First, let $x_0 = 0$ and $x_{n'+1} = 1$, so that $L_{\rho|\sigma}(x_0) = L_{\rho'|\sigma'}(x_0) = 0$ and $L_{\rho|\sigma}(x_{n'+1}) = L_{\rho'|\sigma'}(x_{n'+1}) = 1$. Hence, we can extend the set of elbows E to $E' = E \cup \{x_0, x_{n'+1}\}$.

Pick two consecutive locations x_i, x_{i+1} in E' and consider the line segment $\ell'_\rho(x)$ connecting points $(x_i, L_{\rho'|\sigma'}(x_i))$ and $(x_{i+1}, L_{\rho'|\sigma'}(x_{i+1}))$ as well as the line segment $\ell_\rho(x)$ connecting points $(x_i, L_{\rho|\sigma}(x_i))$ and $(x_{i+1}, L_{\rho|\sigma}(x_{i+1}))$. This is illustrated in **??**.

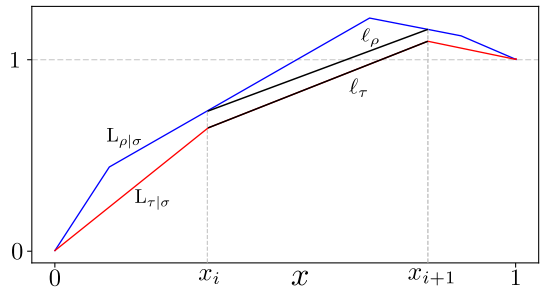


FIG. 8. **Illustration of ??.**

Due to concavity of $L_{\rho|\sigma}$, it is clear that for all $x \in [x_i, x_{i+1}]$, we have $L_{\rho|\sigma}(x) \geq \ell_\rho(x) \geq \ell'_\rho(x) = L_{\rho'|\sigma}(x)$. This argument can be made in all intervals $[x_i, x_{i+1}]$ with $i = 0, \ldots, n'$, so the proof is complete.