Thermodynamic fragments for magic states in quantum computation

Nikolaos Koukoulekidis*
Department of Physics, Imperial College London, London SW7 2AZ, UK

David Jennings
School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK and
Department of Physics, Imperial College London, London SW7 2AZ, UK

(Dated: March 25, 2021)

[To be sharpened] Magic states are key ingredients in schemes to realise universal fault-tolerant quantum computation. Theories of magic states attempt to quantify this computational element via monotones and determine how these states may be efficiently transformed into useful forms. Here we introduce the concept of 'fragments', which generalise the concept of magic monotones and has a natural thermodynamic structure based on majorisation. From this perspective magic can be viewed as a form of free energy within each fragment and is constrained by relative majorisation relations but now on quasi-probability distributions. Notably this approach allows us to incorporate actual physical constraints, for example noise models with particular bias or temperature-dependent features, and study how these constrain general magic distillation protocols. In this context we present general temperature-dependent bounds on distillation rates that any theory of magic must respect. Significantly, this analysis also presents a thermodynamic context which cannot be analysed via traditional methods based on thermodynamic entropies, due to the presence of negativity, and raises novel questions in the context of statistical mechanics.

I. INTRODUCTION AND BACKGROUND

A. Introduction

Magic states are necessary for achieving universal quantum computation within fault-tolerant schemes [1–6]. Identifying magic as a resource for quantum universality has led to several theories which try to provide a framework for its quantification and manipulation [7–10]. The main question that such a theory attempts to answer is:

Given two magic states ρ and ρ' is there a free operation that can convert ρ to ρ'?

We are interested in all resource theories of magic $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ in which free operations cannot generate any amount of resource. Further denote by $\mathcal{D}$ the set of states considered under the theory, that is the union of free and resource states. The structure of such theory is described by a partial order, hereinafter called a *pre-order*, $\prec_{\mathcal{R}}$ between states. We write $\rho' \prec_{\mathcal{R}} \rho$ iff there exists $\mathcal{E} \in \mathcal{O}$ such that $\mathcal{E}(\rho) = \rho'$. Naturally, states may be incomparable under the given theory, meaning that there exists no free operation that converts one to the other. We further call $\mathcal{R}' = (\mathcal{F}', \mathcal{O}')$ a *subtheory* of $\mathcal{R}$ iff $\mathcal{F}' \subseteq \mathcal{F}$ and $\mathcal{O}' \subseteq \mathcal{O}$. The above notation will be used for general resource theories as well.

Formally, the no resource generation condition on the theories translates into two assumptions:

I Free operations send free states into free states, $\mathcal{E} : \mathcal{F} \mapsto \mathcal{F}$, for all $\mathcal{E} \in \mathcal{O}$;

II Resource theory $\mathcal{R}$ is a completely free state preserving theory, in the sense that for any d–dimensional ancilla system and all free operations $\mathcal{E}$, $(\mathbb{1}_d \otimes \mathcal{E})\sigma \in \mathcal{F}$ whenever $\sigma \in \mathcal{F}$.

The first assumption simply states that resources cannot be generated for free and is a minimal requirement for a resource theory. An immediate consequence is that if statistical mixing is included in $\mathcal{O}$, then the resource theory is convex.

Monotones are often used to address the question of state convertibility, although such approaches are usually generic.

The monotonicity condition reflects the no resource generating property of free operations, so that monotones respect the pre-order $\prec_{\mathcal{R}}$ of the theory. A monotone of any general resource theory is a projection of the theory onto the non-negative real numbers, collapsing the pre-order of the theory to the total order defined on the real line. Our contribution is the introduction of a generalised notion of *resource projection* which maps a general resource theory onto a subtheory which in principle still retains a partial structure (as opposed to the real line). Applying this notion on existing magic theories highlights the hidden stochasticity that governs magic state conversions. We show that a magic theory can be subdivided into *fragments* [FIX AND EXPAND]

B. Previous work

The stabiliser theory of magic comprises of the so-called "stabiliser" states (STAB) and operations (SO), while non-stabiliser (resource) states are called magic. The stabiliser operations can be expressed in terms of a

* nk2314@imperial.ac.uk

Stinespring dilation as

$$\mathcal{E}(\rho) = \mathrm{tr}_E[U(\rho \otimes \sigma_E)U^\dagger], \tag{1}$$

for an ancilla stabiliser state σ_E. The motivation of the theory stems from the fact that stabiliser operations are experimentally straightforward to implement and they can be used to detect and correct errors on the stabiliser states due to their construction [3, 11, 12]. The Gottesman-Knill theorem however indicates that stabiliser operations need to be supplemented with magic states in order to achieve universality, justifying the term "magic".

Generalisations of the stabiliser theory appear in the literature intending to include broader classes of operations. The class of stabiliser preserving operations (SPO) is defined as the set of CPTP maps that send stabiliser states into stabiliser states [13]. An important subclass of SPO is the set of completely stabiliser preserving operations (CSPO) [14], which intuitively cannot induce "non-stabiliserness" even when applied to only part of a quantum state, i.e. operations $\mathcal{E}$ such that $(\mathbb{1}_d \otimes \mathcal{E})\sigma \in \mathrm{STAB}$ for all positive dimensions d whenever $\sigma \in \mathrm{STAB}$.

Even though non-stabiliserness is a necessary resource for universality, it has been proven insufficient for magic state distillation [4, 15]. In fact, all states with non-negative Wigner distributions have been proven to be efficiently classically simulable in [2], a result that serves as a generalization of the Gottesman-Knill theorem. The Wigner distribution of a state in odd prime dimensions, formally defined in Section II B, arises as the unique quasi-probability representation of quantum theory that identifies non-contextuality exactly with the states that are efficiently classically simulable [7, 16, 17]. In this framework, the stabiliser states are the only pure states represented with non-negative distributions [18]. However, there exist mixed states with non-negative Wigner distributions that are not mixtures of stabiliser states [19]. Therefore, stabiliser-preserving theories have been extended to a theory that preserves state "Wigner positivity" [9]. Informally, it can be considered as the maximal theory of magic $\mathcal{R}_{\max} = (\mathcal{F}_{\max}, \mathcal{O}_{\max})$, where free states have non-negative Wigner distributions and free operations completely preserve this property.

Things we MUST emphasize:

1. Perhaps a nice lead-in question: "What happens if we view stabiliser states as thermodynamic equilibrium states and magic as a form of free energy?"

2. We have found a scenario in which it is impossible to describe a thermodynamic structure using any entropic approach!

3. We can tackle more 'physicsy' questions like: how much magic can be distilled via available operations with some given fixed-point structure?

4. This allows a diagnostic on the kind of operations needed to do good distillation. I.e. what fixed point structure should they have?

5. We go beyond the concept of monotones and replace a monotone with a σ–fragment.

6. We can get both upper and lower bounds on magic distillation.

II. PRELIMINARIES AND DEFINITIONS

A. Stabilizer Theory

Let $\{|k\rangle\}_{k\in\mathbb{Z}_d}$ be the standard computational basis, defined over $\mathbb{Z}_d = \{0, 1, \dots, d-1\}$, with d an odd prime. Since the field has character d, addition and multiplication on the field are always considered modulo d. The Hilbert space of any system associated with this scheme is $\mathcal{H}_d := \mathrm{span}\{|k\rangle : k \in \mathbb{Z}_d\}$.

The generalised Pauli matrices X, Z can be defined by their respective roles as shift and phase operators,

$$X|k\rangle = |k+1\rangle \tag{2}$$

$$Z|k\rangle = \omega^k |k\rangle, \tag{3}$$

where $\omega := e^{2\pi i/d}$ is the d-th root of unity and addition is modulo d. The Hilbert space $\mathcal{H}_d$ is associated with a phase space $\mathcal{P}_d = \mathbb{Z}_d \times \mathbb{Z}_d$, where every point $\boldsymbol{x} := (x, p)$ corresponds to a displacement operator, defined as

$$D_{\boldsymbol{x}} := \tau^{xp} X^x Z^p, \ (x,p) \in \mathcal{P}_d. \tag{4}$$

The phase factor $\tau := -\omega^{1/2}$ ensures unitarity. For a composite system with product dimension $d = d_1 \dots d_n$ and Hilbert space, $\mathcal{H}_d = \mathcal{H}_{d_1} \otimes \cdots \otimes \mathcal{H}_{d_n}$, the displacement operators are defined as

$$D_{\boldsymbol{x}} := D_{(x_1,p_1)} \otimes \cdots \otimes D_{(x_n,p_n)}, \tag{5}$$

where

$$\boldsymbol{x} := (x_1, p_1, x_2, p_2, \dots, x_n, p_n) \in \mathcal{P}_{d_1} \times \cdots \times \mathcal{P}_{d_n} =: \mathcal{P}_d,$$

denotes the phase space point for an n–copy system.

The displacement operators form the Heisenberg-Weyl group [20, 21] under matrix multiplication modulo phases,

$$\mathrm{HW}_d^n := \{\tau^k D_{\boldsymbol{x}} : k \in \mathbb{Z}_d, \boldsymbol{x} \in \mathcal{P}_d^n\}. \tag{6}$$

The Clifford operations $\mathcal{C}_d^n$ are then defined as the set of unitaries that normalise the Heisenberg-Weyl group. We may define the pure stabiliser states as those states obtained by acting on $|0\rangle$ with Clifford unitaries. Finally, we define STAB as the convex hull of all pure stabiliser states, namely all probabilistic mixtures of states of the form $U|0\rangle\langle 0|U^\dagger$ where U is some Clifford unitary. The stabiliser theory forms a magic resource theory $\mathcal{R} = (\mathrm{STAB}, \mathrm{SO})$ and the free operations are stabiliser operations (SO), defined as compositions of preparation of computational basis states, Clifford unitaries, measurement in the computational basis, and the ability to discard subsystems.

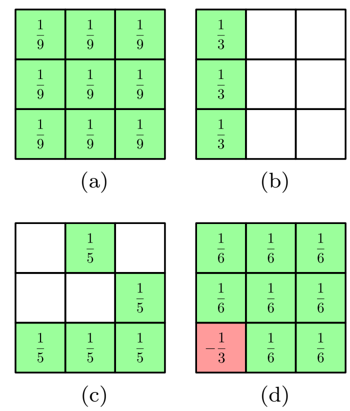

FIG. 1. **Qutrit Wigner distributions of varying magic.** (a) Maximally mixed state $\frac{1}{3}\mathbb{1}$; (b) Stabilizer zero state $|0\rangle\langle 0|$; (c) A non-stabiliser Wigner-positive state; (d) Magic Strange state $|\mathrm{S}\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)$, coined in [7].

B. Wigner representations for quantum states and quantum operations

We define the phase-point operators at every phase space point $\boldsymbol{x} \in \mathcal{P}_d$,

$$A_{\boldsymbol{x}} := \frac{1}{d} \sum_{\boldsymbol{z}\in\mathcal{P}_d} \omega^{\eta(\boldsymbol{x},\boldsymbol{z})} D_{\boldsymbol{z}}, \tag{7}$$

where $\eta(\boldsymbol{x}, \boldsymbol{z})$ is the standard symplectic form, given explicitly by

$$\eta(\boldsymbol{x}, \boldsymbol{z}) := \boldsymbol{z}^T \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \boldsymbol{x}, \tag{8}$$

where $0, \mathbb{1}$ denote the $n \times n$ zero and identity matrices.

The phase-point operators form an orthogonal Hermitian operator basis as shown in Appendix A. Therefore, any quantum state $\rho \in \mathcal{B}(\mathcal{H}_d)$ can be expressed as a linear combination of them,

$$\rho = \sum_{\boldsymbol{z}\in\mathcal{P}_d} \mathrm{W}_\rho(\boldsymbol{z}) A_{\boldsymbol{z}}, \tag{9}$$

where the coefficient vector W_ρ is the Wigner distribution of state ρ,

$$\mathrm{W}_\rho(\boldsymbol{x}) := \frac{1}{d}\mathrm{tr}[A_{\boldsymbol{x}}\rho]. \tag{10}$$

For any quantum state ρ, the Wigner distribution $W_\rho(\boldsymbol{x})$ is readily seen to be a d^2–dimensional quasi-probability distribution over $\mathcal{P}_d$ (see Appendix A for details). In Fig. (1), we show Wigner distributions of different types of qutrit states.

We also have Wigner representations for general quantum channels. We may exploit the channel-state duality and use the normalised Choi-Jamiołkowski state

$$\mathcal{J}(\mathcal{E}) := \frac{1}{d_A}(\mathbb{1}\otimes\mathcal{E}) \sum_{i,j} |ii\rangle\langle jj| \tag{11}$$

to extend the definition of the Wigner distribution to a quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, via the expression

$$\mathrm{W}_{\mathcal{E}}\left(\boldsymbol{y}|\boldsymbol{x}\right) \coloneqq d_A^2 \mathrm{W}_{\mathcal{J}(\mathcal{E})}\left(\bar{\boldsymbol{x}} \oplus \boldsymbol{y}\right) \tag{12}$$

$$= \frac{1}{d_B}\mathrm{tr}_B[A_{\boldsymbol{y}}\mathcal{E}(A_{\boldsymbol{x}})], \tag{13}$$

where $\bar{\boldsymbol{x}} \coloneqq (x, -p)$.

The specific form of Eq. (12) is chosen so that Wigner distributions of operations act as transition matrices for Wigner distributions of states, $\mathrm{W}_{\mathcal{E}(\rho)} = \mathrm{W}_{\mathcal{E}}\mathrm{W}_{\rho}$. In particular, CPTP operations that map between density operators of equal dimensions and have non-negative Wigner distributions correspond to stochastic matrices, as shown in Appendix A

C. Magic monotones

We highlight the defining property of a monotone, as it is useful in later discussion.

Definition 1 (**Resource monotone**). *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a resource theory. A resource monotone $\mathcal{M}$ is a projection from the set of quantum states of the theory onto the real line, so that $\mathcal{M}$ is monotonically decreasing under free operations,*

$$\mathcal{M}(\rho_1) \leq \mathcal{M}(\rho_2) \text{ whenever } \rho_1 \prec_{\mathcal{R}} \rho_2. \tag{14}$$

One of the most fundamental and commonly used magic monotones is the *mana* of a state [7], defined as

$$\mathrm{mana}\left(\rho\right) \coloneqq \ln\left(2\,\mathrm{sn}\left(\rho\right) + 1\right), \tag{15}$$

where the *sum-negativity* [7] is the sum of the negative components in W_ρ,

$$\mathrm{sn}\left(\rho\right) \coloneqq \sum_{\boldsymbol{x}:\mathrm{W}_\rho(\boldsymbol{x})<0} |\mathrm{W}_\rho\left(\boldsymbol{x}\right)|. \tag{16}$$

Mana is an additive[1] magic monotone, so it provides an analytical, necessary condition for many-copy magic state interconvertibility. The magic monotone of *max–thauma* [22] has also been introduced more recently for qudits of odd prime dimension.

D. Majorisation

Majorisation [23] is a collection of powerful tools that has recently found many applications in quantum information theory [24–30]. It describes the [disorder / non-uniformity] of distributions that undergo stochastic transformations.

To formally state majorisation results, we first denote by $\mathrm{S}_d(\boldsymbol{d})$ the set of $(d \times d)$ stochastic matrices that preserve the probability vector $\boldsymbol{d}$. Specifically, for any $\boldsymbol{d}$–*stochastic matrix* $S \in \mathrm{S}_d(\boldsymbol{d})$, all matrix elements are non-negative, all rows sum to 1 and $S\boldsymbol{d} = \boldsymbol{d}$. The set $\mathrm{S}_d(\boldsymbol{d})$ forms a group under matrix multiplication for any $\boldsymbol{d}$ with positive components.

Majorisation describes quantum thermodynamics exactly in the absence of quantum coherence. The use of majorisation in this setting provides vital intuition for our purposes. At any given temperature β^{-1}, the thermal state γ_β is thermodynamically the most disordered state. Thermal operations are defined as operations that cannot extract energy from the Gibbs state, $\mathcal{E}(\gamma_\beta) = \gamma_\beta$. Convertibility between states via thermal operations is equivalent to a stochasticity condition on the energy level populations of the states [31]. Informally, there exists a thermal operation $\mathcal{E}$ such that $\tau = \mathcal{E}(\rho)$ if and only if there exists a matrix $S \in \mathrm{S}_d(\boldsymbol{d})$ such that $\boldsymbol{q} = S\boldsymbol{p}$, where $\boldsymbol{q}, \boldsymbol{p}$ and $\boldsymbol{d}$ and the energy level population vectors of τ, ρ, γ_β respectively.

Based on this setting, we define $\boldsymbol{d}$–majorisation as follows.

Definition 2 (**$\boldsymbol{d}$–majorisation**). *Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d} \in \mathbb{R}^d$, such that the components of $\boldsymbol{d}$ are positive, $\boldsymbol{y}$ is said to $\boldsymbol{d}$–majorise $\boldsymbol{x}$, iff there exists a $\boldsymbol{d}$–stochastic matrix S such that $\boldsymbol{x} = S\boldsymbol{y}$.*

We denote this pre-order by $\boldsymbol{x} \prec_{\boldsymbol{d}} \boldsymbol{y}$. If $\boldsymbol{d} = \frac{1}{d}\mathbf{1}$, the d–dimensional uniform distribution, then $\mathrm{S}_d(\boldsymbol{d})$ is the set of doubly stochastic matrices and we retrieve the familiar notion of majorisation from entanglement theory [24].

The pre-order imposed by $\boldsymbol{d}$–majorisation admits a numerically efficient reformulation in terms of Lorenz curves. Let the vector $\boldsymbol{u}^{\downarrow}$ denote the vector $\boldsymbol{u} \in \mathbb{R}^d$ with its components arranged in non-increasing order.

Definition 3 (**Lorenz curve**). *Let $\boldsymbol{w}, \boldsymbol{d} \in \mathbb{R}^d$, where the components of $\boldsymbol{d}$ are positive with $D = \sum_{i=1}^d d_i$ and denote by $\widetilde{\boldsymbol{w}} \coloneqq (w_1/d_1, \ldots, w_d/[d_d])^T$ the rescaled vector $\boldsymbol{w}$ by $\boldsymbol{d}$.*

Finally, denote by $\pi : \mathbb{Z}_d \mapsto \mathbb{Z}_d$ the permutation that sorts $\widetilde{\boldsymbol{w}}$, $(\widetilde{\boldsymbol{w}}^{\downarrow})_i = w_{\pi(i)}$ for all $i = 1, \ldots, d$.

Consider the piecewise linear curve obtained by joining the points $\{(0,0)\} \cup \{(x_k, \mathrm{L}_{\boldsymbol{w}|\boldsymbol{d}}(k))\}_{k=1,\ldots,d}$, where

$$(x_k, \mathrm{L}_{\boldsymbol{w}|\boldsymbol{d}}(k)) \coloneqq \left(\frac{1}{D}\sum_{i=1}^k d_{\pi(i)}, \sum_{i=1}^k w_{\pi(i)}\right). \tag{17}$$

We define the set of points on this curve, $\mathrm{L}_{\boldsymbol{w}|\boldsymbol{d}}(x)$, $x \in [0,1]$, as the Lorenz curve *of vector $\boldsymbol{w}$ with respect to $\boldsymbol{d}$.*

Components x_k are rescaled by D so that comparison of curves with unequal dimensions is possible. In fact, the Lorenz curves $\mathrm{L}_{\boldsymbol{w}|\boldsymbol{d}}$ and $\mathrm{L}_{\boldsymbol{w}\otimes\boldsymbol{d}|\boldsymbol{d}\otimes\boldsymbol{d}}$, where $\otimes$ denotes the Kronecker product, coincide. Furthermore, a Lorenz

[1] It satisfies the condition $\mathrm{mana}\left(\rho_1 \otimes \rho_2\right) = \mathrm{mana}\left(\rho_1\right) + \mathrm{mana}\left(\rho_2\right)$ which is practical in distillation settings.

curve $\mathrm{L}_{\boldsymbol{w}|\boldsymbol{d}}(x)$ is always concave in x, since it consists of d line segments each with slope $(\widetilde{\boldsymbol{w}}^{\downarrow})_i$ for $i = 1, \ldots, d$ which by definition is a non-increasing sequence. Finally, the points on the interior of the Lorenz curve that connect line segments of different slopes are in general non-differentiable and we call them *elbows*.

A vector $\boldsymbol{y}$ is said to $\boldsymbol{d}$*–majorise* another vector $\boldsymbol{x}$ if and only if the Lorenz curve $\mathrm{L}_{\boldsymbol{y}|\boldsymbol{d}}$ lies above Lorenz curve $\mathrm{L}_{\boldsymbol{x}|\boldsymbol{d}}$, thus reducing $\boldsymbol{d}$–majorisation into a finite set of inequalities.

Theorem 4. *Let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d} \in \mathbb{R}^d$, such that the components of $\boldsymbol{d}$ are positive. Then, $\boldsymbol{x} \prec_{\boldsymbol{d}} \boldsymbol{y}$ if and only if $\mathrm{L}_{\boldsymbol{x}|\boldsymbol{d}}(x) \leq \mathrm{L}_{\boldsymbol{y}|\boldsymbol{d}}(x)$ for all $x \in [0, 1]$ with strict equality at $x = 1$.*

A restatement of the theorem including more equivalent conditions and a proof are provided in Appendix B.

III. STOCHASTIC STRUCTURE OF MAGIC THEORIES

A. Magic fragments

Equipped with the definitions of the Wigner distribution in odd prime dimensions, we can formally recast the maximal magic theory $\mathcal{R}_{\max}$ into a stochasticity setting. The free states correspond to proper probability distributions

$$\mathcal{F}_{\max} \coloneqq \{\rho : \mathrm{W}_{\rho}(\boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \in \mathcal{P}_d\} \tag{18}$$

The free operations should send the set of free states $\mathcal{F}_{\max}$ into itself and completely preserve the non-negativity of the states, in the sense that $\mathcal{E} \in \mathcal{O}_{\max}$ iff $(\mathbb{1}_d \otimes \mathcal{E})\sigma \in \mathrm{STAB}$ for all odd prime dimensions d whenever $\sigma \in \mathcal{F}_{\max}$. It is shown by Wang *et al.* [9] that $\mathcal{O}_{\max}$ coincides with the set of operations $\mathcal{E}$ with stochastic Wigner distributions,

$$\mathcal{O}_{\max} = \{\mathcal{E} : \mathrm{W}_{\mathcal{E}}(\boldsymbol{y}|\boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x}, \boldsymbol{y} \in \mathcal{P}_d\}. \tag{19}$$

Every established magic theory $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ is a subtheory of $\mathcal{R}_{\max}$ in the sense that $\mathcal{F} \subseteq \mathcal{F}_{\max}$ and $\mathcal{O} \subseteq \mathcal{O}_{\max}$, and as such it falls under this new stochasticity setting. For technical simplicity in what follows we assume that $\mathcal{F}$ is a closed set, and note that $\mathcal{F}_{\max}$ is itself a closed set, since it is specified by a finite set of linear constraints of the form $\mathrm{tr}[L\rho] \geq 0$ with operators $L \in \mathcal{B}(\mathcal{H})$ ensuring that the state is positive and Wigner-positive.

Given this context we now define the following key notion, that is central to our analysis.

Definition 5 (**$\boldsymbol{\sigma}$–fragment**)**.** *Given a resource theory of magic $\mathcal{R} = (\mathcal{F}, \mathcal{O})$, the* σ–fragment of $\mathcal{R}$ *is the resource theory $\mathcal{R}_\sigma = (\mathcal{F}_\sigma, \mathcal{O}_\sigma)$, where $\mathcal{F}_\sigma = \{\sigma\}$ and the free operations are restricted to the ones that leave σ invariant,*

$$\mathcal{O}_\sigma \coloneqq \{\mathcal{E} \in \mathcal{O} : \mathcal{E}(\sigma) = \sigma\}. \tag{20}$$

Note that $\mathcal{F} = \bigcup_{\sigma \in \mathcal{F}} \mathcal{F}_\sigma$ trivially. With this basic notion defined, we now show that this union holds for free operations as well and therefore that any resource theory of magic can be faithfully subdivided into σ–fragments, in such a way that any problem of interconversion in the parent magic theory $\mathcal{R}$ can be analysed across the different fragments.

Theorem 6. *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a theory of magic. Every operation in $\mathcal{O}$ leaves at least one free state invariant,*

$$\mathcal{O} = \bigcup_{\sigma \in \mathcal{F}} \mathcal{O}_\sigma. \tag{21}$$

Therefore, $\rho \longrightarrow \tau$ in $\mathcal{R}$ if and only if $\rho \longrightarrow \tau$ in a σ–fragment of $\mathcal{R}$.

Proof. Suppose $\mathcal{E}$ is in a σ–fragment $\mathcal{O}_\sigma$. Then it is also in $\mathcal{O}$, hence $\bigcup_{\sigma \in \mathcal{F}} \mathcal{O}_\sigma \subseteq \mathcal{O}$.

Conversely, suppose $\mathcal{E}$ is in $\mathcal{O}$. The free states are a closed set that is mapped one-to-one to a closed subset $\mathcal{S}$ of the $(d^2 - 1)$–dimensional probability simplex. $\mathcal{S}$ is convex, since any combination of free states is also free and the Wigner distribution is linear. Therefore, $\mathcal{S}$ is convex and compact as a closed convex subset of the bounded compact probability simplex.

We can now view $\mathrm{W}_{\mathcal{E}}$ as a stochastic, continuous mapping from $\mathcal{S}$ to itself, thus Brouwer's fixed point theorem [32] implies that there exists a probability distribution $d_{\boldsymbol{z}}$ for some $\boldsymbol{z} \in \mathcal{P}_d$ that is a fixed point of $\mathrm{W}_{\mathcal{E}}$. This corresponds to a free state $\sigma \coloneqq \sum_{\boldsymbol{z} \in \mathcal{P}_d} d_{\boldsymbol{z}} A_{\boldsymbol{z}} \in \mathcal{F}$. Therefore $\mathcal{E} \in \mathcal{O}_\sigma$, and so $\mathcal{O} = \bigcup_{\sigma \in \mathcal{F}} \mathcal{O}_\sigma$.

The state interconversion result follows immediately. □

The zoo of all magic operation classes is summarised in Fig. (2). It is clear from the diagram that free operations are in general mapped onto strict subsets of stochastic maps. This statement, along with more technical properties of the σ–fragment are discussed in Appendix C. Completely positive-Wigner-preserving operations [9] form the maximal operation class $\mathcal{O}_{\max}$. This theory of magic is thus covered by σ–fragments exactly and every other magic theory is contained within this cover. In particular, the stabiliser operations SO are contained within $\mathcal{O}_{\max}$.

Every quantum circuit aiming at a given magic state conversion $\rho \longrightarrow \tau$ possesses certain symmetries according to Theorem 6 that allow us to study the conversion within only certain σ–fragments that reflect these symmetries. As a simple example, the dephasing channel

$$\Delta(\rho) = \sum_{k \in \mathbb{Z}_d} |k\rangle\langle k| \, \rho \, |k\rangle\langle k| \tag{22}$$

removes coherent phases in the computational basis and therefore leaves exactly all mixtures of computational basis states invariant. A noise channel consisting only of such dephasing operations can be fully analysed in the

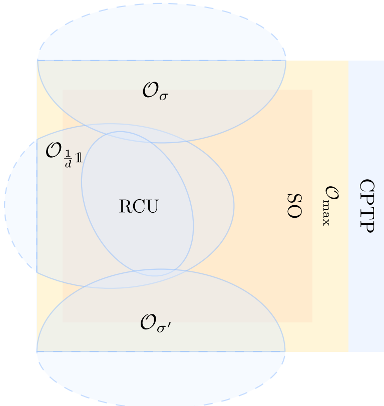

FIG. 2. **Decomposition of a magic theory $\mathcal{R}$ intro σ–fragments.** Examples of magic theories (SO: Stabilizer operations, $\mathcal{O}_{\max}$: Completely positive-Wigner-preserving operations, RCU: Random Clifford Unitaries – subclass of SO) involve operations denoted by the two yellow regions, with other established magic theories between them. We introduce σ–fragments $\mathcal{O}_\sigma$ defined for all free states σ that cover $\mathcal{O}_{\max}$. Each $\mathcal{O}_\sigma$ is extensible to a set of stochastic maps outside the CPTP operations. Within each σ–fragment $\boldsymbol{d}$–majorisation can be used allowing for tractable approach towards the study of magic state interconversion.

σ–fragments for the pure computational basis states σ as proven in Appendix C.

The subdivision of magic theories into σ–fragments is also powerful because the pre-order $\prec_{\mathcal{R}_\sigma}$ of every σ–fragment is described by well-behaved majorisation tools, as we establish in the rest of this section.

B. Majorisation of quasi-probabilities in σ–fragments

We can approach any magic theory through a thermodynamic lens, and in doing so we are provided with valuable insights on the structure of the theory. Firstly, any free state, for example a stabiliser state, can be viewed as a thermal state γ_β. Without loss of generality we can always write a quantum state σ as a thermal state, $\sigma = \gamma_\beta := \frac{1}{\mathcal{Z}_\beta} e^{-\beta H}$ for some $\beta \geq 0$ and Hamiltonian H (either effective or actual)[2]. We can also view the set of free operations $\mathcal{O}_\sigma$ as a counterpart of thermal operations, in the sense that any operation $\mathcal{E}$ in $\mathcal{O}_\sigma$ preserves state σ.

It is then apparent that the pre-order $\prec_{\mathcal{R}_\sigma}$ between the operations in the σ–fragment follows the rules of $\boldsymbol{d}$–majorisation as outlined in Section II D. In particular, the pre-order $\prec_{\mathcal{R}_\sigma}$ of the σ–fragment $\mathcal{R}_\sigma = (\mathcal{F}, \mathcal{O}_\sigma)$ between d–dimensional states corresponds to the majorisation pre-order $\prec_{\mathrm{W}_\sigma}$ between their d^2–dimensional Wigner distributions. For simplicity we shall merge the notation into $\prec_\sigma$, as there is little risk of confusion.

Note that the Wigner components of an n–copy state $\rho^{\otimes n}$ can be calculated directly from W_ρ by convolution of the distribution with itself,

$$\mathrm{W}_{\rho^{\otimes n}} = \mathrm{W}_\rho{}^{\otimes n}, \tag{23}$$

where $\otimes$ can be interpreted as the usual Kronecker product in the last expression. We may use the vector notation $\boldsymbol{w}(\rho) = (w(\rho)_i)_{i=1,\ldots,d^2}$ for the Wigner distribution, in which case we can express the Kronecker product as

$$\mathrm{W}_\rho \otimes \mathrm{W}_\rho = (w(\rho)_i w(\rho)_j)_{i,j=1,\ldots,d^2}. \tag{24}$$

The correspondence between the phase space and pure vector representations of the Wigner distribution is discussed more in Appendix E 1. The vector notation will be useful in the proof of our main result, Theorem 11

Furthermore, we restrict our analysis to σ–fragments, where σ is full-rank. This is justified as majorisation is continuous between fragments, in the sense that the pre-orders $\prec_\sigma$ and $\prec_{\sigma'}$ are equivalent for states σ and σ' which are ϵ–close. [CHECK] Therefore, majorisation analysis is robust under imperfections in the experimental implementation of quantum operations.

In particular, we can always use common noise effects to approximate any σ–fragment where σ is not full-rank by the σ'–fragment where σ' is a noisy approximation of σ. For example, inducing depolarising noise, we can write $\sigma' = (1-\epsilon)\sigma + \epsilon \frac{1}{d}\mathbb{1}$, for some infinitesimal $\epsilon > 0$, so that σ' is full-rank and arbitrarily close to σ. Important examples of such σ–fragments include pure stabiliser states which are rank-1, e.g. the zero state depicted in Fig. (1(b)). Operations in such fragments include important stabiliser operations like the replacement channel, $\mathcal{E}(\rho) = \sigma$ for all states ρ.

Theorem 7. *Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a theory of magic. If $\rho \longrightarrow \tau$ in $\mathcal{R}$ then $\mathrm{W}_\tau \prec_\sigma \mathrm{W}_\rho$ within at least one σ–fragment.*

Proof. Suppose we can convert ρ into τ in the magic theory. Thus there is some $\mathcal{O}_\sigma$ and some $\mathcal{E} \in \mathcal{O}_\sigma$ such that $\mathcal{E}(\rho) = \tau$, and $\mathcal{E}(\sigma) = \sigma$. Therefore, the Wigner distribution of this free operation satisfies $\mathrm{W}_\mathcal{E} \in \mathrm{S}_{d^2}(\mathrm{W}_\sigma)$ and $\mathrm{W}_\mathcal{E}\mathrm{W}_\rho = \mathrm{W}_\tau$. Since σ is full-rank and free, its Wigner distribution is strictly positive in all components, so it directly follows from Definition 2 that $\mathrm{W}_\tau \prec_\sigma \mathrm{W}_\rho$. $\square$

The converse is in general not true, since stochastic matrices do not necessarily correspond to valid quantum operations.

The result of Theorem 7 can be understood as an extension of the idea of a magic monotone, where we replace $\mathcal{M}(\tau) \leq \mathcal{M}(\rho)$ with $\mathrm{W}_\tau \prec_\sigma \mathrm{W}_\rho$. The physical difference between the two expressions is that the majorisation ordering occurs in a specific σ–fragment. Therefore, majorisation constraints can be used to place upper bounds on magic state distillation in a way that allows

---
[2] Technicalities arise for the case where σ is not full rank, but this can be still described via $\beta \to 0$ limiting process.

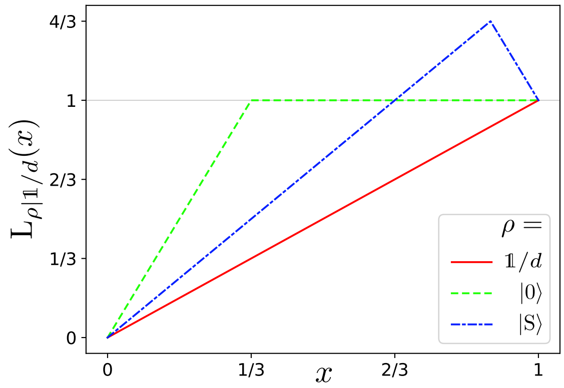

FIG. 3. **Quasi-probability Lorenz curve comparison**. The Lorenz curves are constructed by the Wigner distributions illustrated in Fig. (1) in the unital fragment $\mathcal{O}_{\mathbb{1}/d}$ for $d = 3$. The maximally mixed state curve is simply the line connecting $(0, 0)$ and $(1, 1)$. There is no operation in $\mathcal{O}_{\mathbb{1}/d}$ that can convert $|\mathrm{S}\rangle$ to $|0\rangle$, as their Lorenz curves intersect.

one to incorporate the physics of the allowed operations – specifically, it enables us to bound how much magic can be distilled via quantum operations that, for example, preserve the equilibrium state of the system, or via operations that are symmetric about the Z-axis of the Bloch sphere. We discuss distillation upper bounds in detail in Sections IV and V

This approach can also provide *lower bounds* on distillation, however now more structure about the specific free operations must be included. We briefly discuss distillation lower bounds in Section VI.

[Mention relative majorisation as a possibility]

C. Lorenz curves of quasi-probabilities in σ–fragments

It is straightforward to construct Lorenz curves for Wigner distributions in any σ–fragment. As we have seen in Theorem 7, for any full-rank free state σ we have that W_σ is a strictly positive full-rank probability distribution, and so one can define a corresponding notion of $\boldsymbol{d}$–majorisation on *quasi*–distributions. We write $\mathrm{L}_{\rho|\sigma}(x)$ for the Lorenz curve of W_ρ with respect to W_σ. The vector of ratios $\widetilde{\boldsymbol{w}}(\rho|\sigma)$ used to construct the curve is

$$\widetilde{w}(\rho|\sigma)_i := \frac{w(\rho)_i}{w(\sigma)_i}, \tag{25}$$

and is called the *rescaled* distribution of ρ with respect to σ.

An example of comparison between different Lorenz curves is provided in Fig. (3). The curves in the figure are constructed in the *unital fragment*, i.e. the σ–fragment defined by the maximally mixed state $\sigma = \frac{1}{d}\mathbb{1}$ whose Wigner distribution is the uniform probability distribution.

Normalisation of the Wigner distribution ensures that for all quantum states ρ, $\mathrm{L}_{\rho|\sigma}(x) \geq 0$ and $\mathrm{L}_{\rho|\sigma}(1) = 1$. We stress that $0 \leq \mathrm{L}_{\rho|\sigma}(x) \leq 1$ for all $x \in [0, 1]$ if and only if ρ is a positive Wigner state. As a consequence, checking whether a magic state conversion of the form

$$\rho \xrightarrow{\mathcal{E}\in\mathcal{O}_\sigma} \tau \tag{26}$$

is not possible, reduces to the set of constraints

$$\mathrm{L}_{\rho|\sigma}(x) \geq \mathrm{L}_{\tau|\sigma}(x), \ x \in [0, 1], \tag{27}$$

due to Theorems 4 and 7.

We can refine the number of independent constraints stemming from this inequality. In fact, there are only as many independent constraints as there are elbows in the Lorenz curve of the target state as shown in Proposition 21 in Appendix C. However, in principle any one location x provides a valid constraint leading to some upper distillation bound, while optimising over the location would provide the strictest bound. In particular, we establish the *first elbow constraint* which we use in Sections IV and V explicitly.

Lemma 8. *Consider a magic state interconversion as in Eq. (26), where we denote by (x_0, L_0) and (x_0', L_0') the first elbow coordinates of the initial and target states respectively. If $x_0 < x_0'$, then it holds that*

$$\frac{L_0}{x_0} \geq \frac{L_0'}{x_0'}. \tag{28}$$

Proof. Consider the Lorenz curve constraint at $x = x_0$,

$$\mathrm{L}_{\rho|\sigma}(x_0) \geq \mathrm{L}_{\tau|\sigma}(x_0). \tag{29}$$

Since $x_0 < x_0'$, we can find the target state Lorenz curve coordinate $L_\star'$ at location $x = x_0$ by interpolating between the origin and the target state's first elbow,

$$L_\star' = \frac{x_0}{x_0'} L_0'. \tag{30}$$

We need $L_0 \geq L_\star'$ and rearranging we get Eq. (28). □

We can now turn the attention from the first elbow to the peak of the Lorenz curve to associate it directly with the magic monotone of mana, and equivalently the sum-negativity of a quantum state. This holds independently of the particular σ–fragment one works in and as a result it becomes apparent that mana provides a weaker condition than majorisation for all magic state interconversions.

We first show that the Lorenz curve maximum of state ρ is independent of the σ–fragment and directly related to its sum-negativity.

Lemma 9. *Given a quantum state ρ, the maximum of its Lorenz curve $\mathrm{L}_{\rho|\sigma}$ is independent of the σ–fragment and is equal to $1 + \mathrm{sn}(\rho)$.*

Proof. We may use the vector notation of the Wigner distributions $\boldsymbol{w}(\rho)$ and $\boldsymbol{w}(\sigma)$. We choose the component indexing so that the rescaled distribution

$$\widetilde{\boldsymbol{w}}(\rho|\sigma) := \left(\frac{w(\rho)_1}{w(\sigma)_1}, \dots, \frac{w(\rho)_{d^2}}{w(\sigma)_{d^2}}\right)^T, \tag{31}$$

is sorted, $\widetilde{\boldsymbol{w}} = \widetilde{\boldsymbol{w}}^{\downarrow}$. Note that all components of $\boldsymbol{w}(\sigma)$ are positive, so $\widetilde{w}_i \geq 0$ if and only if $w(\rho)_i \geq 0$ for any $i = 1, \dots, d^2$.

Let $i_\star$ be the index of the smallest non-negative component of $\widetilde{\boldsymbol{w}}^{\downarrow}$. Then, $w(\rho)_i < 0$ if and only if $i > i_\star$, so the maximum of Lorenz curve $\mathrm{L}_{\rho|\sigma}(x)$ takes the value

$$\mathrm{L}_{\rho|\sigma}(x_{i_\star}) = \sum_{i=1}^{i_\star} w(\rho)_i, \tag{32}$$

and is achieved at

$$x_{i_\star} := \sum_{i=1}^{i_\star} w(\sigma)_i. \tag{33}$$

The location of the maximum ($x = x_{i_\star}$) varies from fragment to fragment, but its value is independent of σ,

$$\mathrm{L}_{\rho|\sigma}(x_{i_\star}) = \sum_{\boldsymbol{x}:\mathrm{W}_\rho(\boldsymbol{x})\geq 0} \mathrm{W}_\rho(\boldsymbol{x}) = 1 + \mathrm{sn}(\rho). \tag{34}$$

□

We can therefore view mana as just one feature of the Lorenz curve, namely its maximum value. Conversely, it is now clear that the maximum of the Lorenz curve acts as a valid magic monotone.

Theorem 10. *Given a magic state conversion $\rho \longrightarrow \tau$, the majorisation condition is stronger than the mana condition in every σ–fragment.*

Proof. The maximum of the Lorenz curve of a state ρ is independent of the σ–fragment due to Lemma 9, and can be expressed as an increasing function of mana,

$$\max_{x\in[0,1]} \mathrm{L}_{\rho|\sigma}(x) = 1 + \mathrm{sn}(\rho) = \frac{1}{2}\left(1 + e^{\mathrm{mana}(\rho)}\right). \tag{35}$$

Therefore, the majorisation bound

$$\mathrm{L}_{\rho|\sigma}(x) \geq \mathrm{L}_{\tau|\sigma}(x),\ x \in [0,1] \tag{36}$$

implies the order $\max_{x\in[0,1]} \mathrm{L}_{\rho|\sigma}(x) \geq \max_{x\in[0,1]} \mathrm{L}_{\tau|\sigma}(x)$, hence the mana condition $\mathrm{mana}(\rho) \geq \mathrm{mana}(\tau)$. □

The area $\mathcal{A}_\sigma(\rho)$ between the curve $L_{\rho|\sigma}$ and the line $y = 1$ is also a resource monotone in the σ–fragment. This is clear because for any state conversion like Eq. (26), the Lorenz curve $L_{\tau|\sigma}$ is lower than $L_{\rho|\sigma}$, hence $\mathcal{A}_\sigma(\mathcal{E}(\rho)) \leq \mathcal{A}_\sigma(\rho)$. It cannot be directly expressed in terms of mana, as it depends on positive Wigner components as well.

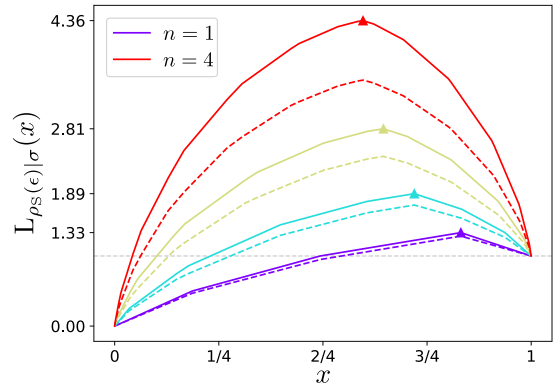

FIG. 4. **Strange state Lorenz curves.** Lorenz curves of $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ for $n = 1, 2, 3, 4$ in the σ–fragment where $\sigma = e^{-\beta H}/\mathcal{Z}_\beta$, with Hamiltonian spectrum $(0, 1, 2)$ over the computational basis and $\beta = 0.5$. Solid lines represent pure Strange states ($\epsilon = 0$); dashed lines represent noisy Strange states ($\epsilon = 0.1$).

In Sections IV and V, we study majorisation constraints on magic distillation arising in different fragments. For this reason, we define the n–copy, ϵ–noisy Strange state,

$$\rho_{\mathrm{S}}(\epsilon)^{\otimes n} := \left[(1-\epsilon)\left|\mathrm{S}\right\rangle\left\langle\mathrm{S}\right| + \epsilon\frac{1}{3}\mathbb{1}\right]^{\otimes n}, \tag{37}$$

in $\mathcal{O}_\sigma$, where the pure magic state $|\mathrm{S}\rangle$ is induced with depolarising noise. Its Wigner distribution is visualised in Fig. (1(d)).

We refer to parameter ϵ as *noise level* to avoid confusion with the error rate δ, a term commonly used in the literature to denote that the distilled state has a marginal overlap of at least $1-\delta$ with the desired magic state. In our case the error rate of the state in Eq. (37) would be $\delta = \frac{2}{3}\epsilon$.

At $\epsilon = 0$, the Strange state is a qutrit magic state of maximal sum-negativity / mana [7] and therefore acts as an ideal distillation target, analogous to a Bell state in bipartite entanglement theory. The Strange state is exceptionally symmetric under Clifford transformations and as a result there exists a *twirling* protocol, discussed in detail in [33, 34], allowing for the conversion of any noisy magic state to the form of Eq. (37) via Cliffords.

In Fig. (4), we illustrate the Lorenz curve $\mathrm{L}_{\rho_{\mathrm{S}}|\sigma}$ [3] of pure and noisy n–copy Strange states in some thermal σ–fragment. Due to Lemma 9, it is clear that the curves peak at $1 + \mathrm{sn}(\rho_{\mathrm{S}}(\epsilon)^{\otimes n})$.

[3] Hereinafter, we may omit obvious variable dependencies like ϵ and n for clarity

IV. MAGIC IN THE UNITAL FRAGMENT

The unital fragment encompasses the circuits which preserve the maximally mixed state ($\mathbb{1}/d$) and so it includes many important families of circuits.

MSD circuits in principle consist of bulk sequences of random Clifford unitaries (RCU) [4], depicted in Fig. (2). Operations in RCU can be expressed as

$$\mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger, \ U_i \in \mathcal{C}_d. \tag{38}$$

Depending on the symmetries of such operations, a Clifford sequence may belong in other σ–fragments as well. In such a case, the majorisation condition (27) needs to be checked in the σ–fragments that reflect all symmetries of the operation sequence.

In general, noisy circuits are well-described by the unital fragment. To see this, consider incorporating noisy channels in the circuit, for example dephasing channels as in Eq. (22) defined in different bases. This process destroys the circuit symmetries, except for the invariance of the maximally mixed state. Dephasing and bit-flip error channels are examples of the many error-inducing channels that respect the unital symmetry. [Expand on significance of unital fragment]

We consider the task of purifying n copies of a noisy Strange state $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ as given in Eq. (37) into a smaller number of copies n' of a less noisy strange state $\rho_{\mathrm{S}}(\epsilon')^{\otimes n'}$, with $\epsilon' < \epsilon$ and $n' \leq n$,

$$\rho_{\mathrm{S}}(\epsilon)^{\otimes n} \longrightarrow \rho_{\mathrm{S}}(\epsilon')^{\otimes n'} \otimes \left(\frac{1}{3}\mathbb{1}\right)^{\otimes(n-n')}, \tag{39}$$

where all copies $n, n', n-n'$ are even. Since the state $\mathbb{1}/3$ is free, tensoring in copies of it does not affect the distillation process. The distillation rate $R := n'/n$ for this process will in general depend on the noise levels, $R = R(\epsilon, \epsilon')$ and our task is to provide it with an upper bound.

The Lorenz curve $\mathrm{L}_{\rho_{\mathrm{S}}|\mathbb{1}/3}$, for some general noise parameter ϵ and number of copies n, is defined at 9^n points between 0 and 1. The exact expressions for the coordinates of these points can take 8 different forms, depending on whether the noise level ϵ is greater or less than $\frac{3}{7}$, the parity of the number of copies n is even or odd and the location relative to the curve peak is on the left hand side (LHS – including the curve peak) or right hand side (RHS) of the curve peak. The full details for the construction of all Lorenz curve forms are provided in Appendix D.

Here we focus on the comparison of the first elbow (which lies in the LHS part) of Lorenz curves with even copies n, n' and low noise levels ($\epsilon' < \epsilon \leq 3/7$).

The Wigner distribution of the 1-copy, ϵ–noisy Strange state can be written as

$$\mathrm{W}_{\rho_{\mathrm{S}}(\epsilon)}(\boldsymbol{x}) = (1-\epsilon)\mathrm{W}_{|\mathrm{S}\rangle\langle\mathrm{S}|}(\boldsymbol{x}) + \epsilon \mathrm{W}_{\frac{1}{3}\mathbb{1}}(\boldsymbol{x}), \tag{40}$$

so we get positive components

$$u(\epsilon) := \frac{1}{6} - \frac{1}{18}\epsilon \tag{41}$$

at the 8 phase space points $\boldsymbol{x} \in \mathcal{P}_3 \setminus \{\boldsymbol{0}\}$ and a negative component

$$-v(\epsilon) := -\left(\frac{1}{3} - \frac{4}{9}\epsilon\right) \tag{42}$$

at the origin $\boldsymbol{x} = \boldsymbol{0}$.

The rescaled distribution in the unital fragment simply is

$$\widetilde{\mathrm{W}}_{\rho_{\mathrm{S}}|\frac{1}{3}\mathbb{1}}(\boldsymbol{x}) = \frac{\mathrm{W}_{\rho_{\mathrm{S}}}(\boldsymbol{x})}{\mathrm{W}_{\frac{1}{3}\mathbb{1}}(\boldsymbol{x})} = d\mathrm{W}_{\rho_{\mathrm{S}}}(\boldsymbol{x}), \tag{43}$$

so ordering the rescaled distribution is equivalent to ordering $\mathrm{W}_{\rho_{\mathrm{S}}}$.

The component values and multiplicities (m_i) in the n–copy case are

$$m_i = 8^{2i}\binom{n}{2i}, \tag{44}$$

$$w(\rho_{\mathrm{S}})_i = u^{2i}(-v)^{n-2i}, \tag{45}$$

$$w(\mathbb{1}/3)_i = \frac{1}{9^n}, \tag{46}$$

where index i runs through $0, \ldots, \frac{n}{2}$. This is derived in Appendix D 2.

It is readily seen that the maximum Wigner component is achieved when $i = 0$, so the corresponding multiplicity and Wigner component are $m_0 = 1$, $w(\rho_{\mathrm{S}})_0 = v^n$. The first elbow coordinates therefore can be expressed as

$$(x_0, L_0) = (m_0\, w(\mathbb{1}/3)_0,\ m_0\, w(\rho_{\mathrm{S}})_0) = \left(\frac{1}{9^n}, v^n\right) \tag{47}$$

The first elbow of the target state is located at $x_0' = \frac{1}{9^{n'}} > x_0$, so we can use the first elbow condition,

$$\frac{L_0}{x_0} \geq \frac{L_0'}{x_0'}, \tag{48}$$

to compute analytical distillation bounds for distillation rate $R = R(\epsilon, \epsilon') := n'/n$ in the unital fragment. We substitute coordinates from Eq. (80) appropriately in Eq. (48) to get

$$R \leq \frac{\ln(3-4\epsilon)}{\ln(3-4\epsilon')}. \tag{49}$$

Specifically for the problem of distilling pure magic states ($\epsilon' = 0$), we obtain an upper bound in the unital fragment given by

$$R \leq 1 + \frac{\ln\left(1 - \frac{4}{3}\epsilon\right)}{\ln 3}. \tag{50}$$

Similar upper bounds on distillation rates for qudits of odd prime dimension include the mana bound [1] and

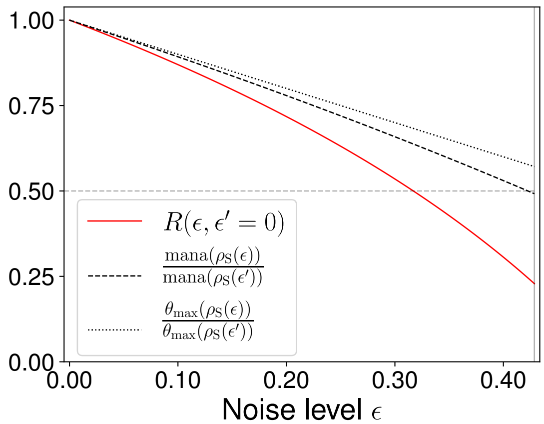

FIG. 5. **Distillation bounds in the unital fragment.** Distillation bounds obtained by majorisation, mana and the line $1 - \epsilon$ (tightest bound of max–thauma) are plotted for $\epsilon' = 0$, up to noise level $\epsilon = 3/7$. Majorisation consistently provides stricter rates than mana and max–thauma.

the max–thauma bound [35] which is defined via a semi-definite program, but possesses properties that allow for an easy comparison with our bound. The mana bound can be directly calculated as

$$R \leq \frac{\mathrm{mana}\,(\rho_S(\epsilon))}{\mathrm{mana}\,(\rho_S(0))} = 1 + \frac{\ln\left(1 - \frac{8}{15}\epsilon\right)}{\ln\frac{5}{3}}. \tag{51}$$

Using the max–thaum properties of super-additivity and vanishing at free states, we obtain

$$\theta_{\max}(\rho_S(\epsilon)) \geq (1-\epsilon)\theta_{\max}(|S\rangle\langle S|) + \epsilon\theta_{\max}\left(\frac{1}{3}\mathbb{1}\right) = (1-\epsilon)\theta_{\max}(\rho_S(0)), \tag{52}$$

so that for a deterministic single-copy distillation process, an initial number of copies $n \geq \theta_{\max}(\tau)/\theta_{\max}(\rho) \geq 1-\epsilon$ is required. Therefore, the max–thauma bound for the pure Strange state distillation process can get as tight as

$$R = \frac{1}{n} = \frac{\theta_{\max}(\rho_S(\epsilon))}{\theta_{\max}(\rho_S(0))} \leq 1 - \epsilon. \tag{53}$$

Both the mana and the max–thauma bounds are looser than the majorization bound we derived via the first elbow constraint as illustrated in Fig. (5).

V. MAGIC BOUNDS IN ARBITRARY STABILISER FRAGMENTS

We now generalise the approach taken in the previous section and consider bounds on magic distillation for an arbitrary stabiliser fragment, $\mathcal{R}_\sigma$ where σ is any quantum state $\sigma \in \mathrm{STAB}$. In other words, we consider those bounds on distillation that apply when the free operations have σ as a fixed point.

These bounds can be interpreted in two different ways: on one hand they can be viewed as a family of upper bounds parameterized by a stabiliser state σ, on the other we can associate σ to actual hardware limitations or to biased noise models in which it is an equilibrium state of some kind. Without loss of generality, we can always write a stabiliser state σ as a Gibbs state $\sigma = \gamma_\beta := \frac{1}{\mathcal{Z}_\beta}e^{-\beta H}$ for some $\beta \geq 0$ and Hamiltonian H as discussed in Section III B.

We focus on the distillation process

$$\rho_S(\epsilon)^{\otimes n} \longrightarrow \rho_S(\epsilon')^{\otimes n'} \otimes \sigma^{\otimes(n-n')}, \tag{54}$$

where the noisy Strange state is given in Eq. (37) and we highlight again that any magic state can be transformed to this form via Clifford operations. Notice that tensoring in copies of γ_β does not affect the process, since the state is free. In the following result, all copies $n, n', n-n'$ are even with $n > n'$, and $\epsilon' = 0$, but the bounds are easily generalised to odd numbers of copies and ϵ' such that $0 < \epsilon' < \epsilon$. Finally, we assume a range of initial noise levels, $\epsilon \leq 3/7$, so that the largest Wigner component of the 1-copy noisy magic state is negative. The numerical value of $3/7$ is higher than corresponding values of relevant existing distillation protocol error thresholds [4, 33].

Given this context, we now provide the following result on bounding the distillation rate $R = R(\epsilon, \epsilon', \beta) := \frac{n'}{n}$ of the process in Eq. (54). We also write $R(\epsilon, \beta) = R(\epsilon, \epsilon' = 0, \beta)$. The bounds depend on the free energy F_β of state σ,

$$F_\beta := \mathrm{tr}[H\sigma] - \beta^{-1}S(\sigma) = -\beta^{-1}\log\mathcal{Z}_\beta, \tag{55}$$

where the von Neumann entropy is $S(\sigma) := -\mathrm{tr}[\sigma\log\sigma]$.

Theorem 11. *Let σ be a qutrit stabiliser state, given by $\sigma = \frac{1}{\mathcal{Z}}e^{-\beta H}$ with H having eigenvalues $E_0 \leq E_1 \leq E_2$, and where β is an (effective) inverse temperature for the state. We define*

$$\beta_\star := \frac{1}{E_2 - E_0}\ln 2 \tag{56}$$

and for $\beta \leq \beta_\star$, we define a threshold noise level

$$\epsilon_\star(\beta) := 3 - \frac{18}{8 - e^{(E_2-E_0)\beta}}. \tag{57}$$

Then any distillation rate $R(\epsilon, \beta)$ in the σ–fragment of a magic theory is bounded as follows:
If $\beta \leq \beta_\star$ and $\epsilon \leq \epsilon_\star$,

$$R(\epsilon, \beta) \leq 1 + \frac{\ln\left(1 - \frac{4}{3}\epsilon\right)}{\beta(E_0 - F_\beta)}. \tag{58}$$

If $\beta \leq \beta_\star$ and $\epsilon_\star < \epsilon$, then

$$R(\epsilon, \beta) \leq 1 + \frac{\ln\left(1 - \frac{1}{3}\epsilon\right) - (E_2 - E_0)(\beta_\star - \beta)}{\beta(E_0 - F_\beta)}. \tag{59}$$

Otherwise if $\beta > \beta_\star$,

$$R(\epsilon, \beta) \leq 1 + \frac{\ln\left(1 - \frac{1}{3}\epsilon\right)}{-\ln 2 + \beta(E_2 - F_\beta)}. \tag{60}$$

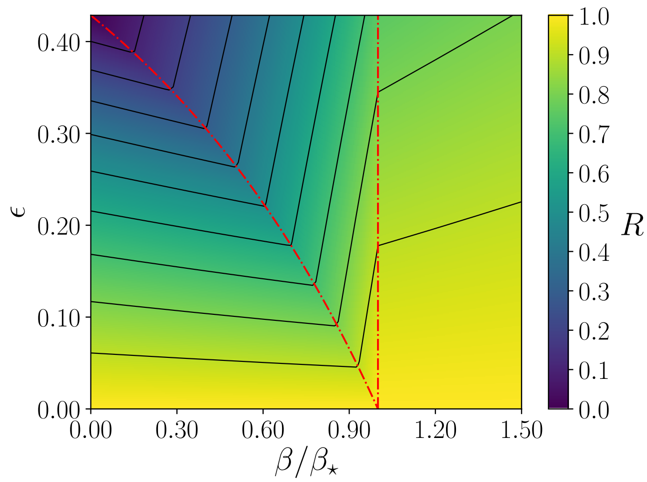

FIG. 6. **Bounds on magic distillation rates** $R(\epsilon, \beta)$ **within stabiliser fragments.** The vertical dashed line is the 'Landauer-like' temperature threshold $\beta_\star$ and the diagonal dashed curve corresponds to the noise threshold $\epsilon_\star(\beta)$ at every $\beta \leq \beta_\star$. The unital fragment corresponds to the $\beta = 0$ line.

A few comments can be given on this result. Firstly, the specific numerical factors in $\epsilon_\star$ are a result of our choice of magic state. Secondly, these bounds are derived based on analysis of only a part of the Lorenz curves and can be improved via a finer analysis. This is apparent by simple numerical calculations on the entirety of the curves. Specifically, the existing bounds follow by considering the dominant terms in the rescaled Wigner distribution and do not, for example, take into account the Lorenz curve's peak structure. Finally, it is striking that we obtain a Landauer-like condition with a characteristic temperature $kT_\star \ln 2 = E_2 - E_0$, where $kT_\star = \beta_\star^{-1}$. It is unclear whether this points to a generic feature that can be directly related to fundamental thermodynamic relations, such as the erasure cost of a single bit being $kT \ln 2$. Given that we work with qutrits this seems surprising, but does deserve further study.

Proof. We provide the full proof here, with some technical details pushed to Appendix E.

Let $\sigma = e^{-\beta H}/\mathcal{Z}_\beta$ be a stabiliser state, where $\beta \geq 0$ and $H = E_0 \left|0\right\rangle\!\left\langle 0\right| + E_1 \left|1\right\rangle\!\left\langle 1\right| + E_2 \left|2\right\rangle\!\left\langle 2\right|$, with $E_k \geq 0$. Its eigen-decomposition can be written as

$$\sigma = \frac{e^{-\beta E_0}}{\mathcal{Z}_\beta} \left|\varphi_0\right\rangle\!\left\langle \varphi_0\right| + \frac{e^{-\beta E_1}}{\mathcal{Z}_\beta} \left|\varphi_1\right\rangle\!\left\langle \varphi_1\right| + \frac{e^{-\beta E_2}}{\mathcal{Z}_\beta} \left|\varphi_2\right\rangle\!\left\langle \varphi_2\right|, \tag{61}$$

where $\{\left|\varphi_k\right\rangle\}$ are pure, orthonormal stabiliser states. Therefore, there exists a Clifford operation that maps σ to γ_β, where

$$\gamma_\beta = \frac{e^{-\beta E_0}}{\mathcal{Z}_\beta} \left|0\right\rangle\!\left\langle 0\right| + \frac{e^{-\beta E_1}}{\mathcal{Z}_\beta} \left|1\right\rangle\!\left\langle 1\right| + \frac{e^{-\beta E_2}}{\mathcal{Z}_\beta} \left|2\right\rangle\!\left\langle 2\right|. \tag{62}$$

[Prove this carefully]

This Clifford operation permutes the Hamiltonian eigenvalues on the phase space, so that the negative component $-v$ of $\rho_S(\epsilon)$ can lie on the same point on the phase space as any of the eigenvalues. For this reason, we impose no order between the eigenvalues, but simply choose E_0 as the eigenvalue that is associated with the state negativity and denote the highest energy by $E_{\max} \coloneqq \max\{E_0, E_1, E_2\}$.

The Wigner distribution of state γ_β can be seen as the ensemble average of the distributions of the computational basis states,

$$\begin{aligned} \mathrm{W}_{\gamma_\beta}(\boldsymbol{x}) &= \sum_{k=0}^{2} \frac{e^{-\beta E_k}}{\mathcal{Z}_\beta} \mathrm{W}_{\left|k\right\rangle\!\left\langle k\right|}(\boldsymbol{x}) & (63)\\ &= \sum_{k=0}^{2} \frac{e^{-\beta E_k}}{\mathcal{Z}_\beta} \delta_{x_0,k} = \frac{e^{-\beta E_{x_0}}}{3\mathcal{Z}_\beta}, & (64) \end{aligned}$$

for all $\boldsymbol{x} \in \mathcal{P}_3$. All Wigner components are strictly positive, therefore the pre-order $\prec_{\gamma_\beta}$ is always well-defined.

Our aim is to obtain a distillation bound for Eq. (54) which depends on variables $n, n', \epsilon, \epsilon'$ as well as β. In the analysis that follows, we again drop obvious variable dependencies for clarity.

We construct the Strange state rescaled distribution

$$\widetilde{\mathrm{W}}_{\rho_S|\gamma}(\boldsymbol{x}) = \frac{\mathrm{W}_{\rho_S}(\boldsymbol{x})}{\mathrm{W}_{\gamma}(\boldsymbol{x})}, \tag{65}$$

which attains four distinct values on the phase space, naturally splitting it into four regions as illustrated in the Fig. (7).

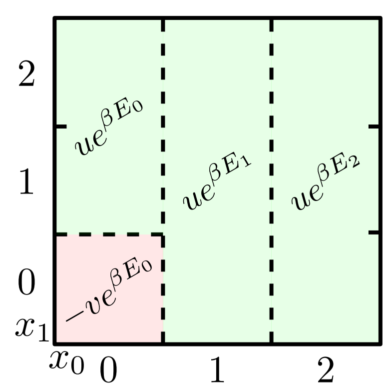

FIG. 7. **Qutrit phase space regions with different rescaled values.** The rescaled distribution attains a unique value in each of the four regions, given by $3\mathcal{Z}\times$ the value depicted in the region, according to Eq. (69).

The component values and multiplicities of the relevant distributions in the four distinct regions are summarised by the following component and multiplicity vectors,

$$\boldsymbol{m} := (1,2,3,3), \tag{66}$$
$$\boldsymbol{w}(\rho_{\mathrm{S}}) := (-v,u,u,u), \tag{67}$$
$$\boldsymbol{w}(\gamma) := \frac{1}{3\mathcal{Z}}\left(e^{-\beta E_0}, e^{-\beta E_0}, e^{-\beta E_1}, e^{-\beta E_2}\right), \tag{68}$$
$$\boldsymbol{w}(\rho_{\mathrm{S}}|\gamma) := 3\mathcal{Z}\left(-ve^{\beta E_0}, ue^{\beta E_0}, ue^{\beta E_1}, ue^{\beta E_2}\right). \tag{69}$$

Using this notation, the component values and multiplicities of the n–copy distributions can be readily provided by Lemma 25 in Appendix E 1. They are parametrised by three independent components i,j,k, with sum $\alpha := i+j+k \leq n$. The n–copy multiplicity is given by

$$m_{ijk} = \frac{n!}{i!j!k!(n-\alpha)!}2^i 3^j 3^k, \tag{70}$$

while the distribution values that correspond to the same index triplet (i,j,k) are given by

$$w(\rho_{\mathrm{S}})_{ijk} = (-v)^{n-\alpha}u^{\alpha}, \tag{71}$$
$$w(\gamma)_{ijk} = (3\mathcal{Z})^{-n}e^{-\beta(n-\alpha)E_0}e^{-\beta(iE_0+jE_1+kE_2)}, \tag{72}$$
$$w(\rho_{\mathrm{S}}|\gamma)_{ijk} = (3\mathcal{Z})^{n}(-v)^{n-\alpha}u^{\alpha}e^{\beta(n-\alpha)E_0}e^{\beta(iE_0+jE_1+kE_2)}. \tag{73}$$

In order to construct the n–copy Lorenz curve $\mathrm{L}_{\rho_{\mathrm{S}}^{\otimes n}|\gamma}$ we need to sort the components of the n–copy rescaled distribution, $w(\rho_{\mathrm{S}}|\gamma)_{ijk}$ in decreasing order. In particular, to find the coordinates of the first elbow (x_0, L_0), we need to evaluate the maximum rescaled component,

$$\begin{aligned}&\boldsymbol{w}(\rho_{\mathrm{S}}|\gamma)_{\max} := \\ &(3\mathcal{Z})^n \max_{i,j,k}\left\{(-v)^{n-\alpha}u^{\alpha}e^{\beta(n-\alpha)E_0}e^{\beta(iE_0+jE_1+kE_2)}\right\},\end{aligned} \tag{74}$$

where $0 \leq i,j,k \leq n$ and $\alpha := i+j+k \leq n$. Notice that for $0 \leq \epsilon \leq 3/7$, we have $v \geq u$. We assume that n is even, so that we need the sum $\alpha = i+j+k$ to be even for the expression to be positive. The following analysis is similar if n is chosen to be odd.

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

$$\begin{aligned}&\boldsymbol{w}(\rho_{\mathrm{S}}|\gamma)_{\max} = \\ &(3\mathcal{Z})^n v^n e^{n\beta E_0} \max_{\substack{\alpha=0,2,\\ \dots,n-2,n}}\left\{\left(\frac{u}{v}e^{\beta(E_{\max}-E_0)}\right)^{\alpha}\right\}.\end{aligned} \tag{75}$$

If the expression $\frac{u}{v}e^{\beta(E_{\max}-E_0)}$ is less than 1 then the maximum occurs at $\alpha=0$, otherwise the maximum occurs at $\alpha = n$. To determine this transition we set

$$\frac{u(\epsilon)}{v(\epsilon)}e^{\beta(E_{\max}-E_0)} := \frac{3-\epsilon}{6-8\epsilon}e^{\beta(E_{\max}-E_0)} = 1. \tag{76}$$

We want to find in which cases there exists a threshold noise level $\epsilon_\star$ at which the transition in Eq. (76) occurs. If $E_{\max} = E_0$, namely if the state negativity lies in the same phase space region as the highest energy, this threshold is constant in temperature and given by $\epsilon_\star = 3/7$. Otherwise, there is a threshold temperature value $\beta_\star$ given by

$$\beta_\star := \frac{1}{E_{\max}-E_0}\ln 2. \tag{77}$$

Below the threshold, $0 \leq \beta \leq \beta_\star$, the transition is well defined and the threshold noise level at which it occurs is given by

$$\epsilon_\star(\beta) := 3 - \frac{18}{8-e^{(E_{\max}-E_0)\beta}}. \tag{78}$$

This encompasses the case $E_{\max} = E_0$. For $\beta > \beta_\star$, we have

$$\frac{3-\epsilon}{6-8\epsilon}e^{\beta(E_{\max}-E_0)} > \frac{3-\epsilon}{6-8\epsilon}2 \geq \frac{1}{2}2 = 1,$$

so there is no transition and we set $\epsilon_\star = 0$.

The maximum rescaled component can then be expressed as

$$\boldsymbol{w}(\rho_{\mathrm{S}}|\gamma)_{\max} = \begin{cases}(3\mathcal{Z})^n v^n e^{n\beta E_0}, & \epsilon \leq \epsilon_\star, \quad \text{(C1)}\\ (3\mathcal{Z})^n u^n e^{n\beta E_{\max}}, & \epsilon > \epsilon_\star. \quad \text{(C2)}\end{cases}$$

Case (C1) corresponds to $(i,j,k) = (0,0,0)$, so the multiplicity is $m_{000} = 1$ and the corresponding Wigner components are $\boldsymbol{w}(\rho_{\mathrm{S}})_{000}, \boldsymbol{w}(\gamma)_{000}$. Case (C2) corresponds to

$$(i,j,k) = \begin{cases}(0,n,0), & \text{if } E_{\max} = E_1,\\ (0,0,n), & \text{if } E_{\max} = E_2.\end{cases} \tag{79}$$

and we have $E_{\max} = E_1$ ($E_{\max} = E_2$), so the multiplicity is $m_{0n0} = 3^n$ ($m_{00n} = 3^n$) and the corresponding Wigner components are $\boldsymbol{w}(\rho_{\mathrm{S}})_{0n0}, \boldsymbol{w}(\gamma)_{0n0}$ ($\boldsymbol{w}(\rho_{\mathrm{S}})_{00n}, \boldsymbol{w}(\gamma)_{00n}$).

The first elbow coordinates can finally be derived as

$$(x_0, L_0) = \begin{cases}\left(\left(\dfrac{e^{-\beta E_0}}{3\mathcal{Z}_\beta}\right)^n, v^n\right), & \text{(C1)}\\[2ex] \left(\left(\dfrac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^n, (3u)^n\right). & \text{(C2)}\end{cases} \tag{80}$$

The Lorenz curves of the initial and target states may each be described by either (C1) or (C2), depending on the physical parameters $\epsilon, \epsilon', \beta$. Specifically, we have three scenarios:

1. (C1) $\to$ (C1) if $E_{\max} = E_0$ or $E_{\max} > E_0$, $\beta < \beta_\star$ and $\epsilon' < \epsilon \leq \epsilon_\star$.
2. (C2) $\to$ (C1) if $E_{\max} > E_0$, $\beta < \beta_\star$ and $\epsilon' \leq \epsilon_\star < \epsilon$.
3. (C2) $\to$ (C2) if $E_{\max} > E_0$, $\beta < \beta_\star$ and $\epsilon_\star \leq \epsilon' < \epsilon$ or $E_{\max} > E_0$, $\beta \geq \beta_\star$.

Note that (C1) $\to$ (C2) is impossible because it would imply $\beta < \beta_\star$ and $\epsilon \le \epsilon_\star \le \epsilon' < \epsilon$, a contradiction.

In all three scenarios, it is simple to check that the initial state's first elbow is always located to the left (closer to 0) of the target state's first elbow, $x_0 \le x_0'$, as proven in Appendix E 2, where the prime ($'$) is used to indicate target state coordinates. Therefore, we can use the first elbow condition,

$$\frac{L_0}{x_0} \ge \frac{L_0'}{x_0'}, \tag{81}$$

to compute analytical distillation bounds for the distillation rate $R = R(\epsilon, \epsilon', \beta) \coloneqq n'/n$ in all three possible scenarios. Involving more elbows gives stricter, but more convoluted necessary distillation constraints.

We substitute coordinates from Eq. (80) appropriately in Eq. (81) to get the following necessary conditions,

$$R \le \begin{cases} \dfrac{\ln\left(1 - \frac{4}{3}\epsilon\right) + \beta(E_0 - F_\beta)}{\ln\left(1 - \frac{4}{3}\epsilon'\right) + \beta(E_0 - F_\beta)}, & \text{(C1)} \to \text{(C1)}, \\[2ex] \dfrac{\ln\left(\frac{1}{2} - \frac{1}{6}\epsilon\right) + \beta(E_{\max} - F_\beta)}{\ln\left(1 - \frac{4}{3}\epsilon'\right) + \beta(E_0 - F_\beta)}, & \text{(C2)} \to \text{(C1)}, \\[2ex] \dfrac{\ln\left(\frac{1}{2} - \frac{1}{6}\epsilon\right) + \beta(E_{\max} - F_\beta)}{\ln\left(\frac{1}{2} - \frac{1}{6}\epsilon'\right) + \beta(E_{\max} - F_\beta)}, & \text{(C2)} \to \text{(C2)}. \end{cases} \tag{82}$$

Equivalently, we require that

$$\epsilon \le \begin{cases} \frac{3}{4} - \frac{3}{4}\left(1 - \frac{4}{3}\epsilon'\right)^R \left(\dfrac{e^{-\beta E_0}}{\mathcal{Z}_\beta}\right)^{1-R}, & \text{(C1)} \to \text{(C1)}, \\[2ex] 3 - 6\left(1 - \frac{4}{3}\epsilon'\right)^R \dfrac{e^{-\beta E_{\max}} e^{R\beta E_0}}{\mathcal{Z}_\beta^{1-R}}, & \text{(C2)} \to \text{(C1)}, \\[2ex] 3 - 6\left(\frac{1}{2} - \frac{1}{6}\epsilon'\right)^R \left(\dfrac{e^{-\beta E_{\max}}}{\mathcal{Z}_\beta}\right)^{1-R}, & \text{(C2)} \to \text{(C2)}. \end{cases} \tag{83}$$

Substituting $\epsilon' = 0$ in Eq. (82) leads to the theorem statement.

□

VI. LOWER MAGIC BOUNDS VIA MAJORISATION

[Add from notes]

VII. FRAGMENTS IN GENERAL RESOURCE THEORIES

[Add from notes]

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. 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 12 (**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 13 (**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 Definition 1, where the pre-order $\prec_{\mathcal{R}'}$ is simply the total order $\le$ 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, $\mathcal{M}(\rho) \coloneqq f(\Pi_s(\rho))$ 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 14 (**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* $\Pi_s(\rho_1) \prec_{\mathcal{R}'} \Pi_s(\rho_2)$;
2. $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2$ *whenever* $\Pi_o(\mathcal{E}) = \Pi_o(\mathcal{E}_1) \circ \Pi_o(\mathcal{E}_2)$.

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

The use of covariant and contravariant in Definitions 12 and 14 refers to the direction of implication between the two pre-orders and operation compositions[4].

VIII. CONCLUSION

[Summary]

[4] 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.

[1] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New Journal of Physics **14**, 113011 (2012).
[2] A. Mari and J. Eisert, Phys. Rev. Lett. **109**, 230503 (2012).
[3] D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. thesis, California Institute of Technology (1997).
[4] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[5] E. Knill, Nature **434** (2005), 10.1038/nature03350.
[6] E. T. Campbell, Phys. Rev. A **83**, 032317 (2011).
[7] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, New Journal of Physics **16**, 013009 (2014).
[8] M. Howard and E. Campbell, Phys. Rev. Lett. **118**, 090501 (2017).
[9] X. Wang, M. M. Wilde, and Y. Su, New Journal of Physics **21**, 103002 (2019).
[10] J. R. Seddon, B. Regula, H. Pashayan, Y. Ouyang, and E. T. Campbell, PRX Quantum **2**, 010345 (2021).
[11] D. Gottesman, in *Encyclopedia of Mathematical Physics*, edited by J.-P. Françoise, G. L. Naber, and T. S. Tsun (Academic Press, Oxford, 2006) pp. 196 – 201.
[12] D. Gottesman, Chaos, Solitons & Fractals **10**, 1749 (1999).
[13] M. Ahmadi, H. B. Dang, G. Gour, and B. C. Sanders, Phys. Rev. A **97**, 062332 (2018).
[14] J. R. Seddon and E. T. Campbell, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **475**, 20190251 (2019).
[15] E. T. Campbell and D. E. Browne, Phys. Rev. Lett. **104**, 030503 (2010).
[16] N. Delfosse, C. Okay, J. Bermejo-Vega, D. E. Browne, and R. Raussendorf, New Journal of Physics **19**, 123024 (2017).
[17] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature **510** (2014), 10.1038/nature13460.
[18] D. Gross, Journal of Mathematical Physics **47**, 122107 (2006).
[19] D. Gross, S. T. Flammia, and J. Eisert, Phys. Rev. Lett. **102**, 190501 (2009).
[20] G. B. Folland, *Harmonic Analysis in Phase Space* (Princeton University Press, 1989).
[21] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2006).
[22] X. Wang, M. M. Wilde, and Y. Su, arXiv e-prints , arXiv:1812.10145 (2018), arXiv:1812.10145 [quant-ph].
[23] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer, 2011).
[24] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[25] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **115**, 210403 (2015).
[26] M. Lostaglio, D. Jennings, and T. Rudolph, Nature Communications **6** (2015), 10.1038/ncomms7383.
[27] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, Nature Communications **5352** (2018), 10.1038/s41467-018-06261-7.
[28] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern, Physics Reports **583** (2015), 0.1016/j.physrep.2015.04.003.
[29] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A **67**, 062104 (2003).
[30] R. O. Vallejos, F. de Melo, and G. G. Carlo, (2021), arXiv:2102.09999.
[31] M. Lostaglio, Reports on Progress in Physics **82**, 114001 (2019).
[32] L. E. J. Brouwer, Mathematische Annalen **71** (1911), 10.1007/BF01456931.
[33] S. Prakash, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **476**, 20200187 (2020).
[34] A. Jain and S. Prakash, Phys. Rev. A **102**, 042409 (2020).
[35] X. Wang, M. M. Wilde, and Y. Su, Phys. Rev. Lett. **124**, 090505 (2020).
[36] A. Vourdas, Reports on Progress in Physics **67**, 267 (2004).
[37] L. Mirsky, Mathematische Nachrichten **20**, 171 (1959).
[38] R. Bhatia, *Matrix Analysis* (Springer, 1997).
[39] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. (Cambridge University Press, USA, 2011).
[40] R. Ash, *Information Theory* (Dover Publications Inc.,

1965).

Appendix A: Properties of Wigner distributions

Here we present important properties of the Wigner distribution that are used throughout the paper.

Proposition 15. *The Wigner distribution of a state* $\rho \in \mathcal{B}(\mathcal{H}_d)$ *is*

(i) Real valued: $\mathrm{W}_\rho \in \mathbb{R}^{d^2}$*;*

(ii) Normalised: $\sum_{\boldsymbol{z}\in\mathcal{P}_d} \mathrm{W}_\rho(\boldsymbol{z}) = 1$*;*

(iii) Bounded: $|\mathrm{W}_\rho(\boldsymbol{x})| \leq \frac{1}{d}$.

(iv) Additive under mixing:
$\mathrm{W}_{p\rho_1+(1-p)\rho_2}(\boldsymbol{x}) = p\mathrm{W}_{\rho_1}(\boldsymbol{x}) + (1-p)\mathrm{W}_{\rho_2}(\boldsymbol{x})$*;*

(v) Multiplicative under tensor products:
$\mathrm{W}_{\rho_A\otimes\rho_B}(\boldsymbol{x}_A \oplus \boldsymbol{x}_B) = \mathrm{W}_{\rho_A}(\boldsymbol{x}_A)\mathrm{W}_{\rho_B}(\boldsymbol{x}_B)$.

Proof. Proof of all properties can be found in the literature [1, 9, 18, 36] except for property (iii) which we prove here.

Let $\{\lambda_i\}_{i\in\mathbb{Z}_d}$ be the (non-negative) eigenvalues of ρ, summing to 1. Let $\{\alpha_{\boldsymbol{x},i}\}_{i\in\mathbb{Z}_d}$ be the eigenvalues of $A_{\boldsymbol{x}}$. For any $\boldsymbol{x}, \alpha_{\boldsymbol{x},i} \in \{-1, 1\}$, due to the hermiticity and unitarity of the phase-point operators. Then,

$$|\mathrm{W}_\rho \boldsymbol{x}| = \frac{1}{d}|\mathrm{tr}[A_{\boldsymbol{x}}\rho]| \leq \frac{1}{d}\left|\sum_i \alpha_{\boldsymbol{x},i}\lambda_i\right| \leq \frac{1}{d}\sum_i \lambda_i = \frac{1}{d}. \tag{A1}$$

The first inequality follows from Theorem 1 of [37] for complex matrices, while the second is the triangle inequality. □

Proposition 16. *The Wigner distribution of a* CPTP *operation* $\mathcal{E} : \mathcal{B}(\mathcal{H}_{d_A}) \mapsto \mathcal{B}(\mathcal{H}_{d_B})$ *is*

(i) Real-valued: $\mathrm{W}_\mathcal{E} \in \mathbb{R}^{d^2} \times \mathbb{R}^{d^2}$*;*

(ii) Normalised: $\sum_{\boldsymbol{z}\in\mathcal{P}_{d_B}} \mathrm{W}_\mathcal{E}(\boldsymbol{z}|\boldsymbol{x}) = 1$ *for any* $\boldsymbol{x} \in \mathcal{P}_{d_A}$*;*

(iii) Bounded: $|\mathrm{W}_\mathcal{E}(\boldsymbol{y}|\boldsymbol{x})| \leq \frac{d_A}{d_B}$*;*

(iv) Transitive: $\mathrm{W}_{\mathcal{E}(\rho)}(\boldsymbol{y}) = \sum_{\boldsymbol{z}\in\mathcal{P}_{d_A}} \mathrm{W}_\mathcal{E}(\boldsymbol{y}|\boldsymbol{z})\mathrm{W}_\rho(\boldsymbol{z})$ *for any* $\boldsymbol{y} \in \mathcal{P}_{d_B}$.

If $d_A = d_B$, and in particular if operation $\mathcal{E}$ maps a Hilbert space onto itself, then the stochasticity condition $|\mathrm{W}_\mathcal{E}(\boldsymbol{y}|\boldsymbol{x})| \leq 1$ is satisfied.

Proof. Proof of all properties can be found in Wang *et al.* [9] except for property (iii) which is a direct consequence of the definition of $\mathrm{W}_\mathcal{E}$ and the corresponding property (iii) in Proposition 15. □

Appendix B: Properties of majorization

1. Simple–majorization equivalence conditions

In the unital fragment, namely the limit of infinite temperature, $\beta = 0$, the free state is the maximally mixed state $\frac{1}{d}\mathbb{1}$ with uniform Wigner distribution $\frac{1}{d}\mathbf{1} = (\frac{1}{d}, \ldots, \frac{1}{d})$. This fragment is governed by simple majorization and we first prove strong equivalences for this type of majorization.

Proposition 17. *Given* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *the following statements are equivalent:*

(i) $\boldsymbol{x} \prec \boldsymbol{y}$*;*

(ii) $\boldsymbol{x} = B\boldsymbol{y}$ *for bistochastic* B*;*

(iii) $\sum_{i=1}^n |x_i - t| \leq \sum_{i=1}^n |y_i - t|$ *for all* $t \in \mathbb{R}$*;*

(iv) $\sum_{i=1}^n (x_i - t)^+ \leq \sum_{i=1}^n (y_i - t)^+$ *for all* $t \in \mathbb{R}$ *and* $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$*, where* $(x)^+ = \max\{x, 0\}$*;*

(v) $L_{\boldsymbol{x}}(k) \leq L_{\boldsymbol{y}}(k)$ *for* $k = 1, \ldots, n-1$ *and* $L_{\boldsymbol{x}}(n) = L_{\boldsymbol{y}}(n)$.

Proof. These simple majorisation properties are well-known and proofs can be found in [23, 31, 38, 39]. The equivalence between (i) and (ii) is the statement of the Hardy, Littlewood and Polya theorem. □

2. Embedding map

Any $\boldsymbol{d}$–majorization problem can be rephrased as a simple majorization problem in a higher dimensional space via the embedding map.

Definition 18. *The embedding map* $\Gamma_{\boldsymbol{d}} : \mathbb{R}^n \mapsto \mathbb{R}^N, N = \sum_{i=1}^n d_i$ *is the function*

$$\Gamma_{\boldsymbol{d}}(\boldsymbol{w}) = \bigoplus_{i=1}^n w_i \, \frac{1}{d_i}\mathbf{1}, \tag{B1}$$

where $\mathbf{1}/d_i$ *is the* d_i*–dimensional uniform distribution. The left inverse* $\Gamma_{\boldsymbol{d}}^{-1} : \mathbb{R}^N \mapsto \mathbb{R}^n$ *is defined to sum up the elements in each block of* $\Gamma_{\boldsymbol{d}}(\boldsymbol{w})$*, so that*

$$\Gamma_{\boldsymbol{d}}^{-1}(\oplus_{i=1}^n w_i \mathbf{1}/d_i) = \boldsymbol{w}. \tag{B2}$$

This is not a right inverse, because $\Gamma_{\boldsymbol{d}}$ *is not surjective.*

The direct sum simply amounts to listing the uniform distributions one after the other. The embedding map maps the Gibbs distribution to the uniform distribution, $\Gamma_{\boldsymbol{d}}(\boldsymbol{d}) = \mathbf{1}/N$. Then, a non-increasing ordering

$\Gamma_{\boldsymbol{d}}(\boldsymbol{z})^{\downarrow}$ in the new space, corresponds to the so-called "β-ordering" of the original vector denoted by the permutation π in Definition 3, mapping $(w_i/d_i) \mapsto (w_i/d_i)^{\downarrow}$ for all $i = 1, \ldots, n$.

3. $\boldsymbol{d}$-majorization equivalence conditions

We take the opportunity here to simply list useful equivalent statements for $\boldsymbol{d}$–majorisation.

Proposition 19. *Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d} \in \mathbb{R}^n$, such that the components of $\boldsymbol{d}$ are positive, the following statements are equivalent:*

(i) $\boldsymbol{x} \prec_{\boldsymbol{d}} \boldsymbol{y}$;

(ii) $\Gamma_{\boldsymbol{d}}(\boldsymbol{x}) \prec \Gamma_{\boldsymbol{d}}(\boldsymbol{y})$;

(iii) $\sum_{i=1}^{n} |x_i - td_i| \leq \sum_{i=1}^{n} |y_i - td_i|$ *for all* $t \in \mathbb{R}$;

(iv) $\sum_{i=1}^{n} (x_i - td_i)^+ \leq \sum_{i=1}^{n} (y_i - td_i)^+$ *for all* $t \in \mathbb{R}$ *and* $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$;

(v) $L_{\boldsymbol{x}|\boldsymbol{d}}(k) \leq L_{\boldsymbol{y}|\boldsymbol{d}}(k)$ *for* $k = 1, \ldots, n-1$ *and* $L_{\boldsymbol{x}|\boldsymbol{d}}(n) = L_{\boldsymbol{y}|\boldsymbol{d}}(n)$.

Proof.

1$\leftrightarrow$ 2 Suppose there exists a $\boldsymbol{d}$–stochastic S such that $\boldsymbol{x} = S\boldsymbol{y}$ and let $B = \Gamma_{\boldsymbol{d}} \circ S \circ \Gamma_{\boldsymbol{d}}^{-1}$. B is a N–dimensional bistochastic matrix, since composition of stochastic matrices is stochastic and $(\Gamma_{\boldsymbol{d}} \circ S \circ \Gamma_{\boldsymbol{d}}^{-1})(\frac{1}{N}\mathbf{1}) = (\Gamma_{\boldsymbol{d}} \circ S)(\boldsymbol{d}) = \Gamma_{\boldsymbol{d}}(\boldsymbol{d}) = \frac{1}{N}\mathbf{1}$. Then, B maps $\Gamma_{\boldsymbol{d}}(\boldsymbol{y})$ to $\Gamma_{\boldsymbol{d}}(\boldsymbol{x})$. Conversely, given B, let $S = \Gamma_{\boldsymbol{d}}^{-1} \circ B \circ \Gamma_{\boldsymbol{d}}$. Similarly, S is a $\boldsymbol{d}$–stochastic matrix that maps $\boldsymbol{y}$ to $\boldsymbol{x}$.

$2 \leftrightarrow 3$, $2 \leftrightarrow 4$, $2 \leftrightarrow 5$ These three statements are equivalent to statements (iii), (iv), (v) in Proposition 17 respectively, for the embedded vectors $\Gamma_{\boldsymbol{d}}(\boldsymbol{x}), \Gamma_{\boldsymbol{d}}(\boldsymbol{y})$, which becomes apparent by rewriting

$$\sum_{i=1}^{n} |x_i - td_i| = \sum_{i=1}^{n} d_i \left| \frac{x_i}{d_i} - t \right| = \sum_{i=1}^{N} |\Gamma_{\boldsymbol{d}}(\boldsymbol{x})_i - t|,$$

$$\sum_{i=1}^{n} (x_i - td_i)^+ = \sum_{i=1}^{N} (\Gamma_{\boldsymbol{d}}(\boldsymbol{x})_i - t)^+,$$

$$L_{\boldsymbol{x}|\boldsymbol{d}}(k) = L_{\Gamma_{\boldsymbol{d}}(\boldsymbol{x})}(k'), \text{with } k = 1, \ldots, n \text{ and } k' = 1, \ldots, N,$$

and similarly for the right hand sides of the inequalities. □

Appendix C: Technical properties of σ–fragments

In this section, we provide some useful technical properties of general σ–fragments.

Proposition 20. *Let $\mathcal{R}_\sigma = (\mathcal{O}_\sigma, \mathcal{F}_\sigma)$ be a σ–fragment of magic theory $\mathcal{R} = (\mathcal{O}, \mathcal{F})$. The following statements hold:*

1. *No σ–fragment is empty.*
2. *If a free operation leaves two states invariant, then it also leaves their mixtures invariant,*

$$\mathcal{O}_\sigma \cap \mathcal{O}_{\sigma'} \subseteq \mathcal{O}_{p\sigma+(1-p)\sigma'} \text{ for any } p \in [0,1].$$

Proof.

1. The identity channel $1_{\mathrm{C}} : \mathcal{D} \mapsto \mathcal{D}$ belongs to every σ–fragment, as $1_{\mathrm{C}} \in \mathcal{O}$ and $1_{\mathrm{C}}\sigma = \sigma$ for all $\sigma \in \mathcal{F}$.
2. Let $\mathcal{E} \in \mathcal{O}_\sigma \cap \mathcal{O}_{\sigma'}$. Then $\mathcal{E} \in$ CPTP and corresponds to stochastic Wigner distribution $\mathrm{W}_{\mathcal{E}}$ such that $\mathrm{W}_{\mathcal{E}}\mathrm{W}_\sigma = \mathrm{W}_\sigma$ and $\mathrm{W}_{\mathcal{E}}\mathrm{W}_{\sigma'} = \mathrm{W}_{\sigma'}$. Then, $\mathrm{W}_{\mathcal{E}}\mathrm{W}_{p\sigma+(1-p)\sigma'} = \mathrm{W}_{p\sigma+(1-p)\sigma'}$ for any $p \in [0,1]$ due to the additive property 15 of the Wigner distribution, implying that state $p\sigma + (1-p)\sigma'$ is also left invariant by $\mathcal{E}$. □

Any free state $\sigma \in \mathcal{F}$ corresponds to a d^2–dimensional probability distribution W_σ and any free operation $\mathcal{E} \in \mathcal{O}$ corresponds to a $d^2 \times d^2$ stochastic matrix (or conditional probability distribution) $\mathrm{W}_{\mathcal{E}}$. Note that these mappings are one-to-one due to the orthogonality of the phase-point operators as an operator basis.

Note further that free states $\mathcal{F}$ are mapped onto a *strict subset* of the set of probability distributions. As a counterexample, the sharp d^2–dimensional probability distribution $(1, 0, \ldots, 0)$ does not correspond to any qudit Wigner distribution because of the boundedness condition in Proposition 15.

Similarly, not all stochastic matrices correspond to completely positive operations. As an example, consider the permutation matrix

$$\Pi_X = \begin{pmatrix} 0&1&0&0&0 \\ 0&0&0&0&1 \\ 0&0&0&1&0 \\ 1&0&0&0&0 \\ 0&0&1&0&0 \end{pmatrix} \otimes \begin{pmatrix} 0&0&1&0&0 \\ 0&0&0&0&1 \\ 0&0&0&1&0 \\ 1&0&0&0&0 \\ 0&1&0&0&0 \end{pmatrix} \in \mathrm{S}_5(\mathrm{W}_{\frac{1}{5}\mathbb{1}}). \quad \text{(C1)}$$

It preserves the uniform distribution $\mathrm{W}_{\frac{1}{5}\mathbb{1}}$, but it does not correspond to any positive (hence quantum) operation.

We now prove a result referenced in Section III C.

Proposition 21. *Let ρ, τ be two quantum states with Lorenz curves $\mathrm{L}_{\rho|\sigma}(x), \mathrm{L}_{\tau|\sigma}(x)$ in the σ–fragment.*

Let t be the number of elbows of $\mathrm{L}_{\tau|\sigma}(x)$ at locations $x_1, \ldots, x_t$.

Then, $\mathrm{L}_{\rho|\sigma}(x) \geq \mathrm{L}_{\tau|\sigma}(x)$ for all $x \in [0,1]$ iff $\mathrm{L}_{\rho|\sigma}(x_i) \geq \mathrm{L}_{\tau|\sigma}(x_i)$ for all $i = 1, \ldots, t$.

Proof. $\mathrm{L}_{\rho|\sigma}(x) \geq \mathrm{L}_{\tau|\sigma}(x)$ for all $x \in [0,1]$ trivially implies $\mathrm{L}_{\rho|\sigma}(x_i) \geq \mathrm{L}_{\tau|\sigma}(x_i)$ for all $i = 1, \ldots, n'$.

Conversely, assume that $\mathrm{L}_{\rho|\sigma}(x_i) \geq \mathrm{L}_{\tau|\sigma}(x_i)$ for all $i = 1, \ldots, r$. First, let $x_0 = 0$ and $x_{n'+1} = 1$, so that $\mathrm{L}_{\rho|\sigma}(x_0) = \mathrm{L}_{\tau|\sigma}(x_0) = 0$ and $\mathrm{L}_{\rho|\sigma}(x_{n'+1}) = \mathrm{L}_{\tau|\sigma}(x_{n'+1}) = 1$. 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 $\ell_\tau(x)$ connecting points $(x_i, \mathrm{L}_{\tau|\sigma}(x_i))$ and $(x_{i+1}, \mathrm{L}_{\tau|\sigma}(x_{i+1}))$ as well as the line segment $\ell_\rho(x)$ connecting points $(x_i, \mathrm{L}_{\rho|\sigma}(x_i))$ and $(x_{i+1}, \mathrm{L}_{\rho|\sigma}(x_{i+1}))$. This is illustrated in Fig. (8).

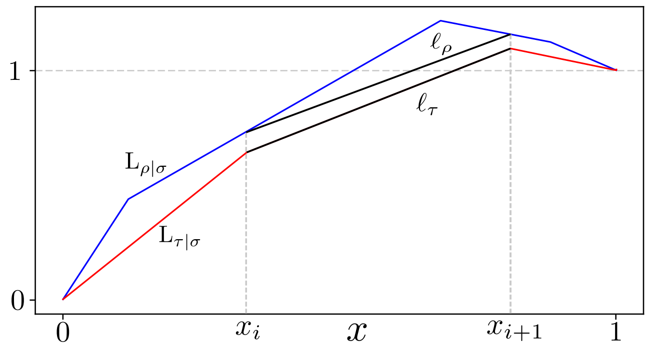

FIG. 8. **Illustration of Proposition 21.**

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

This theorem is of practical importance in calculating the necessary distillation constraints derived via majorization in σ–fragments.

Appendix D: Lorenz curves in the unital fragment

1. Binomial distributions and error bounds

Consider an experiment consisting of n trials of throwing a p–coin, that is a coin with probability p of landing on one side and $1-p$ of landing on the other. We express the sum over an even number of successful trials Φ_+ and the sum over an odd number of successful trials Φ_-,

$$\Phi_+(a;n,p) := \sum_{\ell=0}^{a/2} \binom{n}{2\ell} p^{2\ell}(1-p)^{n-2\ell},$$
$$\text{for even integers } m \in [0,n], \tag{D1}$$
$$\Phi_-(a;n,p) := \sum_{\ell=1}^{(a-1)/2} \binom{n}{2\ell+1} p^{2\ell+1}(1-p)^{n-(2\ell+1)},$$
$$\text{for odd integers } m \in [0,n]. \tag{D2}$$

Note that index m only takes even (odd) values when labelling Φ_+ (Φ_-). In Appendix D 2, we will use Φ_+ and Φ_- to express the elbow coordinates of Lorenz curves in the unital fragment.

We also define the classical entropy of a p–coin as well as the classical relative entropy between a p–coin and a q–coin,

$$S(p) := -p\log p - (1-p)\log(1-p), \tag{D3}$$
$$S(p||q) := p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}. \tag{D4}$$

They are symmetric in the sense that $S(p) = S(1-p)$ and $S(p||q) = S(1-p||1-q)$.

A useful result is the entropic bound on a combination [40].

Lemma 22. *For all $\ell \in [1, np-1]$,*

$$\left[8\ell\left(1-\frac{\ell}{np}\right)\right]^{-\frac{1}{2}} 2^{nS\left(\frac{\ell}{np}\right)} \leq \binom{np}{\ell} \leq \tag{D5}$$
$$\left[2\pi\ell\left(1-\frac{\ell}{np}\right)\right]^{-\frac{1}{2}} 2^{nS\left(\frac{\ell}{np}\right)}. \tag{D6}$$

Proof. For $\ell = 1, 2, np-1, np-2$ check by direct calculation. For all other cases, use Stirling's approximation. [CITE] □

With the help of this lemma, we directly arrive at

Theorem 23. *Given fixed $n > 0$ and p, Φ_+, Φ_- satisfy the following bounds:*

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

Proof. All four statements follow from application of Lemma 22 on the combinatorial coefficient and the defintion of relative entropy given in Eq. (D4) □

2. Strange state Lorenz curve elbow coordinates in the unital fragment

The Wigner distribution of the n–copy qutrit maximally mixed state $(\mathbb{1}/3)^{\otimes n}$ is the uniform distribution

$$\mathrm{W}_{(\mathbb{1}/3)^{\otimes n}} = \Bigg(\overbrace{\frac{1}{9^n}, \dots, \frac{1}{9^n}}^{9^n}\Bigg). \tag{D7}$$

The Wigner distribution of the 1-copy ϵ–noisy Strange state $\rho_{\mathrm{S}}(\epsilon)$ in the unital fragment is a permutation of

$$\mathrm{W}_{\rho_{\mathrm{S}}(\epsilon)} = \Bigg(\overbrace{\frac{1}{6} - \frac{1}{18}\epsilon, \dots, \frac{1}{6} - \frac{1}{18}\epsilon}^{8}, \overbrace{-\frac{1}{3} + \frac{4}{9}\epsilon}^{1}\Bigg) \tag{D8}$$

The two distinct components are plotted Fig. (9) as a function of noise. It is clear that in the unital fragment the Strange state contains Wigner negativities in the regime $0 \leq \epsilon \leq \frac{3}{4}$.

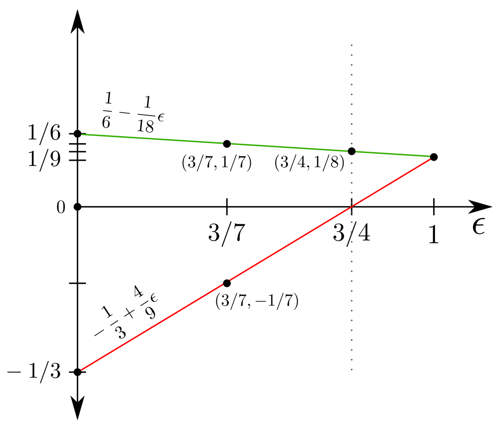

FIG. 9. **Wigner components of the noisy Strange state.** In the interval $0 \leq \epsilon < 3/7$, the negative component is larger than the positive components. At $\epsilon = 3/7$ the Wigner distribution is $(-\frac{1}{7}, \frac{1}{7}, \dots, \frac{1}{7})$. In the interval $3/7 < \epsilon < 3/4$, the positive components are larger than the negative component. In the interval $3/4 \leq \epsilon \leq 1$, there is no negativity.

The Wigner distribution of the n–copy ϵ–noisy Strange state $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ in the unital fragment is given by the convolution $\mathrm{W}_{\rho_{\mathrm{S}}(\epsilon)^{\otimes n}} = \mathrm{W}^{\otimes \mathrm{n}}_{\rho_{\mathrm{S}}(\epsilon)}$. In general, $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ contains $n+1$ distinct components, labelled $0, \dots, n$. We present the distinct Wigner components of $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ along with their multiplicites in Table I. Note that LHS (RHS) refers to elbow coordinates i on the left of and including (right of) the maximum, precisely

$$\text{LHS: } 0 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor; \tag{D9}$$

$$\text{RHS: } \left\lfloor \frac{n}{2} \right\rfloor + 1 \leq i \leq n. \tag{D10}$$

Case			$m_i(n,\epsilon)$	$w_i(n,\epsilon)$
$0 \leq \epsilon < \frac{3}{7}$	n even	LHS	$8^{2i}\binom{n}{2i}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{2i}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{n-2i}$
		RHS	$8^{n-2i}\binom{n}{2i}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{n-2i}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{2i}$
	n odd	LHS	$8^{2i+1}\binom{n}{2i+1}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{2i+1}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{n-2i-1}$
		RHS	$8^{n-2i-1}\binom{n}{2i+1}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{n-2i-1}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{2i+1}$
$\frac{3}{7} \leq \epsilon < \frac{3}{4}$	n even	LHS	$8^{n-2i}\binom{n}{2i}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{n-2i}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{2i}$
		RHS	$8^{2i}\binom{n}{2i}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{2i}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{n-2i}$
	n odd	LHS	$8^{n-2i}\binom{n}{2i}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{n-2i}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{2i}$
		RHS	$8^{2i}\binom{n}{2i}$	$\left(\frac{1}{6} - \frac{1}{18}\epsilon\right)^{2i}\left(-\frac{1}{3} + \frac{4}{9}\epsilon\right)^{n-2i}$

TABLE I. Wigner components $w_i(n,\epsilon)$ of $\rho_{\mathrm{S}}(\epsilon)^{\otimes n}$ along with their multiplicities $m_i(n,\epsilon)$ in decreasing order in i, $0 \leq i \leq n$. The order changes depending on the noise level ϵ, the parity of the number of copies n and the parity of the components (LHS vs RHS). Multiplication $2i$ is considered modulo $(n+1)$.

For example, the distribution of state $\rho_{\mathrm{S}}(0)^{\otimes 2}$ is

$$\mathrm{W}_{\rho_{\mathrm{S}}(0)^{\otimes 2}} = \Bigg(\overbrace{\left(-\frac{1}{3}\right)^2}^{1}, \overbrace{\left(\frac{1}{6}\right)^2, \dots, \left(\frac{1}{6}\right)^2}^{64}, \overbrace{-\frac{1}{3}\cdot\frac{1}{6}, \dots, -\frac{1}{3}\cdot\frac{1}{6}}^{16}\Bigg)$$

Every standard Lorenz curve contains n elbows, labelled by

$$\{(x_i, L_i)\}_{i=-1,0,\dots,n} \ ,$$

where the boundary points $(x_{-1}, L_{-1}) = (0,0)$ and $(x_n, L_n) = (1,1)$ are also included. The maximum is the $(\lfloor n/2 \rfloor)$-th elbow and its coordinates are calculated by collecting all the positive Wigner components,

$$x_{\lfloor n/2 \rfloor} = \frac{1}{2}\left(1 + \left(\frac{7}{9}\right)^n\right), \tag{D11}$$

$$L_{\lfloor n/2 \rfloor} = \frac{1}{2}\left(1 + \left(\frac{15 - 8\epsilon}{9}\right)^n\right). \tag{D12}$$

Expressions for all the elbow coordinates follow from summing up the Wigner components in decreasing order and we present the elbow coordinates for standard Lorenz curves in Table II.

3. Standard Lorenz curve coordinates

We can get explicit expressions for all 9^n points of the standard Lorenz curve $\mathrm{L}_{\rho_{\mathrm{S}}(\epsilon)^{\otimes n}|\mathbb{1}/3}$, in terms of the elbow

Case			x_i	L_i
$0 \le \epsilon < \frac{3}{7}$	n even	LHS	$\Phi_+\left(2i; n, \frac{8}{9}\right)$	$\left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_+\left(2i; n, \frac{12-4\epsilon}{15-8\epsilon}\right)$
		RHS	$x_{\lfloor n/2 \rfloor} + \Phi_-\left(2i; n, \frac{1}{9}\right)$	$L_{\lfloor n/2 \rfloor} - \left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_-\left(2i; n, \frac{3-4\epsilon}{15-8\epsilon}\right)$
	n odd	LHS	$\Phi_-\left(2i; n, \frac{8}{9}\right)$	$\left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_-\left(2i; n, \frac{12-4\epsilon}{15-8\epsilon}\right)$
		RHS	$x_{\lfloor n/2 \rfloor} + \Phi_-\left(2i; n, \frac{1}{9}\right)$	$L_{\lfloor n/2 \rfloor} - \left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_-\left(2i; n, \frac{3-4\epsilon}{15-8\epsilon}\right)$
$\frac{3}{7} \le \epsilon < \frac{3}{4}$	n even	LHS	$\Phi_+\left(2i; n, \frac{1}{9}\right)$	$\left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_+\left(2i; n, \frac{3-4\epsilon}{15-8\epsilon}\right)$
		RHS	$x_{\lfloor n/2 \rfloor} + \Phi_-\left(2i; n, \frac{8}{9}\right)$	$L_{\lfloor n/2 \rfloor} - \left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_-\left(2i; n, \frac{12-4\epsilon}{15-8\epsilon}\right)$
	n odd	LHS	$\Phi_+\left(2i; n, \frac{1}{9}\right)$	$\left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_+\left(2i; n, \frac{3-4\epsilon}{15-8\epsilon}\right)$
		RHS	$x_{\lfloor n/2 \rfloor} + \Phi_+\left(2i; n, \frac{8}{9}\right)$	$L_{\lfloor n/2 \rfloor} - \left(\frac{5}{3} - \frac{8}{9}\epsilon\right)^n \Phi_+\left(2i; n, \frac{12-4\epsilon}{15-8\epsilon}\right)$

TABLE II. Standard Lorenz curves elbow coordinates. The expresion depends on the noise level ϵ, the parity of the number of copies n and the location of the elbow relative to the maximum (LHS vs RHS). Multiplication $2i$ is considered modulo $(n+1)$. Note that the Lorenz curve boundary points are $(x_{-1}, L_{-1}) := (0,0)$ and $(x_n, L_n) = (1,1)$.

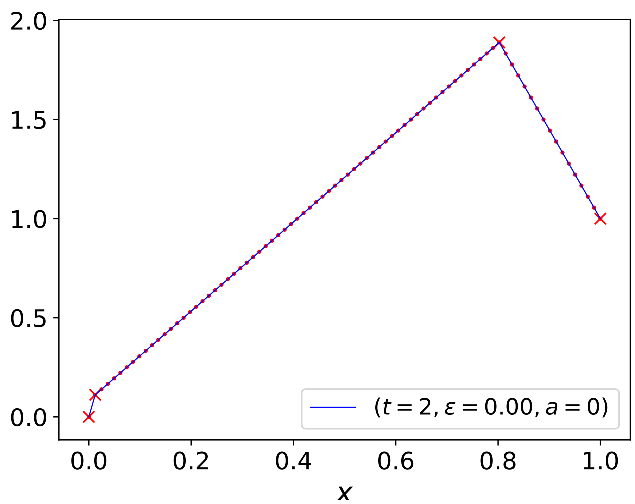

FIG. 10. **All points on the Lorenz curve are uniformly distributed.** Eqs. (D13) and (D14)) capture the coordinates of all points up to the maximum.

coordinates:

$$x_{ij} = \left(1 - \frac{j}{m_i}\right) x_{i-1} + \frac{j}{m_i} x_i, \tag{D13}$$

$$L_{ij} = \left(1 - \frac{j}{m_i}\right) L_{i-1} + \frac{j}{m_i} L_i \tag{D14}$$

for $j = 1, \ldots, m_i$ and $i = 0, \ldots, n$, where multiplicities $m_i = m_i(n, \epsilon)$ are given in Table I.

For visualisation purposes, we plot the Lorenz curve of state $\rho_S(0)^{\otimes 2}$.

Consider the state

$$\rho_S(n', \epsilon, n-n') := \rho_S(\epsilon)^{\otimes n'} \otimes \left(\frac{1}{3}\mathbb{1}\right)^{\otimes(n-n')},$$

Tensoring with the maximally mixed state keeps the Lorenz curve unchanged, but increases the resolution of (the uniformly distributed) points. The new point coordinates are given by:

$$x_{ijk} = (1 - p_{ijk})\, x_{i-1} + p_{ijk} x_i \tag{D15}$$

$$L_{ijk} = (1 - p_{ijk})\, L_{i-1} + p_{ijk} L_i, \tag{D16}$$

where $p_{ijk} = \dfrac{k + (j-1)9^{n-n'}}{9^{n-n'} m_i}$

for $i = 0, \ldots, n'$, $j = 1, \ldots, m_i(n', \epsilon')$ and $k = 1, \ldots, 9^{n-n'}$.

We can unify the indices, by introducing a single index

$$I(i,j,k) := k + \left[(j-1) + \sum_{\ell=0}^{i-1} m_\ell(n', \epsilon')\right] 9^{n-n'}, \tag{D17}$$

so that $I = 1, 2, \ldots, 9^n$. The elbow coordinates correspond to

$$I(i, m_i(n', \epsilon'), 9^{n-n'}) = \sum_{\ell=0}^{i} m_\ell(n', \epsilon'),\ i = 0, \ldots, n'. \tag{D18}$$

The index function I is bijective, i.e.

$$(i,j,k) = (i',j',k') \Leftrightarrow I(i,j,k) = I(i',j',k'). \tag{D19}$$

4. Strange state MSD in the unital fragment

Consider the Strange state MSD process in the unital fragment,

$$\rho_S(n', \epsilon, 0) \xrightarrow{\mathcal{O}_{1/3}} \rho_S(n', \epsilon', n-n'). \tag{D20}$$

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

$$I(i,j,k=1) = j + \sum_{\ell=0}^{i-1} m_\ell(n, \epsilon), \tag{D21}$$

$$I'(i',j',k') = k' + \left[(j'-1) + \sum_{\ell=0}^{i'-1} m_\ell(n', \epsilon')\right] 9^{n-n'}. \tag{D22}$$

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 Proposition 21 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 $i' = 0, \ldots, n'$,

$$j + \sum_{\ell=0}^{i-1} m_\ell(n, \epsilon) = \sum_{\ell=0}^{i'} m_\ell(n', \epsilon'). \tag{D23}$$

As a basic example, consider the process

$$\rho_{\mathrm{S}}(\epsilon)^{\otimes 4} \xrightarrow{\mathcal{O}_{1/3}} \rho_{\mathrm{S}}(\epsilon')^{\otimes 2} \otimes \left(\frac{1}{3}\mathbb{1}\right)^{\otimes 2}. \tag{D24}$$

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:

$$L_{I(i,j)} \geq L'_{I'(0,1,81)} \tag{D25}$$

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 Eq. (D16) 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.

Appendix E: Technical details of bound derivation in stabiliser fragments

1. Component-multiplicity pairs

In general, a 1–copy d–dimensional state ρ is defined exactly by its d^2–dimensional Wigner distribution W_ρ. The distribution W_ρ is usually defined on the phase space, but it can be convenient to define it using pure vector notation. In particular, we introduce a component vector $\boldsymbol{w}(\rho) = (w_i)_{i=1,\dots,D}$ and a multiplicity vector $\boldsymbol{m}(\rho) = (m_i)_{i=1,\dots,D}$, where $D \leq d^2$ which together form a set of component-multiplicity pairs $\{(w_i, m_i)\}_{i=1,\dots,D}$.

Definition 24. *Consider a distribution W and a positive integer $D \leq \dim W$. We call the set of ordered pairs $\{(w_i, m_i)\}_{i=1,\dots,D}$ a* complete set of component-multiplicity pairs, *if W contains m_i components w_i and $\sum_{i=0}^{D} m_i = d^2$.*

Therefore, such a set describes each component of W_ρ exactly once. As an example, two complete sets of pairs for the Strange state are $\{(-1/3, 1), (1/6, 8)\}$ and $\{(-1/3, 1), (1/6, 2), (1/6, 3), (1/6, 3)\}$.

Consider two states ρ_A, ρ_B with Wigner distributions $\mathrm{W}_{\rho_A}, \mathrm{W}_{\rho_B}$ described respectively by complete sets of component-multiplicity pairs

$$\{(w_i, m_i)\}_{i=1,\dots,D_A} \text{ and } \{(w'_j, m'_j)\}_{j=0,\dots,D_B}. \tag{E1}$$

The multiplicative property of the Wigner distribution over a composite phase space $\mathcal{P}_{d_A} \times \mathcal{P}_{d_B}$ shown in Proposition 15,

$$\mathrm{W}_{\rho_A \otimes \rho_B}(\boldsymbol{x}_A \oplus \boldsymbol{x}_B) = \mathrm{W}_{\rho_A}(\boldsymbol{x}_A)\mathrm{W}_{\rho_B}(\boldsymbol{x}_B), \tag{E2}$$

implies that the distribution $\mathrm{W}_{\rho_A \otimes \rho_B}$ is $d_A^2 d_B^2$–dimensional and contains components of the form $w_i w'_j$. Therefore, the set $\{(w_i w'_j, m_i m'_j)\}$ with $i = 1, \dots, D_A$ and $j = 1, \dots, D_B$ is a complete set of component-multiplicity pairs for the distribution of the composite system $\mathrm{W}_{\rho_A \otimes \rho_B}$. This is true because all components are of the form $w_i w'_j$ and

$$\sum_{i=1}^{D_A} \sum_{j=1}^{D_B} m_i m'_j = \sum_{i=1}^{D_A} m_i \sum_{j=1}^{D_B} m'_j = d_A^2 d_B^2.$$

Note that the rescaled distribution is also multiplicative,

$$\widetilde{\mathrm{W}}_{\rho_A \otimes \rho_B | \gamma_A \otimes \gamma_B}(\boldsymbol{x}_A \oplus \boldsymbol{x}_B) = \frac{\mathrm{W}_{\rho_A \otimes \rho_B}(\boldsymbol{x}_A \oplus \boldsymbol{x}_B)}{\mathrm{W}_{\gamma_A \otimes \gamma_B}(\boldsymbol{x}_A \oplus \boldsymbol{x}_B)} = \frac{\mathrm{W}_{\rho_A}(\boldsymbol{x}_A)\mathrm{W}_{\rho_B}(\boldsymbol{x}_B)}{\mathrm{W}_{\gamma_A}(\boldsymbol{x}_A)\mathrm{W}_{\gamma_B}(\boldsymbol{x}_B)} = \widetilde{\mathrm{W}}_{\rho_A|\gamma_A}(\boldsymbol{x}_A)\widetilde{\mathrm{W}}_{\rho_B|\gamma_B}(\boldsymbol{x}_B), \tag{E3}$$

so a complete set of component-multiplicity pairs can be obtained for this distribution in the same fashion as for usual Wigner distributions.

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 $\mathrm{W}_\rho{}^{\otimes n}$.

Lemma 25. *Let W be a distribution defined by a complete set of component-multiplicity pairs $\{(w_i, m_i)\}_{i=1,\dots,D}$ with $D \leq \dim W$ and consider the distribution $W^{\otimes n}$ obtained by taking the 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.

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

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

$$M_{\boldsymbol{k}} = \frac{n!}{k_1! \dots k_D!} \prod_{i=1}^{D} m_i{}^{k_i}, \tag{E4}$$

$$W_{\boldsymbol{k}} = \prod_{i=1}^{D} w_i{}^{k_i}. \tag{E5}$$

Proof. We proceed by induction.

Assume $n = 1$. Let $\boldsymbol{k}_i$ be the vector with its i-th component equal to 1 and 0's elsewhere. The set C_D^1 consists of all vectors of this form, i.e.

$$C_D^1 = \{\boldsymbol{k}_i\}_{i=1,\dots,D}$$

It is also true by direct calculation that

$$(W_{\boldsymbol{k}_i}, M_{\boldsymbol{k}_i}) = (w_i, m_i).$$

Therefore, $\{(W_{\boldsymbol{k}}, M_{\boldsymbol{k}})\}_{\boldsymbol{k}\in C_D^1}$ is a complete set of component-multiplicity pairs for W.

Assume that $\{(W_{\boldsymbol{k}}, M_{\boldsymbol{k}})\}_{\boldsymbol{k}\in C_D^n}$ as given in Eqs. (E4) and (E5) is a complete set of component-multiplicity pairs for the n–copy distribution $W^{\otimes n}$. By construction, the distribution $W^{\otimes(n+1)} = W^{\otimes n}\otimes W$ is multiplicative, so it admits the complete set of component multiplicity pairs

$$\{(W_{\boldsymbol{k}} w_i, M_{\boldsymbol{k}} m_i)\},\ \boldsymbol{k}\in C_D^n \text{ and } i = 1,\ldots,D. \tag{E6}$$

Consider the component sum of the distribution $W^{\otimes(n+1)}$,

$$\begin{aligned}
&\sum_{\boldsymbol{k}\in C_D^n}\sum_{i=1}^{D} M_{\boldsymbol{k}} m_i W_{\boldsymbol{k}} w_i = \sum_{\boldsymbol{k}\in C_D^n} M_{\boldsymbol{k}} W_{\boldsymbol{k}} \sum_{i=1}^{D} m_i w_i = \\
&\sum_{\boldsymbol{k}\in C_D^n} \frac{n!}{k_1!\ldots k_D!}\prod_{i=1}^{D} {m_i}^{k_i} {w_i}^{k_i} \sum_{i=1}^{D} m_i w_i = \\
&\left(\sum_{i=1}^{D} m_i w_i\right)^n \left(\sum_{i=1}^{D} m_i w_i\right) = \left(\sum_{i=1}^{D} m_i w_i\right)^{n+1} = \\
&\sum_{\boldsymbol{q}\in C_D^{n+1}} M_{\boldsymbol{q}} W_{\boldsymbol{q}},
\end{aligned}$$

where in the last expression, vectors $\boldsymbol{q} = (q_1,\ldots,q_D)$ have non-negative integer components that sum to $(n+1)$ and

$$M_{\boldsymbol{q}} = \frac{(n+1)!}{q_1!\ldots q_D!}\prod_{i=1}^{D} {m_i}^{q_i},$$
$$W_{\boldsymbol{q}} = \prod_{i=1}^{D} {w_i}^{q_i}.$$

We have used the multinomial theorem to proceed between lines 2-3 and lines 3-4 of the derivation.

We have achieved a regrouping of the distribution components. Every component $W_{\boldsymbol{q}}$ is of the form $W_{\boldsymbol{k}} w_i$ with $q_i = k_i + 1$ and $q_j = k_j$ for $j \neq i$ and

$$\sum_{\boldsymbol{q}\in C_D^{n+1}} M_{\boldsymbol{q}} = \sum_{\boldsymbol{q}\in C_D^{n+1}} \frac{(n+1)!}{q_1!\ldots q_D!}\prod_{i=1}^{D} {m_i}^{q_i} = \left(\sum_{i=1}^{D} m_i\right)^{n+1} = d^{n+1},$$

which is the dimension of $W^{\otimes(n+1)}$.

Therefore, $\{(W_{\boldsymbol{q}}, M_{\boldsymbol{q}})\}_{\boldsymbol{q}\in C_D^{n+1}}$ is a complete set of component-multiplicity pairs for $W^{\otimes n}$, completing the proof. □

Index vector $\boldsymbol{k}$ has $D-1$ independent components and in the proof of our main theorem in Section V we have $D = 4$, so we simplify the notation by writing the component and multiplicity vectors of the n–copy state distributions as $m_{ijk}, w(\rho_{\mathrm{S}})_{ijk}, w(\gamma)_{ijk}$ and $w(\rho_{\mathrm{S}}|\gamma)_{ijk}$, where i, j, k are the 3 independent index components.

2. First elbow location

Consider a magic state interconversion in a stabilizer σ–fragment, as in Eq. (54), 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 Section V.

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

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

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

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

3. (C2) → (C2). Similarly in this scenario,

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