

Constraints on magic state protocols from the statistical mechanics of Wigner negativity

Nikolaos Koukoulekidis^{1,*} and David Jennings^{1,2}

¹*Department of Physics, Imperial College London, London SW7 2AZ, UK*

²*School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK*

(Dated: February 24, 2022)

Magic states are key ingredients in schemes to realize 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 develop a statistical mechanical framework based on majorization to describe Wigner negative magic states for qudits of odd prime dimension processed under Clifford circuits. We show that majorization allows us to both quantify disorder in the Wigner representation and derive upper bounds for magic distillation. These bounds are shown to be tighter than other bounds, such as from mana and thauma, and can be used to incorporate hardware physics, such as temperature dependence and system Hamiltonians. We also show that a subset of single-shot Rényi entropies remain well-defined on quasi-distributions, are fully meaningful in terms of data processing and can acquire negative values that signal magic. We find that the mana of a magic state is the measure of divergence of these Rényi entropies as one approaches the Shannon entropy for Wigner distributions, and discuss how distillation lower bounds could be obtained in this setting. This use of majorization for quasi-distributions could find application in other studies of non-classicality, and raises non-trivial questions in the context of classical statistical mechanics.

* nk2314@imperial.ac.uk

INTRODUCTION

There has been rapid progress towards the goal of a fault-tolerant quantum computer [1–10]. Nevertheless, many challenges remain and there is increasing need for theory to take into account physical limitations of the hardware involved. The surface code [11–14] is a leading framework for fault-tolerance with very high error thresholds. Within this scheme, Clifford unitaries can be implemented in a robust, fault-tolerant way. However, due to the Eastin-Knill theorem [15], we also know that it is impossible to have a universal set of transversal gates, and although Clifford unitaries can be realized transversally [16, 17], one needs to find ways around the Eastin-Knill restriction. This can be achieved by injecting in quantum states, called magic states, which promote the Clifford group to universal quantum computing [18]. The obstacle to this is that these states are invariably noisy, so protocols involving stabilizer operations must be employed to purify many copies of the magic states and improve the overall performance of the induced quantum gates [18–20]. A key question then arises about the overhead on purifying many copies of a magic state into less noisy forms.

To address this, concrete distillation protocols have been developed, such as the Bravyi-Haah qubit protocol that provides a quadratic reduction in noise per cycle [21]. Such distillation rates have been improved in more recent works [19, 22–24] and there have been proposals of protocols for qudits of odd dimension [25–29] as well as of protocols within a full architectural analysis [30, 31]. There is also analysis of magic protocols from the perspective of magic monotones, which provide upper bounds on distillation rates and find application in the analysis of gate synthesis [32–36]. These frameworks for magic view magic states as resource states with respect to a natural class of quantum operations that are considered cheap, or free, such as stabilizer operations [37–40]. Beyond quantum computing, recent work has also looked at how magic can be used to analyse many-body physics [41] and conformal field theories [42] as well its connection to contextuality [43–48].

Work towards fault-tolerance is increasingly bridging the gap between abstract theory and experiment. Extensive work is being done on error mitigation [49–53], and the incorporation of hardware physics into the theoretical models [54–58]. For example, the XZZX code [59] is a variant of the surface code that incorporates noise bias explicitly and has been shown to attain the hashing bound of random codes [60]. In this work, we develop a framework to analyse magic state distillation protocols where explicit physics of the system is incorporated in the distillation bounds. This is achieved by considering how a given magic state protocol transforms a pair of quantum states – one being a noisy magic state, and the other a stabilizer state that is distinguished by the physics of the system (for example a state at a characteristic temperature, or the maximally mixed state) and acts as a reference state for the protocol. The magic distillation bounds can then be expressed in terms of the physics of the reference state (e.g. free energy changes in the case of temperature). Such bounds are of potential interest in assessing how physical features such as temperature, noise biases or fixed-point structure associated with restricted gate-sets constrain distillation protocols [61–67].

The approach we take is most closely aligned with resource theories of magic, although it differs in key ways. We obtain distillation upper bounds without the use of monotones, but instead use tools from statistical mechanics and recent work in single-shot resource theories [68–74]. Our analysis relies on the discrete Wigner representation of quantum systems, in which all states and operations can be described on a discrete phase space [46, 75]. Crucially, we focus on magic states with negativity in their Wigner representation, which is known to be a necessary condition for universality in the state-injection model [76–80]. However, taking a statistical mechanical perspective in this context raises a problem: the standard Boltzmann entropy is not defined for quasi-distributions. We circumvent this obstacle by making use of a more fundamental tool – majorization theory [81–83]. To our knowledge majorization of quasi-distributions has not been considered in quantum physics before, and therefore these methods could find application in other studies of non-classicality in quantum systems [84–89].

We begin by describing how stabilizer operations correspond to stochastic maps on a discrete phase space, and how majorization can be applied to constrain magic state transformations. We then consider families of magic protocols with increasing complexity. The simplest case we consider is for magic protocols that lead to unital channels. These are a subset of protocols that leave invariant some distinguished state or have some equilibrium fixed-point structure. We then consider protocols that are non-equilibrium processes, but only generate sub-linear correlations that enable a simple description in the thermodynamic limit before giving entropic analysis that can be applied generally.

The relations between the sets of protocols considered are as follows:

$$\begin{aligned} \text{Unital} &\subset \text{local fixed-points} \\ &\subset \text{sub-linear correlations} \subset \text{general protocols.} \end{aligned}$$

We provide explicit bounds for magic protocols that generate unital channels (Eq. (32)), as well as bounds that incorporate the temperature and Hamiltonian of the system (Theorem 7) and discuss extensions to more general scenarios. These bounds are shown to be tighter than other bounds such as those that come from mana and thauma (Fig. 1).

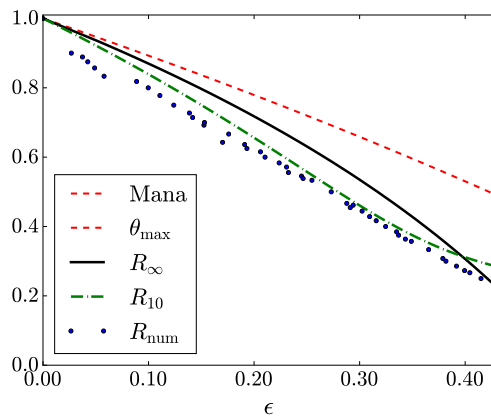


FIG. 1. **Magic distillation bounds from majorization of Wigner quasi-distributions.** The above plot shows distillation upper bounds for qutrits obtained from majorization on unital protocols. The constraints are plotted as a function of the depolarizing error ϵ for a noisy Strange state (Eq. (23)). The bound R_∞ follows from a Lorenz curve analysis presented before Eq. (32), while the bound R_{10} comes from single-shot entropies on the Wigner quasi-distributions, as discussed in the derivation of Eq. (52). Numerical bounds R_{num} are obtained by using the full set of majorization constraints. Bounds R_∞ , R_{10} , R_{num} all provide stricter constraints than mana [90] and max-thauma θ_{max} [91] bounds for noisy Strange states.

We find that the analysis in the presence of negativity displays a range of features that do not appear in classical statistical mechanics, and leads to a picture of Wigner negativity in a quantum circuit being described as non-classical free energy that is processed under stochastic dynamics. This is demonstrated by non-monotonic Lorenz curves (Fig. 3), and by single-shot entropies for general Wigner quasi-distributions that are well-defined, obey data-processing under stabilizer operations and can take on negative values (Eq. (45)).

We show that the mana of a quantum state arises as a residue in the α -parameter for Rényi entropies, and gives the rate of divergence to $-\infty$ as we approach the Shannon entropy of a Wigner distribution (Eq. (47)). Finally, in the Discussion section, we discuss how our approach could be extended to construct explicit lower bounds on distillation that take into account physical constraints.

RESULTS

Phase space representation of quantum mechanics

Central to our construction is the representation of any quantum state or quantum operation on a system of dimension d in terms of quasi-probability representations on a discrete phase space [75, 92]. This construction is a discrete version of Wigner representations in quantum optics [93–95].

Consider a d -dimensional quantum system with Hilbert space \mathcal{H}_d , and let $\{|0\rangle, |1\rangle, \dots, |d-1\rangle\}$ denote the standard computational basis, defined over $\mathbb{Z}_d = \{0, 1, \dots, d-1\}$. On this space, generalised Pauli matrices X, Z can be defined by their respective roles as shift and phase operators, acting on the basis states as follows,

$$X |k\rangle = |k+1\rangle \quad (1)$$

$$Z |k\rangle = \omega^k |k\rangle. \quad (2)$$

Here $\omega := e^{2\pi i/d}$ is the d -th root of unity and addition is taken modulo d . From these we can construct a phase space $\mathcal{P}_d := \mathbb{Z}_d \times \mathbb{Z}_d$ that provides a complete representation of the quantum system. Given a point $\mathbf{z} := (q, p)$ we define a displacement operator,

$$D_{\mathbf{z}} := \tau^{qp} X^q Z^p, \quad (3)$$

where the phase factor $\tau := -\omega^{1/2}$ ensures unitarity. We assume going forward that d is an odd prime [96]. The qubit case $d = 2$ is known to have obstacles to using a Wigner representation to quantify magic in terms of negativity [44, 97], although recent work [98] has shown how to extend the representation to include qubits by enlarging the phase space. For a composite system with dimension $d_{\text{tot}} = d_1 \dots d_n$ and Hilbert space $\mathcal{H}_{d_1} \otimes \dots \otimes \mathcal{H}_{d_n}$, we define displacement

operators as

$$D_{\mathbf{z}} := D_{(q_1, p_1)} \otimes \cdots \otimes D_{(q_n, p_n)}, \quad (4)$$

where now we have

$$\mathbf{z} := (q_1, \dots, q_n, p_1, \dots, p_n) \in \mathcal{P}_{d_1} \times \cdots \times \mathcal{P}_{d_n},$$

to denote the phase space point for the composite system. For simplicity, we assume n copies of a d -dimensional system $d_1 = d_2 = \cdots = d$, and therefore we have that $\mathbf{z} \in \mathbb{Z}_d^{2n}$.

The displacement operators form the Heisenberg-Weyl group [99, 100] under matrix multiplication modulo phases,

$$\text{HW}_d^n := \{\tau^k D_{\mathbf{z}} : k \in \mathbb{Z}_d, \mathbf{z} \in \mathcal{P}_d^n\}. \quad (5)$$

The Clifford group \mathcal{C}_d^n is then defined as the set of unitaries that normalise the Heisenberg-Weyl group [96]. We may define the pure stabilizer states as those states obtained by acting on $|0\rangle$ with Clifford unitaries [92] and, finally, the full set of stabilizer states as the convex hull of all pure stabilizers, namely all probabilistic mixtures of states of the form $U|0\rangle\langle 0|U^\dagger$ where U is Clifford.

Wigner representation for quantum states and quantum operations

In order to provide a complete decomposition of arbitrary quantum states and quantum operations we now define a basis of Hermitian observables that transforms covariantly under the action of the Clifford group. At every point $\mathbf{z} \in \mathcal{P}_d$ we define the phase-point operator

$$A_{\mathbf{z}} := \frac{1}{d} \sum_{\mathbf{y} \in \mathcal{P}_d} \omega^{\eta(\mathbf{z}, \mathbf{y})} D_{\mathbf{y}}, \quad (6)$$

where $\eta(\mathbf{z}, \mathbf{y})$ is the symplectic inner product between any two points $\mathbf{z}, \mathbf{y} \in \mathcal{P}_d$, and is given explicitly by

$$\eta(\mathbf{z}, \mathbf{y}) := \mathbf{y}^T \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \mathbf{z}, \quad (7)$$

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

The phase-point operators form an orthogonal operator basis with respect to the Hilbert-Schmidt inner product, as discussed in Supplementary Note 1. Therefore, any quantum state ρ can be expressed as a linear combination of them, $\rho = \sum_{\mathbf{z} \in \mathcal{P}_d} W_\rho(\mathbf{z}) A_{\mathbf{z}}$, where the coefficients $W_\rho(\mathbf{z})$ give the Wigner distribution of state ρ ,

$$W_\rho(\mathbf{z}) := \frac{1}{d} \text{tr}[A_{\mathbf{z}} \rho]. \quad (8)$$

For any quantum state ρ , the Wigner distribution $W_\rho(\mathbf{z})$ is readily seen to be a d^2 -dimensional quasi-distribution over \mathcal{P}_d . More precisely, $W_\rho(\mathbf{z})$ is a real-valued function on \mathcal{P}_d with the normalisation property $\sum_{\mathbf{z}} W_\rho(\mathbf{z}) = 1$ (see Supplementary Note 1 for details). Moreover, the above Wigner representation has recently been shown to be the only non-contextual quasi-probability representation of stabilizer theories in odd prime dimensions [101].

Any quantum channel \mathcal{E} also admits a Wigner representation. If \mathcal{E} maps some quantum system A of dimension d_A to a quantum system B of dimension d_B and $\mathcal{J}(\mathcal{E}) = (id \otimes \mathcal{E})(|\phi^+\rangle\langle\phi^+|)$, with $|\phi^+\rangle = d_A^{-1/2} \sum_k |kk\rangle$, is its associated Choi state [102], then we can define

$$W_{\mathcal{E}}(\mathbf{y}|\mathbf{z}) := d_A^2 W_{\mathcal{J}(\mathcal{E})}(\bar{\mathbf{z}} \oplus \mathbf{y}), \quad (9)$$

where $\bar{\mathbf{z}} = (q_1, \dots, q_n, -p_1, \dots, -p_n)$ can be viewed as the time-reversed version of \mathbf{z} in the discrete phase space, where momenta are reflected while position coordinates remain unchanged.

Magic theories for quantum computation

A number of magic theories exist, where magic states are computational resource states with respect to a class of quantum operations that are considered free [37]. One natural class of free operations are those obtained from

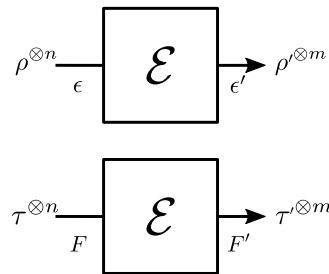


FIG. 2. **Probing magic protocols with reference states.** A magic protocol (top figure) converts n copies of some noisy magic state ρ with error parameter ϵ to m copies of some less noisy state ρ' with $\epsilon' < \epsilon$. The physics of the protocol can be analysed by considering how it would transform some distinguished reference stabilizer state $\tau^{\otimes n}$ (bottom figure). For the case where this reference state is a thermal state at some temperature T our magic distillation bounds are functions of T , the error parameters ϵ, ϵ' , and free energies F, F' . In addition, a quantity ϕ (Eq. (41)) relates the computational and energy bases, and corresponds to the degree to which the Wigner negativity of ρ can be given a sharp energy via the system Hamiltonian.

Clifford unitaries, Pauli measurements and the ability to discard quantum systems. However, there are also other candidates [38–40].

In any theory of magic, one route to bounding distillation rates is through a magic monotone. A magic monotone is a real-valued function of any quantum state $M(\rho)$ that is monotonically non-increasing under the free operations of the magic theory. More precisely $M(\sigma) \leq M(\rho)$ whenever it is possible to convert ρ into σ using free operations.

One prominent magic monotone is the mana of a state [90], defined as

$$\mathcal{M}(\rho) := \log(2\text{sn}(\rho) + 1), \quad (10)$$

where the sum-negativity [90] is the sum of the negative components in W_ρ ,

$$\text{sn}(\rho) := \sum_{\mathbf{z}: W_\rho(\mathbf{z}) < 0} |W_\rho(\mathbf{z})|. \quad (11)$$

Using the fact that $W_\rho(\mathbf{z})$ is a normalized quasi-distribution, we can also write $\mathcal{M}(\rho) = \sum_{\mathbf{z}} |W_\rho(\mathbf{z})|$.

Mana is an additive magic monotone, and the fact that it is non-increasing under free operations provides a constraint on magic state interconvertibility.

In this work we develop bounds that apply to any reasonable magic theory. The central idea is to apply majorization theory to the quasi-distributions that arise for magic states. In the next section we explain how majorization relates to the free operations in a theory of magic.

Stochastic representation of magic protocols

Within the Wigner representation for odd prime dimension, it is well-known that all positively represented states used in Clifford circuits admit an efficient classical simulation [77], so negativity is a necessary resource for universal fault-tolerant quantum computing [76]. Therefore, the free states in any magic theory are required to be a subset of

$$\mathcal{F} := \{\rho : W_\rho(\mathbf{z}) \geq 0 \text{ for all } \mathbf{z} \in \mathcal{P}_d\}. \quad (12)$$

Our focus is on states with negativity, so the particular choice of free states is not critical for our analysis. The remaining component that defines any magic theory is the set of free quantum operations. The most basic assumption we require on free operations is that they send any free state to another free state.

Any magic state protocol will correspond to a quantum channel \mathcal{E} , so from Eq. (9) it admits a Wigner representation $W_{\mathcal{E}}(\mathbf{y}|\mathbf{z})$ that acts as a transition matrix mapping phase space points $\mathbf{z} \rightarrow \mathbf{y}$. The representation obeys the relation

$$W_{\mathcal{E}(\rho)}(\mathbf{y}) = \sum_{\mathbf{z} \in \mathcal{P}_d} W_{\mathcal{E}}(\mathbf{y}|\mathbf{z}) W_\rho(\mathbf{z}), \quad (13)$$

for any ρ . Since the magic protocol sends free states to free states, \mathcal{E} sends all positively represented quantum state to other positively represented quantum states. In such cases, it can be shown [40] that if \mathcal{E} is a free operation then the associated Wigner representation $W_{\mathcal{E}}(\mathbf{y}|\mathbf{z})$ must form a stochastic matrix. In particular, all stabilizer operations correspond to stochastic matrices in the Wigner representation.

We note however, that not all stochastic maps on the phase space correspond to valid quantum operations. The reason is that the maps must also respect the symplectic structure of the phase space, which is an additional non-trivial constraint.

In what follows, we shall assume that we have a magic theory $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ in which the free states \mathcal{F} are represented by non-negative Wigner functions, while the free operations \mathcal{O} are stochastic maps in the Wigner representation. Our analysis of magic in this phase space setting makes use of majorization theory, which we describe in the next section.

Quantifying disorder without entropies

Majorisation [81, 103] is a collection of powerful tools that has found many applications in quantum information theory [74, 104–110]. It describes the disorder of distributions that undergo stochastic transformations, and in its simplest form defines a pre-order on probability distributions. Given two distributions $\mathbf{p} = (p_1, \dots, p_n)$ and $\mathbf{q} = (q_1, \dots, q_n)$ over n outcomes, we say that \mathbf{p} majorizes \mathbf{q} , denoted $\mathbf{p} \succ \mathbf{q}$, if there exists a bistochastic map $A = (A_{ij})$ such that $A\mathbf{p} = \mathbf{q}$, where bistochastic means that $A_{ij} \geq 0$ and $\sum_i A_{ij} = \sum_j A_{ij} = 1$. It can be shown [81] that the condition $\mathbf{p} \succ \mathbf{q}$ over probability distributions is equivalent to $n - 1$ inequalities, which can be checked efficiently.

There is a natural generalisation, which is called \mathbf{d} -majorization [82], or in the context of thermodynamics, thermo-majorization [69]. For a fixed probability distribution $\mathbf{r} = (r_1, \dots, r_n)$ with positive components, we define majorization relative to \mathbf{r} as $\mathbf{p} \succ_{\mathbf{r}} \mathbf{q}$, if and only if there exists a stochastic map A such that $A\mathbf{r} = \mathbf{r}$ and $A\mathbf{p} = \mathbf{q}$. The original majorization condition between probability distributions corresponds to the case $\mathbf{r} = (1/n, \dots, 1/n)$.

In fact, we can further generalize to relative majorization [83, 103, 111–114], defined as an ordering between pairs of vectors and write

$$(\mathbf{p}, \mathbf{r}) \succ (\mathbf{q}, \mathbf{r}') \quad (14)$$

if and only if there is a stochastic map A such that $A\mathbf{r} = \mathbf{r}'$ and $A\mathbf{p} = \mathbf{q}$. We retrieve \mathbf{d} -majorization when $\mathbf{r} = \mathbf{r}'$.

Quasi-probability majorization and non-monotonic Lorenz curves

Our analysis applies majorization to magic states at the level of the associated Wigner distributions. Since these are in general quasi-distributions, it is important to check how majorization is computed for these cases and what differences quasi-distributions bring over genuine probability distributions.

We also make use of the notion of a Lorenz curve of a vector $\mathbf{w} \in \mathbb{R}^n$ relative to some other vector $\mathbf{r} \in \mathbb{R}^n$. Given a vector \mathbf{w} we define \mathbf{w}^\downarrow to be the re-arrangement of the components of \mathbf{w} into decreasing order. Given two n -component vectors \mathbf{w} and \mathbf{r} , we first define $\tilde{\mathbf{w}} = (\tilde{w}_i)$, where $\tilde{w}_i := w_i/r_i$, as the vector of component-wise ratios between \mathbf{w} and \mathbf{r} . We can now define the Lorenz curve of \mathbf{w} relative to \mathbf{r} , denoted $L_{\mathbf{w}|\mathbf{r}}(x)$, as the piece-wise linear function that passes through $(0, 0)$ and the n points

$$(x_k, L_{\mathbf{w}|\mathbf{r}}(x_k)) = \left(\frac{1}{R} \sum_{i=1}^k r_{\pi(i)}, \sum_{i=1}^k w_{\pi(i)} \right), \quad (15)$$

where $R := \sum_{i=1}^n r_i$ and π is the permutation on n objects mapping $\tilde{\mathbf{w}}$ to $\tilde{\mathbf{w}}^\downarrow$. The form of this requires that \mathbf{r} has no zero components, which we shall assume without loss of generality as the rank of a quantum state is not operationally meaningful.

In the usual case where \mathbf{w} and \mathbf{r} are both probability distributions the Lorenz curve is defined on the interval $[0, 1]$, and rises monotonically until it reaches the value 1 at $x = 1$. The value $L_{\mathbf{w}|\mathbf{r}}(x) = 1$ is a global maximum. Moreover, if $\mathbf{w}, \mathbf{w}', \mathbf{r}, \mathbf{r}'$ are all valid probability distributions with \mathbf{r}, \mathbf{r}' having positive components, then

$$(\mathbf{w}, \mathbf{r}) \succ (\mathbf{w}', \mathbf{r}') \text{ if and only if } L_{\mathbf{w}|\mathbf{r}}(x) \geq L_{\mathbf{w}'|\mathbf{r}'}(x),$$

for all $x \in [0, 1]$. This result is shown in [111] and we reproduce it in Supplementary Note 2. It provides a simple way of computing whether relative majorization holds between pairs of probability distributions.

However, if \mathbf{w} is a quasi-probability distribution with negative values, and \mathbf{r} a regular probability distribution things are different. Now the Lorenz curve is no longer monotonically increasing, but is a concave function that breaks through the $L_{\mathbf{w}|\mathbf{r}}(x) = 1$ barrier at an interior point and attains some non-trivial maximum L_\star above the

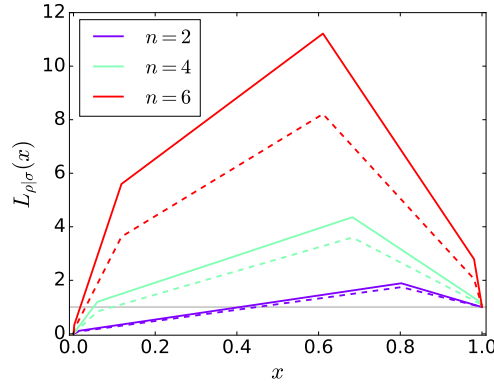


FIG. 3. **A ‘heretical’ family of Lorenz curves.** Traditionally, Lorenz curves for distributions are monotone increasing cumulative functions that reach a maximum value of 1. In contrast, Lorenz curves for magic states break through the value of $L(x) = 1$ due to the presence of negativity in the associated quasi-distributions. The above family of curves correspond to multiple copies of noisy Strange states $\rho := \rho_S(\epsilon)^{\otimes n}$ (Eq. (23)) for $n = 2, 4, 6$ within \mathcal{R}_σ , where $\sigma = \mathbb{1}/3$. Solid lines represent pure Strange states, while dashed lines represent ϵ -noisy Strange states with depolarising error $\epsilon = 0.1$.

value 1, before decreasing monotonically to $L_{\mathbf{w}|\mathbf{r}}(x) = 1$ at the end-point $x = 1$. See Fig. 3 for examples of non-monotonic Lorenz curves for quasi-distributions. Within quantum theory, the breaking of this barrier is associated with the degree of non-classicality.

Relative majorization is usually considered for probability distributions, so we need to verify that the same Lorenz curve conditions apply to quasi-distributions before proceeding with our analysis. This can be done from first principles, but a simpler way to see that this holds is to reduce to the problem of genuine distributions by first ‘masking’ the negativity in the quasi-distribution \mathbf{w} using the reference distribution \mathbf{r} and then applying the conditions for relative majorization. This negativity masking then gives the following result.

Theorem 1. *Let $\mathbf{w} \in \mathbb{R}^n, \mathbf{w}' \in \mathbb{R}^m$ be quasi-distributions, and let $\mathbf{r} \in \mathbb{R}^n, \mathbf{r}' \in \mathbb{R}^m$ be probability distributions with non-zero components. Then, $(\mathbf{w}, \mathbf{r}) \succ (\mathbf{w}', \mathbf{r}')$ if and only if $L_{\mathbf{w}|\mathbf{r}}(x) \geq L_{\mathbf{w}'|\mathbf{r}'}(x)$ for all $x \in [0, 1]$.*

Proof. Since the components of \mathbf{r} are strictly positive, there always exists an $\epsilon > 0$ such that $\mathbf{w}_\epsilon := \epsilon\mathbf{w} + (1 - \epsilon)\mathbf{r}$ is a genuine probability distribution. A similar result holds for \mathbf{w}' and \mathbf{r}' and we choose ϵ sufficiently small so that both \mathbf{w}_ϵ and \mathbf{w}'_ϵ are probability distributions. We now have that $(\mathbf{w}, \mathbf{r}) \succ (\mathbf{w}', \mathbf{r}')$ if and only if $(\mathbf{w}_\epsilon, \mathbf{r}) \succ (\mathbf{w}'_\epsilon, \mathbf{r}')$. This equivalence holds because there exists a stochastic map A such that $A\mathbf{w} = \mathbf{w}'$ and $A\mathbf{r} = \mathbf{r}'$ if and only if

$$A[\epsilon\mathbf{w} + (1 - \epsilon)\mathbf{r}] = \epsilon\mathbf{w}' + (1 - \epsilon)\mathbf{r}' \text{ and } A\mathbf{r} = \mathbf{r}'. \quad (16)$$

In terms of a Lorenz curve condition we have that $(\mathbf{w}_\epsilon, \mathbf{r}) \succ (\mathbf{w}'_\epsilon, \mathbf{r}')$ if and only if $L_{\mathbf{w}_\epsilon|\mathbf{r}}(x) \geq L_{\mathbf{w}'_\epsilon|\mathbf{r}'}(x)$ for all $x \in [0, 1]$. Additionally, we show in Supplementary Note 2 that the Lorenz curve for any quasi-distribution obeys the relation

$$L_{a\mathbf{w}+b\mathbf{r}|\mathbf{r}}(x) = aL_{\mathbf{w}|\mathbf{r}}(x) + bx, \quad (17)$$

for any $a > 0$ and $b \in \mathbb{R}$. This relation implies that $L_{\mathbf{w}_\epsilon|\mathbf{r}}(x) \geq L_{\mathbf{w}'_\epsilon|\mathbf{r}'}(x)$ if and only if $\epsilon L_{\mathbf{w}|\mathbf{r}}(x) + (1 - \epsilon)x \geq \epsilon L_{\mathbf{w}'|\mathbf{r}'}(x) + (1 - \epsilon)x$. Finally, the $(1 - \epsilon)x$ terms cancel on both sides and we get the required relative majorization conditions that $(\mathbf{w}, \mathbf{r}) \succ (\mathbf{w}', \mathbf{r}')$ if and only if $L_{\mathbf{w}|\mathbf{r}}(x) \geq L_{\mathbf{w}'|\mathbf{r}'}(x)$ for all $x \in [0, 1]$, as required. \square

With this theory in place, we turn to magic distillation protocols.

Magic distillation bounds with fixed-points

We can now consider how majorization constrains magic distillation rates starting in this section with an abstract setting based on the fixed-point structure of the physical operations accessible to us. We show that any magic theory can always be decomposed in terms of ‘sub-theories’ with a particular fixed-point structure, which allows the use of \mathbf{d} -majorization. Later, we drop the requirement of fixed-points and instead use the full relative majorization, before giving general entropic constraints.

We begin by defining the following sub-theory of any magic theory \mathcal{R} .

Definition 2. Given a theory of magic $\mathcal{R} = (\mathcal{F}, \mathcal{O})$, we define the sub-theory $\mathcal{R}_\sigma = (\mathcal{F}, \mathcal{O}_\sigma)$, in which the free operations are

$$\mathcal{O}_\sigma := \{\mathcal{E} \in \mathcal{O} : \mathcal{E}(\sigma) = \sigma\}, \quad (18)$$

namely those free operations that leave the distinguished state $\sigma \in \mathcal{F}$ invariant.

This gives a simple way to break up any theory into smaller, more manageable parts, and where the union over all sub-theories returns the parent theory, so we are not discarding any information by breaking up the theory in this way. We make this precise as follows.

Theorem 3. Let $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ be a theory of magic. Then a transformation $\rho_1 \rightarrow \rho_2$ is possible in \mathcal{R} if and only if the transformation $\rho_1 \rightarrow \rho_2$ is possible in at least one sub-theory \mathcal{R}_σ .

The proof of this is straightforward.

Proof. Suppose the interconversion is possible in a \mathcal{R}_σ via some $\mathcal{E} \in \mathcal{O}_\sigma$. But since $\mathcal{O}_\sigma \subseteq \mathcal{O}$ it is also possible in \mathcal{R} . Conversely, suppose the transformation is possible in \mathcal{R} via some $\mathcal{E} \in \mathcal{O}$. The free states \mathcal{F} are a closed, bounded set and moreover the image of \mathcal{F} under the map \mathcal{E} is in \mathcal{F} . By the Brouwer fixed-point theorem [115], this mapping must therefore have a fixed point $\sigma \in \mathcal{F}$, so $\mathcal{E} \in \mathcal{O}_\sigma$ and the interconversion is possible in \mathcal{R}_σ . \square

We can compare with resource monotones, which are also called resource measures. A complete set of measures $\{\mathcal{M}_i\}_i$ is such that $\rho \rightarrow \sigma$ if and only if $\mathcal{M}_i(\rho) \geq \mathcal{M}_i(\sigma)$ for all i . The above set of sub-theories can therefore be viewed as a complete set of ‘‘co-measures’’ for the theory, where the parent resource theory’s complex pre-order of states is mapped to a simpler pre-order for a particular \mathcal{R}_σ . In the next section we show that a given \mathcal{R}_σ can be approximated by a majorization pre-order, and from this we compute magic distillation bounds.

Majorisation of Wigner distributions with fixed-point structure

We consider some $\sigma \in \mathcal{F}$, and its corresponding \mathcal{R}_σ . We are interested in the ability to transform many copies of some noisy magic state ρ towards a more pure form of magic. The state ρ is assumed to have negativity in the Wigner representation, so $W_\rho(\mathbf{z}) < 0$ for some regions of $\mathbf{z} \in \mathcal{P}_d$. The state σ is assumed to have a Wigner distribution with $W_\sigma(\mathbf{z}) > 0$ for all \mathbf{z} in the phase space. This is justified because a non-full rank state σ can be handled as a limiting case in which we first add an infinitesimal fraction of depolarising noise $\epsilon(\mathbb{1}/d)$ and then take $\epsilon \rightarrow 0$.

The free operations within the magic theory \mathcal{R} are represented by stochastic maps, and within \mathcal{R}_σ by stochastic maps that leave $W_\sigma(\mathbf{z})$ invariant. Therefore, a necessary condition for magic state transformations $\rho_1 \rightarrow \rho_2$ within \mathcal{R}_σ will be that

$$W_{\rho_1} \succ_{W_\sigma} W_{\rho_2}, \quad (19)$$

or put another way, that the quasi-distribution W_{ρ_1} is more ordered than the quasi-distribution W_{ρ_2} relative to W_σ . To simplify notation, we denote by $L_{\rho|\sigma}(x)$ the Lorenz curve $L_{W_\rho|W_\sigma}(x)$, and therefore have that

$$\rho_1 \rightarrow \rho_2 \text{ within } \mathcal{R}_\sigma \text{ implies } L_{\rho_1|\sigma}(x) \geq L_{\rho_2|\sigma}(x), \quad (20)$$

for all $x \in [0, 1]$, which restricts the transformations that are possible.

Note that this is not a single numerical constraint, but a family of constraints. For n copies of a qudit system of dimension d the number of terms in W_ρ is d^{2n} , so imposing the Lorenz curve condition corresponds to exponentially many constraints.

Before computing explicit examples, we state some generic aspects of Lorenz curves for magic states, which allow us to interpret previous magic monotones as features of the curves. The first result gives a simple way to see that the sum-negativity/mana of a magic state is a monotone [90].

Theorem 4. Majorization in any \mathcal{R}_σ implies the monotonicity of sum-negativity/mana.

Proof. The sum-negativity of a magic state ρ can be written as $sn(\rho) = \frac{1}{2}(\sum_{\mathbf{z}} |W_\rho(\mathbf{z})| - 1)$. We make use of the L_1 -norm form of relative majorization, (see Supplementary Note 2), which states that $\mathbf{p} \succ_{\mathbf{r}} \mathbf{q}$ if and only if

$$\sum_k |p_k - r_k t| \geq \sum_k |q_k - r_k t|, \quad (21)$$

for all $t \in \mathbb{R}$. Choosing $t = 0$ we get the single condition that $\sum_k |p_k| \geq \sum_k |q_k|$, independent of the choice of \mathbf{r} . Applying this to the Wigner quasi-distributions of two quantum states immediately gives the result. \square

If we have a magic state ρ that has negativity in its Wigner representation, then, as discussed, its Lorenz curve over-shoots 1 and reaches a non-trivial maximum L_\star that depends on the particular state. There is a simple relation between L_\star and sum-negativity/mana, which is provided by the following result.

Theorem 5. *Given a magic state ρ , the maximum L_\star of its Lorenz curve $L_{\rho|\sigma}(x)$ is independent of the \mathcal{R}_σ and equal to $1 + \text{sn}(\rho)$. Moreover, the majorization constraint is stronger than mana in every \mathcal{R}_σ .*

The proof of this is given in Supplementary Note 2. Using the Lorenz curve perspective, it is also simple to construct magic monotones. For example, within a given \mathcal{R}_σ , the area above the horizontal line $y = 1$ is a magic monotone.

Theorem 6. *Given a magic state ρ and a free state σ , let $\mathcal{A}_\sigma(\rho)$ be the area of the region $\{(x, y) : 1 \leq y \leq L_{\rho|\sigma}(x)\}$. Then $\mathcal{A}_\sigma(\rho)$ is a magic monotone for \mathcal{R}_σ .*

Proof. Consider the transformation $\rho_1 \rightarrow \rho_2$ within \mathcal{R}_σ . We have that $L_{\rho_1|\sigma}(x)$ is never below the curve $L_{\rho_2|\sigma}(x)$, and therefore the region above $L = 1$ for ρ_2 is a subset of the corresponding region for ρ_1 . Thus, $\mathcal{A}_\sigma(\rho_1) \geq \mathcal{A}_\sigma(\rho_2)$, so \mathcal{A}_σ is a magic monotone in \mathcal{R}_σ . \square

Note though, in contrast to mana, the area monotone is specific to \mathcal{R}_σ , and its value will vary as we change σ . Therefore its monotonicity depends on the physics of the fixed point and provides a means to analyse magic distillation only under free operations that leave σ invariant. One can also define an area monotone in the more general setting of relative majorization, where one drops the fixed point emphasis.

Magic distillation bounds for unital protocols

We now construct magic state distillation bounds for protocols that generate unital channels, meaning that $\mathcal{E}(\mathbb{1}/d) = \mathbb{1}/d$.

Our approach works for any odd dimension, but for simplicity we consider qutrit magic states ($d = 3$), for which a canonical magic state exists with a simple Wigner distribution. This is the Strange state $|S\rangle$, which is given by

$$|S\rangle := \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle), \quad (22)$$

in the computational basis. Its distribution $W_{|S\rangle\langle S|}(\mathbf{z})$ has a single negative value of $-1/3$ at $\mathbf{z} = (0, 0)$ and the positive value $1/6$ at all other points. We define the ϵ -noisy Strange state as

$$\rho_S(\epsilon) := (1 - \epsilon)|S\rangle\langle S| + \epsilon\frac{1}{3}\mathbb{1}, \quad (23)$$

where ϵ is the depolarizing error parameter. Any magic state ρ can be processed via Clifford operations [29, 116] and put into this canonical form for some $\epsilon \geq 0$.

The Wigner distribution of the single-copy, ϵ -noisy Strange state is given by

$$W_{\rho_S(\epsilon)}(\mathbf{z}) = (1 - \epsilon)W_{|S\rangle\langle S|}(\mathbf{z}) + \epsilon W_{\frac{1}{3}\mathbb{1}}(\mathbf{z}). \quad (24)$$

with a single negative component

$$-v(\epsilon) := -\left(\frac{1}{3} - \frac{4}{9}\epsilon\right) \quad (25)$$

located at the origin $\mathbf{z} = \mathbf{0}$ and positive components

$$u(\epsilon) := \frac{1}{6} - \frac{1}{18}\epsilon \quad (26)$$

at the 8 phase space points $\mathbf{z} \neq \mathbf{0}$. We assume that $\epsilon < 3/4$ to ensure the presence of negativity in the Wigner distribution.

We now consider the problem of purifying n copies of a noisy Strange state $\rho_S(\epsilon)^{\otimes n}$ into a smaller number of copies n' of a less noisy Strange state $\rho_S(\epsilon')^{\otimes n'}$, with $\epsilon' < \epsilon$ and $n' \leq n$. In particular, we compute the Lorenz curves for the transformation

$$\rho_S(\epsilon)^{\otimes n} \longrightarrow \rho_S(\epsilon')^{\otimes n'}, \quad (27)$$

and use them to bound the magic distillation rate $R(\epsilon, \epsilon') := n'/n$.

Due to negativity, the ordering of the components of the rescaled Wigner distribution $W_{\rho_n|\sigma_n}(\mathbf{z}) := W_{\rho_n}(\mathbf{z})/W_{\sigma_n}(\mathbf{z})$, for $\rho_n = \rho_S(\epsilon)^{\otimes n}$, and $\sigma_n = \sigma^{\otimes n} = (\mathbb{1}/3)^{\otimes n}$, depends on whether $-v(\epsilon)$ is raised to an even or odd power. Treating the Wigner distributions as vectors, the component values w_i and associated multiplicities m_i in the n -copy case are found to be

$$m_i = 8^i \binom{n}{i}, \quad (28)$$

$$w_i = u^i (-v)^{n-i}, \quad (29)$$

where index i runs through $0, \dots, n$, and we assume for simplicity that n is even and $\epsilon \leq 3/7$, which implies that $v \geq u$. All details on the analysis are provided in Supplementary Note 3.

The Lorenz curve $L_{\rho_n|\sigma_n}(x)$ reaches a maximum value of

$$\begin{aligned} L_\star &:= L_{\rho_n|\sigma_n}(x_\star) = \sum_{i=0}^{n/2} m_{2i} w_{2i} \\ &= \frac{1}{2} + \frac{1}{2} \left(\frac{15 - 8\epsilon}{9} \right)^n > 1, \end{aligned} \quad (30)$$

which occurs at $x = x_\star$ given by

$$x_\star = \frac{1}{2} + \frac{1}{2} \left(\frac{7}{9} \right)^n. \quad (31)$$

Distillation bounds can be obtained from any part of the Lorenz curve by imposing the constraint that the Lorenz curve of the input state never dips below the curve of the output. One simple constraint can be obtained by considering the initial slopes of the two Lorenz curves.

The coordinates of the first point after the origin for $\rho_S(\epsilon)^{\otimes n}$ are given by $(1/9^n, v(\epsilon)^n)$, while for $\rho_S(\epsilon')^{\otimes n'}$ by $(1/9^{n'}, v(\epsilon')^{n'})$, and by requiring that the initial slope for the input state Lorenz curve is larger than that of the output state (Proposition 12 in Supplementary Note 2) we find that

$$R(\epsilon, \epsilon') \leq R_\infty := \frac{\log(3 - 4\epsilon)}{\log(3 - 4\epsilon')}. \quad (32)$$

The choice of denoting the bound by R_∞ is explained in the derivation of Eq. (52) in relation to single-shot entropies. For the limiting case of pure magic states on the output ($\epsilon' = 0$), this simplifies to

$$R \leq 1 + \frac{\log(1 - \frac{4}{3}\epsilon)}{\log 3}. \quad (33)$$

We compare this to other known distillation bounds coming from mana and thauma. The mana bound can be directly calculated as

$$R \leq \frac{\mathcal{M}(\rho_S(\epsilon))}{\mathcal{M}(\rho_S(0))} = 1 + \frac{\log(1 - \frac{8}{15}\epsilon)}{\log \frac{5}{3}}. \quad (34)$$

The max-thauma [91] is defined as

$$\theta_{\max}(\rho) := \log \min \{2\text{sn}(V) + 1 : V \geq \rho\} \quad (35)$$

and can be calculated numerically via a semi-definite program. For the noisy Strange state, the max-thauma bound coincides with the mana bound, and we find that they are both looser than the majorization bound R_∞ as shown in Fig. 1.

This figure includes numerical estimates of the optimal majorization bounds R_{num} obtained by considering the full Lorenz curve, which show that the bound R_∞ can be further improved. While this is an improvement on prior results, all known distillation protocols have rates much lower than these upper bounds. It remains a major open question to determine what are the best possible rates that can be achieved.

Temperature-dependent bounds for magic distillation protocols

We now show how relative majorization can incorporate the system's Hamiltonian into magic distillation bounds. The way we do this is by considering how the protocol would disturb a reference equilibrium state τ to some different state τ' , if the protocol had been applied to this reference state instead of the actual n -copy magic state.

We consider a magic distillation protocol on multiple identical qudits in a noisy magic state $\rho(\epsilon)$, with noise parameter ϵ , sending

$$\rho(\epsilon)^{\otimes n} \longrightarrow \mathcal{E}(\rho(\epsilon)^{\otimes n}) = \rho(\epsilon')^{\otimes m} \quad (36)$$

with $n \gg 1$, $\epsilon' < \epsilon$ and \mathcal{E} denoting the quantum channel induced by the protocol. We also assume each qudit has a Hamiltonian H and neglect interaction terms. We choose some temperature $T = (k\beta)^{-1}$ where k is Boltzmann's constant and β the inverse temperature. The reference equilibrium state of the n qudits is given by

$$\tau^{\otimes n} = \left(\frac{e^{-\beta H}}{\mathcal{Z}} \right)^{\otimes n}. \quad (37)$$

We may also assume that the reference state τ is not a magic state, and has a strictly positive Wigner distribution $W_\tau(\mathbf{z})$.

A given magic protocol on the n qudits will correspond to a quantum channel \mathcal{E} . We also assume for simplicity that $U_\pi \mathcal{E}(X) U_\pi^\dagger = \mathcal{E}(X)$ for all X and any permutation U_π of the m output subsystems. This is justified because the protocol on the input magic state results in state $\rho(\epsilon')^{\otimes m}$, which is invariant under permutations. Therefore, we are always allowed to symmetrize the output $\mathcal{E}(X)$ by performing a group average over the permutation group for the output systems without changing the performance of the distillation protocol on $\rho(\epsilon)^{\otimes n}$. Thus, we can assume that \mathcal{E} always outputs a symmetric state in general. This means that $\mathcal{E}(\tau^{\otimes n})$ is a symmetric state on m subsystems, so by the quantum de Finetti theorem [117, 118] we have that for $m \gg 1$

$$\mathcal{E}(\tau) \approx \int d\mu(x) \tau_x'^{\otimes m}, \quad (38)$$

where $d\mu(x)$ is a probability measure over a set $\{\tau'_x\}$ of single qudit states.

Free energy and sub-linear correlations in the thermodynamic limit

To keep our analysis simple we make the following physical assumption. We assume that in the asymptotic/thermodynamic limit $n, m \rightarrow \infty$ the correlations generated on the reference equilibrium state are negligible. This implies that $d\mu(x)$ is peaked on a particular state τ' , and $\mathcal{E}(\tau^{\otimes n}) \approx \tau'^{\otimes m}$. This scenario occurs in the context of traditional thermodynamics, and states that the output system is well-described by intrinsic variables that do not scale in m , and correlations are sub-linear in m . In particular, this allows us to compute a free energy per qudit of the output state.

Non-trivial correlations in the thermodynamic limit can also be considered, but leads to a more complex analysis within our majorization framework of Wigner distributions. In this direction, we highlight recent work in majorization in which stochastic independence and correlations are analysed. It has been shown [119] that stochastic independence (no correlations) of independent distributions can be viewed as a resource in an extension of catalytic majorization, and leads to a single-shot operational interpretation of the Shannon entropy [120, 121].

Our bound depends on the von Neumann entropy $S(\rho)$ of a state, and the free energy at a particular temperature. For a state $\tau = e^{-\beta H} / \mathcal{Z}$, the Helmholtz free energy F is given by

$$F := \text{tr}[H\tau] - \beta^{-1} S(\tau) = -\beta^{-1} \log \mathcal{Z}, \quad (39)$$

which is obtained from the internal energy via a Legendre transform [122].

The protocol transforms the equilibrium as $\tau^{\otimes n} \rightarrow \mathcal{E}(\tau^{\otimes n}) = \tau'^{\otimes m}$. The protocol does not generate magic on its own, so we assume that the output state τ' is also a Wigner-positive state. However, this is generally a non-equilibrium state for the system. Despite this it is useful to associate an effective Hamiltonian H' to the output state by considering the change $H \rightarrow H'$ such that equilibrium is restored at the reference temperature T . This Hamiltonian is defined by the expression $\tau' = e^{-\beta H'} / \mathcal{Z}'$, and has free energy $F' = -\beta^{-1} \log \mathcal{Z}'$.

The magic state protocol is now considered by how it transforms the pair of quantum states $(\rho^{\otimes n}, \tau^{\otimes n})$ and is then constrained by the relative majorization condition $(W_\rho, W_\tau) \succ (W_{\rho'}, W_{\tau'})$ that holds due to the protocol being a stochastic map in the Wigner representation.

In Supplementary Note 4 we analyse the Lorenz curves of this condition and obtain the following magic distillation bound that combines computational measures ϵ, ϵ' , with terms that depend on the Hamiltonian and reference temperature T of the physical system. We state the result for the case of qutrits, but a similar analysis works for general odd-dimension qudits.

Theorem 7. *Consider a magic distillation protocol on qutrits that transforms n copies of an ϵ -noisy Strange state into m copies of an ϵ' -noisy Strange state, with depolarising errors $\epsilon' \leq \epsilon \leq 3/7$. We also allow pre/post-processing by local Clifford unitaries.*

Let $T = (k\beta)^{-1}$ be any finite temperature for the physical system and let $H = \sum_{k \in \mathbb{Z}_3} E_k |E_k\rangle\langle E_k|$ be the Hamiltonian of each qutrit subsystem in its eigen-decomposition. Assume that in the thermodynamic limit ($n, m \gg 1$), the protocol applied to the equilibrium state $\tau^{\otimes n} = (e^{-\beta H}/\mathcal{Z})^{\otimes n}$ maps $\tau^{\otimes n} \rightarrow \tau'^{\otimes m}$, where we write $\tau' = e^{-\beta H'}/\mathcal{Z}'$ for some Hermitian H' .

Then the asymptotic magic distillation rate $R = m/n$ is bounded as

$$R \leq \frac{\log\left(1 - \frac{4}{3}\epsilon\right) + \beta(\phi - F)}{\log\left(1 - \frac{4}{3}\epsilon'\right) + \beta(\phi' - F')}, \quad (40)$$

where F is the free energy of τ , and

$$\phi = -\beta^{-1} \log \zeta \quad (41)$$

with ζ given by the expressions

$$\zeta = \sum_{k \in \mathbb{Z}_3} \alpha_k e^{-\beta E_k}, \quad (42)$$

$$\alpha_k = \langle E_k | A_{\mathbf{z}_*} | E_k \rangle, \quad (43)$$

and $W_\tau(\mathbf{z})$ attaining a minimum at $\mathbf{z} = \mathbf{z}_*$. The primed variables are defined similarly for the output system.

The proof of this is provided in Supplementary Note 4, and follows a similar line to the unital protocol bound.

The above bound depends on:

- Quantum computational measures ϵ, ϵ' ,
- Thermodynamic quantities β, F, F' ,
- Intermediate terms ϕ, ϕ' .

The intermediate terms specify how the energy eigenbasis of the system $\{|E_k\rangle\}$ relates to the computational stabilizer basis $\{|k\rangle\}$. In particular, the term ϕ quantifies to what degree a sharp energy value can be associated to the negativity in the Wigner representation. Its form is similar to F , and so ϕ can loosely speaking be viewed as a ‘magic free energy’ term.

The coefficients α_k may be negative for some k , when the Hamiltonian has non-stabilizer eigenstates, but the quantity ϕ is always well-defined, since the function ζ is always positive for τ in the interior of \mathcal{F} . The quantity ϕ can diverge if the state τ acquired zero Wigner components, which occurs for τ on the boundary of the set of Wigner-positive states, and is not defined outside of \mathcal{F} . Finally, if H has a stabilizer eigenbasis then $\phi = E_i$ for some i , independent of the temperature T . Details of this are provided in Supplementary Note 4.

Technical aspects and special cases of free energy bounds

Similar to the unital protocol bounds, the above result is based on only part of the Lorenz curves and so can certainly be tightened with further analysis. The primary role of the pre/post-processing by Clifford unitaries in Theorem 7 is to simplify the form of the bound, and allow it to be expressed in terms of the free energy per particle in a way that does not depend on the parameters $\epsilon, \epsilon', \beta$ in a complex form. In Supplementary Note 4, we give bounds in which one does not include these Clifford changes of basis, thus having a more non-trivial dependency on the parameters.

The analysis makes other simplifying assumptions that could easily be dropped, at the price of more complex expressions. We could perform similar analysis for general qudits, and different choices of magic states, for example. It might also be of interest to consider other choices of reference states τ that are more appropriate to the hardware physics, for example for photonic set-ups [123].

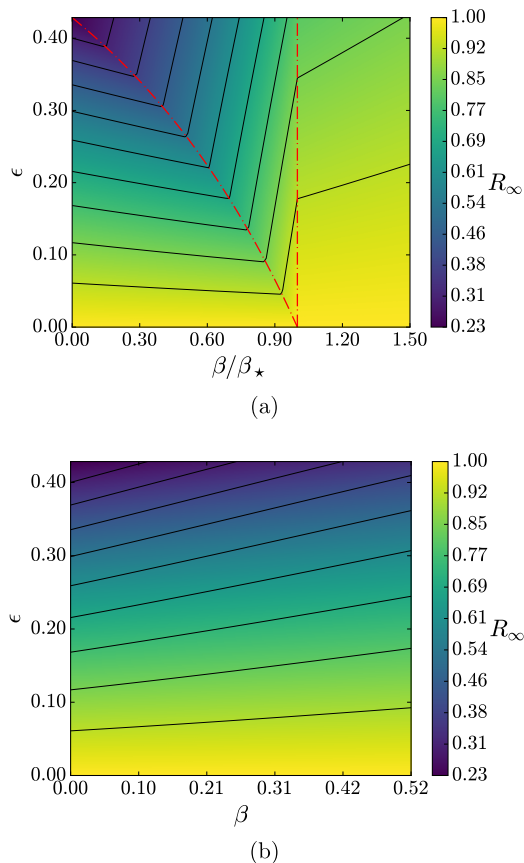


FIG. 4. **Temperature-dependent bounds for magic distillation.** Shown are two contour plots of the bound on $R(\epsilon, \beta)$ for the case $H = H' = \sum_{k \in \mathbb{Z}_3} k |k\rangle\langle k|$ and $\epsilon' = 0$, where β is the inverse temperature and ϵ is the depolarising error of the input magic state. The top figure (a) does not use any changes of Clifford basis, and the form of the bound depends on both the error parameter and temperature. The curved dashed line is $\epsilon_*(\beta)$ and given by Eq. (86) in Supplementary Note 4 and $\beta_* = (kT_*)^{-1}$ is given by $E_2 - \phi = kT_* \ln 2$. In the bottom figure (b) Clifford processing is used resulting in a smoother bound. In both figures the $\beta = 0$ line correspond to the unital bounds.

The assumption which is non-trivial is that we neglect correlations in the reference state in the thermodynamic limit. However, for more general scenarios one could make use of variational tools such as the Bogoliubov inequality [124] for approximating the free energy of a system via product states, to obtain similar bounds.

The simplest special case to consider is where the Hamiltonian H is diagonal in the computational basis, and $H = H'$, implying that the protocol leaves the reference equilibrium state unchanged, hence it corresponds to a Gibbs-preserving map [125]. For the limiting case of $\epsilon' \rightarrow 0$ we obtain Fig. 4 which is a contour plot of the bound on R as a function of inverse temperature β and the depolarising error ϵ for the noisy input magic states. In this figure we show both the bounds with Clifford basis changes and without them. In the more general case, $\Delta F := F' - F \neq 0$, so the protocol, when applied to the reference state, adds/extracts free energy from the system. We demonstrate change of the bounds with respect to a varying output Hamiltonian in Fig. 5.

Extension of single-shot entropies to quasi-distributions

We have considered Wigner distributions of magic states within a statistical mechanical setting in which it was argued that magic can be viewed as a non-classical form of free energy. This was done at the level of majorization constraints that appear due to stabilizer operations being described by stochastic maps in the Wigner representation. In this representation, non-equilibrium states with substantial free energy have Lorenz curves that deviate noticeably from the line $y = x$. We also found that when considered in a distillation protocol, magic could also be linked to physical free energies in a non-trivial way. We can therefore ask: is it possible to make a more direct link between magic and free energy or entropy?

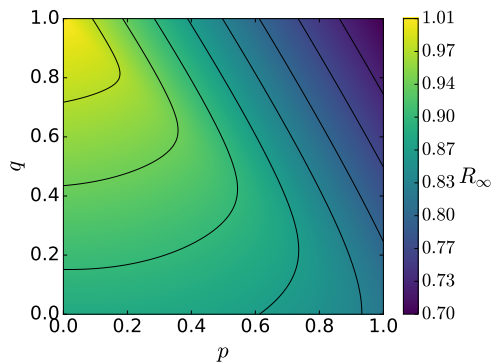


FIG. 5. **Variation of distillation bounds with Hamiltonian.** We illustrate variation of the bound in Eq. (40) with respect to the system Hamiltonian. We fix $(\epsilon, \epsilon', \beta) = (0.1, 0.0, 0.2)$ and look at a family of Hamiltonians for the qutrit systems, given by $H = A_0$ and $H' = (1 - p - q)A_0 + pA_{(1,2)} + q\text{diag}(0, 1, 2)$, with varying parameters p, q , and where A_0 and $A_{(1,2)}$ are phase-point operators.

One obstacle to linking with macroscopic, equilibrium free energy is that the Wigner distribution is generally a quasi-distribution, so the entropy $H(\mathbf{w}) = -\sum_i w_i \log w_i$ of the Wigner distribution is not defined outside the set \mathcal{F} , let alone have a physically meaningful interpretation.

However, it turns out that the analysis for the temperature-dependent bound given by Eq. (40) implicitly made use of the single-shot Rényi divergence [126] $D_\infty(\mathbf{p}||\mathbf{r})$, which is given by

$$D_\infty(\mathbf{p}||\mathbf{r}) = \log \max_i \frac{p_i}{r_i}. \quad (44)$$

It is clear that this does remain well-defined on quasi-distributions, and from the majorization relation also has a clear physical meaning, and in Supplementary Note 4, we show that this divergence has the following properties that are derived via relative majorization.

Theorem 8. *Let τ be in the interior of \mathcal{F} . Then $D_\infty(W_\rho||W_\tau)$ is well-defined for all ρ , and the following hold:*

1. $D_\infty(W_\rho||W_\tau) \geq 0$ for all quantum states ρ .
2. $D_\infty(W_\rho||W_\tau) = 0$ if and only if $\rho = \tau$.
3. $D_\infty(W_{\rho^{\otimes 2n}}||W_{\tau^{\otimes 2n}}) = nD_\infty(W_{\rho^{\otimes 2}}||W_{\tau^{\otimes 2}})$ for any $n \in \mathbb{N}$.
4. $D_\infty(W_\rho||W_\tau) \geq D_\infty(W_{\mathcal{E}(\rho)}||W_{\mathcal{E}(\tau)})$ for any free operation \mathcal{E} such that $\mathcal{E}(\tau)$ is in the interior of \mathcal{F} .

The reason for the power of 2 in the third property is that the maximization is sensitive to the presence of negativity in the Wigner distribution.

We can go further, and show that in fact a range of Rényi entropies remain both well-defined and meaningful in the Wigner representation. The α -Rényi entropy $H_\alpha(\mathbf{p})$ is defined as [126]

$$H_\alpha(\mathbf{p}) := \frac{1}{1-\alpha} \log \sum_i p_i^\alpha. \quad (45)$$

To extend to quasi-distributions we must ensure the following: (a) the entropy is mathematically well-defined on quasi-distributions, and (b) obeys monotonicity under the majorization pre-order. In Supplementary Note 5 we prove the following.

Theorem 9. *If $\alpha = \frac{2a}{2b-1}$ for positive integers a, b with $a \geq b$, then $H_\alpha(\mathbf{w})$ is well-defined on the set of quasi-distributions, and if $\mathbf{w} > \mathbf{w}'$ for two quasi-distributions \mathbf{w}, \mathbf{w}' then $H_\alpha(\mathbf{w}) \leq H_\alpha(\mathbf{w}')$.*

This monotonicity provides a meaningful extension of entropies to quasi-distributions. The unusual form of the α parameter is required so that the entropy function is well-defined and real-valued for all \mathbf{w} . The allowed values of α are dense in $\alpha > 1$, however, while Rényi entropies on probabilities can also be used for $0 < \alpha < 1$, it turns out that this parameter regime is no longer monotonic in terms of the majorization pre-order for quasi-distributions (see Supplementary Note 5 for details).

We can prove the following result, which establishes the equivalence between Wigner negativity and the existence of negative Rényi entropy.

Theorem 10. *A quantum state ρ has Wigner negativity if and only if $H_\alpha(W_\rho) < 0$ for some $\alpha = \frac{2a}{2b-1}$, with positive integers $a \geq b$.*

Therefore, the statistical mechanical description of stabilizer operations provides a setting in which negative entropy is fully meaningful, and quantifies the fact that magic states are more ordered than a perfectly sharp, deterministic classical state $\mathbf{w} = (1, 0, \dots, 0)$ with zero entropy. We note that prior work has shown that negative conditional entropy [127] arises in the context of quantum correlations, but is unrelated to the present negative entropy.

The $\alpha \rightarrow 1$ limit of the Rényi entropy diverges if negativity is present, however it diverges in a well-defined way. We can write the entropy as

$$H_{1+\epsilon}(W_\rho) = -\epsilon^{-1} \log \sum_{\mathbf{z}} |W_\rho(\mathbf{z})|^{1+\epsilon} \quad (46)$$

with $\epsilon \rightarrow 0^+$ through a sequence of rational values. We recall that mana can be written as $\mathcal{M}(\rho) = \log \sum_{\mathbf{z}} |W_\rho(\mathbf{z})|$, so we obtain a second perspective on mana:

$$\mathcal{M}(\rho) = - \lim_{\epsilon \rightarrow 0^+} \epsilon H_{1+\epsilon}(W_\rho). \quad (47)$$

Therefore, the mana of a state ρ is minus the residue of the pole at $\alpha = 1$ for $H_\alpha(W_\rho)$, and quantifies the divergence of the Rényi-entropy as we approach the limiting Shannon entropy $H(W_\rho)$.

We also note that the Rényi entropy can be described as a q -deformation of the Shannon entropy [128], and that $H_\alpha(\mathbf{p})$ can be related to the q -derivative of free energy. For this, the Rényi parameter becomes a temperature term $\alpha = \frac{T_0}{T}$ where T_0 is a reference temperature and derivatives are considered via the limit $T \rightarrow T_0$. Therefore, if we try to push this to the present setting, the presence of Wigner negativity would correspond to a divergence in the first derivative of an effective free energy as $T \rightarrow T_0$. It is interesting to speculate whether this non-classicality could be interpreted in terms of a phase transition [129].

Rényi divergences and general distillation bounds

Every entropy can be obtained from a relative entropy [130] so we now state the extensions for Rényi divergences of Wigner distributions. The α -Rényi divergence $D_\alpha(\mathbf{p}||\mathbf{r})$ is defined as

$$D_\alpha(\mathbf{p}||\mathbf{r}) := \frac{1}{\alpha-1} \log \sum_i p_i^\alpha r_i^{1-\alpha}, \quad (48)$$

where \mathbf{p} can be a quasi-distribution and \mathbf{r} a probability distribution with positive components. In Supplementary Note 5, we also show that the divergence satisfies the following.

Theorem 11. *Let τ be in the interior of \mathcal{F} . If $\alpha = \frac{2a}{2b-1}$ for positive integers a, b with $a \geq b$, then the α -Rényi divergence $D_\alpha(W_\rho||W_\tau)$ is well-defined for all states ρ , and the following hold:*

1. $D_\alpha(W_\rho||W_\tau) \geq 0$ for all quantum states ρ .
2. $D_\alpha(W_\rho||W_\tau) = 0$ if and only if $\rho = \tau$.
3. $D_\alpha(W_{\rho^{\otimes n}}||W_{\tau^{\otimes n}}) = nD_\alpha(W_\rho||W_\tau)$ for any $n \in \mathbb{N}$.
4. $D_\alpha(W_\rho||W_\tau) \geq D_\alpha(W_{\mathcal{E}(\rho)}||W_{\mathcal{E}(\tau)})$ for any free operation \mathcal{E} such that $\mathcal{E}(\tau)$ is in the interior of \mathcal{F} .

We can now apply this to obtain a family of magic distillation bounds based on Rényi divergences.

Theorem 12. *Consider a general magic state distillation protocol on odd prime dimension qudits, that converts a magic state $\rho^{\otimes n} \rightarrow \mathcal{E}(\rho^{\otimes n}) = \rho'^{\otimes m}$ and let τ be any full-rank stabilizer reference state on a qudit. Then, the distillation rate $R := m/n$ is upper bounded as*

$$R \leq \frac{D_\alpha(W_\rho||W_\tau)}{\tilde{D}_\alpha(\rho', \tau')}, \quad (49)$$

where $\alpha = \frac{2a}{2b-1}$ for any positive integers a, b with $a \geq b$ and the average divergence per qudit

$$\tilde{D}_\alpha(\rho', \tau') := \frac{1}{m} D_\alpha(W_{\rho'^{\otimes m}}||W_{\tau_m'}), \quad (50)$$

between the output magic state $\rho'^{\otimes m}$ and $\tau_m' = \mathcal{E}(\tau^{\otimes n})$.

The proof is given in Supplementary Note 5.

While this bound is abstract in its present form, it does suggest future applications by viewing $\tilde{D}_\alpha(\rho', \tau')$ within the context of hypothesis testing [131], as a measure of distinguishability between $\rho'^{\otimes m}$ and the m qudit state τ'_m . However, in order to properly link with hypothesis-testing one would have to first extend such results to quasi-distributions.

If correlations in τ'_m between subsystems can be neglected, then $\tau'_m = \tau'^{\otimes m}$ for some qudit state τ' and

$$\tilde{D}_\alpha(\rho', \tau') = D_\alpha(W_{\rho'} || W_{\tau'}), \quad (51)$$

which gives a generalized version of Theorem 7, with a term such as $D_\alpha(W_\rho || W_\tau)$ that behaves essentially as an α -free energy difference [70] in the Wigner representation.

We can illustrate a basic application of the above bound for magic protocols that give unital maps. In this case it is easy to see that the above bound reduces to the simple form

$$R \leq R_\alpha := \frac{2 \log d - H_\alpha(W_\rho)}{2 \log d - H_\alpha(W_{\rho'})}. \quad (52)$$

In particular for the case of noisy Strange states on qutrits and $\epsilon' = 0$, this bound becomes

$$R \leq R_\alpha = \frac{2(1 - \alpha) \log 3 - \log \left[8 \left(\frac{1}{6} - \frac{\epsilon}{18} \right)^\alpha + \left(-\frac{1}{3} + \frac{4\epsilon}{9} \right)^\alpha \right]}{2(1 - \alpha) \log 3 + \alpha \log 6 - \log(8 + 2^\alpha)}. \quad (53)$$

This upper bound can be numerically minimized over α , for any ϵ . The case R_{10} is shown in Fig. 1 and numerical evidence shows it is close to the optimal bound from majorization.

DISCUSSION

We have described how relative majorization can be used to establish upper bounds on magic distillation protocols that take into account additional physics of the system. Our bounds exploited relatively simple aspects of the Lorenz curves of the quasi-distributions, so it would be of interest to sharpen these bounds and obtain a better handle on the Lorenz curve structure in the $n \rightarrow \infty$ limit. It would also be of interest to analyse what features of single-shot entropies can be extended to quasi-distributions in a sensible form [131, 132]. This raises interesting questions as we no longer have a notion of typicality and the central limit theorem does not apply. That said, for special states such as the Strange state, the asymptotic behaviour is relatively simple, so exact asymptotics for this should be possible. In Supplementary Notes 2, 3 and 4 we give additional results that may be of use for later work.

Because Theorem 3 provides a completeness statement for majorization, we can ask if this approach could additionally provide lower bounds for distillation protocols. In contrast to the upper bounds, it is now essential to include the symplectic constraints on the phase space in the majorization relations. In principle the resultant majorization (and relative majorization) constraints should provide exact specifications on what can be obtained via free operations. Interestingly, this route does not require the use of error correcting codes, but instead is built on group-theoretic features of the phase space. The Clifford group action on a quantum system corresponds to the action of the affine symplectic group $G := SL(2, \mathbb{Z}_d) \times \mathbb{Z}_d^2$ on the discrete phase space [96, 100]. Therefore, any convex mixture of Clifford unitaries will correspond to a convex mixture of these group actions. The Hardy-Littlewood theorem [133] tells us that majorization is obtained from convex mixtures of arbitrary permutations, a result that has also been generalized to convex mixtures of elements of a general group G . G -majorization has been studied in the classical literature and a range of results are known about it [134–137]. One can therefore consider sub-groups of the affine symplectic group and analyse the resulting majorization/relative majorization that follows from its action on quasi-distributions. However if the sub-group considered is too simple then it may lead to trivial distillation rates. An example of this is when we restrict to the set of Weyl covariant channels [138], which are represented by a convex mixture of displacements on the phase space. Initial work by the authors have shown that the majorization conditions for Weyl covariant channels can be solved exactly in terms of discrete Fourier transforms, however it is found that the resultant distillation rates are trivial.

One notable structure that could be exploited for G -majorization is when G is a finite reflection group. For this, the G -majorization pre-order is guaranteed to be described by a finite list of conditions [134], just as is the case for the relative majorization ordering. Therefore one route to concrete lower bounds on distillation rates is to consider stochastic maps obtained by reflection sub-groups of $SL(2, \mathbb{Z}_d) \times \mathbb{Z}_2$, or other simple sub-groups, for which the majorization relation reduces to a finite set of conditions. Beyond this, another promising route is to formulate the majorization conditions for a sub-set of operations as a semi-definite program and then attempt to estimate realisable

distillation rates by exploiting dualities and appropriate coarse-grainings. This approach has recently been applied by one of the authors to estimate the set of quantum states accessible by covariant channels in the resource theory of asymmetry [139] where one allows depolarizing noise that is then minimized. It would be of interest to see if a similar approach could be followed for magic state distillation.

Beyond magic state distillation, the topic of G -majorization has been extensively studied, but to our knowledge there has not been work on relative G -majorization. This would correspond to transformations that are not unital. Physically, this regime would correspond to a form of thermo-majorization obtained from looking at the action of the Clifford group at a micro-canonical level, and then reducing to a small subsystem [122]. While this seems like a painful thing to consider, there is motivation for this beyond the aim of magic protocols: in the case of classical statistical mechanics on a phase space this is precisely the situation, albeit in the continuum limit. Statistical mechanics of actual systems obey Hamiltonian dynamics, thus they automatically respect a symplectic form [122, 140]. Therefore, the pre-order of statistical mechanical states with respect to phase space dynamics preserving a Gibbs state must correspond to a symplectic majorization condition. Of course, technical features arise in the continuum limit when considering distributions on an unbounded phase space, but this could be remedied by either considering a compact phase space (e.g. for a particle on a ring) or by first studying the discrete case. Such scenarios also arise in the Quantum Hall Effect [141], which is another regime where these techniques could be of use.

Lastly, we note that it would be of interest to consider the possibility of applying the approach taken here to other scenarios in which one wishes to distinguish classical from non-classical behaviour based on quasi-probability representations [75, 85–89, 142].

METHODS

Mathematical methods used for proofs

In order to derive our results we make use of three broad subfields.

The first one is the phase space representation of quantum systems as described in the first section of our Results. Generally, there are various possible representations [75], but we make use of the distinguished Wigner representation [92] for odd-dimensional systems.

We then develop a theory of majorization [81–83] for quasi-distributions so as to impose a partial order between magic states.

Finally, we extend the theory of classical Rényi entropies and divergences [126] for probability distributions to quasi-distributions and show that they possess sufficient properties for the derivation of our entropic bounds.

Precise definitions of all the above concepts and explanations of their incorporation to the study of magic distillation are provided in the Results section as the concepts become relevant.

DATA AVAILABILITY

No datasets were generated or analysed during the current study.

CODE AVAILABILITY

The code used to produce the figures in this work is available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

We thank Earl Campbell, Matteo Lostaglio, Philippe Faist and Nicole Yunger Halpern for helpful comments and discussions. NK is supported by the EPSRC Centre for Doctoral Training in Controlled Quantum Dynamics. DJ is supported by the Royal Society and a University Academic Fellowship.

COMPETING INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

DJ proposed the initial idea. Derivations were done by NK. Analysis and writing of the manuscript were jointly done by NK and DJ.

REFERENCES

-
- [1] Campbell, E. T., Terhal, B. M. & Vuillot, C. Roads towards fault-tolerant universal quantum computation. *Nature* **549**, 172–179 (2017).
- [2] Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86**, 5188–5191 (2001).
- [3] Raussendorf, R., Browne, D. E. & Briegel, H. J. Measurement-based quantum computation on cluster states. *Phys. Rev. A* **68**, 022312 (2003).
- [4] Nickerson, N. H., Fitzsimons, J. F. & Benjamin, S. C. Freely scalable quantum technologies using cells of 5-to-50 qubits with very lossy and noisy photonic links. *Phys. Rev. X* **4**, 041041 (2014).
- [5] Nikahd, E., Sedighi, M. & Saheb Zamani, M. Nonuniform code concatenation for universal fault-tolerant quantum computing. *Phys. Rev. A* **96**, 032337 (2017).
- [6] Chao, R. & Reichardt, B. W. Fault-tolerant quantum computation with few qubits. *npj Quantum Inf.* **4**, 1–8 (2018).
- [7] Lin, C., Yang, G., Luo, Q. & Li, X. Pieceable fault tolerant conversion between 5-qubit code and 7-CSS code. *Quantum Inf. Process.* **19**, 243 (2020).
- [8] Lin, C. & Yang, G. Concatenated pieceable fault-tolerant scheme for universal quantum computation. *Phys. Rev. A* **102**, 052415 (2020).
- [9] Chamberland, C. *et al.* Building a fault-tolerant quantum computer using concatenated cat codes (2020).
- [10] Bourassa, J. E. *et al.* Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer. *Quantum* **5**, 392 (2021).
- [11] Bravyi, S. B. & Kitaev, A. Y. Quantum codes on a lattice with boundary (1998).
- [12] Freedman, M. H. Quantum computation and the localization of modular functors. *Found. Comput. Math.* **1**, 183–204 (2001).
- [13] Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *J. Math. Phys.* **43**, 4452–4505 (2002).
- [14] Raussendorf, R. & Harrington, J. Fault-tolerant quantum computation with high threshold in two dimensions. *Phys. Rev. Lett.* **98**, 190504 (2007).
- [15] Eastin, B. & Knill, E. Restrictions on transversal encoded quantum gate sets. *Phys. Rev. Lett.* **102**, 110502 (2009).
- [16] Calderbank, A. R. & Shor, P. W. Good quantum error-correcting codes exist. *Phys. Rev. A* **54**, 1098–1105 (1996).
- [17] Steane, A. Multiple-particle interference and quantum error correction. *Proc. Math. Phys. Eng. Sci.* **452**, 2551–2577 (1996).
- [18] Bravyi, S. & Kitaev, A. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A* **71**, 022316 (2005).
- [19] Jones, C. Multilevel distillation of magic states for quantum computing. *Phys. Rev. A* **87**, 042305 (2013).
- [20] O’Gorman, J. & Campbell, E. T. Quantum computation with realistic magic-state factories. *Phys. Rev. A* **95**, 032338 (2017).
- [21] Bravyi, S. & Haah, J. Magic-state distillation with low overhead. *Phys. Rev. A* **86**, 052329 (2012).
- [22] Haah, J., Hastings, M. B., Poulin, D. & Wecker, D. Magic state distillation with low space overhead and optimal asymptotic input count. *Quantum* **1**, 31 (2017).
- [23] Hastings, M. B. & Haah, J. Distillation with sublogarithmic overhead. *Phys. Rev. Lett.* **120**, 050504 (2018).
- [24] Litinski, D. Magic State Distillation: Not as Costly as You Think. *Quantum* **3**, 205 (2019).
- [25] Campbell, E. T., Anwar, H. & Browne, D. E. Magic-state distillation in all prime dimensions using quantum reed-muller codes. *Phys. Rev. X* **2**, 041021 (2012).
- [26] Anwar, H., Campbell, E. T. & Browne, D. E. Qutrit magic state distillation. *New J. Phys.* **14**, 063006 (2012).
- [27] Dawkins, H. & Howard, M. Qutrit magic state distillation tight in some directions. *Phys. Rev. Lett.* **115**, 030501 (2015).
- [28] Krishna, A. & Tillich, J.-P. Towards low overhead magic state distillation. *Phys. Rev. Lett.* **123**, 070507 (2019).
- [29] Prakash, S. Magic state distillation with the ternary golay code. *Proc. Math. Phys. Eng. Sci.* **476**, 20200187 (2020).
- [30] Chamberland, C. & Noh, K. Very low overhead fault-tolerant magic state preparation using redundant ancilla encoding and flag qubits. *npj Quantum Inf.* **6**, 91 (2020).
- [31] Chamberland, C. & Cross, A. W. Fault-tolerant magic state preparation with flag qubits. *Quantum* **3**, 143 (2019).
- [32] Campbell, E. T. & Howard, M. Unifying gate synthesis and magic state distillation. *Phys. Rev. Lett.* **118**, 060501 (2017).
- [33] Campbell, E. T. & Howard, M. Unified framework for magic state distillation and multiqubit gate synthesis with reduced resource cost. *Phys. Rev. A* **95**, 022316 (2017).
- [34] Prakash, S., Jain, A., Kapur, B. & Seth, S. Normal form for single-qutrit clifford+ t operators and synthesis of single-qutrit gates. *Phys. Rev. A* **98**, 032304 (2018).

- [35] Seddon, J. R., Regula, B., Pashayan, H., Ouyang, Y. & Campbell, E. T. Quantifying quantum speedups: Improved classical simulation from tighter magic monotones. *PRX Quantum* **2**, 010345 (2021).
- [36] Leone, L., Oliviero, S. F. E. & Hamma, A. Rényi entropy of magic (2021).
- [37] Chitambar, E. & Gour, G. Quantum resource theories. *Rev. Mod. Phys.* **91**, 025001 (2019).
- [38] Ahmadi, M., Dang, H. B., Gour, G. & Sanders, B. C. Quantification and manipulation of magic states. *Phys. Rev. A* **97**, 062332 (2018).
- [39] Seddon, J. R. & Campbell, E. T. Quantifying magic for multi-qubit operations. *Proc. Math. Phys. Eng. Sci.* **475**, 20190251 (2019).
- [40] Wang, X., Wilde, M. M. & Su, Y. Quantifying the magic of quantum channels. *New J. Phys.* **21**, 103002 (2019).
- [41] Sarkar, S., Mukhopadhyay, C. & Bayat, A. Characterization of an operational quantum resource in a critical many-body system. *New J. Phys.* **22**, 083077 (2020).
- [42] White, C. D., Cao, C. & Swingle, B. Conformal field theories are magical. *Phys. Rev. B* **103**, 075145 (2021).
- [43] Bermejo-Vega, J., Delfosse, N., Browne, D. E., Okay, C. & Raussendorf, R. Contextuality as a resource for models of quantum computation with qubits. *Phys. Rev. Lett.* **119**, 120505 (2017).
- [44] Howard, M., Wallman, J., Veitch, V. & Emerson, J. Contextuality supplies the ‘magic’ for quantum computation. *Nature* **510**, 351–355 (2014).
- [45] Zurel, M., Okay, C. & Raussendorf, R. Hidden variable model for universal quantum computation with magic states on qubits. *Phys. Rev. Lett.* **125**, 260404 (2020).
- [46] Okay, C., Zurel, M. & Raussendorf, R. On the extremal points of the Λ -polytopes and classical simulation of quantum computation with magic states (2021).
- [47] Schmid, D., Selby, J. H., Pusey, M. F. & Spekkens, R. W. A structure theorem for generalized-noncontextual ontological models (2020).
- [48] Schmid, D., Selby, J. H., Wolfe, E., Kunjwal, R. & Spekkens, R. W. Characterization of noncontextuality in the framework of generalized probabilistic theories. *PRX Quantum* **2**, 010331 (2021).
- [49] Jones, N. C. *et al.* Layered architecture for quantum computing. *Phys. Rev. X* **2**, 031007 (2012).
- [50] Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
- [51] Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
- [52] Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
- [53] McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308 (2017).
- [54] Kandala, A. *et al.* Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).
- [55] Holmes, A. *et al.* Resource optimized quantum architectures for surface code implementations of magic-state distillation. *Microprocess. Microsyst.* **67**, 56–70 (2019).
- [56] Colless, J. I. *et al.* Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **8**, 011021 (2018).
- [57] Song, C. *et al.* Quantum computation with universal error mitigation on superconducting quantum processor (2018).
- [58] Bravyi, S., Sheldon, S., Kandala, A., McKay, D. C. & Gambetta, J. M. Mitigating measurement errors in multiqubit experiments. *Phys. Rev. A* **103**, 042605 (2021).
- [59] Bonilla Ataides, J. P., Tuckett, D. K., Bartlett, S. D., Flammia, S. T. & Brown, B. J. The XZZX surface code. *Nat. Commun.* **12**, 2172 (2021).
- [60] Bennett, C. H., DiVincenzo, D. P., Smolin, J. A. & Wootters, W. K. Mixed-state entanglement and quantum error correction. *Phys. Rev. A* **54**, 3824–3851 (1996).
- [61] Tuckett, D. K. *et al.* Tailoring surface codes for highly biased noise. *Phys. Rev. X* **9**, 041031 (2019).
- [62] Aliferis, P. & Preskill, J. Fault-tolerant quantum computation against biased noise. *Phys. Rev. A* **78**, 052331 (2008).
- [63] Stephens, A. M., Munro, W. J. & Nemoto, K. High-threshold topological quantum error correction against biased noise. *Phys. Rev. A* **88**, 060301 (2013).
- [64] Li, Y. A magic state’s fidelity can be superior to the operations that created it. *New J. Phys.* **17**, 023037 (2015).
- [65] Babbush, R. *et al.* Encoding electronic spectra in quantum circuits with linear t complexity. *Phys. Rev. X* **8**, 041015 (2018).
- [66] Guillaud, J. & Mirrahimi, M. Repetition cat qubits for fault-tolerant quantum computation. *Phys. Rev. X* **9**, 041053 (2019).
- [67] Fowler, A. G. & Gidney, C. Low overhead quantum computation using lattice surgery (2019).
- [68] Janzing, D., Wocjan, P., Zeier, R., Geiss, R. & Beth, T. Thermodynamic cost of reliability and low temperatures: Tightening Landauer’s principle and the second law. *Int. J. Theor. Phys.* **39**, 2717–2753 (2000).
- [69] Horodecki, M. & Oppenheim, J. Fundamental limitations for quantum and nanoscale thermodynamics. *Nat. Commun.* **4**, 2059 (2013).
- [70] Brandão, F., Horodecki, M., Ng, N., Oppenheim, J. & Wehner, S. The second laws of quantum thermodynamics. *PNAS* **112**, 3275–3279 (2015).
- [71] Vinjanampathy, S. & Anders, J. Quantum thermodynamics. *Contemp. Phys.* **57**, 545–579 (2016).

- [72] Goold, J., Huber, M., Riera, A., del Rio, L. & Skrzypczyk, P. The role of quantum information in thermodynamics—a topical review. *J. Phys. A Math.* **49**, 143001 (2016).
- [73] Lostaglio, M. An introductory review of the resource theory approach to thermodynamics. *Reports on Progress in Physics* **82**, 114001 (2019).
- [74] Gour, G., Jennings, D., Buscemi, F., Duan, R. & Marvian, I. Quantum majorization and a complete set of entropic conditions for quantum thermodynamics **9**, 5352 (2018).
- [75] Ferrie, C. & Emerson, J. Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations. *J. Phys. A Math.* **41**, 352001 (2008).
- [76] Veitch, V., Ferrie, C., Gross, D. & Emerson, J. Negative quasi-probability as a resource for quantum computation. *New J. Phys.* **14**, 113011 (2012).
- [77] Mari, A. & Eisert, J. Positive Wigner functions render classical simulation of quantum computation efficient. *Phys. Rev. Lett.* **109**, 230503 (2012).
- [78] Gottesman, D. Stabilizer codes and quantum error correction. Ph.D. thesis, California Institute of Technology (1997).
- [79] Knill, E. Quantum computing with realistically noisy devices. *Nature* **434**, 39–44 (2005).
- [80] Campbell, E. T. Catalysis and activation of magic states in fault-tolerant architectures. *Phys. Rev. A* **83**, 032317 (2011).
- [81] Marshall, A. W., Olkin, I. & Arnold, B. C. Inequalities: Theory of Majorization and Its Applications (Springer, 2011).
- [82] Veinott, A. F. Least d-majorized network flows with inventory and statistical applications. *Management Science* **17**, 547–567 (1971).
- [83] Ruch, E. & Mead, A. The principle of increasing mixing character and some of its consequences. *Theoretica chimica acta* **41**, 95–117 (1976).
- [84] Fine, A. Hidden variables, joint probability, and the bell inequalities. *Phys. Rev. Lett.* **48**, 291–295 (1982).
- [85] Allahverdyan, A. E. & Danageozian, A. Excluding joint probabilities from quantum theory. *Phys. Rev. A* **97**, 030102 (2018).
- [86] Arvidsson-Shukur, D. R. M. *et al.* Quantum advantage in postselected metrology. *Nat. Commun.* **11**, 3775 (2020).
- [87] Yunger Halpern, N., Swingle, B. & Dressel, J. Quasiprobability behind the out-of-time-ordered correlator. *Phys. Rev. A* **97**, 042105 (2018).
- [88] Lostaglio, M. Quantum fluctuation theorems, contextuality, and work quasiprobabilities. *Phys. Rev. Lett.* **120**, 040602 (2018).
- [89] Levy, A. & Lostaglio, M. Quasiprobability distribution for heat fluctuations in the quantum regime. *PRX Quantum* **1**, 010309 (2020).
- [90] Veitch, V., Mousavian, S. A. H., Gottesman, D. & Emerson, J. The resource theory of stabilizer quantum computation. *New J. Phys.* **16**, 013009 (2014).
- [91] Wang, X., Wilde, M. M. & Su, Y. Efficiently computable bounds for magic state distillation. *Phys. Rev. Lett.* **124**, 090505 (2020).
- [92] Gross, D. Hudson’s theorem for finite-dimensional quantum systems. *J. Math. Phys.* **47**, 122107 (2006).
- [93] Wigner, E. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.* **40**, 749–759 (1932).
- [94] Vourdas, A. Quantum systems with finite hilbert space. *Reports on Progress in Physics* **67**, 267–320 (2004).
- [95] Kenfack, A. & Życzkowski, K. Negativity of the Wigner function as an indicator of non-classicality. *J. opt. B, Quantum semiclass. opt.* **6**, 396–404 (2004).
- [96] Appleby, D. M. Symmetric informationally complete–positive operator valued measures and the extended clifford group. *J. Math. Phys.* **46**, 052107 (2005).
- [97] Mermin, N. D. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.* **65**, 3373–3376 (1990).
- [98] Raussendorf, R., Bermejo-Vega, J., Tyhurst, E., Okay, C. & Zurek, M. Phase-space-simulation method for quantum computation with magic states on qubits. *Phys. Rev. A* **101**, 012350 (2020).
- [99] Folland, G. B. Harmonic Analysis in Phase Space (Princeton University Press, 1989).
- [100] Bengtsson, I. & Życzkowski, K. Geometry of Quantum States: An Introduction to Quantum Entanglement (Cambridge University Press, 2006).
- [101] Schmid, D., Du, H., Selby, J. H. & Pusey, M. F. The only noncontextual model of the stabilizer subtheory is Gross’s (2021).
- [102] Watrous, J. The Theory of Quantum Information (Cambridge University Press, 2018).
- [103] Blackwell, D. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics* **24**, 265 – 272 (1953).
- [104] Nielsen, M. A. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.* **83**, 436–439 (1999).
- [105] Ćwikliński, P., Studziński, M., Horodecki, M. & Oppenheim, J. Limitations on the evolution of quantum coherences: Towards fully quantum second laws of thermodynamics. *Phys. Rev. Lett.* **115**, 210403 (2015).
- [106] Lostaglio, M., Jennings, D. & Rudolph, T. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nat. Commun.* **6**, 6383 (2015).
- [107] Gour, G., Müller, M. P., Narasimhachar, V., Spekkens, R. W. & Yunger Halpern, N. The resource theory of informational nonequilibrium in thermodynamics. *Physics Reports* **583**, 1–58 (2015).
- [108] Horodecki, M., Horodecki, P. & Oppenheim, J. Reversible transformations from pure to mixed states and the unique measure of information. *Phys. Rev. A* **67**, 062104 (2003).
- [109] Puchała, Z., Rudnicki, L. & Życzkowski, K. Majorization entropic uncertainty relations. *J. Phys. A Math.* **46**, 272002 (2013).

- [110] Vallejos, R. O., de Melo, F. & Carlo, G. G. Principle of majorization: Application to random quantum circuits. *Phys. Rev. A* **104**, 012602 (2021).
- [111] Ruch, E., Schraner, R. & Seligman, T. H. The mixing distance. *J. Chem. Phys.* **69**, 386–392 (1978).
- [112] Renes, J. M. Relative submajorization and its use in quantum resource theories. *J. Math. Phys.* **57**, 122202 (2016).
- [113] Buscemi, F. & Gour, G. Quantum relative lorenz curves. *Phys. Rev. A* **95**, 012110 (2017).
- [114] Rethinasamy, S. & Wilde, M. M. Relative entropy and catalytic relative majorization. *Phys. Rev. Research* **2**, 033455 (2020).
- [115] Brouwer, L. E. J. Über abbildung von mannigfaltigkeiten. *Mathematische Annalen* **71** (1911).
- [116] Jain, A. & Prakash, S. Qutrit and ququint magic states. *Phys. Rev. A* **102**, 042409 (2020).
- [117] Hudson, R. L. & Moody, G. R. Locally normal symmetric states and an analogue of de finetti’s theorem. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **33**, 343–351 (1976).
- [118] Christandl, M., König, R., Mitchison, G. & Renner, R. One-and-a-half quantum de finetti theorems. *Commun. Math. Phys.* **273**, 473–498 (2007).
- [119] Lostaglio, M., Müller, M. P. & Pastena, M. Stochastic independence as a resource in small-scale thermodynamics. *Phys. Rev. Lett.* **115**, 150402 (2015).
- [120] Müller, M. P. & Pastena, M. A generalization of majorization that characterizes shannon entropy. *IEEE Trans. Inf. Theory* **62**, 1711–1720 (2016).
- [121] Boes, P., Eisert, J., Gallego, R., Müller, M. P. & Wilming, H. Von neumann entropy from unitarity. *Phys. Rev. Lett.* **122**, 210402 (2019).
- [122] Pathria, R. K. & Beale, P. D. *Statistical Mechanics* (Elsevier, 2011).
- [123] Bombin, H. *et al.* Interleaving: Modular architectures for fault-tolerant photonic quantum computing (2021).
- [124] Bogolubov, J., N. N. On model dynamical systems in statistical mechanics. *Physica* **32**, 933–944 (1966).
- [125] Faist, P., Oppenheim, J. & Renner, R. Gibbs-preserving maps outperform thermal operations in the quantum regime. *New J. Phys.* **17**, 043003 (2015).
- [126] Rényi, A. On measures of entropy and information. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability* **1**, 547–561 (1960).
- [127] Rio, L. d., Åberg, J., Renner, R., Dahlsten, O. & Vedral, V. The thermodynamic meaning of negative entropy. *Nature* **474**, 61–63 (2011).
- [128] Baez, J. C. Rényi entropy and free energy (2011).
- [129] Domb, C. *Phase transitions and critical phenomena* (Elsevier, 2000).
- [130] Gour, G. & Tomamichel, M. Entropy and relative entropy from information-theoretic principles. *IEEE Trans. Inf. Theory* **67**, 6313–6327 (2021).
- [131] Tomamichel, M. A Framework for Non-Asymptotic Quantum Information Theory. Ph.D. thesis, ETH Zurich (2012).
- [132] Renner, R. Security of Quantum Key Distribution. Ph.D. thesis, Swiss Federal Institute of Technology Zurich (2005).
- [133] Hardy, G. *et al.* *Inequalities* (Cambridge University Press, 1952).
- [134] Giovagnoli, A. & Wynn, H. P. G-majorization with applications to matrix orderings. *Linear Algebra Appl.* **67**, 111–135 (1985).
- [135] Steerneman, A. G. M. G-majorization, group-induced cone orderings, and reflection groups. *Linear Algebra Appl.* **127**, 107–119 (1990).
- [136] Giovagnoli, A. & Wynn, H. P. Cyclic majorization and smoothing operators. *Linear Algebra Appl.* **239**, 215–225 (1996).
- [137] Eaton, M. L. & Perlman, M. D. Reflection Groups, Generalized Schur Functions, and the Geometry of Majorization. *The Annals of Probability* **5**, 829 – 860 (1977).
- [138] Fukuda, M. & Holevo, A. S. On weyl-covariant channels (2006).
- [139] Alexander, R., Gvirtz-Chen, S. & Jennings, D. Infinitesimal reference frames suffice to determine the asymmetry properties of a quantum system (2021).
- [140] Arnold, V. I. Symplectic geometry and topology. *J. Math. Phys.* **41**, 3307–3343 (2000).
- [141] Klitzing, K. v., Dorda, G. & Pepper, M. New method for high-accuracy determination of the fine-structure constant based on quantized hall resistance. *Phys. Rev. Lett.* **45**, 494–497 (1980).
- [142] Barnett, S. M. & Radmore, P. M. *Methods in Theoretical Quantum Optics*. Oxford science publications (Clarendon Press, 1997).