Entropy and reversible catalysis

H. Wilming

Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland,
Leibniz Universität Hannover, Appelstraße 2, 30167 Hannover, Germany

I show that non-decreasing entropy provides a necessary and sufficient condition to convert the state of a physical system into a different state by a reversible transformation that acts on the system of interest and a further "catalyst" whose state has to remain invariant exactly in the transition. This statement is proven both in the case of finite-dimensional quantum mechanics, where von Neumann entropy is the relevant entropy, and in the case of systems whose states are described by probability distributions on finite sample spaces, where Shannon entropy is the relevant entropy. The results give an affirmative resolution to the (approximate) "catalytic entropy conjecture" introduced by Boes et al. [PRL 122, 210402 (2019)]. They provide a complete single-shot characterization without external randomness of von Neumann entropy and Shannon entropy. I also compare the results to the setting of phenomenological thermodynamics and show how they can be used to obtain a quantitative single-shot characterization of Gibbs states in quantum statistical mechanics.

A central question in quantum information theory is which quantum states on some physical system may be transformed into which other states on the same (or a different) physical system by a given set of operations. This question underlies quantum resource theories, such as entanglement [1–4], thermodynamics [5–10] or asymmetry [11–13] (see Ref. [14] for a review on quantum resource theories). Common to most resource theories is that they allow for probabilistic mixing of operations, i.e., to use a source of classical randomness (such as a coin toss) to decide on the operation that is implemented. This use of randomness immediately implies that the resource theory is *convex*, which greatly simplifies the mathematical analysis. While often a natural assumption due to the commonplace access to (quasi-)randomness and classical communication, one may ask what happens if either one considers the cost of classical or quantum randomness explicitly or simply disallows the use of classical or quantum randomness. In the most extreme limit one would then end up only allowing the use of unitary operations and the question of which states can be inter-converted becomes trivial: namely all those states which are unitarily invariant, or in other words, all states with the same spectrum (including multiplicities).

In many resource theories the use of a *catalyst* is operationally well motivated and may greatly enrich the set of possible state transitions [9, 15–22]. A catalyst is a system which remains invariant in a given process, but may or may not, depending on the resource theory, build up correlations to other systems. Since its state does not change in a process, no resources are used up and the catalyst may be used again to facilitate further state transitions on other systems – this is conceptually similar to catalysts in chemistry or the ubiquitous "periodically working machines" in thermodynamics. One may therefore wonder what happens if we consider only unitary operations together with the possibility for catalysts. Interestingly, Ref. [22] found that in this setting the von Neumann entropy plays a special role due to its sub-additivity property (see below). It was in fact conjectured that the von Neumann entropy is the *only* constraint for state-transitions once one allows for arbitrary small errors on the system (but not the catalyst). This conjecture was called "catalytic entropy conjecture" and can also be formulated in the classical setting as a conjecture connecting Shannon entropy to catalytic permu-
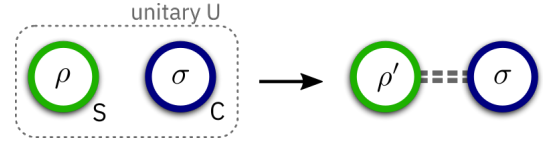

FIG. 1. A catalytic transition: A unitary operation U is applied to systems S and C in the state $\rho \otimes \sigma$. The resulting reduced state on S is (arbitrarily close to) ρ', while the reduced state on C is preserved exactly, but correlated to S (indicated by the dashed lines). The main result of this paper shows that a matching σ and U can be found if and only if $H(\rho') \geq H(\rho)$.

tations of probability distributions. The ideas behind this conjecture already have found application in the context of thermodynamics and fluctuation theorems [23, 24].

The special role of von Neumann entropy is surprising for several reasons: Typically, the Shannon or von Neumann entropy appears in settings involving many weakly correlated systems due to the phenomenon of *typicality* [25–27]. It was therefore long believed that the von Neumann entropy only plays a special role in asymptotic settings, such as the thermodynamic limit in physics or the limit of many identically and independently distributed signals in information theory. In particular, in resource theories involving free randomness and allowing for catalysts (that however may not become correlated to the system of interest), state transitions are usually characterized by an infinite set of constraints [9, 18–20]. It was only very recently first conjectured and then proven that these conditions may collapse to the von Neumann entropy (or similar quantities, such as relative entropy or free energy) if one allows for catalysts that become correlated to the system [28–34]. However, these settings still made use of external randomness – either by allowing for classical randomness explicitly or allowing free access to systems such as heat baths, which can be seen as sources of randomness. It is thus interesting that the catalytic entropy conjecture posits that von Neumann entropy plays such a special role in situations that neither allow for the use of external randomness, nor require an asymptotic limit. This paper provides an affirmative resolution of the catalytic entropy conjecture.

The catalytic entropy conjecture can also be formulated in

full analogy in the classical case by replacing finite dimensional density matrices with probability distributions on finite sample spaces, unitary operations with permutations (reversible transformations on the sample space) and von Neumann entropy with Shannon entropy. In the main text of this paper we will only consider the quantum case and only briefly comment on how to proof the classical version of our main result. A full proof of the classical result is given in Appendix C. While the proof of the classical result also implies the quantum result, it has the drawback of requiring a larger catalyst in general.

Setting and main result. Throughout we consider a system S described by density matrices ρ and ρ' on a Hilbert-space of dimension d. In the following we write $H(\rho)$ for von Neumann entropy, which is defined as

$$H(\rho) = -\mathrm{Tr}[\rho \log(\rho)]. \tag{1}$$

Von Neumann entropy is continuous in ρ [35, 36] and has several useful properties, such as unitary invariance ($H(\rho) = H(U\rho U^\dagger)$ for any unitary U), additivity ($H(\rho \otimes \sigma) = H(\rho) + H(\sigma)$) and sub-additvity: $H(\rho_{12}) \leq H(\rho_1) + H(\rho_2)$, where ρ_{12} is a bipartite quantum state with marginals ρ_1 and ρ_2. Finally, denote by $D(\rho, \rho') := \frac{1}{2}\|\rho - \rho'\|_1$ the trace-distance between two density-matrices. We then define catalytic state transitions formally as follows (see also Fig. 1).

Definition 1 (Approximate catalytic transformation). *Consider two finite-dimensional density matrices ρ and ρ' on the same system S. We write $\rho \rightarrow_\epsilon \rho'$ if there exists a finite-dimensional density matrix σ on a system C and a unitary U on SC such that*

$$\mathrm{Tr}_S[U\rho \otimes \sigma U^\dagger] = \sigma \tag{2}$$

and

$$D\big(\mathrm{Tr}_C[U\rho \otimes \sigma U^\dagger], \rho'\big) \leq \epsilon. \tag{3}$$

The following main result of this paper then shows that the set of states that are reachable from a given state ρ is given exactly by the set of states with higher von Neumann entropy:

Theorem 2 (Catalytic transformations characterize von Neumann entropy). *The following are equivalent:*

i) $\rho \rightarrow_\epsilon \rho'$ *for all* $\epsilon > 0$.

ii) $H(\rho') \geq H(\rho)$.

Before coming to the proof of this statement, let us first discuss the formal similarity between the theorem and the corresponding set-up in thermodynamics and then give an application for quantum statistical mechanics.

Analogy with Thermodynamics. Let us consider the following idealized, but ubiquitous setting of a thermodynamic work-process acting on three systems: A system S composed of a working substance and all other parts that may change in the process (such as heat baths), a collection of systems C composed of all systems that take part in the process but return to their initial state when the process has finished (gears,

pistons etc.), and finally an idealized work-storage device W with vanishing entropy (e.g., a purely mechanical device such as a suspended weight in a uniform gravitational field). Since the system C is cyclic, it does not contribute any energy or entropy to the process. Similarly, W only contributes or absorbs the work necessary for the process, but does not act as a source or sink of entropy. Let the initial and final states of S be labelled by a and b, respectively. By the Second Law of Thermodynamics, the considered process can only be possible if the entropy of S is non-decreasing: $H(b) \geq H(a)$ (we use the same letter for thermodynamic entropy as for von Neumann or Shannon entropy). Conversely, and importantly, however, it is commonly assumed in setting up the thermodynamic framework that for any two states a, b of S there is some such idealized work-process connecting the two states (for a very clear, recent exposition see [37]). Indeed, this is required to be able to consistently define a thermodynamic entropy. Therefore, the states a and b can be connected by such a process if and only if $H(b) \geq H(a)$, in complete analogy with Theorem 2, where the role of W is played by the external laboratory implementing the unitary U. In the thermodynamic setting, the system C is usually left implicit, but is clearly physically necessary and its precise design depends on the states a and b. While in the thermodynamic case the source of the increase of entropy is left unspecified, in the microscopic case discussed here it is due to the build up of correlations between S and C. Explicitly, we have $H(\rho') - H(\rho) = I(S' : C)$, where $I(S' : C)$ denotes the mutual information between S and C after the unitary evolution.

Quantitative single-shot characterization of Gibbs states. Let us now illustrate the use of the main result with an application in quantum statistical mechanics. We first have to establish some background material. A *passive state* is any quantum state ρ whose energy (w.r.t. some fixed Hamiltonian H) cannot be lowered by a unitary transformation. Equivalently, we may say that no work can be extracted from a passive state using unitary operations. Even though ω is passive, it is not necessarily *completely passive*, meaning that a positive amount of energy

$$\overline{W}(\rho) := \lim_{n \to \infty} \frac{1}{n} \sup_U \left[\mathrm{Tr}[\rho^{\otimes n} \hat{\mathsf{H}}_n] - \mathrm{Tr}[U\rho^{\otimes n} U^\dagger \hat{\mathsf{H}}_n] \right]$$

may be extracted *per copy* if many copies of ρ are available. Here, $\hat{\mathsf{H}}_n := \sum_{j=1}^n \mathsf{H}_j$ with $\mathsf{H}_j \equiv \mathsf{H}$ is the total Hamiltonian on n copies. Let us assume for simplicity that the ground state of H is unique. It has been proven that [38]

$$\overline{W}(\rho) = \mathrm{Tr}[\rho\,\mathsf{H}] - \mathrm{Tr}[\omega_{\beta(\rho)}(\mathsf{H})\,\mathsf{H}], \tag{4}$$

where $\omega_\beta(\mathsf{H}) := \exp(-\beta\mathsf{H})/\mathrm{Tr}[\exp(-\beta\mathsf{H})]$ denotes a Gibbs state and $\beta(\rho) \geq 0$ is chosen such that $H(\omega_{\beta(\rho)}(\mathsf{H})) = H(\rho)$ (if $H(\rho) = 0$ one has to take the limit $\beta \to +\infty$). Thus, the only completely passive states are ground states and Gibbs states with positive temperature [39, 40].

The above characterization of Gibbs states relies on a thermodynamic limit. It was observed in Ref. [23], and discussed in detail for the particular case of three-level systems, that the energy of passive states may nevertheless in general be

reduced using catalytic transitions instead of unitary operations unless the state in question is a Gibbs state. This may be surprising, since the catalyst, by virtue of not changing its state, cannot compensate for energetic changes. It was an open problem to determine how much work can be extracted from an arbitrary state using a single catalytic transformation [41]. Let us denote this quantity by

$$W_{\text{cat.}}(\rho) := \sup_{\rho \to_0 \rho'} \left(\text{Tr}[\rho \mathsf{H}] - \text{Tr}[\rho' \mathsf{H}] \right). \quad (5)$$

Then Theorem 2 immediately leads to the following corollary, which shows that catalysts allow to extract the same amount of energy from a single passive states as on average from asymptotically many copies:

Corollary 3. $W_{\text{cat.}}(\rho) = \overline{W}(\rho)$.

Proof. By Theorem 2 and continuity we have

$$W_{\text{cat.}}(\rho) = \sup_{\rho' : H(\rho') \geq H(\rho)} \left(\text{Tr}[\rho \mathsf{H}] - \text{Tr}[\rho' \mathsf{H}] \right). \quad (6)$$

By concavity of von Neumann entropy, the optimizer must have $H(\rho') = H(\rho)$. But by Gibbs variational principle, $\omega_{\beta(\rho)}(\mathsf{H})$ is the state with lowest energy among all states with entropy $H(\rho)$. Hence $W_{\text{cat.}}(\rho) = \text{Tr}[\rho \mathsf{H}] - \text{Tr}[\omega_{\beta(\rho)}(\mathsf{H})\mathsf{H}] = \overline{W}(\rho)$. $\square$

Proof of the Theorem 2. The direction i) $\Rightarrow$ ii) follows directly from sub-additivity, unitary invariance and continuity of von Neumann entropy: For a given ϵ, denote by ρ'_ϵ the final state on S. Then we have

$$H(\rho'_\epsilon) + H(\sigma) \geq H(U(\rho \otimes \sigma)U^\dagger) = H(\rho) + H(\sigma) \quad (7)$$

and hence $H(\rho'_\epsilon) \geq H(\rho)$. By continuity we thus find $H(\rho') \geq H(\rho)$.

The converse direction ii) $\Rightarrow$ i) requires several Lemmas. First, we collect a combination of some standard results on typicality and majorization. We write $\rho \succeq \rho'$ if ρ *majorizes* ρ', meaning that there exists a probability distribution q_i over unitaries V_i such that $\rho' = \sum_i q_i V_i \rho V_i^\dagger$. Similarly, we write $\boldsymbol{a} \succeq \boldsymbol{b}$ for two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ if there exists a probability distribution q_i over permutation-matrices π_i such that $\boldsymbol{a} = \sum_i q_i \pi_i \boldsymbol{b}$.

Lemma 4 (Typicality and majorization). *Let ρ and ρ' be two finite-dimensional density matrices of dimension d with $H(\rho) < H(\rho')$. Then for any $\epsilon > 0$ and large enough n there exists a state $\rho'_{\epsilon,n}$ such that $\rho^{\otimes n} \succeq \rho'_{\epsilon,n}$ and $D(\rho'_{\epsilon,n}, \rho'^{\otimes n}) \leq \epsilon$. Moreover, the error ϵ can be bounded as*

$$\epsilon \leq O(\exp(-n \Delta H^2 / 4)) \quad (8)$$

with $\Delta H := H(\rho') - H(\rho)$ and $\rho'_{\epsilon,n}$ may be chosen to have the same eigenbasis as $\rho'^{\otimes n}$.

A proof-sketch of Lemma 4 is given in Appendix A. It is clear that the given error bound is not optimal for every choice of ρ and ρ', since, for example, $\epsilon = 0$ is possible if $\rho \succeq \rho'$.

However, Ref. [42] shows that the given error bound is essentially optimal up to constants as a bound that does not take into account detailed information about ρ and ρ'. In Appendix D we use this to give an estimate of the size of the catalyst.

The next Lemma will be essential to construct a candidate catalyst by making use of Lemma 4. It is based on the *Schur-Horn theorem*, which states that for any $d \times d$ Hermitian matrix its vector of eigenvalues $\boldsymbol{\lambda}$ majorizes the vector of diagonal elements in every orthonormal basis. Conversely, every vector that is majorized by $\boldsymbol{\lambda}$ may be obtained as the diagonal elements in a suitable orthonormal basis. In particular if $\boldsymbol{p}$ and $\boldsymbol{p}'$ denote the ordered vectors of eigenvalues of two density matrices ρ and ρ', respectively, then $\rho \succeq \rho'$ if and only if $\boldsymbol{p} \succeq \boldsymbol{p}'$. In the following we denote by $\mathcal{D}_{\rho'}$ the *dephasing channel* in the eigenbasis of ρ' that acts as

$$\mathcal{D}_{\rho'}[\rho] = \sum_i |i\rangle\langle i|\rho|i\rangle\langle i|, \quad (9)$$

where the $|i\rangle$ constitute an orthonormal eigenbasis of ρ'.

Lemma 5 (Basic lemma). *Let $H(\rho) < H(\rho')$ and $\mathcal{D}_{\rho'}$ the dephasing channel in the eigenbasis of ρ'. Then for any $\epsilon > 0$ there exists an $n \in \mathbb{N}$ and a unitary U such that for any $1 \leq k \leq n$:*

$$D\big(\rho', \mathcal{D}_{\rho'}[\chi_k]\big) \leq \epsilon, \quad \chi := U\rho^{\otimes n}U^\dagger, \quad (10)$$

where $\chi_k := \text{Tr}_{\{1,\dots,n\}\setminus\{k\}}[\chi]$. The error ϵ scales as in Lemma 4.

Proof. We make use of the state $\rho'_{\epsilon,n}$ guaranteed by Lemma 4. By the Schur-Horn theorem there exists a unitary U such that

$$\rho'_{\epsilon,n} = \mathcal{D}_{\rho'}^{\otimes n}[\chi]. \quad (11)$$

However, we have that (writing $\overline{k} := \{1, \dots, n\} \setminus \{k\}$)

$$\mathcal{D}_{\rho'}[\chi_k] = \text{Tr}_{\overline{k}}[(\mathcal{D}_{\rho'} \otimes \mathbf{1}_{\overline{k}}[\chi]] = \text{Tr}_{\overline{k}}[\mathcal{D}_{\rho'}^{\otimes n}[\chi]] = \text{Tr}_{\overline{k}}[\rho'_{\epsilon,n}]$$

by locality of quantum mechanics. But since the trace-distance is non-increasing under partial traces, we then find

$$D\big(\rho', \mathcal{D}_{\rho'}[\chi_k]\big) \leq D\big(\rho'^{\otimes n}, \rho'_{\epsilon,n}\big) \leq \epsilon. \quad (12)$$

$\square$

The construction of the unitary U guaranteed by the Schur-Horn theorem is explained in Appendix B. The final Lemma that we require provides a way for us to get rid of unwanted coherences (arising from Lemma 5) in the final state without correlating the catalyst internally (which would spoil the catalyst).

Lemma 6 (No propagation of correlations for mixed unitary channels). *Consider a mixed unitary quantum channel $\mathcal{C}[\cdot] = \sum_i p_i V_i \cdot V_i^\dagger$ acting on a system S, where the p_i denote probabilities and the V_i are unitary operators. Dilate $\mathcal{C}$ using an auxiliary system C and state $\sigma = \sum_i p_i |i\rangle\langle i|$ using the unitary $V := \sum_i V_i \otimes |i\rangle\langle i|$ as*

$$\mathcal{C}[\rho] = \text{Tr}_2\left[V\rho \otimes \sigma V^\dagger\right]. \quad (13)$$

Finally apply the dilation to a state $\rho_{S\overline{S}}$ on S and a further system $\overline{S}$. Then

$$\mathrm{Tr}_S\left[(V \otimes \mathbf{1}_{\overline{S}})\rho_{S\overline{S}} \otimes \sigma(V^\dagger \otimes \mathbf{1}_{\overline{S}})\right] = \rho_{\overline{S}} \otimes \sigma. \quad (14)$$

That is, the dilating system C is catalytic and remains uncorrelated to $\overline{S}$.

Proof. The result immediately from the unitary invariance of the (partial) trace:

$$\mathrm{Tr}_S\left[(V \otimes \mathbf{1}_{\overline{S}})\rho_{S\overline{S}} \otimes \sigma(V^\dagger \otimes \mathbf{1}_{\overline{S}})\right] \quad (15)$$

$$= \sum_i p_i \mathrm{Tr}_S\left[(V_i \otimes \mathbf{1}_{\overline{S}})\rho_{S\overline{S}}(V_i^\dagger \otimes \mathbf{1}_{\overline{S}})\right] \otimes |i\rangle\langle i|$$

$$= \sum_i p_i \rho_{\overline{S}} \otimes |i\rangle\langle i| = \rho_{\overline{S}} \otimes \sigma. \quad (16)$$

$\square$

We are now in position to prove ii) $\Rightarrow$ i) of Theorem 2. The proof proceeds in two parts. First we construct a catalyst σ_1 for the exact transition from ρ to the equal mixture

$$\overline{\chi} := \frac{1}{n}\sum_{k=1}^n \chi_k \quad (17)$$

of the states $\chi_k = \mathrm{Tr}_{\overline{k}}[\chi]$, where χ is the state from Lemma 5. Then we use a second catalyst R in state σ_2 to implement the dephasing map and obtain $\mathcal{D}_{\rho'}[\overline{\chi}]$, which is ϵ-close to the target ρ'. The part R of the catalyst thus effectively acts as a source of randomness. By Lemma 6 and the fact that the dephasing map is a mixed unitary channel, this second part can be done in such a way that the two parts of the catalyst remain uncorrelated. Therefore, the two-step process is still catalytic when both parts of the catalyst are considered as one joint catalyst in state $\sigma_1 \otimes \sigma_2$. As a side comment, we mention that the results of [43] imply that σ_2 only needs to have a dimension of the order of $\sqrt{d}$. Furthermore, note that by perturbing ρ' arbitrarily slightly, we can always ensure that $H(\rho) < H(\rho')$ since we allow for arbitrarily small errors and von Neumann entropy is continuous. We thus only need to prove that we can do the transition $\rho \to_{\epsilon=0} \overline{\chi}$ in the case $H(\rho') > H(\rho)$. To show this, we make use of a trick that was used in recent work by Shiraishi and Sagawa [34]: We denote by $S = S_1$ the system and by $S_2, \ldots, S_n$ and A subsystems of the first part of the catalyst. The S_i all have the same Hilbert-space dimension as S and A has Hilbert-space dimension n (see Fig 2 for an illustration of the structure of the catalyst). Then define

$$\sigma_1 = \frac{1}{n}\sum_{k=1}^n \rho^{\otimes k-1} \otimes \chi_{1,\ldots,n-k} \otimes |k\rangle\langle k|_A, \quad (18)$$

where $\chi_{1,\ldots,i}$ denotes the reduced density matrix of χ consisting of the subsystems 1 to i and we define χ_0 and $\rho^{\otimes 0}$ to be the trivial state 1. We now apply the following sequence of unitaries on $\rho \otimes \sigma_1$ (see Fig. 3):

1. $U \otimes |n\rangle\langle n| + \mathbf{1} \otimes \sum_{k=1}^{n-1} |k\rangle\langle k|$, with U the unitary from Lemma 5.
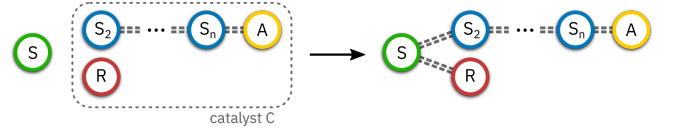

FIG. 2. The structure of the constructed catalyst C: It contains subsystems $S_2, \ldots, S_n$, which are copies of the target system S together with an auxiliary system A of dimension n as well as a catalytic source of randomness R. The dashed lines indicate possible correlations. The source of randomness is utilized to dilate the decoherence-channel $\mathcal{D}_{\rho'}$ on S in such a way that it does not become correlated to the systems $S_2, \ldots, S_n$ and A in the process.

2. The cyclic shift of sub-systems $S_i \to S_{i+1}$ with $S_n \to S_1$.

3. The cyclic shift on A, acting as $|i\rangle \to |i+1\rangle$ with $|n+1\rangle = |1\rangle$.

After the three steps the catalyst is back to its initial state. The state on the system, on the other hand, is given by $\overline{\chi}$ and, after applying the dephasing map using the system R, we find

$$D\left(\frac{1}{n}\sum_{k=1}^n \mathcal{D}_{\rho'}[\chi_k], \rho'\right) \leq \frac{1}{n}\sum_{k=1}^n D(\mathcal{D}_{\rho'}[\chi_k], \rho') \leq \epsilon$$

by the triangle inequality and Lemma 5.

The classical case. As mentioned above, we can also formulate a the classical version of Theorem 2. To do that, we can define a catalytic transition $\boldsymbol{p} \to_\epsilon \boldsymbol{p}'$ between two probability vectors $\boldsymbol{p}, \boldsymbol{p}' \in \mathbb{R}^d$ as in the quantum case, but with the catalysts' density matrix replaced by a further probability vector $\boldsymbol{q} \in \mathbb{R}^{d_C}$ and the unitary U replaced by a permutation acting on the canonical basis vectors of $\mathbb{R}^d \otimes \mathbb{R}^{d_C}$. The corresponding theorem, which is fully proven in Appendix C, can then be stated as follows:

Theorem 7. *Let $\boldsymbol{p}, \boldsymbol{p}' \in \mathbb{R}^d$ be two probability vectors with Shannon entropies $H(\boldsymbol{p})$ and $H(\boldsymbol{p}')$, respectively. The following are equivalent:*

1. *For all $\epsilon > 0$ we have $\boldsymbol{p} \to_\epsilon \boldsymbol{p}'$.*

2. $H(\boldsymbol{p}) \leq H(\boldsymbol{p}')$.

Let me briefly comment on the main difference in the proof as compared to the quantum case: The essential construction of the catalyst is quite similar to the quantum case, however clearly we cannot make use of the Schur-Horn theorem, since we do not have access to unitary operations. The proof therefore proceeds by building into the catalyst a source of randomness, which instead of being used to dephase the system, is already correlated with the first part of the catalyst $S_2 \cdots S_n A$ from the beginning and can be used to implement the transition $\boldsymbol{p}^{\otimes n} \to_{\varepsilon=0} \boldsymbol{p}'_{\epsilon,n}$ by a random permutation in the case that the auxiliary system A is in state n. This source of randomness in general needs to have a dimension of the order of d^n in contrast to the the quantum case, which only requires a dimension of the order of $\sqrt{d}$.
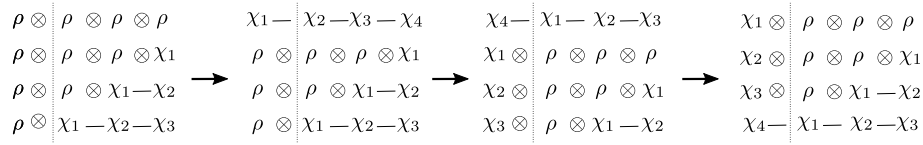
$$\begin{array}{cccc}
\rho \otimes \rho \otimes \rho \otimes \rho & \chi_1 - \chi_2 - \chi_3 - \chi_4 & \chi_4 - \chi_1 - \chi_2 - \chi_3 & \chi_1 \otimes \rho \otimes \rho \otimes \rho \\
\rho \otimes \rho \otimes \rho \otimes \chi_1 & \rho \otimes \rho \otimes \rho \otimes \chi_1 & \chi_1 \otimes \rho \otimes \rho \otimes \rho & \chi_2 \otimes \rho \otimes \rho \otimes \chi_1 \\
\rho \otimes \rho \otimes \chi_1 - \chi_2 & \rho \otimes \rho \otimes \chi_1 - \chi_2 & \chi_2 \otimes \rho \otimes \rho \otimes \chi_1 & \chi_3 \otimes \rho \otimes \chi_1 - \chi_2 \\
\rho \otimes \chi_1 - \chi_2 - \chi_3 & \rho \otimes \chi_1 - \chi_2 - \chi_3 & \chi_3 \otimes \rho \otimes \chi_1 - \chi_2 & \chi_4 - \chi_1 - \chi_2 - \chi_3
\end{array}$$

FIG. 3. Illustration of the steps in the first part of the proof for $n = 4$. The different columns denote the systems $S_1 \cdots S_n$ with $S_1 = S$ and the different rows indicate the state $|k\rangle$ of the auxiliary system A. The hyphen indicates that the corresponding subsystems may be correlated. This procedure is essentially identical to the one in Ref. [34]. In the proof of the classical result (see Appendix C), each of the subsystems with marginals called χ_i here is further correlated to the second part R of the catalyst which is not shown.

Conclusion and open problems. We have seen that von Neumann entropy and Shannon uniquely characterize state transitions that allow for the use of a catalyst but are otherwise reversible. There are several natural open problems left for future work: First, in Ref. [22], an exact form of the quantum catalytic entropy conjecture was conjectured, where no error is allowed in the transition from ρ to ρ' at the expense of the additional constraint that the rank of ρ' has to be at least as large as that of ρ. This form of the conjecture seems much more difficult to prove and probably requires methods that go beyond standard typicality results (which always yield asymptotic statements with vanishingly small, but finite error).

A second open problem is to investigate whether a similar result also holds for other standard entropic quantities, such as mutual information or relative entropy.

Finally, it will be worthwhile to explore the consequences of the given results for applications. Some further immediate applications of an affirmative resolution of the catalytic entropy conjecture in the context of (quantum) thermodynamics have been explored in Refs. [22, 24], but more applications can be expected. In particular, it would be interesting to see whether there are useful applications in the context of (potentially entanglement-assisted) communication scenarios.

Note added: After the publication of the first preprint of this work, Refs. [44, 45] appeared that show similar single-shot characterizations of standard entropic quantities in the context of quantum teleportation and quantum entanglement.

[1] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

[2] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Physical Review A **54**, 3824 (1996).

[3] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Physical Review Letters **78**, 2275 (1997).

[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Reviews of Modern Physics **81**, 865 (2009).

[5] E. H. Lieb and J. Yngvason, Phys. Rep. **310**, 1 (1999).

[6] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, Int. J. Th. Phys. **39**, 2717 (2000).

[7] M. Horodecki and J. Oppenheim, Nature Comm. **4**, 2059 (2013).

[8] F. G. S. L. Brandao, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Phys. Rev. Lett. **111**, 250404 (2013).

[9] F. G. S. L. Brandao, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, PNAS **112**, 3275 (2015).

[10] M. Lostaglio, Reports on Progress in Physics **82**, 114001 (2019).

[11] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Rev. Mod. Phys. **79**, 555 (2007).

[12] I. Marvian and R. W. Spekkens, Phys. Rev. A **90**, 062110 (2014).

[13] I. Marvian and R. W. Spekkens, Nature Comm. **5**, 3821 (2014).

[14] E. Chitambar and G. Gour, Reviews of Modern Physics **91**, 025001 (2019).

[15] D. Jonathan and M. B. Plenio, Physical Review Letters **83**, 3566 (1999).

[16] S. Daftuar and M. Klimesh, Physical Review A **64**, 042314 (2001).

[17] W. van Dam and P. Hayden, Physical Review A **67**, 060302 (2003).

[18] G. Aubrun and I. Nechita, Communications in Mathematical Physics **278**, 133 (2007).

[19] M. Klimesh, "Inequalities that collectively completely characterize the catalytic majorization relation," (2007), arXiv:0709.3680v1.

[20] S. Turgut, J. Phys. A **40**, 12185 (2007).

[21] N. H. Y. Ng, L. Mančinska, C. Cirstoiu, J. Eisert, and S. Wehner, New J. Phys. **17**, 085004 (2015).

[22] P. Boes, J. Eisert, R. Gallego, M. P. Müller, and H. Wilming, Physical Review Letters **122**, 210402 (2019).

[23] C. Sparaciari, D. Jennings, and J. Oppenheim, Nature Communications **8** (2017), 10.1038/s41467-017-01505-4.

[24] P. Boes, R. Gallego, N. H. Y. Ng, J. Eisert, and H. Wilming, Quantum **4**, 231 (2020).

[25] C. E. Shannon, The Bell System Technical Journal **27**, 379 (1948).

[26] J. A. T. Thomas M. Cover, *Elements of Information Theory* (Wiley John + Sons, 2006).

[27] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2009).

[28] R. Gallego, J. Eisert, and H. Wilming, New J. Phys. **18**, 103017 (2016).

[29] H. Wilming, R. Gallego, and J. Eisert, Entropy **19**, 241 (2017).

[30] M. Lostaglio, M. P. Müller, and M. Pastena, Phys. Rev. Lett. **115**, 150402 (2015).

[31] M. P. Müller and M. Pastena, IEEE Trans. Inf. Th. **62**, 1711 (2016).

[32] M. P. Müller, Physical Review X **8**, 041051 (2018).

[33] S. Rethinasamy and M. M. Wilde, Physical Review Research **2**, 033455 (2020).

[34] N. Shiraishi and T. Sagawa, Physical Review Letters **126**, 150502 (2021).

[35] M. Fannes, Communications in Mathematical Physics **31**, 291 (1973).

[36] K. M. R. Audenaert, Journal of Physics A: Mathematical and Theoretical **40**, 8127 (2007).

[37] P. Kammerlander and R. Renner, (2020), arXiv:2002.08968 [math-ph].

[38] R. Alicki and M. Fannes, Physical Review E **87**, 042123 (2013).

[39] W. Pusz and S. Woronowicz, Commun. Math. Phys. **58**, 273 (1978).

[40] A. Lenard, J. Stat. Phys. **19**, 575 (1978).

[41] In [23] it was argued that by concatenating arbitrarily many catalytic transitions the bound in Corollary 3 can be achieved. However, concatenating two catalytic transitions builds up correlations between the catalysts and thus is in general not catalytic as a whole. Our main result shows that one can nevertheless always find a catalytic transformation implementing the two steps in one.

[42] T. Holenstein and R. Renner, IEEE Transactions on Information Theory **57**, 1865 (2011).

[43] P. Boes, H. Wilming, R. Gallego, and J. Eisert, Physical Review X **8**, 041016 (2018).

[44] T. V. Kondra, C. Datta, and A. Streltsov, Physical Review Letters **127**, 150502 (2021).

[45] P. Lipka-Bartosik and P. Skrzypczyk, Physical Review Letters **127**, 080502 (2021).

[46] P. Boes, N. H. Y. Ng, and H. Wilming, "The variance of relative surprisal as single-shot quantifier," arXiv:2009.08391.

[47] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer New York, 2011).

Appendix A: Proof-sketch of Lemma 4

We here provide a sketch of the proof of Lemma 4. For any $\delta > 0$ and state $\rho = \sum_i p_i \, |i\rangle\langle i|$ define the typical subspace $\Pi_\delta^{\rho^{\otimes n}}$ as the subspace spanned by those states $|i_1\rangle \otimes \cdots \otimes |i_n\rangle$ such that

$$\left| \frac{1}{n} \sum_{j=1}^n \log(1/p_{i_j}) - H(\rho) \right| < \delta. \quad \text{(A1)}$$

We further identify $\Pi_\delta^{\rho^{\otimes n}}$ with the projector onto the subspace. It follows from Hoeffding's inequality that $\rho^{\otimes n}$ is approximated by

$$\hat{\rho}_{\delta,n} := \frac{\Pi_\delta^{\rho^{\otimes n}} \rho^{\otimes n} \Pi_\delta^{\rho^{\otimes n}}}{\mathrm{Tr}[\Pi_\delta^{\rho^{\otimes n}} \rho^{\otimes n}]} \quad \text{(A2)}$$

with an error in trace-distance of the order $\exp(-n\delta^2)$ and this error is essentially optimal up to constants for general density matrices [42] (but of course may deviate significantly for particular density matrices). Note that constructing $\hat{\rho}_{\delta,n}$ simply consists of the following steps:

1. Take $\rho^{\otimes n}$ with eigenvalues of the form $q_{i_1 \ldots i_n} = p_{i_1} \cdots p_{i_n}$ and write it out in its eigenbasis.

2. Go through all diagonal elements $q_{i_1 \ldots i_n}$ and if

$$\left| -\frac{1}{n} \log(q_{i_1 \ldots i_n}) - H(\rho) \right| \geq \delta$$

   replace it with 0. Otherwise keep it.

3. Renormalize the density matrix.

Furthermore, the following properties are well-known [27]:
i) For any $\varepsilon > 0$ and sufficiently large n, we have

$$(1-\varepsilon)\mathrm{e}^{n(H(\rho)-\delta)} \leq \mathrm{Tr}[\Pi_\delta^{\rho^{\otimes n}}] \leq \mathrm{e}^{n(H(\rho)+\delta)}. \quad \text{(A3)}$$

ii) We have the operator inequality

$$\mathrm{e}^{-n(H(\rho)+\delta)}\Pi_\delta^{\rho^{\otimes n}} \leq \Pi_\delta^{\rho^{\otimes n}} \rho^{\otimes n} \Pi_\delta^{\rho^{\otimes n}} \leq \mathrm{e}^{-n(H(\rho)-\delta)}\Pi_\delta^{\rho^{\otimes n}}. \quad \text{(A4)}$$

iii) The approximation $\hat{\rho}_{\delta,n}$ is diagonal in the same basis as $\rho^{\otimes n}$, i.e. we can choose $\mathcal{D}_{\hat{\rho}_{\delta,n}} = \mathcal{D}_\rho^{\otimes n}$.

These conditions ensure that the approximation $\hat{\rho}_{\delta,n}$ majorizes the approximation $\hat{\rho}'_{\delta,n}$ if

$$\Delta H + \frac{\log(1-\epsilon)}{2n} > 2\delta. \quad \text{(A5)}$$

Thus, if we choose $\delta = \Delta H/4$ and n large enough, $\hat{\rho}_{\delta,n} \succeq \hat{\rho}'_{\delta,n}$. By the Schur-Horn theorem, there now exists a unitary U such that

$$\hat{\rho}'_{\delta,n} = \mathcal{D}_{\hat{\rho}'_{\delta,n}}[U\hat{\rho}_{\delta,n}U^\dagger] = \mathcal{D}_{\rho'}^{\otimes n}[U\hat{\rho}_{\delta,n}U^\dagger] =: \mathcal{C}[\hat{\rho}_{\delta,n}]. \quad \text{(A6)}$$

An explicit construction of U given $\hat{\rho}_{\delta,n}$ and $\hat{\rho}'_{\delta,n}$ is provided in the next section.

We now set

$$\rho'_{\epsilon,n} := \mathcal{C}[\rho^{\otimes n}] = \mathcal{D}_{\rho'}^{\otimes n}[U\rho^{\otimes n}U^\dagger]. \quad \text{(A7)}$$

Then, since the trace-distance is non-increasing under completely positive, trace-preserving maps and using the triangle inequality, we get

$$D\big(\rho'_{\epsilon,n}, \rho'^{\otimes n}\big) \leq D\big(\rho'_{\epsilon,n}, \hat{\rho}'_{\delta,n}\big) + D\big(\hat{\rho}'_{\delta,n}, \rho'^{\otimes n}\big) \quad \text{(A8)}$$

$$= D\big(\mathcal{C}[\rho^{\otimes n}], \mathcal{C}[\hat{\rho}_{\delta,n}]\big) + D\big(\hat{\rho}'_{\delta,n}, \rho'^{\otimes n}\big)$$

$$\leq D\big(\rho^{\otimes n}, \hat{\rho}_{\delta,n}\big) + D\big(\hat{\rho}'_{\delta,n}, \rho'^{\otimes n}\big) \quad \text{(A9)}$$

$$\leq O(\exp(-n\delta^2)) + O(\exp(-n\delta^2)) \quad \text{(A10)}$$

$$= O(\exp(-n\Delta H^2/4)). \quad \text{(A11)}$$

Note that in particular we have $\rho'_{\epsilon,n} = \mathcal{D}_{\rho'}^{\otimes n}[\rho'_{\epsilon,n}]$ and that the unitary U is the same unitary that is also used in the construction of the catalyst, i.e., for the state $\chi = U\rho^{\otimes n}U^\dagger$.

Appendix B: How to explicitly construct unitaries implementing majorization and the complexity of our constructions

We now discuss in more detail the explicit construction of the relevant unitary operator and catalyst state for a given catalytic transition $\rho \to \rho'$. Essentially without loss of generality we will assume that ρ and ρ' are diagonal in the same basis, since we can always adjust the final eigenbasis by a unitary without affecting the catalyst.

First, we remind the reader the construction of the catalyst and unitary given in the main text consists of two steps. In the first step, a catalyst and unitary is constructed whose effect is to bring the system to a state whose diagonal coincides with the final state. In the second step a second, uncorrelated part of the catalyst, which is maximally mixed, is used to decohere the system in the eigenbasis of the final state. The construction for the second step only depends on the eigenbasis of ρ' and consists of dilating the decoherence-channel to a unitary operation. An explicit and optimal (in terms of Hilbert-space dimension) construction can be found in Ref. [43] and I will omit discussing this second step in more detail.

For the first step, the state of the catalyst is given by:

$$\sigma_1 = \frac{1}{n} \sum_{k=1}^n \rho^{\otimes k-1} \otimes \chi_{1,\ldots,n-k} \otimes |k\rangle\langle k|_A \quad \text{(B1)}$$

$$= \frac{1}{n} \sum_{k=1}^n \rho^{\otimes k-1} \otimes \mathrm{Tr}_{n-k+1\ldots n}[U\rho^{\otimes n}U^\dagger] \otimes |k\rangle\langle k|_A. \quad \text{(B2)}$$

Note that the state σ_1 is constructed from three ingredients: the counting-register A, copies of the input state ρ and (reduced states of) the state $\chi = U\rho^{\otimes n}U^\dagger$. Here, U is the unitary that implements the majorization $\rho^{\otimes n} \succeq \rho'_{\epsilon,n}$ in the sense that χ coincides with $\rho'_{\epsilon,n}$ in the eigenbasis of $\rho'^{\otimes n}$. The construction of $\rho'_{\epsilon,n}$ is made explicit (up to the precise choice of n as

a function of ϵ) in the previous section. The final unitary that is applied to system and catalyst can be written as

$$W = P\left[U \otimes |n\rangle\langle n|_A + \sum_{k=1}^{n-1} \mathbf{1} \otimes |k\rangle\langle k|\right], \qquad \text{(B3)}$$

where P is the joint permutation matrix representing steps 2. and 3. in the main text.

It therefore remains to discuss how to explicitly construct the unitary U (which is the same as in the previous section). For the convenience of the reader, I now explain this construction for the general case of two density matrices $\omega \succeq \omega'$. To obtain U one then applies this procedure to $\hat{\rho}_{\delta,n} \succeq \hat{\rho}'_{\delta,n}$ (for sufficiently large n). That is, given that $\omega \succeq \omega'$ we now look for the unitary U such that $U\omega U^\dagger$ coincides on the diagonal with ω' (in the eigenbasis of ω'). For simplicity we assume again that ω and ω' are diagonal in the same basis $\{|i\rangle\}$ (otherwise we simply add a unitary rotation at the end), so that

$$\omega = \sum_i w_i |i\rangle\langle i|, \quad \omega' = \sum_i w'_i |i\rangle\langle i|. \qquad \text{(B4)}$$

We further assume that $w_1 \geq w_2 \geq \ldots$ and $w'_1 \geq w'_2 \geq \ldots$, which again can be ensured by a unitary transformation. The condition $\omega \succeq \omega'$ now is equivalent to the condition $\boldsymbol{w} \succeq \boldsymbol{w}'$ as mentioned in the main text, where $\boldsymbol{w} = (w_1, \ldots, w_d)^\top$ and $\boldsymbol{w}' = (w'_1, \ldots, w'_d)^\top$ are probability vectors. We now need to introduce the notion of a *T-transform* on a pair of indicies (j, k) as the stochastic matrix acting as

$$(T_{(j,k)}(t)\boldsymbol{w})_i = \begin{cases} w_i & \text{if} \quad i \neq j, k \\ tw_j + (1-t)w_k & \text{if} \quad i = j \\ tw_k + (1-t)w_j & \text{if} \quad i = k. \end{cases} \qquad \text{(B5)}$$

That is, the T-transform transposes two entries of a probability vector with probability $(1-t)$ and does nothing to the remaing entries. Note that for $t = 1$ the T-transform does nothing, while for $t = 0$ it interchanges the two entries, and for $t = 1/2$ it mixes them equally. Importantly, if $\boldsymbol{w} \succeq \boldsymbol{w}'$, then $\boldsymbol{w}'$ can be obtained by applying at most $d-1$ suitable T-transforms to the probability vector $\boldsymbol{w}$. Concretely, we can use the following algorithm involving a loop over a counting variable i which starts with $i = 1$ [47]:

1. If $i = 1$ set $\boldsymbol{w}^{(i)} := \boldsymbol{w}$.

2. If $\boldsymbol{w}^{(i)} = \boldsymbol{w}'$ stop. Otherwise continue:

3. Set $k^{(i)}$ to be the smallest integer such that $w'_{k^{(i)}} > w^{(i)}_{k^{(i)}}$ and $j^{(i)}$ the largest integer such that $w'_{j^{(i)}} < w^{(i)}_{j^{(i)}}$.

4. Set $\delta = \min\left(w^{(i)}_{j^{(i)}} - w'_{j^{(i)}}, w'_{k^{(i)}} - w^{(i)}_{k^{(i)}}\right)$ and $t^{(i)} = 1 - \delta/(w^{(i)}_{j^{(i)}} - w^{(i)}_{k^{(i)}})$.

5. Set $\boldsymbol{w}^{(i+1)} := T_{(j^{(i)}, k^{(i)})}(t^{(i)})\boldsymbol{w}^{(i)}$.

6. Increase $i \mapsto i + 1$. Go to step 1.

It is shown in Ref. [47] that the algorithm terminates in at most $d - 1$ steps. Collecting the different parameters $j^{(i)}, k^{(i)}$ and $t^{(i)}$, we find

$$\boldsymbol{w}' = T_{(j^{(d-1)}, k^{(d-1)})}(t^{(d-1)}) \cdots T_{(j^{(1)}, k^{(1)})}(t^{(1)})\boldsymbol{w}. \quad \text{(B6)}$$

For the benefit of the reader, below I list Wolfram Mathematica code that performs a single step of the loop. I.e., it calculates the parameters j, k, t to go one step from a vector $\boldsymbol{w}$ to a vector $\boldsymbol{q}$ and applies the corresponding T-transform to $\boldsymbol{w}$.

```
(* Helper Function: Returns position of last entry in list
     x for which pattern pat is true *)
LastPosition[x_, pat_] :=
   Length[x] + 1 - FirstPosition[Reverse[x],
       pat];

(* T-Transform applied to vector w*)
T[w_, {j_, k_, t_}] :=
   t w + (1 - t) Permute[w, Cycles[{{j, k}}]];

(* Computes values for {j,k,t} to make one step in
     majorization space from w to the direction of q *)
OneStep[w_, q_] := Module[{j, k, delta, t},
   If[w != q,
   k = FirstPosition[q - w, _?Positive][[1]];
   j = LastPosition[w - q, _?Positive][[1]];
   delta = Min[{w[[j]] - q[[j]], q[[k]] -
       w[[k]]}];
   t = 1 - delta/(w[[j]] - w[[k]]);
   {j, k, t}
   , w]
   ];

(* Apply T-transform to go to obtain next vector in loop *)
T[w, OneStep[w, q]]
```

Importantly, the T transform can be implemented unitarily as a rotation matrix on the subspace spanned by $|j\rangle$ and $|k\rangle$. I.e., if we consider the unitary matrix $U_{(j,k)}(t)$ that acts as

$$U_{(j,k)}(t)|i\rangle = \begin{cases} |i\rangle & \text{if} \quad i \neq j, k \\ \sqrt{t}\,|j\rangle + \sqrt{1-t}\,|k\rangle & \text{if} \quad i = j \\ \sqrt{t}\,|k\rangle - \sqrt{1-t}\,|j\rangle & \text{if} \quad i = k, \end{cases} \quad \text{(B7)}$$

then the density matrix $U_{(j,k)}(t)\omega U_{(j,k)}(t)^\dagger$ has as diagonal entries the vector $T_{(j,k)}(t)\boldsymbol{w}$. Correspondingly, we find that if we set

$$U = U_{(j^{(d-1)}, k^{(d-1)})}(t^{(d-1)}) \cdots U_{(j^{(1)}, k^{(1)})}(t^{(1)}), \quad \text{(B8)}$$

then the density matrix $U\omega U^\dagger$ has as as diagonal elements the vector

$$T_{(j^{(d-1)}, k^{(d-1)})}(t^{(d-1)}) \cdots T_{(j^{(1)}, k^{(1)})}(t^{(1)})\boldsymbol{w} = \boldsymbol{w}'. \quad \text{(B9)}$$

as we wished for. This finishes showing how to explicitly construct the majorization unitary.

Let us finally comment on the difficulty of numerically constructing classical descriptions of all the objects in practice for a given transition $\rho \to \rho'$. First it is clear that if the dimension

d of the Hilbert-space of ρ is very large, then even finding the sepctrum of ρ, and hence computing the entropy of ρ is computationally hard. For example, if ρ would be the thermal state of a quantum many-body system, specified by some local Hamiltonian, then computing either the spectrum of ρ or the entropy of ρ is typically completely intractable. However, this is not specific to the problem at hand but true for any problem that involves the spectrum or von Neumann entropy of the density matrix of a many-body system (and hence many problems in quantum information theory).

Nevertheless, even when d is not large, it is computationally difficult to compute the catalyst and unitary used in the proof in this paper. This is due to the fact that the construction of the catalyst involves a large number n of copies of ρ to construct the states $\hat{\rho}_{\delta,n}$ and $\hat{\rho}'_{\delta,n}$ from which U is computed. In other words, the Hilbert-space dimension of the catalyst constructed here grows exponentially with n (prohibiting to write down concrete matrix representations of all the involved objects) and taking large n is necessary to achieve good precision ϵ in general. It is important to realize, however, that the catalyst and unitaries constructed here will often not be the optimal ones. To illustrate this, consider the initial state ρ with eigenvalues $(1/2, 1/2, 0)$ and final state ρ' with eigenvalues $(2/3, 1/6, 1/6)$, both diagonal in the basis $\{|1\rangle, |2\rangle, |3\rangle\}$. It is possible to implement the transition $\rho \to \rho'$ *exactly* with a small catalyst of dimension $d_C = 2$, namely σ with eigenvalues $(2/3, 1/3)$ and eigenbasis $\{|1\rangle_C, |2\rangle_C\}$. The unitary U that implements this transition acts as

$$U|2\rangle \otimes |1\rangle_C = |1\rangle \otimes |2\rangle_C, \qquad (B10)$$
$$U|1\rangle \otimes |2\rangle_C = |2\rangle \otimes |1\rangle_C, \qquad (B11)$$
$$U|3\rangle \otimes |1\rangle_C = |2\rangle \otimes |2\rangle_C, \qquad (B12)$$
$$U|2\rangle \otimes |2\rangle_C = |3\rangle \otimes |1\rangle_C \qquad (B13)$$

and $U|i\rangle \otimes |j\rangle_C = |i\rangle \otimes |j\rangle_C$ on the remaining basis vectors. Nevertheless, the construction presented in this paper will only yield an approximate transition $\rho \mapsto \rho'_\epsilon$ with arbitrarily small ϵ provided that n (and hence the catalyst dimension) is large. It is therefore an interesting open problem to find *optimal* constructions in terms of Hilbert-space dimension of catalysts for given transitions $\rho \to \rho'$. Indeed, it is currently unknown whether in the case $H(\rho) < H(\rho')$ one can achieve perfect conversion with $\epsilon = 0$ with a finite-dimensional catalyst (see also Section D).

Appendix C: The classical result

In this section we formulate catalytic transitions in the classical setting and prove the corresponding characterization of Shannon entropy. We continue to speak of systems, and a system S is described by a probability vector $\boldsymbol{p} \in \mathbb{R}^{|S|}$, where $|S|$ denotes the number of distinct states of S. Joint states of two systems S and S' are described by probability vectors in $\mathbb{R}^{|S_1|} \otimes \mathbb{R}^{|S_2|} \simeq \mathbb{R}^{|S_1||S_2|}$. If $\boldsymbol{p}_{SS'} \in \mathbb{R}^{|S|} \otimes \mathbb{R}^{|S'|}$ is a probability vector describing a bipartite state, we continue to write

$$\boldsymbol{p}_S := \mathrm{Tr}_{S'}[\boldsymbol{p}_{SS'}] \qquad (C1)$$

for the marginal on S, whose entries are given by

$$(\boldsymbol{p}_S)_i = \sum_{j=1}^{|S'|} (\boldsymbol{p}_{SS'})_{i,j}. \qquad (C2)$$

Instead of unitary matrices acting on $\mathbb{C}^d$, we now continue permutation matrices π acting on $\mathbb{R}^d$ by re-shuffling the basis vectors:

$$(\pi\boldsymbol{p})_i = \sum_{j=1}^{d} \pi_{i,j} p_j, \qquad (C3)$$

where π is a permutation matrix and $\boldsymbol{p} = (p_1, \ldots, p_d)^\top$. For two vectors $\boldsymbol{p}, \boldsymbol{p}' \in \mathbb{R}^d$ we have $\boldsymbol{p} \succeq \boldsymbol{p}'$ if and only if there exists a probability distribution q_α of permutations $\pi^{(\alpha)}$ such that

$$\boldsymbol{p}' = \sum_{\alpha} q_\alpha \pi^{(\alpha)} \boldsymbol{p}. \qquad (C4)$$

The Shannon entropy of a probability vector $\boldsymbol{p} = (p_1, \ldots, p_d)^\top$ is defined as

$$H(\boldsymbol{p}) = -\sum_{j=1}^{d} p_j \log(p_j). \qquad (C5)$$

Finally, the trace-distance is replaced by the total variation distance given by

$$D(\boldsymbol{p}, \boldsymbol{p}') := \frac{1}{2} \sum_{j=1}^{d} |p_j - p'_j|. \qquad (C6)$$

A formal definition of catalytic transitions in the classical case can now be given in full analogy to the quantum case:

Definition 8 (Classical approximate catalytic transitions). *For two probability vectors $\boldsymbol{p}, \boldsymbol{p}' \in \mathbb{R}^{|S|}$ on a system S we write $\boldsymbol{p} \to_\epsilon \boldsymbol{p}'$ if there exists a finite-dimensional probability vector $\boldsymbol{q} \in \mathbb{R}^{|C|}$ on a system C and permutation matrix π on $\mathbb{R}^{|S|} \otimes \mathbb{R}^{|C|}$ such that*

$$\mathrm{Tr}_S[\pi\boldsymbol{p} \otimes \boldsymbol{q}] = \boldsymbol{q} \text{ and } D\big(\mathrm{Tr}_C[\pi\boldsymbol{p} \otimes \boldsymbol{q}], \boldsymbol{p}'\big) \leq \epsilon. \qquad (C7)$$

For completeness, we also state again the theorem from the main text we want to prove.

Theorem 9. *Let $\boldsymbol{p}, \boldsymbol{p}' \in \mathbb{R}^d$ be two probability vectors with Shannon entropies $H(\boldsymbol{p})$ and $H(\boldsymbol{p}')$, respectively. The following are equivalent:*

1. *For all $\epsilon > 0$ we have $\boldsymbol{p} \to_\epsilon \boldsymbol{p}'$.*

2. $H(\boldsymbol{p}) \leq H(\boldsymbol{p}')$.

We now prove Theorem 9. The implication i) $\Rightarrow$ ii) follows as before from additivity, sub-additivity, continuity as well as invariance under permutations of Shannon entropy. We therefore only need to prove the converse direction. As all the results about typicality transfer unchanged (up to notation) we
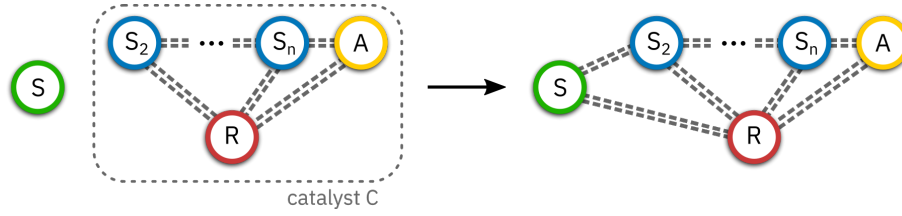
FIG. 4. The structure of the constructed catalyst C in the classical case: It contains subsystems $S_2, \ldots, S_n$, which are copies of the target system S together with an auxiliary system A and a catalytic source of randomness R. The dashed lines indicate possible correlations. The source of randomness is now utilized to fuel a random permutation on $S_1 \cdots S_n$ to make use of majorization. The catalyst and global permutation π are constructed in such a way that the correlations between all sub-systems of the catalyst remain invariant.

will freely use them in the proof. In particular Lemma 4 implies that for sufficiently large n we have

$$\boldsymbol{p}^{\otimes n} \succeq \boldsymbol{p}'_{\epsilon,n} \tag{C8}$$

with $D\big(\boldsymbol{p}'^{\otimes n}, \boldsymbol{p}'_{\epsilon,n}\big) \leq \epsilon$ and the same error scaling as before. Let $(q_\alpha, \pi^{(\alpha)})$ be the corresponding distribution over permutations required for the majorization in (C8). We again introduce $n-1$ systems $S_2 \cdots S_n$, an auxiliary system A with $|A| = n$ and a source of randomness R whose dimension now coincides with the number of permutations in the collection $(q_\alpha, \pi^{(\alpha)})$ (at most d^n by Caratheodory's theorem). Fig. 4 illustrates the structure of the catalyst. We denote the canonical basis vectors on A by $\boldsymbol{a}_k$ and those on R by $\boldsymbol{r}_\alpha$ for clarity of notation. Finally, define

$$\boldsymbol{x}^{(\alpha)}_{1,\ldots,i} := \operatorname{Tr}_{S_{i+1}\cdots S_n}\left[\pi^{(\alpha)} \boldsymbol{p}^{\otimes n}\right] \tag{C9}$$

with $\boldsymbol{x}^{(\alpha)}_0 = 1$ the trivial state. We then define the catalyst

state

$$\boldsymbol{q} := \frac{1}{n}\sum_{k=1}^{n}\sum_\alpha q_\alpha \boldsymbol{p}^{\otimes k-1} \otimes \boldsymbol{x}^{(\alpha)}_{1,\ldots,n-k} \otimes \boldsymbol{a}_k \otimes \boldsymbol{r}_\alpha \tag{C10}$$

and the permutation matrix on $S_1 \cdots S_n A R$ with $S = S_1$ given by

$$\pi := \sum_\alpha \pi^{(\alpha)} \otimes (\boldsymbol{a}_n \boldsymbol{a}_n^\top) \otimes (\boldsymbol{r}_\alpha \boldsymbol{r}_\alpha^\top) + \boldsymbol{1}_{S_1 S_2 \cdots S_n} \otimes P_A \otimes \boldsymbol{1}_R, \tag{C11}$$

where P_A denotes the orthonormal projector onto $\operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n-1}\}$. Note that $\boldsymbol{r}_\alpha \boldsymbol{r}_\alpha^\top$ is simply the orthonormal projector onto the span of $\boldsymbol{r}_\alpha$ and similarly for $\boldsymbol{a}_k \boldsymbol{a}_k^\top$. It is worth to take a pause to understand the structure of the catalyst state $\boldsymbol{q}$: If A is in state $\boldsymbol{a}_k$ it consists of $k-1$ copies of $\boldsymbol{p}$ together with the first $n-k$ marginals of the state $\boldsymbol{p}'_{\epsilon,n}$. However, these latter marginals are correlated with the source of randomness R whose marginal is such that it is in state α with probability q_α. Thus, the source of randomness is correlated with all the other systems, but in a well controlled way. The permutation π has the effect of applying $\pi^{(\alpha)}$ to the systems $S_1 \cdots S_n$ if system A is in state n and system R is in state α. If these conditions are not met, it does nothing.

The final procedure now again consists of three steps (see again Fig. 3 in the main text):

1. Apply π to the initial state $\boldsymbol{p} \otimes \boldsymbol{q}$, obtaining the state

$$\pi \boldsymbol{p} \otimes \boldsymbol{q} = \frac{1}{n}\sum_\alpha q_\alpha \pi^{(\alpha)} \boldsymbol{p}^{\otimes n} \otimes \boldsymbol{a}_n \otimes \boldsymbol{r}_\alpha + \frac{1}{n}\sum_{k=1}^{n-1}\sum_\alpha q_\alpha \boldsymbol{p}^{\otimes k} \otimes \boldsymbol{x}^{(\alpha)}_{1,\ldots,n-k} \otimes \boldsymbol{a}_k \otimes \boldsymbol{r}_\alpha \tag{C12}$$

2. Apply the cyclic shift $\mathbb{S}$ which acts as $S_k \mapsto S_{k+1}$ among the subsystems S_i.

3. Apply the cyclic shift $\boldsymbol{a}_k \mapsto \boldsymbol{a}_{k+1}$ on the system A.

Observe that all steps correspond to permutation matrices, and therefore their composition is also a permutation matrix. Noting that $\operatorname{Tr}_{S_1}[\mathbb{S} \cdot] = \operatorname{Tr}_{S_n}[\cdot]$, the final state on the catalyst is then given by

$$\mathrm{Tr}_{S_n}[\frac{1}{n}\sum_\alpha q_\alpha \pi^{(\alpha)}\boldsymbol{p}^{\otimes n}] \otimes \boldsymbol{a}_1 \otimes \boldsymbol{r}_\alpha + \frac{1}{n}\sum_{k=1}^{n-1}\sum_\alpha q_\alpha \mathrm{Tr}_{S_n}[\boldsymbol{p}^{\otimes k} \otimes \boldsymbol{x}^{(\alpha)}_{1,\dots,n-k}] \otimes \boldsymbol{a}_{k+1} \otimes \boldsymbol{r}_\alpha \tag{C13}$$

$$= \frac{1}{n}\sum_\alpha q_\alpha \boldsymbol{x}^{(\alpha)}_{1,\dots,n-1} \otimes \boldsymbol{a}_1 \otimes \boldsymbol{r}_\alpha + \frac{1}{n}\sum_{k=2}^{n}\sum_\alpha q_\alpha \mathrm{Tr}_{S_n}[\boldsymbol{p}^{\otimes k-1} \otimes \boldsymbol{x}^{(\alpha)}_{1,\dots,n-k+1}] \otimes \boldsymbol{a}_k \otimes \boldsymbol{r}_\alpha \tag{C14}$$

$$= \frac{1}{n}\sum_\alpha q_\alpha \boldsymbol{x}^{(\alpha)}_{1,\dots,n-1} \otimes \boldsymbol{a}_1 \otimes \boldsymbol{r}_\alpha + \frac{1}{n}\sum_{k=2}^{n}\sum_\alpha q_\alpha \boldsymbol{p}^{\otimes k-1} \otimes \boldsymbol{x}^{(\alpha)}_{1,\dots,n-k} \otimes \boldsymbol{a}_k \otimes \boldsymbol{r}_\alpha \tag{C15}$$

$$= \frac{1}{n}\sum_{k=2}^{n}\sum_\alpha q_\alpha \boldsymbol{p}^{\otimes k-1} \otimes \boldsymbol{x}^{(\alpha)}_{1,\dots,n-k} \otimes \boldsymbol{a}_k \otimes \boldsymbol{r}_\alpha \tag{C16}$$

$$= \boldsymbol{q}. \tag{C17}$$

Similarly, it is easy to check that the final state on the system S is given by

$$\mathrm{Tr}_{S_2\cdots S_n AR}[\pi \boldsymbol{p} \otimes \boldsymbol{q}] = \frac{1}{n}\sum_{k=1}^{n}\mathrm{Tr}_{\bar{k}}[\boldsymbol{p}'_{\epsilon,n}], \tag{C18}$$

where $\mathrm{Tr}_{\bar{k}}$ again denotes taking the marginal of subsystem S_k. Again we find

$$D\big(\boldsymbol{p}', \frac{1}{n}\sum_{k=1}^{n}\mathrm{Tr}_{\bar{k}}[\boldsymbol{p}'_{\epsilon,n}]\big) \leq \epsilon. \tag{C19}$$

This finishes the proof.

Appendix D: Size of the catalyst

It is interesting to ask how the construction of the catalyst scales in certain limiting cases. The dimension of the catalyst is essentially controlled by the number of copies n required to achieve a certain precision ϵ in the construction of the catalyst, which in turn only depends on the error for majorization from Lemma 4. This error leads to a scaling of the required number of copies n as

$$n = O\left(\frac{\log(1/\epsilon)}{\Delta H^2}\right) \tag{D1}$$

and hence the dimension of the catalyst diverges as $d_c = O(\exp(\log(1/\epsilon)/\Delta H^2))$. In particular, the catalyst dimension may in general diverge in a state transformation between two states with almost the same entropy. In Ref. [46] a lower bound for exact catalytic transitions was proven for state-transition with a fixed change of the *variance of surprisal* $V(\rho) = \mathrm{Tr}[\rho(-\log(\rho) - H(\rho))^2]$ but a small change of entropy. It was found that the dimension of the catalyst has to grow at least as as $O(\exp(\Delta H^{-1/8}))$, which is compatible with the above estimate. It is worth emphasizing, however, that these diverging catalysts sizes are only needed when ρ does not majorize ρ' already — if $\rho \succeq \rho'$ the catalyst can be very small, namely of dimension d in the classical case and $\sqrt{d}$ in the quantum case [43].