Extracting work from absence of correlations

Markus P. Müller[1] and Michele Pastena[1]
[1]Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 19, D-69120 Heidelberg, Germany
(Dated: December 2, 2014)

As Landauer's Principle and Szilard's engine illustrate, the reduction of a system's entropy costs work, while its increase can be used to extract work from a heat bath. One consequence in standard thermodynamics is that correlations have work value: the total entropy of two correlated systems is less than the sum of their local entropies, and thus work can be extracted if this correlation is consumed. In this paper, we show that the situation is surprisingly different for microscopic and quantum systems far from the thermodynamic limit: quite the contrary, *absence of correlations* can be used to extract work. Recently, it has been shown that the possible state transitions in the microscopic regime are severely limited by an infinite family of "second laws". We show that stochastic independence, if consumed as a resource, allows to overcome these limitations, to extract additional work reliably, and to achieve all state transitions that are otherwise only possible in the thermodynamic limit. Our result also yields an operational non-asymptotic characterization of von Neumann (neg-)entropy in terms of a majorization relation which generalizes the trumping relation from entanglement theory.

There has recently been a surge of interest in studying thermodynamics for small classical and quantum systems beyond the thermodynamic limit [1–10]. It has been shown that thermodynamics in the microscopic regime differs significantly from its macroscopic counterpart: for example, the work extractable from a quantum state is in general smaller than the work needed to create that state [3]; additional quantum systems can be used as "catalysts" to facilitate state transitions; and the second law is replaced by an infinite family of constraints that must be satisfied by any spontaneous process [4].

In this work, we demonstrate another crucial difference between the two regimes, namely, that *stochastic independence* (that is, *absence* of correlations) can be used to extract work in the microscopic regime. Quite the contrary, this is impossible in the thermodynamic limit, that is, the regime described by the standard laws of thermodynamics. To see this, imagine a large ensemble of quantum systems, each of them in a quantum state ρ_{AB}, on a tensor product of Hilbert spaces A, B of dimensions $d_A, d_B < \infty$. If this state is correlated, that is $\rho_{AB} \neq \rho_A \otimes \rho_B$, then

$$I(\rho_{AB}) > I(\rho_A) + I(\rho_B), \qquad (1)$$

with $I(\rho) := \log d_\rho - H(\rho)$ the *negentropy* of ρ, where $H(\rho) := -\mathrm{tr}(\rho \log \rho)$ is von Neumann entropy and d_ρ the Hilbert space dimension. It is well-known [11] that a quantum state whose negentropy is $I(\rho)$ (with all energy levels fully degenerate) can be used to draw $I(\rho) k_B T$ of average work per particle from a heat bath at temperature T, where k_B is Boltzmann's constant. Likewise, this amount of work has to be spent to create a large number of systems in state ρ from maximally mixed states. These insights are illustrated by the well-known thought experiments of Landauer erasure and the Szilard engine. Thus, the subadditivity of von Neumann entropy, and its consequence (1), show that we can always draw more work ($I(\rho_{AB}) k_B T$) from a correlated state than from the corresponding uncorrelated state ($(I(\rho_A) + I(\rho_B)) k_B T$).

In other words, if we have a thermodynamic transition from a product state $\rho_A \otimes \rho_B$ to a correlated state ρ_{AB}, while keeping the local reduced states constant, then this transition must always consume work, which manifests itself in several research results on the impact of correlations in thermodynamics [12–14]. However, we will now show that the analogous process in the microscopic regime, as depicted in Figure 1, allows for a work gain.

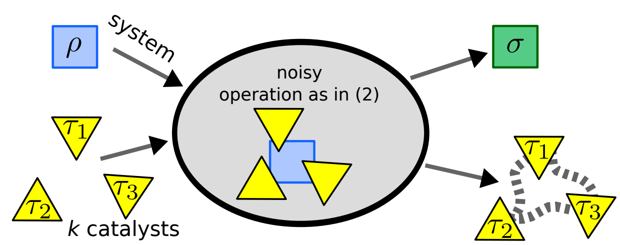

Figure 1: We model the use of stochastic independence as a "fuel" by considering state transitions $\rho \to \sigma$ in conjunction with k "catalysts", initially uncorrelated, in states $\tau_1, \ldots, \tau_k$. After the thermodynamic process, the catalysts must be retained exactly in their initial states, but we allow correlations to build up, resulting in a joint state $\tau_{1,2,\ldots,k}$ that still has $\tau_1, \ldots, \tau_k$ as local marginals. This is a non-cyclic process, consuming stochastic independence as a resource. It turns out that state transitions $\rho \to \sigma$ are possible to arbitrary accuracy if and only if $H(\rho) \leq H(\sigma)$ for the von Neumann entropy.

Thermodynamics as a resource theory. A fruitful method to study the thermodynamics of small quantum systems is to reformulate it as a resource theory. This approach, which has also been applied to entanglement theory [15], models the fact that there is often only a limited amount of control that is in principle possible when interacting with a quantum system. Given the corresponding set of restrictions, one can ask for the ultimate limits of state transitions that are possible within the set of allowed operations, and answer general questions about the interconvertibility of resources.

More concretely, we work within the *resource theory of nonuniformity* [16, 17]. Given any quantum system A in the state ρ_A, one is allowed to add on a maximally mixed state $\gamma_E = \mathbf{1}/d_E$ on any ancilla quantum system E, apply any unitary U on the total system, and trace out arbitrary subsystems. The possible operations can then all be written in the form

$$\mathcal{E}(\rho) := \mathrm{Tr}_{E'}\left[U_{AE}\left(\rho_A \otimes \gamma_E\right)U_{AE}^{\dagger}\right], \qquad (2)$$

where $A \otimes E = A' \otimes E'$ denotes two decompositions of the total Hilbert space, such that $\mathcal{E}$ is a map from quantum system A to A'. Maps of the form (2) are called *noisy operations*. Energy is not directly modelled within this resource theory; all Hamiltonians are implicitly assumed to be fully degenerate. However, as suggested in [5, 6], one can still consider work extraction indirectly by first extracting as many pure qubits as possible via maps of the form (2), and then using these pure qubits in a Szilard engine (or in ways described in [11]) to extract work from a heat bath.

This can be quantified by considering *sharp states* [17] σ_I, which for $I \in \mathbb{N}$ are just a collection of I pure qubits: $\sigma_I = |\psi\rangle\langle\psi|^{\otimes I}$, with $|\psi\rangle \in \mathbb{C}^2$ an arbitrary pure state. If I is not an integer, but of the form $I = \log(\ell/k)$ with $k, \ell \in \mathbb{N}$ coprime, $\ell \geq k$, and $\log \equiv \log_2$, we set

$$\sigma_I = \mathrm{diag}(\underbrace{1/k, \ldots, 1/k}_{k}, \underbrace{0, \ldots, 0}_{\ell-k}),$$

and one can draw $Ik_BT\ln 2$ of work from a sharp state σ_I, which can be done reliably with success probability close to unity [18]. Let us write $\rho \xrightarrow{\text{noisy}} \sigma$ if there exists a noisy operation $\mathcal{E}$ as in (2) such that $\mathcal{E}(\rho) = \sigma$. For work extraction, we are thus asking for the largest I with $\rho \xrightarrow{\text{noisy}} \sigma_I$. The answer turns out to be

$$I_0(\rho) := \log d_\rho - H_0(\rho), \qquad (3)$$

where $H_0(\rho) := \log \mathrm{rank}(\rho)$. This is in general *smaller* then the (von Neumann) negentropy $I(\rho)$ – therefore, the amount of work that can reliably be extracted from a single copy of ρ, $I_0(\rho)k_BT\ln 2$, is in general less than the amount of work per particle that can be extracted in the thermodynamic limit, which is $I(\rho)k_BT\ln 2$. If one allows a small probability $\varepsilon > 0$ of failure for the work extraction process, then $I_0(\rho)$ has to be replaced by a suitably "smoothed" [19] version $I_0^\varepsilon(\rho)$, basically maximizing I_0 over states in an ε-close vicinity of ρ (for details see Fig. S2 in [3] and Lemma 79 and Fig. 9 in [17]). For small $\varepsilon > 0$, the extractable work $I_0^\varepsilon(\rho)$ will generally still be much smaller than $I(\rho)$. For several concrete examples of states, see [5].

A general way to decide whether $\rho \xrightarrow{\text{noisy}} \sigma$ for given states ρ, σ is via majorization [20–23]. For classical probability distributions $p = (p_1, \ldots, p_m)$ and $q = (q_1, \ldots, q_m)$, we say that p majorizes q, and write $p \succ q$, if and only if $\sum_{i=1}^k p_i^\downarrow \geq \sum_{i=1}^k q_i^\downarrow$ for all $k = 1, 2, \ldots, m$, where $p_1^\downarrow \geq p_2^\downarrow \geq \ldots$ denotes the components of p in non-increasing order. For quantum states ρ and σ, we write $\rho \succ \sigma$ if and only if $\lambda(\rho) \succ \lambda(\sigma)$, where $\lambda(\rho)$ and $\lambda(\sigma)$ are the probability distributions of eigenvalues of ρ and σ. As shown in [16], we have $\rho \xrightarrow{\text{noisy}} \sigma$ for states ρ, σ of the same Hilbert space dimensions if and only if $\rho \succ \sigma$. If ρ and σ have different dimensionalities, we can multiply them with suitably sized maximally mixed states such that the resulting states have the same size, and check majorization for those. From this, one can easily prove (3).

In addition to the previously specified allowed transformations in the resource theory of nonuniformity, one usually adds one further possibility: namely, to have an additional quantum state ("catalyst") τ that is part of the process, but not altered by it [4]. In other words, instead of asking whether there is a noisy operation $\mathcal{E}$ with $\mathcal{E}(\rho) = \sigma$, one asks whether there is an operation of this kind and a state τ such that $\mathcal{E}(\rho \otimes \tau) = \sigma \otimes \tau$.

For the special case of equal dimensionalities of ρ and σ, we say that ρ *trumps* σ, and write $\rho \succ_T \sigma$, if there is a finite-dimensional quantum state τ such that $\rho \otimes \tau \succ \sigma \otimes \tau$. In [21], it was shown that there are states such that $\rho \not\succ \sigma$ but $\rho \succ_T \sigma$; in other words, in some cases, the presence of an additional catalyst τ enables transitions $\rho \to \sigma$ that are otherwise impossible.

Catalysis does not increase the amount of extractable work [4] as given in (3), but it expands the set of possible state transitions. To characterize those, we define the Burg entropy [24] $H_{\mathrm{Burg}}(\rho) := \mathrm{tr}\log\rho$, and the Rényi entropies of order $\alpha \in \mathbb{R} \setminus \{0, 1\}$ as

$$H_\alpha(\rho) := \frac{\mathrm{sgn}(\alpha)}{1-\alpha}\log\mathrm{tr}(\rho^\alpha).$$

The cases $\alpha \in \{-\infty, 0, 1, +\infty\}$ are defined via suitable limits, resulting in $H_0(\rho) = \lim_{\alpha\searrow 0} H_\alpha(\rho) = \log\mathrm{rank}(\rho)$, $H_1(\rho) = -\mathrm{tr}(\rho\log\rho) = H(\rho)$, $H_\infty(\rho) = -\log\lambda_{\max}$ and $H_{-\infty} = \log\lambda_{\min}$, where $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalues of ρ. Analogous definitions apply to classical discrete probability distributions if we interpret them as diagonal density matrices. As shown in [22, 23], these entropies fully characterize the possible state transitions in the presence of a catalyst: *Suppose ρ and σ are quantum states with $\lambda^\downarrow(\rho) \neq \lambda^\downarrow(\sigma)$. Then $\rho \succ_T \sigma$ if and only if $H_{\mathrm{Burg}}(\rho) < H_{\mathrm{Burg}}(\sigma)$ and $H_\alpha(\rho) < H_\alpha(\sigma)$ for all $\alpha \in \mathbb{R} \setminus \{0\}$.*

This can be interpreted as an "infinite family of second laws [4]": not only von Neumann entropy H, but all the Rényi and Burg entropies must increase during all possible state transitions.

Main results. We now return to the question of the work value of correlations between states. As depicted in Figure 1, we consider k independent states $\tau_1, \ldots, \tau_k$ that are collectively in a product state, hence uncorrelated. We try to obtain a transition from a given state ρ to another state σ (which may be a sharp state used for work extraction) within the resource theory of nonuniformity. However, we additionally allow that the "catalysts" τ_i become correlated with each other during the process. That is, these states may collectively go from $\tau_1 \otimes \ldots \otimes \tau_k$ to a correlated state $\tau_{1,\ldots,k}$, while exactly retaining their individual states, such that $\tau_{1,\ldots,k}$ still has local reduced states $\tau_1, \ldots, \tau_k$.

We have previously argued by means of (1) that this transition necessarily *costs* energy in the thermodynamic limit, but we will now see that this is not so in the microscopic regime. In the special case of equal dimensionalities of ρ and σ, the transition in Figure 1 is possible if and only if

$$\rho \otimes \left(\tau_1 \otimes \ldots \otimes \tau_k\right) \succ \sigma \otimes \tau_{1,2,\ldots,k}. \tag{4}$$

Let us say that that ρ *c-trumps* σ, and write $\rho \succ_c \sigma$, if and only if there exists $k \in \mathbb{N}_0$ and a k-partite quantum state $\tau_{1,2,\ldots,k}$ such that (4) holds. Trumping is a special case of c-trumping with $k = 1$, and majorization is a special case with $k = 0$. Clearly, c-trumping is transitive: $\rho \succ_c \sigma$ and $\sigma \succ_c \omega$ implies that $\rho \succ_c \omega$. We also have

$$\rho \succ \sigma \;\;\Rightarrow\;\; \rho \succ_T \sigma \;\;\Rightarrow\;\; \rho \succ_c \sigma.$$

If there are pairs of states ρ and σ such that $\rho \not\succ_T \sigma$ but $\rho \succ_c \sigma$, then this shows that allowing correlations to build up in the catalysts enables state transitions that are otherwise impossible. Our main result shows that this is indeed the case:

Theorem 1. *Suppose that ρ and σ do not have identical sets of eigenvalues. Then $\rho \succ_c \sigma$ if and only if $\mathrm{rank}(\rho) \le \mathrm{rank}(\sigma)$ and $H(\rho) < H(\sigma)$, for H the von Neumann entropy.*

Moreover, we can always choose $k = 3$ in (4).

It is easy to see that the inequalities for rank and H are necessary for c-trumping: the Rényi entropies H_α are subadditive [25, 26] only in the two cases $\alpha = 0$ and $\alpha = 1$. Using subadditivity together with additivity and Schur concavity [17, 20] of H_α, (4) implies for $\alpha \in \{0, 1\}$

$$\begin{aligned} H_\alpha(\rho) + \sum_{i=1}^k H_\alpha(\tau_i) \;&\le\; H_\alpha(\sigma) + H_\alpha(\tau_{1,2,\ldots,k}) \\ &\le\; H_\alpha(\sigma) + \sum_{i=1}^k H_\alpha(\tau_i). \end{aligned}$$

For $\alpha = 0$ this shows that $\mathrm{rank}(\rho) \le \mathrm{rank}(\sigma)$, and for $\alpha = 1$ we get $H(\rho) \le H(\sigma)$, where equality could hold only if we had $\tau_{1,2,\ldots,k} = \tau_1 \otimes \ldots \otimes \tau_k$, which would imply $\rho \succ_T \sigma$ and consequently $H(\rho) < H(\sigma)$.

Before we turn to the complete proof, we discuss some physical implications. Since the rank is discontinuous, the rank inequality itself is not physically meaningful: we can always approximate σ to arbitrary accuracy by another full-rank state σ'. The crucial inequality is $H(\rho) \le H(\sigma)$, which resembles the usual second law of thermodynamics: entropy has to grow.

Thus, having access to stochastically independent additional states that are allowed to become correlated removes the infinite family of "second laws" given by the Burg and Rényi entropies, and replaces them by a condition that resembles the standard second law which otherwise only governs the thermodynamic limit. Moreover, it allows to extract additional work:

Theorem 2. *Dispensing stochastic independence as an additional resource allows to extract $I(\rho) k_B T \ln 2$ of work from a single copy of a given state ρ reliably by a process depicted in Figure 1. That is, building up additional correlations allows to outperform the standard law of microscopic work extraction in (3).*

Proof. We want a transition from ρ to a sharp state σ_I with $I = \log(\ell/k)$ as large as possible. Choose $i, j \in \mathbb{N}$ such that $d_\rho \cdot i = \ell \cdot j$, then $\rho \otimes \gamma_i$ and $\sigma_I \otimes \gamma_j$, with $\gamma_j = \mathbf{1}/j$ the maximally mixed state on $\mathbb{C}^j$, have the same dimensionalities. We can go from ρ to $\tilde{\sigma}_I$ arbitrarily close to σ_I by a process as in Figure 1 if and only if $\rho \otimes \gamma_i$ c-trumps $\tilde{\sigma}_I \otimes \gamma_j$. According to Theorem 1, this is equivalent to $H(\rho \otimes \gamma_i) \le H(\sigma_I \otimes \gamma_j)$, or $\log(d_\rho i) - I(\rho) \le \log(\ell j) - I$. Thus, we can achieve $I = I(\rho)$ but not more. $\square$

As Bennett [27] suggested, and as formulated in the resource theory of nonuniformity, we can think of an "information battery" in the form of a reservoir of pure bits which may be used for work extraction. Our result adds a twist to this idea: it is not only the purity of bits that can be used for work extraction, but also absence of correlations between them. Stochastic independence can be used as a "fuel"; but for this, it is necessary that the bits or catalysts $\tau_1, \ldots, \tau_k$ themselves are neither fully thermalized (that is, maximally mixed) nor completely pure, as we show in Appendix C. Quite surprisingly, since $I(\tau_1 \otimes \ldots \otimes \tau_k) \le I(\tau_{1,\ldots,k})$, a process like this necessarily *increases* the purity of the "stochastic independence battery" as measured by von Neumann negentropy, in contrast to the original proposal where the bits dispense purity and become more mixed.

Theorem 1 also provides an operational characterization of negentropy, or rather of its order of comparison on density matrices: for two states ρ, σ we have $I(\rho) \ge I(\sigma)$ if and only if σ can be obtained to arbitrary accuracy from ρ by a process as given in Figure 1.

Sketch of proof of Theorem 1. We have to show that $\mathrm{rank}(\rho) \le \mathrm{rank}(\sigma)$ and $H(\rho) < H(\sigma)$ implies the existence of a catalyst $\tau_{1,\ldots,k}$ satisfying (4). Since majorization depends only on the states' eigenvalues, and since unitaries are free operations in the resource theory of nonuniformity, we can diagonalize all states and work with classical probability distributions. Thus, it is sufficient to prove the following theorem, where we use the notation $\mathrm{rank}(p)$ for the number of non-zero entries of the probability distribution $p = (p_1, \ldots, p_m)$.

Theorem 3. *Let $p, q \in \mathbb{R}^m$ be probability distributions with $p^\downarrow \ne q^\downarrow$. Then there exists $k \in \mathbb{N}_0$ and a finite k-partite* *probability distribution* $r_{1,2,\ldots,k}$ *such that*

$$p \otimes (r_1 \otimes r_2 \otimes \ldots \otimes r_k) \succ q \otimes r_{1,2,\ldots,k}$$

if and only if $\mathrm{rank}(p) \leq \mathrm{rank}(q)$ *and* $H(p) < H(q)$.

The given inequalities are necessary by the same argument as in the quantum case. Sufficiency will be established by two lemmas that will both use the following notation. If $q = q_A \in \mathbb{R}^m$ is any probability distribution on A, we introduce a second system B of size $n+1$ and a joint probability distribution $q_{AB} \in \mathbb{R}^{m(n+1)}$ by

$$q_{AB} := \begin{pmatrix} q_1 - a_1 & \frac{a_1}{n} & \frac{a_1}{n} & \ldots & \frac{a_1}{n} \\ q_2 - a_2 & \frac{a_2}{n} & \frac{a_2}{n} & \ldots & \frac{a_2}{n} \\ \vdots & \vdots & \vdots & & \vdots \\ q_m - a_m & \frac{a_m}{n} & \frac{a_m}{n} & \ldots & \frac{a_m}{n} \end{pmatrix}, \qquad (5)$$

where $0 \leq a_i \leq q_i$ and $n \in \mathbb{N}$, using the matrix notation $(q_{AB})_{i,j} = q_{AB}(i,j)$. The marginal on A is given by summing over the rows, yielding $q_A = (q_1, \ldots, q_m) = q$, and the marginal on B is obtained by summing over the columns,

$$q_B = \left(1 - a, \frac{a}{n}, \ldots, \frac{a}{n}\right) \in \mathbb{R}^{n+1},$$

where $a = \sum_{i=1}^m a_i$. This notation will allow for a crucial observation on catalysis and correlation: there are distributions p, q such that $p \not\succ_T q$, i.e. there is no catalyst r with $p_A \otimes r_B \succ q_A \otimes r_B$, but nevertheless

$$p_A \otimes q_B \succ_T q_{AB}, \qquad (6)$$

such that there is another system C and catalyst c_C with $p_A \otimes (q_B \otimes c_C) \succ q_{AB} \otimes c_C$. That is, a "catalytic" transition from p to q on A is possible and allows to retrieve the total catalyst $q_B \otimes c_C$ unaltered as the marginal on BC, but in correlation with A. An example is given by $p = \left(\frac{91}{100}, \frac{1}{20}, \frac{1}{25}\right)$, $q = \left(\frac{17}{20}, \frac{7}{50}, \frac{1}{100}\right)$ and q_{AB} as in (5) with $n = 1$ and all $a_i = \frac{1}{120}$. In $m = 3$ dimensions, majorization and trumping are equivalent [21]; thus $p_1 + p_2 < q_1 + q_2$ implies $p \not\succ_T q$, but one can verify the Rényi and Burg entropy conditions for (6).

The following lemma shows that we can get many more examples of this kind by checking the Rényi entropies of orders $\alpha \geq 1$, and may thus be of independent interest in itself:

Lemma 4. *Let* $p, q \in \mathbb{R}^m$ *be distributions such that* q *has full rank,* $q \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$ *and* $H_\alpha(p) < H_\alpha(q)$ *for all* $\alpha \in [1, +\infty]$. *Then there is some* $a > 0$ *and* $n \in \mathbb{N}$ *such that* $p_A \otimes q_B \succ_T q_{AB}$, *where* $a_i := a/m$ *in (5).*

In the appendix, we prove this statement by showing that $H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB})$ is true for all $\alpha \neq 0$ in the limit $n \to \infty$ (and similarly for H_{Burg}), and by applying a standard uniformity argument which shows that we can actually find a fixed finite $n \in \mathbb{N}$ such that this inequality is simultaneously true for all α.

The same proof strategy can be applied to obtain the following lemma:

Lemma 5. *Let* $p, q \in \mathbb{R}^m$ *be distributions such that* q *has full rank,* $H(p) < H(q)$, *and* $q \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$. *Then there exists* $\delta > 0$ *and* $n \in \mathbb{N}$ *such that* $H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB})$ *for all* $\alpha \in [1, +\infty]$, *where* $a_i := q_i - \delta$ *in (5).*

In order to prove Theorem 3, we may assume that $q \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$, and we remove common zeroes from p and q until q has full rank. Lemma 5 gives us an extension q_{AB} of $q = q_A$ such that $H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB})$ for all $\alpha \in [1, +\infty]$. Applying Lemma 4 to these two states yields an extension q_{ABC} of q_{AB} such that $p_A \otimes q_B \otimes q_C \succ_T q_{ABC}$, and thus a catalyst c_D such that $p_A \otimes q_B \otimes q_C \otimes c_D \succ q_{ABC} \otimes c_D$. Majorization is preserved by tensor products [6], hence we can multiply this relation on both sides with another copy $q_E = q = q_A$ of q. Relabeling $A \leftrightarrow E$ on the right-hand side does not change the entries of the total probability vector and the majorization order, hence

$$p_A \otimes (q_E \otimes q_B \otimes q_C \otimes c_D) \succ q_A \otimes (q_{EBC} \otimes c_D).$$

This proves Theorem 3. Regarding CD as a single system, we see that we have $k = 3$ catalysts in total.

Conclusions. We have shown that stochastic independence can be used as a "fuel" to extract additional work reliably from small quantum systems, with the von Neumann entropy characterizing the possible state transitions. This is in contrast to the thermodynamic limit, where building up correlations always has a positive work cost. We thus provide another example of a resource which, similarly as e.g. quantum coherence [28], has crucial impact on thermodynamics in the microscopic regime, but whose effect vanishes in the thermodynamic limit.

From the mathematical side, we have defined a majorization relation which generalizes majorization and trumping, and which turns out to uniquely characterize Shannon and von Neumann negentropies. This definition captures the idea, detailed in several lemmas and examples, that catalytic state transitions can be enhanced by allowing correlations to build up.

Our work opens up a number of interesting questions. Are $k = 2$ catalysts always enough for c-trumping? Can we give any bound on their sizes, similarly as for standard catalysis [29]? Is there a generalization of our result to the resource theory of athermality, where quantum states are allowed to carry non-trivial Hamiltonians? It also seems worthwhile to look for concrete physical situations where local states τ_i of large quantum systems, interacting with other systems in a heat bath, are forced to remain constant (say, due to local conservation laws). Our result suggests that there could be a tendency to build up correlations, similarly as there is a tendency to thermalize if the purity of the local states is allowed to decrease. This is particularly interesting due to the fact that the transition from product to correlated states is often regarded as an instance of an arrow of time.

Acknowledgments. We are grateful to Fernando

Brandão, Gilad Gour, and Varun Narasimhachar for comments.

[1] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth, *Thermodynamic Cost of Reliability and Low Temperatures: Tightening Landauer's Principle and the Second Law*, Int. J. Theor. Phys. **39**(12), 2717–2753 (2000).

[2] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, *The Resource Theory of Quantum States Out of Thermal Equilibrium*, Phys. Rev. Lett. **111**, 250404 (2013).

[3] M. Horodecki and J. Oppenheim, *Fundamental limitations for quantum and nanoscale thermodynamics*, Nat. Comm. **4**, 2059 (2013).

[4] F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, *The second laws of quantum thermodynamics*, arXiv:1305.5278.

[5] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, *Inadequacy of von Neumann entropy for characterizing extractable work*, New J. Phys. **13**, 053015 (2011).

[6] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, *A Quantitative Landauer's Principle*, arXiv:1211.1037.

[7] J. Åberg, *Truly work-like work extraction via single-shot analysis*, Nat. Comm. **4**, 1925 (2013).

[8] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, *Laws of thermodynamics beyond the von Neumann regime*, arXiv:1207.0434.

[9] P. Skrzypczyk, A. J. Short, and S. Popescu, *Extracting work from quantum systems*, arXiv:1302.2811.

[10] P. Skrzypczyk, A. J. Short, and S. Popescu, *Thermodynamics for individual quantum systems*, arXiv:1307.1558.

[11] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, *Thermodynamics of quantum informational systems – Hamiltonian description*, Open Syst. Inf. Dyn. **11**(3), 205–217 (2004).

[12] W. H. Zurek, *Quantum discord and Maxwell's demons*, Phys. Rev. A **67**, 012320 (2003).

[13] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, *The thermodynamic meaning of negative entropy*, Nature **474**, 61–63 (2011).

[14] M. Perarnau-Llobet, K. V. Hovhannisyan, M. Huber, P. Skrzypczyk, N. Brunner, and A. Acín, *Extractable work from correlations*, arXiv:1407.7765.

[15] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys. **81**, 865 (2009).

[16] M. Horodecki, P. Horodecki, and J. Oppenheim, *Reversible transformations from pure to mixed states and the unique measure of information*, Phys. Rev. A **67**, 062104 (2003).

[17] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern, *The resource theory of informational nonequilibrium in thermodynamics*, arXiv:1309.6586.

[18] C. Browne, A. J. P. Garner, O. C. O. Dahlsten, and V. Vedral, *Guaranteed Energy-Efficient Bit Reset in Finite Time*, Phys. Rev. Lett. **113**, 100603 (2014).

[19] M. Tomamichel, *A Framework for Non-Asymptotic Quantum Information Theory*, PhD thesis, ETH Zurich, 2012.

[20] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer, 2011.

[21] D. Jonathan and M. B. Plenio, *Entanglement-Assisted Local Manipulation of Pure Quantum States*, Phys. Rev. Lett. **83**, 3566–3569 (1999).

[22] M. Klimesh, *Inequalities that Collectively Completely Characterize the Catalytic Majorization Relation*, arXiv:0709.3680.

[23] S. Turgut, *Necessary and Sufficient Conditions for the Trumping Relation*, J Phys. A: Math. Theor. **40**, 12185–12212 (2007).

[24] J. P. Burg, *Maximum entropy spectral analysis*, in: *Proceedings 37th Annual Meeting of the Society of Exploration Geophysicists*, Oklahoma City, OK (1967).

[25] J. Aczél, Z. Daróczy, *On Measures of Information and Their Characterizations*, Academic Press, New York, 1975.

[26] N. Linden, M. Mosonyi, and A. Winter, *The structure of Rényi entropic inequalities*, Proc. R. Soc. A **469**(2158), 20120737 (2013).

[27] C. H. Bennett, *The Thermodynamics of Computation – a Review*, Int. J. Theor. Phys. **21**(12), 905–940 (1982).

[28] M. Lostaglio, D. Jennings, and T. Rudolph, *Thermodynamic laws beyond free energy relations*, arXiv:1405.2188.

[29] N. H. Y. Ng, L. Mančinska, C. Cirstoiu, J. Eisert, and S. Wehner, *Limits to catalysis in quantum thermodynamics*, arXiv:1405.3039.

[30] T. van Erven and P. Harremoës, *Rényi Divergence and Kullback-Leibler Divergence*, IEEE Trans. Inf. Th. **60**(7), 3797–3820 (2014).

[31] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley & Sons, New Jersey, 2006.

Appendix A: Basic definitions and notation

A *probability distribution* is a vector $p \in \mathbb{R}^m$ with entries $p_i \geq 0$ such that $\sum_{i=1}^m p_i = 1$. The *rank* of a probability distribution, $\mathrm{rank}(p)$, is defined as the number of nonzero entries of p. Thus $1 \leq \mathrm{rank}(p) \leq m$, and we say that p has "full rank" if $\mathrm{rank}(p) = m$. In what follows, "$\log$" denotes the natural logarithm, such that $\exp(\log(x)) = x$ for all $x \in \mathbb{R}$ (in contrast to the main text, we are not using the binary logarithm in the appendix). We say that a function $f : I \to \mathbb{R}$ with $I \subset \mathbb{R}$ is *increasing* if $x < y \Rightarrow f(x) \leq f(y)$ for all $x, y \in I$, and that it is *strictly increasing* if $x < y \Rightarrow f(x) < f(y)$ (analogous definitions apply to *decreasing* / *strictly decreasing*).

We start by defining the *Rényi entropies*, following the conventions of [4], as well as the *Burg entropy* [24].

Definition 6. *Let $p \in \mathbb{R}^m$ be any probability distribution. For $\alpha \in \mathbb{R} \setminus \{0,1\}$, we define the* Rényi entropy of order α *as*

$$H_\alpha(p) := \frac{\operatorname{sgn}(\alpha)}{1-\alpha} \log \sum_{i=1}^m p_i^\alpha,$$

and for $\alpha \in \{-\infty, 0, 1, +\infty\}$, we set

$$H_0(p) := \log \operatorname{rank}(p), \qquad H_1(p) \equiv H(p) = -\sum_{i=1}^m p_i \log p_i, \qquad H_\infty(p) := -\log \max_i p_i, \qquad H_{-\infty}(p) := \log \min_i p_i.$$

The Burg entropy is defined as

$$H_{\mathrm{Burg}}(p) := \sum_{i=1}^m \log p_i.$$

If p does not have full rank then $H_\alpha(p) = H_{\mathrm{Burg}}(p) = -\infty$ for all $\alpha < 0$.

Note that this choice of definition ensures continuity of H_α in α except at $\alpha = 0$, in the sense that

$$\lim_{\alpha\to\infty} H_\alpha(p) = H_\infty(p), \qquad \lim_{\alpha\to 1} H_\alpha(p) = H_1(p), \qquad \lim_{\alpha\to -\infty} H_\alpha(p) = H_{-\infty}(p), \qquad \lim_{\alpha\searrow 0} H_\alpha(p) = H_0(p).$$

However, $\lim_{\alpha\nearrow 0} H_\alpha(p)$ exists only if p has full rank, in which case it equals $-\log m = -H_0(\alpha)$. It is elementary to check that [4]

$$\lim_{\alpha\searrow 0} \frac{1-\alpha}{\alpha}\big(H_\alpha(p) - \log m\big) = \lim_{\alpha\nearrow 0} \frac{1-\alpha}{\alpha}\left(-H_\alpha(p) - \log m\right) = \frac{1}{m} H_{\mathrm{Burg}}(p) + \log m. \tag{A1}$$

Furthermore, Rényi entropy satisfies

$$H_\alpha(p) \in \begin{cases} [0, \log m] & \text{if } \alpha \geq 0 \\ [-\infty, -\log m] & \text{if } \alpha < 0, \end{cases}$$

and for every $\alpha \neq 0$, the maximal value $\operatorname{sgn}(\alpha)\log m$ is attained if and only if $p = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$, cf. [30]. The corresponding statement for the Burg entropy is $H_{\mathrm{Burg}}(p) \leq -m \log m$, with equality if and only if $p = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$.

Definition 7 (Majorization [20])**.** *Let $p, q \in \mathbb{R}^m$ be probability distributions. We say that p* majorizes *q, and write $p \succ q$, if and only if*

$$\sum_{i=1}^k p_i^\downarrow \geq \sum_{i=1}^k q_i^\downarrow \qquad \text{for all } k = 1, \ldots, m,$$

where $p_1^\downarrow \geq p_2^\downarrow \geq \ldots \geq p_m^\downarrow$ denotes the entries of p in descending order (and similarly for q). Furthermore, we say that p trumps *q, and write $p \succ_T q$, if and only if there exists $n \in \mathbb{N}$ and a probability distribution $c \in \mathbb{R}^n$ such that*

$$p \otimes c \succ q \otimes c.$$

Klimesh [22] and Turgut [23] have proven that the trumping relation is closely related to the Rényi and Burg entropies. In our notation, their result is as follows.

Lemma 8. *Let $p, q \in \mathbb{R}^m$ be probability distribution such that $p^\downarrow \neq q^\downarrow$. Then $p \succ_T q$ if and only if*

$$H_\alpha(p) < H_\alpha(q) \text{ for all } \alpha \in \mathbb{R} \setminus \{0\}, \qquad \text{and} \qquad H_{\mathrm{Burg}}(p) < H_{\mathrm{Burg}}(q).$$

Since we are interested in catalysis, we will in the following deal with multipartite (mostly bipartite) probability distributions. In the bipartite case, we use the following notation. We denote the first system by A (of size $m \in \mathbb{N}$), and the second by B (of size $n \in \mathbb{N}$). Joint distributions on AB will be denoted as matrices with entries $(p_{AB})_{i,j} := p(a = i, b = j)$. For example, if $p = p_A = (p_1, \ldots, p_m)$ and $q = q_B = (q_1, \ldots, q_n)$, then

$$p_A \otimes q_B = \begin{pmatrix} p_1 q_1 & p_1 q_2 & p_1 q_3 & \cdots & p_1 q_n \\ p_2 q_1 & p_2 q_2 & p_2 q_3 & \cdots & p_2 q_n \\ \vdots & \vdots & \vdots & & \vdots \\ p_m q_1 & p_m q_2 & p_m q_3 & \cdots & p_m q_n \end{pmatrix}$$

In general, the marginal distributions on A resp. B can be obtained by summing over the rows resp. columns of p_{AB}. There is a specific bipartite probability distribution that will be important in what follows. If we have any probability distribution $q \equiv q_A = (q_1, \ldots, q_m) \in \mathbb{R}^m$, we consider the specific extension

$$q_{AB} := \begin{pmatrix} q_1 - a_1 & \frac{a_1}{n} & \frac{a_1}{n} & \ldots & \frac{a_1}{n} \\ q_2 - a_2 & \frac{a_2}{n} & \frac{a_2}{n} & \ldots & \frac{a_2}{n} \\ \vdots & \vdots & \vdots & & \vdots \\ q_m - a_m & \frac{a_m}{n} & \frac{a_m}{n} & \ldots & \frac{a_m}{n} \end{pmatrix} \qquad \text{for any choice of } a_i \in [0, q_i] \text{ and } n \in \mathbb{N}. \tag{A2}$$

This is an $m \times (n+1)$ matrix, and a bipartite probability distribution with marginal q_A on A. Clearly

$$q_B = \left(1 - a, \frac{a}{n}, \ldots, \frac{a}{n}\right) \in \mathbb{R}^{n+1}, \qquad \text{where } a = \sum_{i=1}^m a_i.$$

Appendix B: Detailed mathematical proof of the main theorem

As described in the main text, we need two lemmas. The first one is as follows.

Lemma 9. *Let $p, q \in \mathbb{R}^m$ be probability distributions such that q has full rank, $H(p) < H(q)$, and $q \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$. Then there exists some $\delta \in (0, \min_i q_i)$ and $N \in \mathbb{N}$ such that for $a_i := q_i - \delta$ and q_{AB} as in (A2), the following statement is true for all $n \geq N$:*

$$H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB}) \qquad \textit{for all } \alpha \in [1, +\infty].$$

Proof. Note that $p \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$ because $H(p) < H(q) < \log m$. In the following, we will always assume that $\alpha > 1$, $\alpha \in \mathbb{R}$ (unless stated otherwise). With the given choice of a_i, we get $a = \sum_{i=1}^m a_i = 1 - m\delta$. Consider the following expression:

$$\Delta_n^{(\alpha)} := H_\alpha(q_{AB}) - H_\alpha(q_B) - H_\alpha(p_A) = \frac{1}{1-\alpha} \log \frac{m\delta^\alpha + n^{1-\alpha} \sum_{i=1}^m (q_i - \delta)^\alpha}{\left(\sum_{i=1}^m p_i^\alpha\right)\left(m^\alpha \delta^\alpha + (1 - m\delta)^\alpha n^{1-\alpha}\right)}.$$

We use the expression on the right-hand side to define $\Delta_n^{(\alpha)}$ also for non-integer $n \geq 1$. We have to show that this expression is positive for all α if n is large enough. In fact, in the limit,

$$\lim_{n \to \infty} \Delta_n^{(\alpha)} = \log m - H_\alpha(p) > 0 \qquad \text{for all } \alpha > 1, \tag{B1}$$

which is however only a pointwise statement. We furthermore need the fact that

$$\Delta_n^{(\alpha)} \text{ is strictly increasing in } n \text{ if } \alpha \in (1, \infty). \tag{B2}$$

To see this, simply take the derivative:

$$\frac{\partial}{\partial n} \Delta_n^{(\alpha)} = n^{\alpha-2} \delta^\alpha (b_2 - b_1), \text{ where } b_1 = \frac{m}{n^{\alpha-1} m\delta^\alpha + \sum_{i=1}^m (q_i - \delta)^\alpha} > 0, \qquad b_2 = \frac{m^\alpha}{n^{\alpha-1} m^\alpha \delta^\alpha + (1 - m\delta)^\alpha} > 0.$$

Since the (non-uniform) probability distribution with the m entries $\frac{q_i - \delta}{1 - m\delta}$ satisfies $H_\alpha\left(\left(\frac{q_i - \delta}{1 - m\delta}\right)_i\right) < \log m$, we have $\sum_{i=1}^m \left(\frac{q_i - \delta}{1 - m\delta}\right)^\alpha > m^{1-\alpha}$, and thus

$$\frac{b_1}{b_2} = \frac{n^{\alpha-1} m\delta^\alpha + m^{1-\alpha}(1 - m\delta)^\alpha}{n^{\alpha-1} m\delta^\alpha + \sum_{i=1}^m \left(\frac{q_i - \delta}{1 - m\delta}\right)^\alpha (1 - m\delta)^\alpha} < 1,$$

hence $b_1 < b_2$, which proves (B2). Furthermore, for $\alpha = 1$, we have

$$\Delta_n^{(1)} := H(q_{AB}) - H(q_B) - H(p_A) = m\delta \log m - \sum_{i=1}^m (q_i - \delta) \log \frac{q_i - \delta}{1 - m\delta} - H(p_A),$$

and this expression is independent of n. Since $\lim_{\delta\searrow 0}\Delta_n^{(1)} = H(q_A) - H(p_A) > 0$, there exists some $\delta \in (0, \min_i q_i)$ such that with this choice of δ, we have $\Delta_n^{(1)} > 0$. So let us choose and fix this δ for all that follows. By continuity, for $n = 1$, there exists some $\varepsilon > 0$ such that $\Delta_{n=1}^{(\alpha)} > 0$ for all $1 \leq \alpha \leq 1+\varepsilon$, and due to (B2)

$$\Delta_n^{(\alpha)} > 0 \qquad \text{for all } n \in \mathbb{N} \text{ and } 1 \leq \alpha \leq 1+\varepsilon. \tag{B3}$$

Furthermore, if n is large enough, then we have the exact equality

$$\Delta_n^{(\infty)} := H_\infty(q_{AB}) - H_\infty(q_B) - H_\infty(p_A) = \log m - H_\infty(p_A) > 0.$$

Therefore, a standard compactness argument for the interval $[1+\varepsilon, +\infty]$, together with (B1) and (B2), shows that there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, we have $\Delta_n^{(\alpha)} > 0$ for all α in that interval. Together with (B3), this proves the claim. $\square$

Lemma 10. *Let $p, q \in \mathbb{R}^m$ be probability distributions such that q has full rank, $q \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$, and $H_\alpha(p) < H_\alpha(q)$ for all $\alpha \in [1, +\infty]$. Then there exists some $a \in (0, m \cdot \min_i q_i)$ such that for q_{AB} as given in (A2) with $a_i := a/m$, we have*

$$p_A \otimes q_B \succ_T q_{AB} \qquad \text{for all } n \geq N.$$

Proof. First consider the case that p has full rank. Note that $p \neq \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$ since $H_1(p) < H_1(q) < \log m$. We will use the criterion in Lemma 8 to prove trumping. First note that

$$H_{\text{Burg}}(q_{AB}) = \sum_{i=1}^m \log\left(q_i - \frac{a}{m}\right) + mn\log\frac{a}{mn}, \qquad H_{\text{Burg}}(p_A \otimes q_B) = (n+1)\sum_{i=1}^m \log p_i + m\left(\log(1-a) + n\log\frac{a}{n}\right).$$

It is then elementary to see that the inequality $H_{\text{Burg}}(p_A \otimes q_B) < H_{\text{Burg}}(q_{AB})$ is equivalent to

$$\sum_{i=1}^m \log p_i + n\underbrace{\left(\sum_{i=1}^m \log p_i + m\log m\right)}_{(*)} + m\log(1-a) < \sum_{i=1}^m \log\left(q_i - \frac{a}{m}\right).$$

Since $\sum_{i=1}^m \log p_i = H_{\text{Burg}}(p) < -m\log m$, the factor $(*)$ is negative. Hence this inequality is true if n is large enough; in other words, there exists $N(a) \in \mathbb{N}$ (which may depend on the choice of a) such that

$$H_{\text{Burg}}(p_A \otimes q_B) < H_{\text{Burg}}(q_{AB}) \qquad \text{for all } n \geq N(a). \tag{B4}$$

For all $\alpha \in [-\infty, +\infty]$, define the quantity

$$\tilde{\Delta}_n^{(\alpha)} := H_\alpha(q_{AB}) - H_\alpha(q_B) - H_\alpha(p_A).$$

If $\alpha = 0$ this equals 0; for general finite $\alpha \neq 1$, it is

$$\tilde{\Delta}_n^{(\alpha)} = \frac{\operatorname{sgn}(\alpha)}{1-\alpha}\log\frac{\sum_{i=1}^m \left(q_i - \frac{a}{m}\right)^\alpha + n^{1-\alpha}a^\alpha m^{1-\alpha}}{\left(\sum_{i=1}^m p_i^\alpha\right)\left((1-a)^\alpha + n^{1-\alpha}a^\alpha\right)} \qquad (\alpha \in \mathbb{R}\setminus\{1\}).$$

First we prove the following:

$$\tilde{\Delta}_n^{(\alpha)} \text{ is } \begin{cases} \text{eventually constant in } n & \text{if } \alpha = -\infty \\ \text{increasing in } n & \text{if } -\infty < \alpha < 1 \\ \text{constant in } n & \text{if } \alpha = 1 \\ \text{decreasing in } n & \text{if } 1 < \alpha < +\infty \\ \text{eventually constant in } n & \text{if } \alpha = +\infty. \end{cases} \tag{B5}$$

By "eventually constant", we mean that there is some $N \in \mathbb{N}$ such that for all $n \geq N$, we have $\tilde{\Delta}_n^{(\alpha)} = \tilde{\Delta}_N^{(\alpha)}$. This is the case for $\alpha = -\infty$ and $\alpha = +\infty$, because in this case, all entropies only depend on the minimal resp. maximal entries of q_{AB} resp. q_B; if n is large, the location of these extrema is fixed, and direct calculation shows that all n-dependency cancels out. The special case $\alpha = 0$ is easily checked directly, too. For $\alpha = 1$, direct calculation shows that

$$\tilde{\Delta}_n^{(1)} = -\sum_{i=1}^m \left(q_i - \frac{a}{m}\right)\log\left(q_i - \frac{a}{m}\right) + a\log m + (1-a)\log(1-a) - H(p) \tag{B6}$$

which is independent of n. For the remaining cases $\alpha \in \mathbb{R} \setminus \{0,1\}$, differentiate:

$$\frac{\partial}{\partial n}\tilde{\Delta}_n^{(\alpha)} = \left(\frac{a}{n}\right)^{\alpha} \operatorname{sgn}(\alpha)(c_1 - c_2), \text{ where } c_1 = \frac{m^{1-\alpha}}{\sum_{i=1}^m \left(q_i - \frac{a}{m}\right)^{\alpha} + n^{1-\alpha}a^{\alpha}m^{1-\alpha}} > 0, \ \ c_2 = \left((1-a)^{\alpha} + n^{1-\alpha}a^{\alpha}\right)^{-1} > 0.$$

Similarly as in the proof of Lemma 9, one can show that

$$\sum_{i=1}^m \left(q_i - \frac{a}{m}\right)^{\alpha} \begin{cases} \geq m^{1-\alpha}(1-a)^{\alpha} & \text{if } \alpha < 0 \text{ or } \alpha > 1 \\ \leq m^{1-\alpha}(1-a)^{\alpha} & \text{if } 0 < \alpha < 1, \end{cases}$$

and thus

$$\frac{c_1}{c_2} = \frac{m^{1-\alpha}(1-a)^{\alpha} + n^{1-\alpha}a^{\alpha}m^{1-\alpha}}{\sum_{i=1}^m \left(q_i - \frac{a}{m}\right)^{\alpha} + n^{1-\alpha}a^{\alpha}m^{1-\alpha}} \begin{cases} \leq 1 & \text{if } \alpha < 0 \text{ or } \alpha > 1 \\ \geq 1 & \text{if } 0 < \alpha < 1. \end{cases}$$

This yields the sign of $c_1 - c_2$ and thus of $\frac{\partial}{\partial n}\tilde{\Delta}_n^{(\alpha)}$ for all values of α and proves (B5). By direct calculation, the large-n limit of $\tilde{\Delta}_n^{(\alpha)}$ evaluates to

$$\lim_{n\to\infty} \tilde{\Delta}_n^{(\alpha)} = \begin{cases} -\log m - H_\alpha(p) & \text{if } \alpha \in [-\infty, 0) \\ \log m - H_\alpha(p) & \text{if } \alpha \in (0,1) \\ \text{expression (B6) above} & \text{if } \alpha = 1 \\ H_\alpha\left(\left(\frac{q_i - a/m}{1-a}\right)_i\right) - H_\alpha(p) & \text{if } \alpha \in (1, +\infty] \end{cases} \tag{B7}$$

which is discontinuous at $\alpha = 0$ and $\alpha = 1$.

So far, $a \in (0, m \cdot \min_i q_i)$ was arbitrary; now we are going to fix the value of a. Define $p_{\max} := \max_i p_i$, $q_{\max} := \max_i q_i$. Since $H_\infty(p) < H_\infty(q)$, we have $p_{\max} > q_{\max}$. Thus, there exists some $a_\infty > 0$ such that

$$H_\infty\left(\left(\frac{q_i - \frac{a}{m}}{1-a}\right)_i\right) - H_\infty(p) = \log p_{\max} - \log \frac{q_{\max} - \frac{a}{m}}{1-a} > 0 \qquad \text{for all } 0 \leq a \leq a_\infty.$$

Therefore, since $H_\infty = \lim_{\alpha\to\infty} H_\alpha$, there exists some $\alpha_\infty \in \mathbb{R}$ such that

$$H_\alpha\left(\left(\frac{q_i - \frac{a_\infty}{m}}{1-a_\infty}\right)_i\right) - H_\alpha(p) > 0 \qquad \text{for all } \alpha \geq \alpha_\infty. \tag{B8}$$

For every $\alpha \in [1, \alpha_\infty]$, choose some $a_\alpha \in (0, a_\infty)$ such that

$$H_\alpha\left(\left(\frac{q_i - \frac{a_\alpha}{m}}{1-a_\alpha}\right)_i\right) - H_\alpha(p) > 0.$$

Such an a_α always exists due to $H_\alpha(q) - H_\alpha(p) > 0$ and continuity of H_α. Choose $\varepsilon_\alpha > 0$ such that

$$H_\beta\left(\left(\frac{q_i - \frac{a_\alpha}{m}}{1-a_\alpha}\right)_i\right) - H_\beta(p) > 0 \qquad \text{for all } \beta \in [1, \alpha_\infty] \cap (\alpha - \varepsilon_\alpha, \alpha + \varepsilon_\alpha). \tag{B9}$$

Furthermore, note that

$$\text{for all } \alpha \in [0, +\infty], \text{ the function } a \mapsto H_\alpha\left(\left(\frac{q_i - \frac{a}{m}}{1-a}\right)_i\right) \text{ is decreasing on the interval } [0, m \cdot \min_i q_i].$$

This follows from the observation that for $a, b \in [0, m \cdot \min_i q_i]$ with $a \leq b$, the probability distribution $[(q_i - b/m)/(1-b)]_i$ majorizes the probability distribution $[(q_i - a/m)/(1-a)]_i$, and the Rényi entropies H_α with $\alpha \geq 0$ are Schur concave [17, 20].

Thus, (B9) is also satisfied if a_α is replaced by any $a \in (0, a_\alpha]$. The intervals $(\alpha - \varepsilon_\alpha, \alpha + \varepsilon_\alpha)$ constitute an open cover of $[1, \alpha_\infty]$. Since this is a compact interval, there is a finite subcover, indexed by $\alpha_1, \ldots, \alpha_m \in [1, \alpha_\infty]$, with $m \in \mathbb{N}$ finite. Now set $a' := \min_{i=1,\ldots,m} a_{\alpha_i}$. Combining (B8) and (B9), we get

$$H_\alpha\left(\left(\frac{q_i - \frac{a}{m}}{1-a}\right)_i\right) - H_\alpha(p) > 0 \qquad \text{for all } \alpha \in [1, +\infty], \ \ 0 \leq a \leq a'.$$

Due to (B5) and (B7), we thus obtain

$$\tilde{\Delta}_n^{(\alpha)} \geq \lim_{n\to\infty} \tilde{\Delta}_n^{(\alpha)} = H_\alpha\left(\left(\frac{q_i - \frac{a}{m}}{1-a}\right)_i\right) - H_\alpha(p) > 0 \qquad \text{for all } \alpha \in (1,+\infty],\;\; a \in (0,a'],\;\; n \in \mathbb{N}$$

(recall that $\tilde{\Delta}_n^{(\alpha)}$ depends on the choice of a). Due to (B6), we have $\lim_{a\searrow 0} \tilde{\Delta}_n^{(1)} = H(q) - H(p) > 0$, so there exists $a \in (0,a')$ such that $\tilde{\Delta}_{n=1}^{(1)} > 0$ for this choice of a. **We now fix this value of a for all that follows.** Due to continuity, there exists $\varepsilon > 0$ such that $\tilde{\Delta}_{n=1}^{(\alpha)} > 0$ for all $\alpha \in [1-\varepsilon, 1]$. According to (B5), this implies that $\tilde{\Delta}_n^{(\alpha)} > 0$ for all $\alpha \in [1-\varepsilon,1]$ and all $n \in \mathbb{N}$. In summary, we have achieved that

$$\tilde{\Delta}_n^{(\alpha)} > 0 \qquad \text{for all } n \in \mathbb{N},\;\; \alpha \in [1-\varepsilon, +\infty]. \tag{B10}$$

Next we consider $\alpha \in (0, 1-\varepsilon)$. Since $\tilde{\Delta}_n^{(0)} = 0$ for all n is not useful as a special case, we define another quantity

$$\bar{\Delta}_n^{(\alpha)} := \begin{cases} \frac{1-\alpha}{|\alpha|}\tilde{\Delta}_n^{(\alpha)} & \text{if } \alpha \in \mathbb{R}\setminus\{0\} \\ \frac{1}{m(n+1)}\left(H_{\mathrm{Burg}}(q_{AB}) - H_{\mathrm{Burg}}(p_A \otimes q_B)\right) & \text{if } \alpha = 0. \end{cases}$$

The resulting quantity is continuous in α, also at $\alpha = 0$ due to (A1). Using that $H_{\mathrm{Burg}}\left(\left(\frac{q_i - a/m}{1-a}\right)_i\right) < -m\log m$, it is straightforward to check that

$$\frac{\partial}{\partial n}\bar{\Delta}_n^{(0)} = \frac{m\log\frac{1-a}{m} - \sum_{i=1}^m \log\left(q_i - \frac{a}{m}\right)}{m(n+1)^2} > 0,$$

hence $\bar{\Delta}_n^{(0)}$ is strictly increasing in n. The large-n limit is

$$\lim_{n\to\infty}\bar{\Delta}_n^{(0)} = -\frac{1}{m}H_{\mathrm{Burg}}(p) > 0.$$

Considering only $\alpha \in [0, 1-\varepsilon]$, the $\bar{\Delta}_n^{(\alpha)}$ are an increasing sequence of continuous functions on this compact interval, converging pointwise to a strictly positive value due to (B5) and (B7). Thus, a standard compactness argument proves that there exists some $N' \in \mathbb{N}$ such that $\bar{\Delta}_n^{(\alpha)} > 0$ for all $n \geq N'$ and $\alpha \in [0, 1-\varepsilon]$, hence

$$\tilde{\Delta}_n^{(\alpha)} > 0 \qquad \text{for all } n \geq N',\;\; \alpha \in (0, 1-\varepsilon]. \tag{B11}$$

Now we come to the case $\alpha < 0$. According to (B5) and (B7), there exists $N'' \in \mathbb{N}$ such that for all $n \geq N''$, it holds $\tilde{\Delta}_n^{(-\infty)} = -\log m - H_\infty(p) > 0$. Due to continuity, there is some $\alpha_{-\infty} \in \mathbb{R}$ such that $\tilde{\Delta}_{N''}^{(\alpha)} > 0$ for all $\alpha \in [-\infty, \alpha_{-\infty}]$, and thus (again due to (B5))

$$\tilde{\Delta}_n^{(\alpha)} > 0 \qquad \text{for all } n \geq N'',\;\; \alpha \in [-\infty, \alpha_{-\infty}]. \tag{B12}$$

Finally we treat the range $\alpha \in (\alpha_\infty, 0)$. Arguing as above, the $\bar{\Delta}_n^{(\alpha)}$ are an increasing sequence of continuous functions on the compact interval $[\alpha_{-\infty}, 0]$, converging pointwise to a strictly positive value. By compactness there exists $N''' \in \mathbb{N}$ such that $\bar{\Delta}_n^{(\alpha)} > 0$ for all $n \geq N'''$, and thus

$$\tilde{\Delta}_n^{(\alpha)} > 0 \qquad \text{for all } n \geq N''',\;\; \alpha \in [\alpha_{-\infty}, 0). \tag{B13}$$

Combining (B4), (B10), (B11), (B12), and (B13), and setting $N := \max\{N(a), N', N'', N'''\}$, we get

$$H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB}) \text{ for all } \alpha \in \mathbb{R}\setminus\{0\} \qquad \text{and} \qquad H_{\mathrm{Burg}}(p_A \otimes q_B) < H_{\mathrm{Burg}}(q_{AB}) \qquad \text{for all } n \geq N.$$

Clearly $(p_A \otimes q_B)^\downarrow \neq q_{AB}^\downarrow$, because otherwise we would have $H(p_A \otimes q_B) = H(q_{AB})$. Thus, Lemma 8 proves that $p_A \otimes q_B \succ_T q_{AB}$.

We have proven the statement of the lemma in the case that p has full rank. Now consider the case that $\mathrm{rank}(p) < m$. Since q and thus q_{AB} has full rank, we only have to show that $H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB})$ for all $\alpha \in (0, +\infty)$. To this end, we can simply repeat the proof above with a few small changes. First, the cases of Burg entropy and Rényi entropy for $\alpha < 0$ can be ignored. Second, the proof of (B10) remains valid, but the proof of (B11) has to be changed: instead of $\bar{\Delta}_n^{(\alpha)}$, we have to consider the quantity $\tilde{\Delta}_n^{(\alpha)}$ directly, which now satisfies $\tilde{\Delta}_n^{(0)} = \log m - H_0(p) > 0$ for all n. The rest of the argumentation remains unchanged, proving the statement of the lemma also for the case that p does not have full rank. $\square$

Combining the previous two lemmas yields a first formulation of our main result.

Lemma 11. *Let $p, q \in \mathbb{R}^m$ be probability distributions such that q has full rank. If $H(p) < H(q)$ then there exists $k \in \mathbb{N}$ (in fact, we can always choose $k = 3$) and a k-partite state $r_{1,2,\ldots,k}$ with marginals $r_1, r_2, \ldots, r_k$ such that*

$$p \otimes \big(r_1 \otimes r_2 \otimes \ldots \otimes r_k\big) \succ q \otimes r_{1,2,\ldots,k}.$$

Proof. The special case that $q = \big(\frac{1}{m}, \ldots, \frac{1}{m}\big)$ is trivial: in this case $p \succ q$, and we can simply set $k = 0$ (no catalyst), or alternatively $k = 1$ with an arbitrary catalyst.

So suppose $q \neq \big(\frac{1}{m}, \ldots, \frac{1}{m}\big)$. We first apply Lemma 9 to conclude that there exists some extension q_{AB} of $q = q_A$ such $H_\alpha(p_A \otimes q_B) < H_\alpha(q_{AB})$ for all $\alpha \in [1, +\infty]$. Clearly the extension q_{AB} given in that lemma has full rank, but is not a uniform distribution. Therefore, we can apply Lemma 10 to the two states $p_A \otimes q_B$ and q_{AB}, and obtain the existence of an extension q_{ABC} (introducing a third system C) of q_{AB} such that

$$(p_A \otimes q_B) \otimes q_C \succ_T q_{ABC}.$$

By definition of catalysis, there is an additional system D and a catalyst (probability distribution) c_D on D such that

$$p_A \otimes q_B \otimes q_C \otimes c_D \succ q_{ABC} \otimes c_D.$$

Since the majorization relation is invariant with respect to tensor product with another probability distribution, we obtain

$$p_A \otimes q_B \otimes q_C \otimes c_D \otimes q_E \succ q_{ABC} \otimes c_D \otimes q_E,$$

where $q_E = q = q_A$ is another copy of q (note however that q_B and q_C are in general *not* copies of $q = q_A$). Swapping systems A and E on the right-hand side does not alter the probability values and the majorization order, thus

$$p_A \otimes (q_E \otimes q_B \otimes q_C \otimes c_D) \succ q_A \otimes (q_{EBC} \otimes c_D).$$

If we regard CD as a single system (which we may, since the marginal of $q_{EBC} \otimes c_D$ on CD is $q_C \otimes c_D$), we see that we have $k = 3$ subsystems in addition to system A. $\square$

Theorem 12. *Let $p, q \in \mathbb{R}^m$ be probability distributions with $p^\downarrow \neq q^\downarrow$. Then there exists $k \in \mathbb{N}_0$ and a finite k-partite probability distribution $r_{1,2,\ldots,k}$ such that*

$$p \otimes (r_1 \otimes r_2 \otimes \ldots \otimes r_k) \succ q \otimes r_{1,2,\ldots,k} \tag{B14}$$

if and only if $\mathrm{rank}(p) \leq \mathrm{rank}(q)$ *and* $H(p) < H(q)$.

Proof. Suppose there exists a catalyst $r_{1,2,\ldots,k}$ with the stated properties. Then we can apply additivity and subadditivity [25, 26] as well as Schur concavity [20] of the Rényi entropies of orders $\alpha = 0$ and $\alpha = 1$ (Hartley and Shannon entropy) and obtain

$$H_\alpha(p) + \sum_{i=1}^k H_\alpha(r_i) \leq H_\alpha(q) + H_\alpha(r_{1,2,\ldots,k}) \leq H_\alpha(q) + \sum_{i=1}^k H_\alpha(r_i).$$

Since $H_0(p) = \log \mathrm{rank}(p)$, this shows that $\mathrm{rank}(p) \leq \mathrm{rank}(q)$. For Shannon entropy $H = H_1$, we obtain equality in the second inequality of this expression (subadditivity) if and only if $r_{1,2,\ldots,k} = r_1 \otimes r_2 \otimes \ldots \otimes r_k$; this follows inductively from the fact that the mutual information of two random variables is zero if and only if the joint bipartite probability distribution factorizes [31]. So if we had $H(p) = H(q)$ then $p \otimes (r_1 \otimes r_2 \otimes \ldots \otimes r_k) \succ q \otimes (r_1 \otimes r_2 \otimes \ldots \otimes r_k)$, or $p \succ_T q$. But then Lemma 8 would prove that $H(p) < H(q)$, which is a contradiction.

Conversely, suppose that $p, q \in \mathbb{R}^m$ are probability distributions that are not equal up to permutation and satisfy $\mathrm{rank}(p) \leq \mathrm{rank}(q)$ and $H(p) < H(q)$. Without loss of generality we may assume that $p^\downarrow = p$ and $q^\downarrow = q$, i.e. that the entries of p and q are in descending order. Let $\ell := \mathrm{rank}(q)$, then $\ell \leq m$ and $q = \tilde{q} \oplus 0_{m-\ell}$, where $\tilde{q} = (q_1, \ldots, q_\ell) \in \mathbb{R}^\ell$ has full rank, and $0_{m-\ell} = (0, \ldots, 0) \in \mathbb{R}^{m-\ell}$ is the zero vector of dimension $m - \ell$. Since $\mathrm{rank}(p) \leq \mathrm{rank}(q) = \ell$, we can also write $p = \tilde{p} \oplus 0_{m-\ell}$, where $\tilde{p} \in \mathbb{R}^\ell$ does not necessarily have full rank. Then (B14) for some probability distribution $r_{1,2,\ldots,k}$ is equivalent to

$$\tilde{p} \otimes (r_1 \otimes r_2 \otimes \ldots \otimes r_k) \succ \tilde{q} \otimes r_{1,2,\ldots,k}.$$

Since $H(\tilde{p}) = H(p) < H(q) = H(\tilde{q})$, and since $\tilde{q}$ has full rank, Lemma 11 applies and shows that a probability distribution $r_{1,2,\ldots,k}$ exists that satisfies this relation. $\square$

Appendix C: Pure and maximally mixed catalysts are useless

Suppose that one of the catalysts τ_i is a pure or maximally mixed state, and we have c-trumping as in (4). We may relabel the catalysts such that, without loss of generality, $i = k$.

In the case that $\tau_i = \tau_k$ is a pure state, this implies that $\tau_{1,2,\ldots,k}$ is a product state. Thus, (4) becomes

$$\rho \otimes (\tau_1 \otimes \tau_2 \otimes \ldots \tau_{k-1}) \otimes \tau_k \succ \sigma \otimes \tau_{1,2,\ldots,(k-1)} \otimes \tau_k.$$

By considering the sets of eigenvalues of both sides of this relation, it is clear that the relation is equivalent to

$$\rho \otimes (\tau_1 \otimes \tau_2 \otimes \ldots \tau_{k-1}) \succ \sigma \otimes \tau_{1,2,\ldots,(k-1)},$$

that is, we can simply disregard τ_k without altering the c-trumping relation: pure catalysts are useless.

An analogous conclusion holds for maximally mixed catalysts. To this end, suppose that $\tau_k = \mathbf{1}_d/d$, the maximally mixed state on $\mathbb{C}^d$. We argue within the resource theory of nonuniformity as introduced in [16], using notation from [17]. We write $\rho \overset{\text{noisy}}{\longrightarrow} \sigma$ for two quantum state ρ and σ, both living on Hilbert spaces of possibly different but finite dimensions, if and only if there exists a "noisy operation" that maps ρ to σ. A noisy operation is a map of the form

$$\rho_A \mapsto \mathrm{Tr}_E \left[U_{AB} \left(\rho_A \otimes \gamma_B \right) U_{AB}^\dagger \right],$$

where B is a d-dimensional quantum system, and $\gamma_B = \mathbf{1}_d/d$ the maximally mixed state on B; moreover, $AB = DE$ are two bipartitions of the same Hilbert space, such that the resulting density matrix is a state on D. In the case where A and D have the same dimension, we have $\rho \overset{\text{noisy}}{\longrightarrow} \sigma$ to arbitrary accuracy if and only if $\rho \succ \sigma$. The former statement means that for every $\varepsilon > 0$ there is a state σ_ε such that $\|\sigma - \sigma_\varepsilon\|_1 < \varepsilon$ and $\rho \overset{\text{noisy}}{\longrightarrow} \sigma_\varepsilon$. Let us write $\rho \longrightarrow \sigma$ if this is the case.

Adding maximally mixed states and taking partial traces are both noisy operations. Thus, if τ_k is the maximally mixed state, (4) implies

$$\rho \otimes (\tau_1 \otimes \ldots \otimes \tau_{k-1}) \longrightarrow \rho \otimes (\tau_1 \otimes \ldots \otimes \tau_{k-1} \otimes \tau_k) \longrightarrow \sigma \otimes \tau_{1,\ldots,k} \longrightarrow \sigma \otimes \tau_{1,\ldots,(k-1)}.$$

Hence $\rho \otimes (\tau_1 \otimes \ldots \otimes \tau_{k-1}) \succ \sigma \otimes \tau_{1,\ldots,(k-1)}$, and again we can remove τ_k without altering the c-trumping relation.