Von Neumann entropy from unitarity

Paul Boes,[1] Jens Eisert,[1] Rodrigo Gallego,[1] Markus P. Müller,[2,3] and Henrik Wilming[1,4]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria*
[3]*Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada*
[4]*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*

The von Neumann entropy is a key quantity in quantum information theory and, roughly speaking, quantifies the amount of quantum information contained in a state when many identical and independent (*i.i.d.*) copies of the state are available, in a regime that is often referred to as being asymptotic. In this work, we provide a new operational characterization of the von Neumann entropy which neither requires an *i.i.d.* limit nor any explicit randomness. We do so by showing that the von Neumann entropy fully characterizes single-shot state transitions in unitary quantum mechanics, as long as one has access to a catalyst — an ancillary system that can be re-used after the transition — and an environment which has the effect of dephasing in a preferred basis. Building upon these insights, we formulate and provide evidence for the *catalytic entropy conjecture*, which states that the above result holds true even in the absence of decoherence. If true, this would prove an intimate connection between single-shot state transitions in unitary quantum mechanics and the von Neumann entropy. Our results add significant support to recent insights that, contrary to common wisdom, the standard von Neumann entropy also characterizes single-shot situations and opens up the possibility for operational single-shot interpretations of other standard entropic quantities. We discuss implications of these insights to readings of the third law of quantum thermodynamics and hint at potentially profound implications to holography.

In quantum information theory it is common to distinguish tasks as falling in one of two regimes: Either one deals with situations in which many identically and independently distributed (*i.i.d.*) quantum systems appear. This regime is usually referred to as the *asymptotic regime*. Such tasks include, for example, Schumacher compression [1], entanglement distillation [2] and quantum hypothesis testing [3, 4]. Or, in sharp contrast, one deals with situations that only involve a single quantum system, the so-called *single-shot* regime. Examples of protocols that have been analyzed in the single-shot setting include the decoupling of quantum systems [5], hypothesis testing [6], and state transitions in quantum thermodynamics [7]. Common wisdom has it that different quantities characterize these two regimes. In the first regime, the von Neumann entropy (vNE) or quantities directly related to it prevail, such as the standard quantum relative entropy or mutual information, while in the second regime quantities such as quantum Rényi divergences [8–11] and smoothed versions of the above [12, 13] become important.

This common wisdom is, however, recently being challenged [14–19], as it has been shown that the vNE determines possible single-shot state transitions in quantum mechanics — under unitary evolutions — provided that three assumptions hold [18]: i) one can prepare a suitable *catalyst*, i.e. an auxiliary system that does not change its state during the process but might become correlated with the system on which the transition is performed; ii) one has access to an environment, or source of randomness, that is modelled as a large system in the maximally mixed state; iii) one has full control over system, catalyst and the environment, in the sense that one can implement any unitary on the joint system. Now, while introducing a catalyst as described by i) is operationally justified since it can subsequently be reused to perform transitions on further new systems, assumption ii) assigns an undesirably special role to maximally mixed systems, while assumption iii) is in conflict with the common experience that environments cannot practically be accessed with full degree of control.

In this work, we provide an operational characterization of the von Neumann entropy in terms of single-shot state transitions that does without assumptions ii) and iii). This may be seen as remarkable that a characterization is possible without resorting to ii) and iii) whatsoever. Instead, our characterization builds upon two natural classes of dynamics in quantum mechanics: controlled unitary evolution and uncontrolled decoherence to some given preferred basis. We also present applications of this characterization related to notions of cooling in quantum thermodynamics in a way as is usually discussed in the context of quantum readings of the *third law of thermodynamics* and discuss possible implications of our results for recent work on the decoupling of systems and the AdS/CFT correspondence in the context of *holography*. Finally, we formulate, and provide evidence for, a conjecture, which, if true, shows that the von Neumann entropy *can be derived directly from unitary quantum mechanics alone* as it fully characterizes catalytic, single-shot state transitions.

Main result. We will now present our main result and then discuss its implications. To state the result, let $\mathcal{D}$ be the quantum channel that decoheres a system in a given orthonormal basis $\{|j\rangle\}$ of its Hilbert space, according to

$$\mathcal{D}[\sigma] = \sum_j \langle j|\sigma|j\rangle \; |j\rangle\langle j| \, .$$

Density matrices diagonal in $\{|j\rangle\}$ will be called *quasi-classical*. Our main result can be stated as follows.

Theorem 1 (Single-shot characterization of the von Neumann entropy)**.** *Let ρ and ρ' be two density matrices of the same finite dimension and with different spectra. Then the following two statements are equivalent:*

i) $S(\rho') > S(\rho)$ *and* $\operatorname{rank}(\rho') \geq \operatorname{rank}(\rho)$.

ii) There exists a finite-dimensional, quasi-classical density matrix σ and a unitary U such that

$$\mathrm{Tr}_2\left[U(\rho\otimes\sigma)U^\dagger\right]=\rho', \tag{1}$$

$$\mathcal{D}\left[\mathrm{Tr}_1\left[U(\rho\otimes\sigma)U^\dagger\right]\right]=\sigma. \tag{2}$$

The proof is presented in Appendix A. Note first that the choice of basis $\{|j\rangle\}$ is irrelevant, since any basis change can be included in U. Furthermore, if one has $S(\rho')>S(\rho)$ but $\mathrm{rank}(\rho')<\mathrm{rank}(\rho)$, then by Theorem 1 the transition is not possible exactly. However, it can be done to arbitrary precision, since any state can be arbitrarily well approximated by a state with full rank. From a physical point of view, the condition on the rank is therefore not important.

To interpret this result, one can imagine a situation in which only a small region of space, say, the laboratory, can be controlled unitarily with high degree of precision while any system outside this region is decohered very quickly in some given basis. This is a common situation in current experimental devices. Given these constraints, the goal is to transform a quantum system from ρ to ρ' by acting unitarily on this system together with an ancillary system in a quasi-classical state that one can "borrow" from the environment so long as, upon being returned to the environment, it decoheres back to its initial state and can hence be used to aid further transitions. Then, Theorem 1 says that the vNE fully characterizes possible transitions in this natural setup (see Fig. 1 for a comparison of results and settings).

Note finally that, in general, the auxiliary system is clearly necessary to implement the transition $\rho\to\rho'$ since otherwise we would act unitarily on ρ' and therefore could not change its spectrum. The same restriction would arise if we demanded that the auxiliary system is returned *uncorrelated* from the system. Thus, σ truly acts like a catalyst by enabling transitions that would otherwise be impossible and can, after decohering in the environment, catalyse further transitions $\rho\to\rho'$ on further independent copies of ρ. At the same time, the statement provides a new perspective to the crucial role of correlations between the system and the catalyst.

Applications to notions of cooling and the third law. We now discuss an application of Theorem 1 to one of the key problems in quantum thermodynamics. Namely, we analyze how it can be used as a protocol for *cooling to very low temperatures* beyond the *i.i.d.* setting. This is a situation usually captured in readings of the *third law of thermodynamics* or *Nernst's Unattainability Principle (UP)*, bounding achievable rates to cooling. Specifically, in this context, we consider the reading of the problem of preparing systems in a state which is arbitrarily close to being pure. Let us for simplicity take as an initial system two uncorrelated qubits $\rho=\varrho\otimes\varrho$ with $S(\varrho)<1/2$ (even the generalization to other systems is obvious). Theorem 1 then implies that it is possible to implement a transition satisfying (1) and (2) so that the final state is $\rho'=\varrho'\otimes\mathbf{1}_2$, where $\mathbf{1}_k$ represents a maximally mixed state of dimension k and ϱ is any full-rank state with $S(\varrho')=\epsilon$ for arbitrarily small $\epsilon>0$, i.e. arbitrarily close, in trace distance, to a pure state. This is reminiscent of protocols of *algorithmic cooling* [20–23] which take a large number n of "warm" qubits ϱ and distill from them $n_c=n(1-S(\varrho))$ "cold" qubits having each a smallest eigenvalue $\lambda_{\min}=\mathcal{O}(\exp(-n))$ (see in particular Ref. [20]). The advantage of our protocol employing a catalyst is that we can obtain *arbitrarily cold systems* using a small number of copies, $n=2$ in this case, in contrast to the asymptotic *i.i.d.* setting considered in algorithmic cooling. Furthermore, the fact that the protocol of Theorem 1 is catalytic allows one to repeat the protocol for $n/2$ copies of ρ using a single ancillary system. Taking $S(\varrho)\approx 1/2$ we obtain $n_c\approx n/2$ qubits which are arbitrarily close to a pure state. This coincides with the bound given by algorithmic cooling which in this case is $n_c=n(1-S(\varrho))\approx n/2$ and that is the ultimate bound for any entropy non-decreasing protocol. Hence, our protocol not only distills arbitrarily cold qubits with few copies, but also has an optimal efficiency —in terms of the rate of almost pure qubits— when applied sequentially in the asymptotic limit. At the same time, however, our protocol establishes correlations among the cold qubits produced. Hence, although they can be used individually for further applications, it would be wrong to conclude that using our results one can prepare an arbitrary number $(\varrho')^{\otimes n}$ of uncorrelated quasi-pure states using the same catalyst over and over (see Appendix B for further discussion of this point). This again stresses the importance of correlations in the scheme.

The fact that one can produce systems in a state ϱ' which is arbitrarily close to a pure state might, moreover, at first glance seem to be in contradiction with the third law of thermodynamics as formulated in the UP. The UP states that infinite

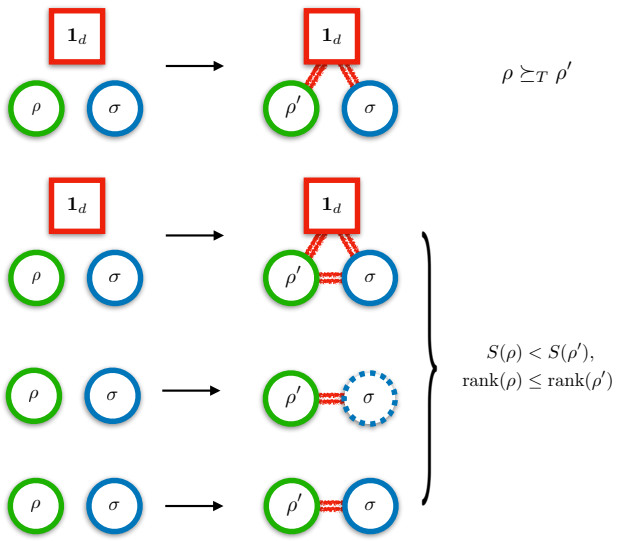

FIG. 1. Comparison of various settings and results. *Top:* State-transitions implementable using a source of randomness and an uncorrelated catalyst σ are characterized by the trumping relations. *Middle Top:* State transitions allowing for source of randomness and a correlated catalyst are characterized by entropy and rank [18]. *Middle Bottom:* By Theorem 1, state transitions using a correlated catalyst and a dephasing environment that acts on the catalyst (dashed boundary) are also characterized by entropy and rank. *Bottom:* State transitions using a correlated catalyst alone are characterized by entropy and rank. This is the content of Conjecture 1.

time is required to cool down a system to its ground state (see, e.g., Refs. [24–27] for recent approaches to quantum readings of the UP and their relation with pure state preparation). However, we note that preparing an arbitrarily pure ϱ' requires also an arbitrarily large catalyst σ and might also require a very large environment to implement the dephasing map $\mathcal{D}$, which in turn ensures that it cannot be prepared in finite time.

Relation to previous work. Let us now briefly discuss the relation of our results to previous work. To begin with, we note that one can use previous results to fully characterize the possible state transitions $\rho \to \rho'$ for the special case in which the catalyst is constrained to be a maximally mixed state. Specifically, one can recast recent results [28, 29] as the statement that there exist d and U such that

$$\mathrm{Tr}_2[U(\rho \otimes \mathbf{1}_d)U^\dagger] = \rho', \tag{3}$$

$$\mathcal{D}[\mathrm{Tr}_1[U(\rho \otimes \mathbf{1}_d)U^\dagger]] = \mathbf{1}_d, \tag{4}$$

if and only if ρ *majorizes* ρ', denoted by $\rho \succeq \rho'$ [28]. Clearly, the above is a special case of Eqs. (1) and (2). Majorization captures the state transitions that are possible under random unitary evolution and hence the above establishes the intuitive result that every random unitary evolution can be implemented with a sufficiently large source of randomness without affecting the latter's state.

To compare this result with Theorem 1 it should be noted that $\rho \succeq \rho'$ is, as a constraint, much stronger than $S(\rho') > S(\rho)$. Indeed one can see that Rényi entropies S_α, defined as

$$S_\alpha(\rho) = \frac{1}{1-\alpha}\log \mathrm{Tr}(\rho^\alpha) \quad (\alpha \in \mathbb{R}\backslash\{1\}), \tag{5}$$

cannot decrease for transitions $\rho \to \rho'$ with $\rho \succeq \rho'$, where the vNE is given by the particular case of $S \equiv S_1 := \lim_{\alpha\to 1} S_\alpha$. The infinite set of conditions given by the Rényi entropies

$$S_\alpha(\rho') \geq S_\alpha(\rho) \;\; \forall\, \alpha \in \mathbb{R} \tag{6}$$

become both necessary and sufficient for the existence of a further auxiliary system σ such that $\rho \otimes \sigma \succeq \rho' \otimes \sigma$ — an important relation known as *trumping* [30, 31] in quantum information theory. The trumping constraints lie, in strength, strictly between those imposed by majorization and the vnE alone.

Lastly, in Ref. [18] it is shown that by allowing for correlations between both systems it is possible to collapse the infinite set of conditions for the trumping conditions to essentially the vNE. In particular, it is shown that condition i) in Theorem 1 is equivalent to the existence of σ and U so that $\rho \otimes \sigma \succeq \rho'\sigma$, where $\rho'\sigma$ denotes a density matrix such that $\mathrm{Tr}_2(\rho'\sigma) = \rho'$ and $\mathrm{Tr}_1(\rho'\sigma) = \sigma$. This statement differs from Theorem 1 in that one needs to make use of a maximally mixed system over which one has full unitary control, while Theorem 1 includes external randomness only in the form of an uncontrolled dephasing map (see Fig. 1 for comparison).

Catalytic entropy conjecture. The discussion above raises the natural question whether an external environment, being modelled as a maximally mixed state or a dephasing map as above, is at all necessary to implement all transitions which do not decrease the vNE. This is what we capture in the following conjecture.

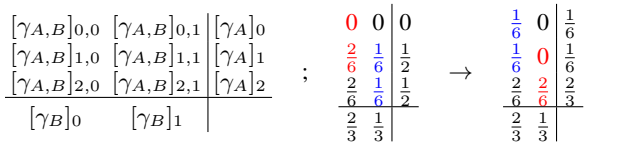

FIG. 2. Given an arbitrary bipartite state on A, B denoted $\gamma_{A,B}$, the table at the left-hand side indicates the meaning of each entry, where $[\gamma_{A,B}]_{i,j} := \langle i, j \,|\, \gamma_{A,B} \,|\, i, j \rangle$ on a given computational basis of AB. The two tables at the right hand side indicate a particular transition of the form $\rho \otimes \sigma \to U(\rho \otimes \sigma)U := \rho'\sigma$. In this case we take ρ and σ to be of dimension 3 and 2 respectively, and both diagonal in the computational basis. The unitary U is simply a classical permutation which swaps the red entries with the blue entries. Note that the final state satisfies $\mathrm{Tr}_2(\rho'\sigma) = \sigma$ since the bottom row remains unchanged, as demanded by condition (b). The column sums on the right-hand side of each table represent $\rho = \mathrm{diag}(0, 1/2, 1/2)$ and $\rho' = \mathrm{diag}(1/6, 1/6, 2/3)$. Since $S_\infty(\rho)$ is determined by the largest eigenvalue of ρ, this example realizes a catalytic transition $\rho \to \rho'$ with $S_\infty(\rho) > S_\infty(\rho')$ and hence excludes the possibility that catalytic state transitions are constrained by the trumping relations.

Conjecture 1 (Catalytic entropy conjecture)**.** *Let ρ and ρ' be two density matrices of the same finite dimension and with different spectra. Then the following two statements are equivalent:*

(a) $S(\rho') > S(\rho)$ *and* $\mathrm{rank}(\rho') \geq \mathrm{rank}(\rho)$.

(b) There exists a density matrix σ and a unitary U such that

$$\mathrm{Tr}_2\left[U(\rho \otimes \sigma)U^\dagger\right] = \rho' \text{ and } \mathrm{Tr}_1\left[U(\rho \otimes \sigma)U^\dagger\right] = \sigma. \tag{7}$$

The implication (b) $\Rightarrow$ (a) follows directly from the subadditivity of the vNE and S_0, hence the real content of the conjecture is that (a) are the only constraints on transitions of the form (b). If true, this conjecture would therefore imply that the von Neumann entropy characterizes catalytic state-transitions in unitary quantum mechanics in full generality, without the need to introduce noise or *i.i.d.* limits (see Fig. 1).

Let us now discuss why we believe this conjecture to be true. To begin with, it is easy to generate counterexamples that rule out the possibility that transitions of the form (b) are constrained by the aforementioned trumping relations. In Fig. 2 we provide such a counterexample together with a method to construct further examples. But in fact, we can rule out more general constraints than (6) with the help of the following lemma.

Lemma 2 (Weak solution to catalytic entropy conjecture)**.** *Let ρ and ρ' be two density matrices of the same, finite dimension and with different spectra. Then the following two statements are equivalent:*

(I) $S(\rho') > S(\rho)$ *and* $\mathrm{rank}(\rho') \geq \mathrm{rank}(\rho)$.

(II) There exists a density matrix σ, a unitary U and some finite dimension d such that

$$\mathrm{Tr}_2\left[U(\rho \otimes \mathbf{1}_d \otimes \sigma)U^\dagger\right] = \rho' \otimes \mathbf{1}_d, \tag{8}$$

$$\mathrm{Tr}_1\left[U(\rho \otimes \mathbf{1}_d \otimes \sigma)U^\dagger\right] = \sigma. \tag{9}$$

This result, which is proven in Appendix A, supports the conjecture in two ways: Firstly, it shows that the catalytic entropy conjecture is true up to an additional maximally mixed system that remains uncorrelated to the system of interest, but not to the catalyst. It can also be seen as an instance of the full catalytic entropy conjecture for the specific states $\rho \otimes \mathbf{1}_d$ and $\rho' \otimes \mathbf{1}_d$. Secondly, and more importantly, it allows us to prove the following corollary:

Corollary 3 (Characterization of entropy functions)**.** *Let f be a function from the set of density matrices to the real numbers such that for every transition of the form* (b) *between full-rank density matrices, $f(\rho') > f(\rho)$. Then exactly one of the following two statements is true:*

1. $S(\rho') > S(\rho) \Leftrightarrow f(\rho') > f(\rho)$,
2. *f is non-additive or discontinuous.*

Corollary 3 follows from Lemma 2 by showing that any such function f has to be a linear function of the vNE (see Appendix D for a proof). Thus, for full-rank density matrices, if Conjecture 1 was false, any additional constraint on transitions of the form (b) would have to be given by exotic entropic functions that are not additive or are discontinuous. For instance, this corollary immediately implies that none of the functions $S_\alpha,\ \alpha \neq 0, 1$, can be a monotone for transitions of the form (b) since they all satisfy none of the two conditions in the corollary.

Discussion and open questions. In this work, we have provided a new operational characterization of von Neumann entropy which adds significant support to recent proposals that, contrary to common wisdom, the standard von Neumann entropy characterizes not only the *i.i.d.* limit but also single-shot protocols in quantum information theory. We have done so by showing that the von Neumann entropy fully determines the possibility of single-shot state transitions in unitary quantum mechanics, as long as one has access to a catalyst and environmental dephasing in a preferred basis. Furthermore, we have formulated the *catalytic entropy conjecture* which essentially states that the above result holds true even in the absence of decoherence. We have also presented evidence for the truth of this conjecture by ruling out alternatives.

Our work suggests that there might be a novel, hitherto unexplored sector of quantum information theory in which operations on *single* copies of a quantum state are characterized directly in terms of standard entropic quantities like vNE. For example, one may ask what happens in Theorem 1 or Conjecture 1 if we introduce another reference system R that is initially correlated or entangled with the system 1 (let us denote system 1 by A for now, and let C be the catalytic system 2). Applying a unitary $U_{A,C}$ on the system and catalyst, denoting the new states of the systems by R', A' and C', we obtain $R' = R$, by construction $C' = C$ and $S(A') \geq S(A)$ since A becomes correlated with C. Furthermore, the mutual information $I(R : A) = S(R) + S(A) - S(R, A)$ satisfies $I(R' : A') \leq I(R : A)$. Are these necessary conditions also *sufficient* for the existence of a transformation of that form — in particular, can A retain almost all of its correlations with R under correlating-catalytic transformations? A positive answer to this or other similar questions would yield a new single-shot interpretation of the standard mutual information which could potentially be useful in the context of *decoupling* [5, 32–34] or merging of quantum states.

The results also hint at the insight that entanglement in single many-body systems can well be captured in terms of the von-Neumann entropy. Ideas on *single-copy entanglement* have been considered in situations where each specimen consists of a many-body system, already naturally featuring asymptotically many constituents [35]. Then it can be unreasonable to capture entanglement of subsystems in yet another asymptotic limit of many copies of identical quantum many-body systems. The results laid out here give substance to the intuition that even in single specimens of quantum many-body systems, entanglement can in this context be quantified in terms of the familiar von-Neumann entanglement entropy.

Results of this kind would also have implications in the context of *holographic approaches* to quantum gravity, as in the AdS/CFT correspondence (see, for example, Refs. [36–43]). In these approaches, standard von Neumann (entanglement) entropies of boundary regions turn out to correspond to geometric quantities of a dual gravity theory in the bulk. In fact, it is exactly the mutual information that we have just discussed which is believed to be directly related to geometric quantities like area also in other (non-AdS/CFT) approaches to emergent spacetime [44]. To shed some light on this correspondence, it is therefore natural to consider operational interpretations of entropy in the boundary theory, and to "dualize" them to obtain corresponding interpretations of geometric quantities in the bulk. A difficulty in doing so, however, is that the protocols on the boundary theory either involve many copies of the state (which seems unphysical given that there is a unique spacetime), or they lead to quantification in terms of single-shot entropies (see, e.g., Ref. [40]) which do not always have a direct dual interpretation. The proven and conjectured results of this paper could therefore resolve this difficulty, by supplying direct single-shot interpretation of standard entropic quantities which might ultimately shed some light on the operational basis of geometric quantities. It is the hope that the present work stimulates such endeavors.

Acknowledgements. We acknowledge funding from DFG (GA 2184/2-1, CRC 183, EI 519/14-1, EI 519/9-1, FOR 2724), the ERC (TAQ) and the Studienstiftung des deutschen Volkes. HW further acknowledges contributions from the Swiss National Science Foundation via the NCCR QSIT as well as project No. 200020_165843. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

[1] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[2] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[3] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99 (1991).
[4] T. Ogawa and H. Nagaoka, in *Asymptotic Theory of Quantum Statistical Inference* (World Scientific Pub Co Pte Lt, 2005) pp. 28–42.
[5] C. Majenz, M. Berta, F. Dupuis, R. Renner, and M. Christandl, Phys. Rev. Lett. **118**, 080503 (2017).
[6] M. Mosonyi and T. Ogawa, Commun. Math. Phys. **334**, 1617 (2015).
[7] F. G. S. L. Brandao, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Phys. Rev. Lett. **111**, 250404 (2013).
[8] N. Datta, IEEE Trans. Inf. Theory **55**, 2816 (2009).
[9] M. Berta, K. P. Seshadreesan, and M. M. Wilde, J. Math. Phys. **56**, 022205 (2015).
[10] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, J. Math. Phys. **54**, 122203 (2013).
[11] M. M. Wilde, A. Winter, and D. Yang, Comm. Math. Phys. **331**, 593 (2014).
[12] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, ETH Zurich (2005).
[13] N. D. R. Renner, IEEE Tr. Inf. Th. **55**, 2807 (2009).
[14] M. P. Müller and M. Pastena, IEEE Trans. Inf. Th. **62**, 1711 (2016).
[15] M. Lostaglio, M. P. Müller, and M. Pastena, Phys. Rev. Lett. **115**, 150402 (2015).
[16] R. Gallego, J. Eisert, and H. Wilming, New J. Phys. **18**, 103017 (2016).
[17] H. Wilming, R. Gallego, and J. Eisert, Entropy **19**, 241 (2017).
[18] M. P. Müller, (2017), 1707.03451.
[19] A. M. Alhambra, L. Masanes, J. Oppenheim, and C. Perry, ArXiv:1709.06139.
[20] L. J. Schulman and U. V. Vazirani, Proceedings of the thirty-first annual ACM symposium on Theory of computing - STOC '99 (1999), 10.1145/301250.301332.
[21] L. J. Schulman, T. Mor, and Y. Weinstein, Phys. Rev. Lett. **94**, 120501 (2005).
[22] P. O. Boykin, T. Mor, V. Roychowdhury, F. Vatan, and R. Vrijen, PNAS **99**, 3388 (2002).
[23] S. Raeisi and M. Mosca, Phys. Rev. Lett. **114**, 100404 (2015).
[24] A. Levy, R. Alicki, and R. Kosloff, Phys. Rev. E **85**, 061126 (2012).
[25] J. Scharlau and M. P. Müller, Quantum **2**, 54 (2018).
[26] L. Masanes and J. Oppenheim, Nature Comm. **8**, 14538 (2017).
[27] H. Wilming and R. Gallego, Phys. Rev. X **7**, 041033 (2017).
[28] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern, Phys. Rep. **583**, 1 (2015).
[29] P. Boes, H. Wilming, R. Gallego, and J. Eisert, (2018), 1804.03027.
[30] M. Klimesh, (2007), 0709.3680v1.
[31] S. Turgut, J. Phys. A **40**, 12185 (2007).
[32] M. Horodecki, J. Oppenheim, and A. Winter, Nature **436**, 673 (2005).
[33] P. Hayden, "Decoupling: A building block for quantum information theory," (2011).
[34] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, Commun. Math. Phys. **328**, 251 (2014).
[35] J. Eisert and M. Cramer, Phys. Rev. A **72**, 042112 (2005).
[36] L. Susskind, J. Math. Phys. **36**, 6377 (1995).
[37] J. M. Maldacena, Int. J. Theor. Phys. **38**, 1113 (1999).
[38] S. Ryu and T. Takayanagi, Phys. Rev. Lett. **96**, 181602 (2006).
[39] P. Hayden, M. Headrick, and A. Maloney, Phys. Rev. D **87**, 046003 (2013).
[40] B. Czech, P. Hayden, N. Lashkari, and B. Swingle, J. High En. Phys. **2015**, 157 (2015).
[41] N. Lashkari and M. Van Raamsdonk, JHEP **2016**, 153 (2016).
[42] H. Casini, E. Testé, and G. Torroba, Phys. Rev. Lett. **118**, 261602 (2017).
[43] A. Jahn, M. Gluza, F. Pastawski, and J. Eisert, (2017), arXiv:1711.03109.
[44] C. Cao, S. M. Carroll, and S. Michalakis, Phys. Rev. D **95**, 024031 (2017).
[45] A. Horn, Am. J. Math. **76**, 620 (1954).
[46] J. Schwinger, Proc. Natl. Ac. Sc. **46**, 570 (1960).
[47] R. F. Werner, J. Phys. A **34**, 7081 (2001).

Appendix A: Proof of Theorem 1 and Lemma 2

In this section we prove Theorem 1 and Lemma 2. The proofs of both results rely on the following recent result from Ref. [18].

Theorem 4 (Correlating-catalytic majorization [18])**.** *Let ρ, ρ' be two density matrices on the same, finite-dimensional Hilbert space $\mathcal{H}_A$ such that $S(\rho) < S(\rho')$ and $\operatorname{rank}(\rho) \leq \operatorname{rank}(\rho')$. Then there exists a density matrix τ on a finite-dimensional Hilbert space $\mathcal{H}_B$ and a bipartite density matrix $\rho'\tau$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ such that*

$$\rho \otimes \tau \succeq \rho'\tau, \quad \operatorname{Tr}_B[\rho'\tau] = \rho', \quad \operatorname{Tr}_A[\rho'\tau] = \tau.$$

Another result that will be used frequently is the Schur-Horn-Theorem.

Theorem 5 (Schur-Horn [45])**.** *For a matrix H, let $\lambda(H)$ be the vector of its eigenvalues and $\operatorname{diag}(H)$ the vector of its diagonal entries. If H is Hermitian, then the following are equivalent:*

- $\lambda(H) \succeq \operatorname{diag}(H)$,
- *there exists a unitary matrix U such that*

$$U\hat{\lambda}(H)U^\dagger = H,$$

where $\hat{\lambda}(H)$ is the diagonal matrix with diagonal $\lambda(H)$.

In particular, the Schur-Horn theorem implies that, if $\rho \succeq \rho'$, then there exist unitaries U, V such that

$$\rho' = V\left(\mathcal{D}_J[U\rho U^\dagger]\right)V^\dagger. \tag{A1}$$

Here and in the following, in contrast to the main text, we explicitly denote the choice of basis $J = \{|j\rangle\}$ in the notation for the decoherence map, $\mathcal{D} = \mathcal{D}_J$. If we choose J as the eigenbasis of ρ' then V is the identity map. We are now in position to prove Theorem 1.

Proof of Theorem 1. We begin with proving that i) implies ii). Thus, assume that $S(\rho) < S(\rho')$ and $\mathrm{rank}(\rho) \leq \mathrm{rank}(\rho')$. Then Theorem 4 together with (A1) implies that there exists a unitary $W_{A,B}$ and two bases J_A and J_B such that

$$(\mathcal{D}_{J_A} \otimes \mathcal{D}_{J_B}) \left[W_{A,B}(\rho \otimes \tau) W_{A,B}^\dagger \right] = \rho' \tau.$$

From locality of quantum mechanics and the Schur-Horn theorem we thus find that

$$\rho' \tilde{\tau} := (D_{J_A} \otimes \mathbb{I}) \left[W_{A,B}(\rho \otimes \tau) W_{A,B}^\dagger \right]$$

is a quantum state with the properties $\mathrm{Tr}_B[\rho' \tilde{\tau}] = \rho'$ and $\tilde{\tau} = \mathrm{Tr}_A[\rho' \tilde{\tau}] \succeq \tau$. Here, $\mathbb{I}$ denotes the identity super-operator.

As a second step, we show that we can realize any dephasing map on a system A using an ancillary system in a maximally mixed state. To see this, let R be a system of the same dimension d as A and let $\{U_k\}_{k=1}^d$ be a unitary operator basis on A, meaning a collection of d unitaries U_k such that

$$\mathrm{Tr}\left[U_j U_k^\dagger \right] = d \delta_{j,k}. \tag{A2}$$

Such a set of operators exists on every finite-dimensional Hilbert-space [46, 47]. Then, define the unitary

$$V_{A,R} = \sum_{j=1}^d |j\rangle\langle j|_A \otimes (U_j)_R,$$

where we recall that $J = \{|j\rangle\}$. Now, it is easy to check that for any $\rho = \rho_A$,

$$\mathrm{Tr}_R \left[V_{A,R}(\rho \otimes \mathbf{1}_d) V_{A,R}^\dagger \right] = \mathcal{D}_J[\rho].$$

In a third step, we now show that we can use this dilation of the dephasing map to construct a catalyst for Theorem 1. To do so, let

$$\sigma := \tau \otimes \mathbf{1}_d$$

and define the unitary

$$U_{A,B,R} = (V_{A,R} \otimes \mathbf{1}_B)(W_{A,B} \otimes \mathbf{1}_R).$$

From the previous discussion and the construction of the dephasing unitary $V_{A,R}$, we know that

$$\mathrm{Tr}_R \left[U_{A,B,R}(\rho \otimes \sigma) U_{A,B,R}^\dagger \right] = \rho' \tilde{\tau}.$$

Thus, what is left to be proven is that σ is indeed a valid catalyst, i.e., does not change in the course of the process except from building up coherences. We will show that it undergoes the transition

$$\sigma = \tau \otimes \mathbf{1}_d \to \tilde{\tau} \otimes \mathbf{1}_d.$$

To show this, first note that the dephasing dilation implemented by $V_{A,R}$ leaves the state $\mathbf{1}_d$ of R locally unchanged. But this means that we only have to show that R does not become correlated with B in the dephasing step, since it follows from locality that the marginal on R remains unchanged and the marginal on B evolves from τ to $\tilde{\tau}$. To see that B and R remain uncorrelated, we simply compute the action of the dephasing unitary $V_{A,R}$ on B, R, to get

$$\begin{aligned}
&\mathrm{Tr}_A \left[U_{A,B,R}(\rho \otimes \sigma) U_{A,B,R}^\dagger \right] \\
&= \sum_{j,k} \mathrm{Tr}_A \left[|j\rangle\langle j|_A W_{A,B}(\rho \otimes \tau) W_{A,B}^\dagger |k\rangle\langle k|_A \right] \otimes \frac{U_j U_k^\dagger}{d_R} \\
&= \sum_j \langle j|_A W_{A,B}(\rho \otimes \tau) W_{A,B}^\dagger |j\rangle_A \otimes \mathbf{1}_d \\
&= \tilde{\tau} \otimes \mathbf{1}_d,
\end{aligned}$$

where we have dropped identities for notational convenience. This proves that i) implies ii).

Let us now prove that ii) implies i). In the following let $\alpha \in \{0, 1\}$. Since $S_0(\rho) = \log(\mathrm{rank}(\rho))$, both S_0 and $S_1 = S$ are subadditive and additive. Since the final state on the catalyst, which we now call σ', satisfies $\mathcal{D}_J[\sigma'] = \sigma$, it follows that $\sigma' \succeq \sigma$ and thus $S_\alpha(\sigma') \leq S_\alpha(\sigma)$. Furthermore, from additivity and subadditivity we get

$$\begin{aligned}
S_\alpha(\rho) + S_\alpha(\sigma) &= S_\alpha(\rho \otimes \sigma) = S_\alpha(\rho' \sigma') \\
&\leq S_\alpha(\rho') + S_\alpha(\sigma') \leq S_\alpha(\rho') + S_\alpha(\sigma).
\end{aligned}$$

For $\alpha = 0$, this proves $\mathrm{rank}(\rho) \leq \mathrm{rank}(\rho')$. For $\alpha = 1$, equality, i.e. $S(\rho) = S(\rho')$, is only possible if $S(\rho'\sigma') = S(\rho') + S(\sigma')$, and it is well-known that this implies $\rho'\sigma' = \rho' \otimes \sigma'$. Thus $\rho \otimes \sigma \succ \rho' \otimes \sigma' \succ \rho' \otimes \sigma$, and so $\rho \succ_T \rho'$ for the trumping relation, which together with $S(\rho) = S(\rho')$ implies that ρ and ρ' have the same spectrum, i.e. are unitarily equivalent [30, 31]. This contradicts the assumptions of the theorem. We must thus have $S(\rho) < S(\rho')$, which completes the proof. □

Let us now turn to the proof of Lemma 2, which builds on the proof of Theorem 1.

Proof of Lemma 2. For this proof we re-use all the notation from the proof of the implication i)⇒ii) of Theorem 1. In particular note that the final state $\tilde{\tau}$ on the B-subsystem of the catalyst only needs to be dephased in a basis J_B to be returned exactly, since, by construction, $\mathrm{diag}(\tilde{\tau}) = \lambda(\tau)$. Using the dephasing construction already used in the proof of Theorem 1 we can include a further system R_2 in the maximally mixed state into the system and use the dephasing ttunitary $V_{R_2 B}$ at the end of the process to dephase system B. The only property that we still need to prove is that this does not introduce correlations between A and R_2. However, this is exactly the same calculation that shows that there are no correlations between B and R at the end in the proof of Theorem 1. We only have to exchange R for R_2 and B for A. This finishes the proof. □

Appendix B: Catalytic cooling

Let us first present in detail how to prepare almost pure states with a protocol that uses Theorem 1. Using this theorem, we have that, given system Q_1 in state $\rho_{Q_1} = \varrho \otimes \varrho$ with $2S(\varrho) < 1$, one can find U and a catalyst C in state σ so that

$$\gamma_{Q_1C} = (\mathcal{D}_J \circ \mathcal{U}_1)[\rho_{Q_1} \otimes \sigma_C] \tag{B1}$$

where $\mathcal{D}_J$ is the map locally dephasing the system C and leaving Q_1 untouched (formally $\mathbb{I}_{Q_1} \otimes \mathcal{D}_J$), and $\mathcal{U}_1[\bullet] = U \bullet U^\dagger$. Also, we denote by γ_{Q_1C} a bipartite state on Q_1C which, according to Theorem 1, fulfills $\mathrm{Tr}_{Q_1}(\gamma_{Q_1C}) = \sigma_C$ and $\mathrm{Tr}_C(\gamma_{Q_1C}) = \rho'_{Q_1} = \varrho' \otimes \mathbf{1}_2$, where ϱ' can be any full-rank state, but in the following we are interested in the case where ρ' is arbitrarily close to a pure state.

This protocol can be iterated on an arbitrary number n of subsystems $Q_1, \ldots, Q_n$, taking initially $\rho_{Q_1,\ldots,Q_N} = \rho_{Q_1} \otimes \cdots \otimes \rho_{Q_n}$ as input, where $\rho_{Q_i} = \varrho \otimes \varrho$ for all i. We define the unitary channels $\mathcal{U}_i$ which apply the unitary U to systems Q_iC and act trivially in the rest of the subsystems, that is,

$$\mathcal{U}_i[\bullet] = U_{Q_iC} \otimes \mathbb{I}_{|Q_iC} \bullet U^\dagger_{Q_iC} \otimes \mathbb{I}_{|Q_iC}. \tag{B2}$$

Then, applying these unitary channels, each followed by a dephasing map on C, one obtains

$$\gamma_{Q_1,\ldots,Q_nC} = \mathcal{D}_J \circ \mathcal{U}_n \circ \cdots \circ \mathcal{D}_J \circ \mathcal{U}_1[\rho_{Q_1,\ldots,Q_n} \otimes \sigma] \tag{B3}$$

where, due to Theorem 1, we have

$$\mathrm{Tr}_{|Q_i}(\gamma_{Q_1,\ldots,Q_nC}) = \varrho' \otimes \mathbf{1}_2 \ \ \forall\, i,$$
$$\mathrm{Tr}_{|C}(\gamma_{Q_1,\ldots,Q_nC}) = \sigma.$$

Hence, with this protocol we have prepared $n/2$ subsystems whose marginal ϱ' is arbitrarily close to a pure state. Note, however, that the resulting state of the compound $\gamma_{Q_1,\ldots,Q_n}$ displays correlations between its parts, hence, although each subsystem in state ϱ' can be individually used —for instance as a pure state input of a quantum computation— the whole compound $\gamma_{Q_1,\ldots,Q_n}$ deviates from the state

$$\tilde{\gamma}_{Q_1,\ldots,Q_n} \coloneqq \rho'_{Q_1} \otimes \cdots \otimes \rho'_{Q_n}.$$

This can be seen for instance by comparing the minimum eigenvalue $\lambda_{\min}$ of both states in the limit of large n, which gives

$$\lim_{n\to\infty} \frac{\lambda_{\min}(\tilde{\gamma}_{Q_1,\ldots,Q_n})}{\lambda_{\min}(\gamma_{Q_1,\ldots,Q_n})} \leq \lim_{n\to\infty} \frac{\lambda_{\min}(\tilde{\gamma}_{Q_1,\ldots,Q_n})}{\lambda_{\min}(\gamma_{Q_1,\ldots,Q_nC})} \tag{B4}$$
$$\leq \lim_{n\to\infty} \frac{\lambda_{\min}(\tilde{\gamma}_{Q_1,\ldots,Q_n})}{\lambda_{\min}(\rho_{Q_1,\ldots,Q_n} \otimes \sigma)} \tag{B5}$$
$$= \lim_{n\to\infty} \frac{(\frac{1}{2}\lambda_{\min}(\varrho'))^n}{\lambda_{\min}(\varrho)^{2n}\lambda_{\min}(\sigma)} \tag{B6}$$
$$= 0 \tag{B7}$$

where (B4) follows simply because tracing out one subsystem can only increase the minimum eigenvalue; (B5) follows due to (B3). To see this note that map $\mathcal{D}_J \circ \mathcal{U}_n \circ \cdots \circ \mathcal{D}_J$ can be implemented as a global unitary on $Q_1, \ldots, Q_n$ together with a source of randomness of sufficiently large dimension d which is responsible of the dephasing. That is, there exists V so that

$$\mathrm{Tr}(V \rho_{Q_1,\ldots,Q_n} \otimes \sigma \otimes \mathbf{1}_d V^\dagger) = \gamma_{Q_1,\ldots Q_nC}.$$

This implies in turn that $\rho_{Q_1,\ldots,Q_n} \otimes \sigma \succeq \gamma_{Q_1,\ldots Q_nC}$ (see for instance Ref. [28]) and that

$$\lambda_{\min}(\rho_{Q_1,\ldots,Q_n} \otimes \sigma) \leq \lambda_{\min}(\gamma_{Q_1,\ldots Q_nC}).$$

Eq. (B6) follows from simple algebra. Lastly, (B7) follows from the fact that $\lambda_{\min}(\varrho')$ is arbitrarily small while $\lambda_{\min}(\sigma) > 0$. To see the latter we recall the result of Appendix F.1 from Ref. [27], which shows that any transition of the form (B3) employing a catalyst σ without full rank with spectrum $\{\sigma_i\}$, can be also be implemented with a full-rank catalyst $\tilde{\sigma}$ with spectrum $\{\sigma_i | \sigma_i > 0\}$. In other words, we can assume without loss of generality that σ is full rank.

Appendix C: The classical case

In the following, we denote the marginals of a probability distribution r on $X \times Y$ by r_X resp. r_Y, such that $r_X(x) = \sum_{y\in Y} r(x,y)$ and $r_Y(y) = \sum_{x\in X} r(x,y)$. This is the classical analogue of the partial trace.

Conjecture 2 (Classical catalytic entropy conjecture)**.** *Let p and p' be two different probability distributions on a finite space of events X. Then the following two statements are equivalent:*

i) $S(p) \leq S(p')$, where S is the Shannon entropy.

ii) For every $\epsilon > 0$, there exists a probability distribution q on a finite space Y and a permutation P on $X \times Y$ such that

$$[P(p \otimes q)]_Y = q, \quad \|[P(p \otimes q)]_X - p'\|_1 \leq \epsilon. \tag{C1}$$

There are two reasons for which we only conjecture approximability of p' to arbitrary accuracy instead of perfect achievability. Firstly, in order to drop the rank condition from condition i); secondly, to account for the case in which p and p' differ by irrational amounts. In this case, permutations only realize the transition $p \to p'$ approximately.

Note that, since the statement of the catalytic entropy conjecture is unitarily invariant on the input states, and permutations are special cases of unitary operations, a proof of the classical catalytic entropy conjecture would essentially also prove the quantum version. The converse, however, is not necessarily true: it is apriori possible that only the quantum formulation holds. Nevertheless, as in the quantum case, one can show that the Shannon entropy is essentially the unique additive monotone. This follows from the following classical

version of Lemma 2. It uses the notation $\mathrm{rank}(p)$ to denote the number of non-zero entries of a discrete probability distribution p.

Lemma 6 (Weak solution to catalytic entropy conjecture (classical))**.** *Let p and p' be two different probability distributions of the same, finite dimension and with rational entries. Then the following two statements are equivalent:*

(I) $S(p') > S(p)$ *and* $\mathrm{rank}(p') \geq \mathrm{rank}(p)$.

(II) There exists a probability distribution q on a finite sample space Z, a d-dimensional sample space Y, and a permutation P on $X \times Y \times Z$ such that

$$[P(p \otimes \mathbf{1}_d \otimes q)]_Z = q, \quad [P(p \otimes \mathbf{1}_d \otimes q)]_{X,Y} = p' \otimes \mathbf{1}_d.$$

Here, $\mathbf{1}_d^\top = (1/d, \ldots, 1/d)$ denotes the uniform distribution on Y.

Proof. We only consider the non-obvious direction, i.e. we show that (I)⇒(II). According to Ref. [18], if condition (I) is satisfied, then there exists a probability distribution $\tilde{q}$ on some sample space $\tilde{Z}$ such that $p \otimes \tilde{q} \succeq p'\tilde{q}$. Since p, p' are rational, and so are $\tilde{q}$ and $p'\tilde{q}$, the majorization relation implies that this transition can be realized exactly with a random permutation. In other words, there exists a $\tilde{d}$-dimensional ancilla A in the state $\mathbf{1}_{\tilde{d}}$ and the global permutation $P = \sum_{i=1}^{\tilde{d}} \Pi_i \otimes P_i$, where Π_i denotes the rank-one projector onto the standard basis $\{\mathbf{e}_i\}$ of A, that is, $\Pi_i(q) = q_i \mathbf{e}_i$, such that

$$\frac{1}{\tilde{d}} \sum_{i=1}^{\tilde{d}} P_i(p \otimes \tilde{q}) = p'\tilde{q}. \tag{C2}$$

Next, choose $d = \tilde{d}$ as the dimension of Y and consider the permutation

$$P' = \sum_{i=1}^{d} (\Pi_i)_Y \otimes (\pi^i)_A, \tag{C3}$$

where π is a permutation defined by $\pi \mathbf{e}_j = \mathbf{e}_{j+1 \bmod d}$. Applying both of these permutations to the total system yields

$$\begin{aligned} &P'P\left[(\mathbf{1}_{\tilde{d}})_Y \otimes (\mathbf{1}_{\tilde{d}})_A \otimes p \otimes \tilde{q}\right] \\ =&P'\left[(\mathbf{1}_{\tilde{d}})_Y \otimes \left(\sum_i^d (\mathbf{e}_i)_A/d \otimes P_i(p \otimes \tilde{q})\right)\right] \\ =&\sum_{i,j=1}^{d} (\mathbf{e}_j)_Y/d \otimes (\mathbf{e}_{i+j \bmod d})_A/d \otimes P_i(p \otimes \tilde{q}). \end{aligned}$$

From the last expression, we see that summing over Y leaves A uncorrelated from both X and $\tilde{Z}$, since $\sum_j \Pi_{i+j}/d^2 = \mathbf{1}_d$, and summing over A leaves Y and X uncorrelated. Hence, by identifying $Z = A \times \tilde{Z}$ and $q = (\mathbf{1}_d)_A \otimes \tilde{q}$, the statement of the lemma follows. □

Appendix D: S is the only continuous additive monotone

Here we give a proof of Corollary 3. This corollary follows immediately from the following lemma, which itself has Lemma 2 as its key ingredient.

Lemma 7 (Properties of real and additive functions)**.** *Let f be a real function on the set of all finite-dimensional density matrices which is* continuous *(on all subsets of density matrices of fixed dimension) and* additive*, i.e. $f(\rho \otimes \sigma) = f(\rho) + f(\sigma)$. Furthermore, suppose that f is a* monotone *with respect to transitions of the form (b) of Conjecture 1, i.e. satisfaction of condition (b) implies that $f(\rho) \leq f(\rho')$. Then there exist a constant $a \geq 0$ and dimension-dependent constants $b_n \in \mathbb{R}$, such that*

$$f(\rho) = a \cdot S(\rho) + b_n,$$

with n the Hilbert space dimension of ρ, and $b_{m,n} = b_m + b_n$.

Proof. For any density matrix ρ of dimension n, define the negentropy $I(\rho) := \log n - S(\rho)$. Let ρ, ρ' be full-rank density matrices of possibly different dimensions n, n' such that $I(\rho) = I(\rho')$, then

$$S(\rho \otimes \mathbf{1}_{n'}) = \log n - I(\rho) + \log n' = S(\rho' \otimes \mathbf{1}_n).$$

Let $\epsilon > 0$, and let σ_ϵ be any full-rank state of size nn' such that $\|\sigma_\epsilon - \rho \otimes \mathbf{1}_{n'}\| < \epsilon$ and $S(\sigma_\epsilon) < S(\rho \otimes \mathbf{1}_{n'})$, then $S(\sigma_\epsilon) < S(\rho' \otimes \mathbf{1}_n)$, hence Lemma 2 implies that there is some $d \in \mathbb{N}$ such that $\sigma_\epsilon \otimes \mathbf{1}_d \to \rho' \otimes \mathbf{1}_n \otimes \mathbf{1}_d$, where "$\to$" denotes that a transition of the form (b) is possible. Thus

$$f(\sigma_\epsilon \otimes \mathbf{1}_d) \leq f(\rho' \otimes \mathbf{1}_n \otimes \mathbf{1}_d),$$

and additivity of f yields $f(\sigma_\epsilon) \leq f(\rho' \otimes \mathbf{1}_n)$. Since $\lim_{\epsilon \to 0} \sigma_\epsilon = \rho \otimes \mathbf{1}_{n'}$, and since f is continuous, this implies that $f(\rho \otimes \mathbf{1}_{n'}) \leq f(\rho' \otimes \mathbf{1}_n)$. Reversing the roles of ρ and ρ' in the above argumentation gives the converse inequality, and hence $f(\rho \otimes \mathbf{1}_{n'}) = f(\rho' \otimes \mathbf{1}_n)$. Define the new real function $j(\tau) := f(\mathbf{1}_n) - f(\tau)$, where n is the dimension of the density matrix τ, then j is also additive, and it vanishes on the maximally mixed states. Thus $j(\rho) = j(\rho')$.

In summary, we have shown that j is constant on the level sets of I. Thus, there is a real function $g : [0, \infty) \to \mathbb{R}$ such that $j(\rho) = g(I(\rho))$ for all ρ. Let $x, y \in [0, \infty)$ with $x < y$, and let ρ_x, ρ_y be finite-dimensional full-rank density matrices of dimensions n_x, n_y with $I(\rho_x) = x$ and $I(\rho_y) = y$. Then

$$\begin{aligned} g(x+y) &= g(I(\rho_x) + I(\rho_y)) = g(I(\rho_x \otimes \rho_y)) \\ &= j(\rho_x \otimes \rho_y) = j(\rho_x) + j(\rho_y) \\ &= g(I(\rho_x)) + g(I(\rho_y)) = g(x) + g(y). \end{aligned}$$

Furthermore, $S(\rho_y \otimes \mathbf{1}_{n_x}) < S(\rho_x \otimes \mathbf{1}_{n_y})$, hence there is some $d \in \mathbb{N}$ such that $\rho_y \otimes \mathbf{1}_{n_x} \otimes \mathbf{1}_d \to \rho_x \otimes \mathbf{1}_{n_y} \otimes \mathbf{1}_d$, therefore $j(\rho_y \otimes \mathbf{1}_{n_x} \otimes \mathbf{1}_d) \geq j(\rho_x \otimes \mathbf{1}_{n_y} \otimes \mathbf{1}_d)$, and additivity implies $j(\rho_y) \geq j(\rho_x)$. It follows that $g(y) \geq g(x)$.

We thus see that g is both *additive* and *non-decreasing*, and it is well-known (and easy to verify) that this implies that

$g(x) = ax$ for some $a \geq 0$, i.e. $j(\rho) = aI(\rho)$. Going back to the definition of f, this gives us

$$f(\rho) = aS(\rho) + b_n,$$

with n the dimension of ρ and $b_n := f(\mathbf{1}_n) - a \log n$. Finally, additivity of f and $\mathbf{1}_{m,n} = \mathbf{1}_m \otimes \mathbf{1}_n$ imply $b_{m,n} = b_m + b_n$. □

Note that $b_{m,n} = b_m + b_n$ does not automatically entail that b_m is proportional to $\log m$ (and thus to S_0): there are other well-known examples of functions on the integers which are additive in this sense.