Resource theory for work and heat

Carlo Sparaciari,[1,*] Jonathan Oppenheim,[1,†] and Tobias Fritz[2,3,‡]

[1]*Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom*
[2]*Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany*
[3]*Perimeter Institute for Theoretical Physics, 31 Caroline St N, Waterloo, Ontario N2L 2Y5, Canada*

Several recent results on thermodynamics have been obtained using the tools of quantum information theory and resource theories. So far, the resource theories utilized to describe thermodynamics have assumed the existence of an infinite thermal reservoir, by declaring that thermal states at some background temperature come for free. Here, we propose a resource theory of quantum thermodynamics without a background temperature, so that no states at all come for free. We apply this resource theory to the case of many noninteracting systems and show that all quantum states are classified by their entropy and average energy, even arbitrarily far away from equilibrium. This implies that thermodynamics takes place in a two-dimensional convex set that we call the *energy-entropy diagram*. The answers to many resource-theoretic questions about thermodynamics can be read off from this diagram, such as the efficiency of a heat engine consisting of finite reservoirs, or the rate of conversion between two states. This allows us to consider a resource theory which puts work and heat on an equal footing, and serves as a model for other resource theories.

I. INTRODUCTION

To make precise statements about thermodynamics, particularly at the quantum scale, we need to precisely define what thermodynamics is. In particular, we have to specify what an experimenter is allowed to do to take a system from one state into another. This specification defines a resource theory, something which has been used to successfully describe thermodynamic phenomena occurring at the microscopic scale [1–16]. Any line of research which specifies what the rules of thermodynamics are can be said to define a resource theory [17–26]. These theories typically consist of a state space and a set of allowed operations that can be performed on the states (see, e.g., Refs. [27–31] for reviews).

The resource theories developed so far for quantum thermodynamics are based on assuming that thermal states (Gibbs states) at a fixed background temperature come for free. In these theories, states are classified by their free energy, and this quantity also equals the amount of work that can be extracted from many copies of a given state [2,5,16,20,22]. However, declaring those thermal states to be free boils down to assuming the existence of an infinite thermal reservoir. This cannot always be taken for granted [32,33]. In some applications, such as many types of engines, the system under consideration operates on such short timescales that it must be considered a closed system. In other applications, the environment is finite and its state changes due to the interaction with the system, for example, when a power plant dumps large amounts of heat into the environment. It seems therefore imperative to develop thermodynamics as a resource theory *without* assuming the existence of an infinite thermal reservoir. This is the aim of the present paper.

The set of allowed operations in our resource theory is much broader than the one a typical experimentalist can implement, as any energy-preserving unitary is allowed. Therefore, our theory primarily delineates fundamental limitations to what is possible in "real life." However, we suspect that our abstract results can actually be implemented using a more realistic set of operations, as in the resource theory of Thermal Operations [14], where only changing the energy levels of the system and thermal contact with a heat bath is allowed. But for the time being, our results should be seen as upper limits, only achievable in idealized conditions. Moreover, the results that we present within our theory are concerned with the asymptotic regime, i.e., the limit of many noninteracting identical systems. This follows the abstract approach to resource theories developed in Ref. [34]. Although this is a limitation, we believe that one needs to understand the asymptotic structure of a resource theory first before analyzing the single-shot regime, and this is what we achieve here for thermodynamics.

In addition to providing a general framework for thermodynamics in the absence of an infinite bath, we prove in Theorem 1 that two quantum states are asymptotically equivalent under energy-preserving unitaries if and only if they have same entropy and average energy. Because of this equivalence, we interpret the specification of entropy and average energy of a state as the description of a thermodynamic *macrostate*. Thermodynamics in the asymptotic limit is then studied by only considering the set of macrostates, which we call the *energy-entropy diagram*. This diagram is a complete description of thermodynamics in the asymptotic limit. All of this takes place arbitrarily far away from equilibrium.

We then use our methods to study rates of conversion between two states of a closed system and to propose a definition of the work and heat exchanged while interconverting two states using a finite thermal reservoir and a battery. The resulting expressions for work and heat, Eqs. (21) and (22), recover the standard ones in the limit of an infinite thermal reservoir.

---

*carlo.sparaciari.14@ucl.ac.uk
†j.oppenheim@ucl.ac.uk
‡fritz@mis.mpg.de

II. FRAMEWORK AND ALLOWED OPERATIONS

The systems we consider in our resource theory consist of n copies of a single d-level system described by $\mathbb{C}^d$ (a *qudit*)

with fixed Hamiltonian H; both d and H are parameters of the theory while n varies. We assume the total Hamiltonian H_{tot} of the total n-copy system to be the sum of single-qudit Hamiltonians H, each of them acting on a different copy, which makes the different copies noninteracting. The resource objects of our theory are quantum states on such an n-qudit system, for arbitrary $n \in \mathbb{N}$. The allowed operations for turning one such state into another are all the global unitaries U acting on the total system which are energy preserving, $[U, H_{\text{tot}}] = 0$. Thus, we assume perfect control over our closed system, and the sole limitation is set by the first law of thermodynamics, requiring conservation of energy. The class of operations is purposely broad, as we are interested in fundamental limitations imposed by the laws of nature, as opposed to the ones imposed by our limited control over macroscopic systems. Because of the unitary nature of the operations, the state transformations we consider are reversible by definition. For simplicity, we assume the Hamiltonian to be fixed throughout, without any possibility of changing it. In Theorem 1, we also permit the use of an ancilla system of sublinear size and energy spectrum, which can be initialized in an arbitrary state and gets discarded at the end. While the ancilla allows us to act more freely over the main system, it does not modify the physical assumptions made so far. Indeed, since we work in the thermodynamic limit, the sublinearity of the ancilla makes its energetic and entropic contribution (per single copy of the system) vanishingly small. Finally, when talking about rates of conversion, we permit discarding subsystems that are decoupled from the rest.

Our resource theory can describe both the thermodynamics of closed systems (where the thermal reservoir is absent), and open systems interacting with a finite thermal reservoir (where the size of the system is comparable to the one of the reservoir). Thus, our framework extends the one of Thermal Operations [1–5,20], in which one can add an arbitrary number of thermal states $\tau_\beta = Z_\beta^{-1} e^{-\beta H}$ at a given temperature β^{-1}, with $Z_\beta = \text{Tr}[e^{-\beta H}]$ being the partition function of the system. Indeed, adding arbitrary many thermal states is equivalent to the system being in contact with a reservoir with infinite heat capacity. By not allowing this possibility, we obtain a theory which can describe, among other things, systems in contact with a finite reservoir in which thermal states are themselves a valuable resource, and recover the case of an infinite reservoir in the limit. In the following, we use the notation log to denote the logarithm in base 2.

III. ASYMPTOTIC EQUIVALENCE OF QUANTUM STATES

Our resource theory clearly has many conserved quantities: Since the allowed operations are all unitaries, one can be converted into another only if they live on the same number of qudits and have the same spectrum. Moreover, since our unitaries are energy preserving, the states must have the same *distribution* over the energy levels, or equivalently the same moments of energy. This makes our theory very restrictive.

However, at the asymptotic level the situation is quite different: As it turns out, two quantum states can be interconverted if and only if they have the same entropy and average energy [35]. In particular, we say that a state ρ is *asymptotically equivalent* to another state σ, and write $\rho \asymp \sigma$ if the equivalent conditions of the following theorem hold.

Theorem 1. For states ρ and σ on any quantum system of dimension d with given Hamiltonian H, the following are equivalent:

  (a) The states have equal entropy and average energy,

$$S(\rho) = S(\sigma), \quad E(\rho) = E(\sigma). \tag{1}$$

  (b) There exists an ancillary system A of $O(\sqrt{n \log n})$ many qudits whose Hamiltonian H_A satisfies $\|H_A\| \leqslant O(n^{2/3})$ with state η as well as an energy-preserving unitary U such that

$$\|\text{Tr}_A[U(\rho^{\otimes n} \otimes \eta)U^\dagger] - \sigma^{\otimes n}\|_1 \xrightarrow{n \to \infty} 0. \tag{2}$$

In the above theorem, we have that $\|X\|_1 = \text{Tr}[\sqrt{X^\dagger X}]$ is the trace norm, $E(\rho) = \text{Tr}[H\rho]$ is the average energy, and $S(\rho) = -\text{Tr}[\rho \log \rho]$ is the von Neumann entropy, which was shown to be equivalent to a thermodynamic entropy [36], and coincides, in the thermodynamic limit, with the Boltzmann entropy (see, for example, Ref. [5] in a similar context as the present work). We prove this theorem in the Appendix, and also show that the bounds on the size and Hamiltonian of the ancilla can equivalently be taken to be $o(n)$. This sublinearity is essential: The amount of entropy and energy that can be exchanged between the system and the ancilla tends to 0 as $n \to \infty$, when measured per copy of the system. This sketches the reason why asymptotic equivalence implies equal entropy and average energy.

To show the opposite direction, we need to specify a protocol which achieves (2), turning $\rho^{\otimes n}$ into something close (in trace norm) to $\sigma^{\otimes n}$. Concretely, our ancilla is composed of three subsystems, each of them playing a different role in our protocol. The first subsystem provides a source of randomness, used to modify the spectrum of the state $\rho^{\otimes n}$; its state is maximally mixed, and its Hamiltonian is trivial. The second subsystem is used as a register and allows us to dilate slightly irreversible operations on the global system to reversible ones. Its initial state is pure, and the Hamiltonian is again trivial. The third subsystem makes the transformation energy preserving and allows for introducing and removing coherence in the energy eigenbasis. It has a nontrivial Hamiltonian, and its state is in a uniform superposition of the energy eigenstates. Overall, the ancilla satisfies the properties listed in (b). The full details are given in the Appendix.

Thus we can interconvert asymptotically, using the set of allowed operations, between states with the same entropy S and average energy E, in a reversible manner. Consequently, we can classify any quantum state asymptotically in terms of these two quantities only. Such a passage from quantum to macroscopic states is at the core of thermodynamics in the guise of the passage from *microstates* to *macrostates*. Our result seems to capture this, despite being built on the idealized assumption of noninteracting copies. From now on, we identify the many-copy limit that one takes when considering asymptotic equivalence with the standard macroscopic limit of thermodynamics.

IV. THE ENERGY-ENTROPY DIAGRAM

Theorem 1 shows that it is only the energy $E(\rho)$ and entropy $S(\rho)$ of a state ρ that determines its behavior under many-copies transformations. Hence, as far as the many-copies level is concerned, we can *identify* a state ρ with the pair of values
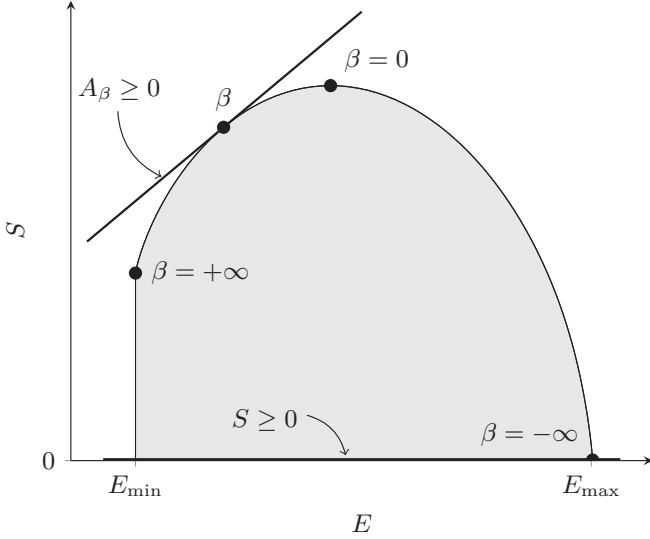
FIG. 1. The energy-entropy diagram representing the state space of a quantum system with Hamiltonian H, degenerate in the ground state (vertical line). The physical points are inside the gray area. Each point (E,S) represents an equivalence class of microstates, i.e., a single macrostate. Inequality (7), postulating the non-negativity of entropy, is satisfied by the points above the E axis. For a given β, Ineq. (8) is satisfied by those points below the drawn line which is tangent to the physical region and goes through the point $(E(\tau_\beta), S(\tau_\beta))$ with slope β.

$(E(\rho), S(\rho)) \in \mathbb{R}^2$. In order to understand thermodynamics as a resource theory asymptotically, we therefore need to ask which pairs of numbers $x = (x_E, x_S) \in \mathbb{R}^2$ arise from a state in this manner. We call this set the *energy-entropy diagram*. The energy-entropy diagram depends on the system Hamiltonian H and can be characterized as follows:

Proposition 2. The values $(E(\rho), S(\rho))$ form a closed convex subset of $\mathbb{R}^2$ as in Fig. 1, where the lower boundary is the line $x_S = 0$ and the upper boundary is the curve $\beta \mapsto (E(\tau_\beta), S(\tau_\beta))$ traced by the thermal states $\tau_\beta = Z_\beta^{-1} e^{-\beta H}$ for $\beta \in [-\infty, +\infty]$.

Here, we use Cartesian coordinates $x = (x_E, x_S)$ when speaking about a point in $\mathbb{R}^2$ that may or may not belong to the energy-entropy diagram, and we also write $x(\rho) := (E(\rho), S(\rho))$ for the point associated to a specific state ρ. Thus we have the components $x(\rho)_E = E(\rho)$ and $x(\rho)_S = S(\rho)$, and the curve of thermal states is given by $\beta \mapsto x(\tau_\beta)$.

Proof. It is a standard fact that the states of maximal entropy for a given energy are precisely the thermal states. One way to see that this also holds for $\beta < 0$ is by using the fact that it holds for $\beta > 0$ and reversing the sign of the Hamiltonian.

It remains to show that also every smaller value of the entropy is achievable for a given energy. Clearly $S = 0$ is achievable, namely by considering a pure state in a suitable superposition of energy levels which has the desired expectation value of energy. Moreover, the set of all states of a given energy is a convex subset of all the density matrices, and in particular it is path connected; since the map $\rho \mapsto S(\rho)$ is continuous, it follows that also its image under S is path connected, so that also all intermediate entropy values can be achieved for the given energy.

Convexity follows from the alternative characterization in terms of linear inequalities that we will derive as Proposition 4. ∎

Remark 3. The curve of thermal states is parametrized by $\beta \mapsto (E(\tau_\beta), S(\tau_\beta))$, where

$$E(\tau_\beta) = \text{Tr}\left[H \frac{e^{-\beta H}}{Z_\beta}\right] = -\frac{d \log Z_\beta}{d\beta}, \tag{3}$$

$$S(\tau_\beta) = -\text{Tr}\left[\frac{e^{-\beta H}}{Z_\beta} \log\left(\frac{e^{-\beta H}}{Z_\beta}\right)\right] = \beta E(\tau_\beta) + \log Z_\beta. \tag{4}$$

Differentiating with respect to β and collecting terms results in the fundamental thermodynamic relation $dS(\tau_\beta) = \beta \, dE(\tau_\beta)$. For the energy-entropy diagram, this implies that the parameter β is precisely equal to the slope of the tangent at each point on the curve of thermal states.

By virtue of being convex and topologically closed, one can describe the energy-entropy diagram also in a dual way by writing down all the linear inequalities that bound it. These inequalities are most conveniently stated in terms of the quantities

$$A_\beta(x) := \beta x_E - x_S + \log Z_\beta, \tag{5}$$

so that $A_\beta(\tau_\beta) = 0$. Note that $A_\beta(x)$ is a linear function of x.

For a given value of β, we call A_β the *β-athermality* [37], since it vanishes on the thermal state $x(\tau_\beta)$, and we think of $A_\beta(x)$ as a measure of how far x is from being equal to $x(\tau_\beta)$. The β-athermality differs from the free energy $x_E - \beta^{-1} x_S$ only by an additional factor of β and an additive constant. One of the reasons that we prefer using (5) over the free energy is that on a state ρ, we can also neatly write it as the relative entropy distance to the thermal state,

$$A_\beta(x(\rho)) = D(\rho \| \tau_\beta), \tag{6}$$

as one can see by writing out the definition of relative entropy and plugging in $\tau_\beta = Z_\beta^{-1} e^{-\beta H}$. This again justifies the term "β-athermality." For $\beta = 0$, we obtain the negentropy $A_0(x(\rho)) = \log d - S(\rho)$.

We can now state the characterization of the energy-entropy diagram by linear inequalities.

Proposition 4. The energy-entropy diagram is the set of all points $x = (x_S, x_E) \in \mathbb{R}^2$ such that

$$x_S \geqslant 0, \tag{7}$$

$$A_\beta(x) \geqslant 0 \quad \forall \beta \in \mathbb{R}. \tag{8}$$

Proof. All the inequalities hold for an achievable point $x(\rho) = (E(\rho), S(\rho))$, since both the entropy and the relative entropy (6) are non-negative.

In the other direction, we need to show that if $x \in \mathbb{R}^2$ satisfies all the claimed inequalities, then it lies in the energy-entropy diagram. So by assumption, we have $x_S \geqslant 0$ and

$$\beta x_E - x_S \geqslant \beta E(\tau_\beta) - S(\tau_\beta)$$

for all $\beta \in \mathbb{R}$. Taking $\beta \to +\infty$ and $\beta \to -\infty$ shows that we also must have $E_{\min} \leqslant x_E \leqslant E_{\max}$. Since there is a unique thermal state at every given energy in this range, there is a unique β such that $E(\tau_\beta) = x_E$. Using this β, we obtain from

the previous inequality

$$x_S \leqslant S(\tau_\beta) + \beta[x_E - E(\tau_\beta)] = S(\tau_\beta),$$

so that x lies indeed below the curve traced by the thermal states in Fig. 1. ∎

Remark 5. In Propositions 2 and 4, and in some of our upcoming results, we also consider negative temperatures, that is, $\beta < 0$. This is due to the fact that our systems are finite; for infinite systems, there is no E_{max} and no thermal state at $\beta = 0$, meaning that the diagram is unbounded to the right and bounded by the thermal curve which increases forever. While we expect our methods to still apply in this case, our present proofs only hold for finite systems.

While thermal states at a negative temperature do not usually arise as a result of thermalisation, they still play an important roles in multiple physical effects (such as, for instance, in lasers, where coherent light amplification is obtained through population inversion) [38]. The main difference between thermal states at $\beta > 0$ and $\beta < 0$, in our theory, is that the former are completely passive states from which we cannot extract energy by means of unitary operations, while the latter are active states, from which energy can be extracted.

V. THERMODYNAMIC VARIABLES AND THE CONVEX CONE OF MACROSTATES

In this section, we introduce an additional macroscopic quantity to characterize the state of a thermodynamic system, referred to as *system size* or *amount of substance*. With this parameter, we can fully characterize thermodynamic transformations on any number of copies of the system, and later also allow for discarding subsystems, so that the number of systems involved changes. So let us consider what happens at the many-copies level asymptotically.

Proposition 6. For any $n \in \mathbb{N}$, the energy-entropy diagram of the n-copy Hamiltonian $H^{(n)}$ equals the energy-entropy diagram of H, scaled up by a factor of n.

Proof. Since $E(\rho^{\otimes n}) = nE(\rho)$, and similarly for S, it is clear that the energy-entropy diagram of H, when scaled by n, is contained in the energy-entropy diagram of $H^{(n)}$. The converse follows from Proposition 2, because any thermal state of $H^{(n)}$ is an nth tensor power of a thermal state of H. ∎

Together with Theorem 1, this also implies that for every n-system state ρ there is a single-system state σ such that $\rho \asymp \sigma^{\otimes n}$ (that is, ρ is asymptotically equivalent to $\sigma^{\otimes n}$), although ρ itself may be arbitrarily far from being a product state.

In order to keep track of $n = n(\rho)$, the number of copies of the system on which a state ρ lives, it is useful to consider the *triple* of numbers $(E(\rho), S(\rho), n(\rho)) \in \mathbb{R}^3$, for which we also write $y(\rho)$. Each component of this triple is an additive function of ρ, and therefore

$$y(\rho \otimes \sigma) = y(\rho) + y(\sigma). \tag{9}$$

By Theorem 1, the three components of $y(\rho)$ provide a complete classification of single-system and multisystem states in thermodynamics—with given single-system Hamiltonian H—up to asymptotic equivalence.

Now we could consider the set of all points $y = (y_E, y_S, y_n)$ that are of the form $y = y(\rho)$ for some state ρ, and call it the *energy-entropy-size diagram* associated to the Hamiltonian H. But all of our results are only up to asymptotic equivalence, so that we effectively only consider states ρ with $n(\rho) \gg 1$. Equivalently, we can also work with small values of n, but then forget that n is required to be an integer by pretending that the system size can be an arbitrary non-negative real number. We then still use the symbol n, although it now plays the role of an "amount of substance," just as in the ideal gas law $pV = nRT$. In principle, converting between number of microsystems and amount of substance involves rescaling by the Avogadro constant. This is another perspective on what we are doing here, except that we choose to measure the amount of substance with the unit in which the Avogadro constant is equal to 1. Based on this intuition, we thus use Definition 7.

Definition 7. The convex cone Therm(H) consists of all points $y \in \mathbb{R}^3$ that are of the form $y = n(x_E, x_S, 1)$ for some $n \in \mathbb{R}_{\geqslant 0}$ and (x_E, x_S) in the energy-entropy diagram.

In other words, Therm(H) is the convex cone that we obtain by taking the energy-entropy diagram in $\mathbb{R}^2$ and applying the standard "homogenization" trick for turning a convex set into a convex cone by adding an additional coordinate [39, p. 31]. We call a point $y \in $ Therm(H) *normalized* if $y_n = 1$. Every nonzero $y \in $ Therm(H) is a unique scalar multiple of a normalized point, so that for most purposes it is sufficient to consider normalized points only (see Sec. VI).

Remark 8. If we slice Therm(H) at constant third coordinate $n \in \mathbb{N}$ by considering all points of the form $(x_E, x_S, n) \in $ Therm(H), then this set is precisely the energy-entropy diagram of $H^{(n)}$ thanks to Proposition 6.

Remark 9. Taking every (multisystem) state ρ to be represented by a point $y(\rho) \in $ Therm(H) is a standard construction of thermodynamics: It corresponds to passing from the microstate to the macrostate. The thermodynamic variables of a macrostate are precisely [40] energy E, entropy S, and system size n, and the macrostate is specified completely by these three numbers. If one identifies the passage from microstate to macrostate with the information-theoretic many-copies limit, then our Theorem 1 offers a mathematically rigorous explanation for *why* the macroscopic variables are exactly these three and no others. There *are* many other extensive quantities that are invariant under energy-preserving unitaries—take the Rényi entropies or the variance of energy as examples. These quantities would indeed be relevant also macroscopically if we had required an *exact* conversion of $\rho^{\otimes n}$ into $\sigma^{\otimes n}$ for some n, possibly together with a sublinear ancilla. But our definition of many-copy equivalence allows for approximate conversions that become closer to exact as $n \to \infty$. This is a more permissive notion of asymptotic equivalence, under which correspondingly fewer quantities are invariant, namely only the ones that are asymptotically continuous [41]. In the language of Refs. [28,34], allowing such approximate conversions introduces an "epsilonification."

One may wonder how it is possible that the passage from microstate to macrostate within our idealized theory yields results that are so close to the standard one. For example, the class of allowed operations considered in our model is extremely wide, and moreover our results are valid only in the many-copy limit. This is at least partly explained by noting that the many-copy limit is a faithful enough description (at the macroscopic level) since the interactions between particles are

small (they have an area scaling), compared to the extensive quantities (which have a volume scaling), as the number of particles grows to infinity. Moreover, results such as the von Neumann's quantum ergodic theorem [42] help explain this phenomena.

Returning to technical developments, we extend the β-athermalities from $\mathbb{R}^2$ to $\mathbb{R}^3$ by setting

$$A_\beta(y) := \beta y_E - y_S + y_n \log Z_\beta. \qquad (10)$$

On the energy-entropy diagram, which is embedded in $\mathrm{Therm}(H)$ as the set of all normalized points, this coincides with our previous definition of $A_\beta(x)$, and from there we have extended linearly.

On an actual state ρ, we can again express the β-athermality as a relative entropy distance,

$$A_\beta(y(\rho)) = D\big(\rho \| \tau_\beta^{\otimes n(\rho)}\big),$$

where the $n(\rho)$ appears because one needs to consider the thermal state on a suitable number of copies of the system.

The characterization of the energy-entropy diagram by linear inequalities extends easily to $\mathrm{Therm}(H)$.

Proposition 10. The convex cone $\mathrm{Therm}(H)$ is the set of all $y = (y_E, y_S, y_n) \in \mathbb{R}^3$ such that $y_S \geqslant 0$ and $A_\beta(y) \geqslant 0$ for all $\beta \in (-\infty, +\infty)$.

Proof. Any point in $\mathrm{Therm}(H)$ satisfies these inequalities thanks to Proposition 4 together with the fact that $A_\beta(\lambda y) = \lambda A_\beta(y)$ for all $\lambda > 0$, so that it is sufficient to consider normalized points $y \in \mathrm{Therm}(H)$ only. Conversely, suppose that $y \in \mathbb{R}^3$ satisfies all these inequalities. Then from $y_S \geqslant 0$ and $A_0(y) = y_n \log d - y_S \geqslant 0$ we conclude $y_n \geqslant 0$. If it is the case that $y_n = 0$, then we conclude $y_S = 0$, and then also $y_E = 0$ from considering $A_\beta(y) \geqslant 0$ in the two limits $\beta \to \pm\infty$. Otherwise, we have $y_n > 0$, and the point $y_n^{-1}(y_E, y_S)$ satisfies all the inequalities necessary to lie in the energy-entropy diagram by Proposition 4, and therefore $y \in \mathrm{Therm}(H)$. ∎

What this says is that there are two kinds of additive resource monotones that are relevant to thermodynamics:

(1) The entropy function $\rho \mapsto S(\rho)$;

(2) The β-athermality functions $\rho \mapsto A_\beta(\rho)$ indexed by $\beta \in (-\infty, +\infty)$.

In the terminology of Ref. 34, Sec. 7, these are *extremal* monotones. There are two more extremal monotones that one obtains by considering A_β as $\beta \to \pm\infty$, which results in the two functions

$$\rho \mapsto E(\rho) - n(\rho) E_{\min}, \qquad \rho \mapsto n(\rho) E_{\max} - E(\rho).$$

It follows from Ref. [34, Corollary 7.9] that every other additive (and suitably continuous) monotone is a non-negative linear combination or integral of these extremal ones.

VI. MACROSCOPIC THERMODYNAMICS AS A GENERAL PROBABILISTIC THEORY

As we will illustrate in the upcoming sections, pretty much any resource-theoretic question about macroscopic thermodynamics can be formulated and answered within the convex cone picture that we have developed. However, since the cone $\mathrm{Therm}(H) \subseteq \mathbb{R}^3$ may be a bit challenging to visualize, it helps the intuition to represent any macrostate $y = (y_S, y_E, y_n) \in \mathrm{Therm}(H)$ by the corresponding normalized macrostate $x :=$
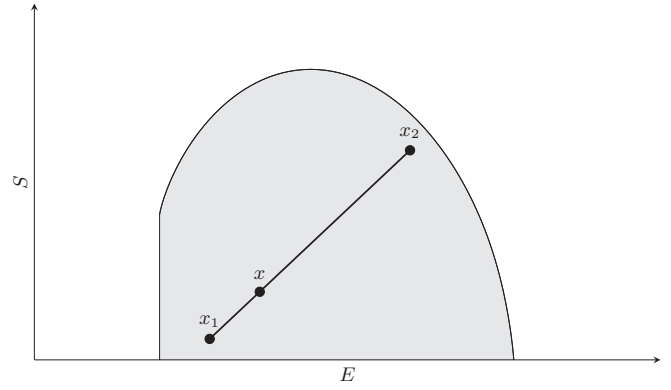
FIG. 2. Combining two systems in normalized macrostates x_1 and x_2 results in a total system in the normalized macrostate x, which is a convex combination of x_1 and x_2, where the coefficient of x_1—proportional to the distance between x and x_2—is equal to the size of the first system relative to the total system, and similarly for x_2.

$y_n^{-1}(y_E, y_S)$ in the energy-entropy diagram, equipped with a weight of y_n. In this picture, combining systems as in (9) corresponds to taking a convex combination of normalized macrostates, in the sense that

$$\frac{y(\rho \otimes \sigma)}{n(\rho \otimes \sigma)} = \frac{n(\rho)}{n(\rho \otimes \sigma)} \frac{y(\rho)}{n(\rho)} + \frac{n(\sigma)}{n(\rho \otimes \sigma)} \frac{y(\sigma)}{n(\sigma)}, \qquad (11)$$

where $\frac{y(\rho \otimes \sigma)}{n(\rho \otimes \sigma)}$, $\frac{y(\rho)}{n(\rho)}$, and $\frac{y(\sigma)}{n(\sigma)}$ are all normalized macrostates. Normalizing by dividing by system size turns the energy and entropy coordinates, which are extensive quantities, into intensive quantities. While extensive quantities combine across subsystems additively, the associated intensive ones combine across subsystems via convex combinations with coefficients given by the relative subsystem sizes (Fig. 2). This implies that everything that we do with the convex cone can alternatively be done directly in the energy-entropy diagram, by simply normalizing the macrostates and keeping track of system size separately.

At the purely mathematical level, all of this is nicely analogous to the issue of normalization of density matrices: It is usually more intuitive to assume the normalization, and therefore one often normalizes explicitly; but it is occasionally also advantageous to use un-normalized density matrices in order to keep track of the normalization, which represents a "probability to occur," analogous to our system size coordinate. Conversely, it is often useful to decompose a given normalized density matrix into a convex combination of other ones, such as pure states; many of the puzzling features of quantum theory can be attributed to the fact that such a decomposition is highly nonunique [43]. The same applies to thermodynamics: It may occasionally be useful to write a normalized macrostate x as a convex combination of other ones, or equivalently to decompose a given $y \in \mathrm{Therm}(H)$ into a sum $y = y_1 + y_2$ for $y_i \in \mathrm{Therm}(H)$. Of particular interest are decompositions into normalized macrostates that are extreme points of the energy-entropy diagram. Again such decompositions are highly nonunique, and we will argue that this nonuniqueness is among the essential features of thermodynamics and underlies, e.g., the possibility of constructing heat engines (Sec. IX). What
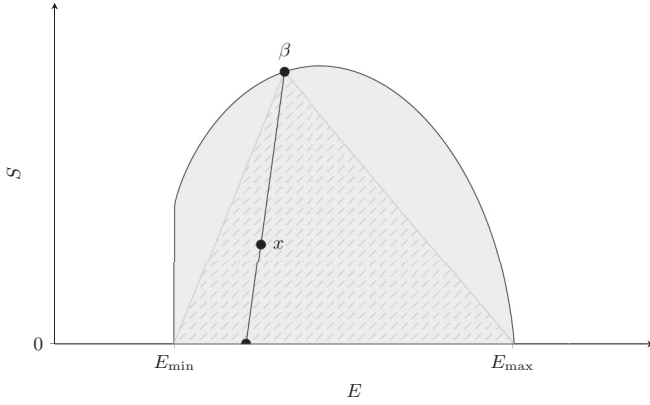
FIG. 3. Decomposing a normalized macrostate into a combination of the thermal macrostate at temperature β^{-1} and a pure macrostate, together with the set of all states that have such a decomposition at the given β (hatched).

this means is that macroscopic thermodynamics is, purely mathematically, an example of a *general probabilistic theory* [44–46]. The physical meaning, however, is very different from how one usually thinks of a general probabilistic theory such as quantum theory, and in fact already applies at the level of *classical* thermodynamics, and in fact already applies at the level of *classical* thermodynamics.

In more detail, Proposition 2 implies that the extreme points of the energy-entropy diagram are the following:

(1) The thermal macrostates $x(\tau_\beta)$, for $\beta \in [-\infty, +\infty]$, which are all different unless $H = 0$.

(2) The pure macrostate $x(|E_{\min}\rangle\langle E_{\min}|)$, which coincides with the ground state $x(\tau_\infty)$ in the case of nondegeneracy, and the pure macrostate $x(|E_{\max}\rangle\langle E_{\max}|)$, which may similarly coincide with the maximally excited state $x(\tau_{-\infty})$.

By first writing every normalized macrostate as a convex combination of a thermal state and a pure macrostate (Fig. 3), and then decomposing the pure macrostate further into a combination of ground states and maximally excited states, we can represent every normalized macrostate as a convex combination of a suitable thermal state $x(\tau_\beta)$ with $x(|E_{\min}\rangle\langle E_{\min}|)$ and $x(|E_{\max}\rangle\langle E_{\max}|)$. For a given β, Fig. 3 shows the region of normalized macrostates x that have a decomposition of this form for fixed β. So we can write any $y(\rho) \in \text{Therm}(H)$ as

$$\frac{y(\rho)}{n(\rho)} = c_\beta x(\tau_\beta) + c_{\min} x(|E_{\min}\rangle\langle E_{\min}|)$$
$$+ c_{\max} x(|E_{\max}\rangle\langle E_{\max}|), \quad (12)$$

for suitable weights $c_\beta, c_{\min}, c_{\max} \in [0,1]$. If we choose rational approximations for these coefficients and suitably rescale the system size such that the product of each coefficient with $n(\rho)$ is an integer, then we even obtain an asymptotic equivalence,

$$\rho \; \asymp \; \tau_\beta^{\otimes c_\beta n(\rho)} \otimes |E_{\min}\rangle\langle E_{\min}|^{\otimes c_{\min} n(\rho)}$$
$$\otimes |E_{\max}\rangle\langle E_{\max}|^{\otimes c_{\max} n(\rho)}. \quad (13)$$

In this way, any state looks macroscopically like a combination of a number of thermal states (at one temperature), ground states, and maximally excited states. The nonuniqueness in this decomposition lies in the possibility of choosing the

temperature β^{-1}; the number of states of each kind in the decomposition will vary with that temperature. In practice, one can fix β first and then determine the coefficients by equating energy, entropy, and system size of the two sides of the equation. If these coefficients turn out to be non-negative, then one has found a feasible decomposition. This is an instance of decomposing a state in a general probabilistic theory into pure states.

In summary, at the mathematical level, macroscopic thermodynamics is a general probabilistic theory. Some of the essential features of thermodynamics are intimately related to the nonuniqueness of decompositions of states into extremal states—the same phenomenon that is behind many of the mysterious aspects of quantum theory. But although the mathematics is an instance of the formalism of general probabilistic theories, the physical meaning is very different from the standard interpretation of the latter [44].

VII. THE MAXIMAL EXTRACTABLE WORK

We now use the methods developed in the previous sections to analyze the maximal work that can be extracted from many copies of a given state ρ. Following the standard definition of *work*, we introduce a separate subsystem called *battery* consisting of ℓ copies of the pure ground state,

$$\omega_{\text{in}} = \rho^{\otimes n} \otimes |E_{\min}\rangle\langle E_{\min}|^{\otimes \ell},$$

where ℓ is yet to be determined. Extracting work from $\rho^{\otimes n}$ means that one devises a protocol which turns ω_{in} into a final state in which the battery system is in the maximally excited state,

$$\omega_{\text{fin}} = \sigma^{\otimes n} \otimes |E_{\max}\rangle\langle E_{\max}|^{\otimes \ell},$$

where σ is a state to be determined in such a way that the work, which is the energy exchanged with the battery per copy of ρ,

$$W = \frac{\ell}{n}(E_{\max} - E_{\min}) = E(\rho) - E(\sigma)$$

is maximized. By Theorem 1, such a protocol exists if and only if ω_{in} and ω_{fin} have the same energy and entropy. Since the battery subsystem has no entropy, this means that we need to have $S(\rho) = S(\sigma)$; and moreover, in order to maximize ℓ, we should choose σ such that its energy is minimized. Having minimal energy for given entropy $S(\rho)$ means that we draw the horizontal line in the energy-entropy diagram through the macrostate $x(\rho)$, and determine the point where it hits the boundary; then we need to choose σ to be a representative of this macrostate. In the regime where $S(\rho)$ is greater than the ground-state degeneracy, we therefore have $\sigma = \tau_{\tilde\beta}$, where the temperature $\tilde\beta^{-1} > 0$ is such that $S(\tau_{\tilde\beta}) = S(\rho)$. In this case, the maximal amount of work that we can extract asymptotically per copy of ρ is given by

$$W_{\max}(\rho) = E(\rho) - E(\tau_{\tilde\beta}), \quad (14)$$

in agreement with existing results [16,47].

In conclusion, the maximal total amount of pure energy that can be extracted from $\rho^{\otimes n}$ can be neatly read off the energy-entropy diagram: It is given by the horizonal distance between the macrostate $x(\rho)$ and the curve of thermal states, multiplied by the system size n.

VIII. WORK AND HEAT

As we have seen in the previous section, coupling a system to a battery allows for the exchange of energy with the battery. So in contrast to the situation of Theorem 1, this means that now only entropy is a conserved quantity, and one can move between any two macrostates which are on the same horizontal line in the energy-entropy diagram.

So what can we do in order to move between states that are not even on the same horizontal line? In analogy with adjoining a battery, it is natural to do this by adding a thermal reservoir of finite size, with which the original system can then exchange *heat*. So while energy exchanged with the battery is what we consider work, energy exchanged with the thermal reservoir is our definition of heat. Assuming an environment to consist of a reservoir plus a battery is also motivated by the decomposition of any macrostate into a thermal and a pure part as in the previous section.

Again using our previous results, we will compute the work and heat required to transform any given state ρ into any desired state σ, as a function of the initial and final temperature of the reservoir. When the size of the reservoir tends to infinity, or equivalently when its initial and final temperature coincide, the work and heat exchanged specialize to the standard ones in terms of the free energy.

Getting to the technicalities, each subsystem consists of any number of microsystems with Hamiltonian H as before. We assume that the thermal reservoir consists initially of m copies of the thermal state at some temperature β_1^{-1}, and the battery of ℓ copies of a pure ground state. The initial state of the total system is therefore given by

$$\omega_{\text{in}} = \rho^{\otimes n} \otimes \tau_{\beta_1}^{\otimes m} \otimes |E_{\min}\rangle\langle E_{\min}|^{\otimes \ell}. \tag{15}$$

Since ρ and σ may have different entropy and average energy, turning the former into the latter means that we also have to modify the reservoir and battery states. In fact, in order to convert ρ into σ, we apply Theorem 1 to the final state of all three subsystems, which we assume to be close to

$$\omega_{\text{fin}} = \sigma^{\otimes n} \otimes \tau_{\beta_2}^{\otimes m} \otimes |E_{\max}\rangle\langle E_{\max}|^{\otimes \ell}, \tag{16}$$

where now the reservoir is in a thermal state at a possibly different temperature β_2^{-1}, and the battery is in the maximally excited state.

It is worth noting that, in general, the final state of the reservoir does not have to be thermal, since the interaction with the system might have driven the environment out of equilibrium. However, if the final state of the reservoir is athermal, one would be able to extract additional work from it, while keeping the entropy of this system unchanged (as we have previously shown in Sec. VII). According to our definition of a battery (as a work-exchanging device), we have that all the possible work associated with the system transformation $\rho \to \sigma$ should be exchanged with it, and none should be locked inside the battery. For this reason, it seems natural to ask the final state of the reservoir to be thermal.

Now in order for Theorem 1 to apply, we need to consider the asymptotic limit, that is, when n, m, and $\ell \gg 1$. In this case, we can convert ω_{in} into ω_{fin} using the set of allowed operations if and only if they have the same average energy and entropy; conservation of system size is already guaranteed

to hold. This gives two equations that we can solve for m and ℓ, resulting in

$$\frac{m}{n} = \frac{S(\sigma) - S(\rho)}{S(\tau_{\beta_1}) - S(\tau_{\beta_2})}, \tag{17}$$

and a somewhat more complicated expression for $\frac{\ell}{n}$. So in order for m to be non-negative, we should have $\beta_1 < \beta_2$ if $S(\rho) > S(\sigma)$ and vice versa (assuming that $\beta_1, \beta_2 > 0$). Physically, this implies that when we dump entropy from the system into the thermal reservoir, we increase its temperature, and vice versa, as we would expect in the case of a finite-size thermal reservoir. We refer to our thermal reservoir as being of finite size because, even if it is composed of $m \to \infty$ copies, the reservoir size is finite *relative to* the system size n, in contrast to the case analyzed in Ref. [5]. Similarly, a positive value $\ell > 0$ means that we achieve an extraction of work from the system, while if ℓ comes out negative, then we can make it positive by taking the initial state of the battery to be $|E_{\max}\rangle^{\otimes \ell}$, and the final one $|E_{\min}\rangle^{\otimes \ell}$, which corresponds to an injection of work into the system. For simplicity, we focus on the case that $m, \ell > 0$ with ω_{in} and ω_{fin} as above, while the other cases are analogous.

We can now evaluate the work extracted and heat provided during the state transformation. We identify these two quantities with, respectively, the energy difference between the final and initial ℓ copies of pure states, and with the energy difference between the initial and final m copies of thermal states. Thus, work is the energy stored inside the pure states, and heat is the energy exchanged with the thermal states. Using the result of Eq. (17), we obtain the following expressions for the work extracted and the heat provided per copy of ρ and σ,

$$W_{\beta_1,\beta_2}(\rho \to \sigma) = \frac{\ell}{n}(E_{\max} - E_{\min})$$
$$= [E(\rho) - E(\sigma)]$$
$$- \frac{E(\tau_{\beta_1}) - E(\tau_{\beta_2})}{S(\tau_{\beta_1}) - S(\tau_{\beta_2})}[S(\rho) - S(\sigma)] \tag{18}$$

$$Q_{\beta_1,\beta_2}(\rho \to \sigma) = \frac{m}{n}[E(\tau_{\beta_1}) - E(\tau_{\beta_2})]$$
$$= \frac{E(\tau_{\beta_1}) - E(\tau_{\beta_2})}{S(\tau_{\beta_1}) - S(\tau_{\beta_2})}[S(\sigma) - S(\rho)]. \tag{19}$$

These quantities depend on the initial and final system state, but also on the initial and final temperature of the reservoir. Our definition of work and heat is consistent with the first law of thermodynamics, since we have $\Delta E(\rho \to \sigma) = Q_{\beta_1,\beta_2}(\rho \to \sigma) - W_{\beta_1,\beta_2}(\rho \to \sigma)$, where $\Delta E = E(\rho) - E(\sigma)$ is the average energy difference between the final and initial state of the system, independently of β_1 and β_2.

These equations for work and heat are similar to the standard ones. In fact, work is given by the free energy difference between ρ and σ, for an external effective temperature β_{eff}^{-1} depending on the initial and final temperatures of the thermal reservoir,

$$\beta_{\text{eff}}(\beta_1, \beta_2) = \frac{S(\tau_{\beta_1}) - S(\tau_{\beta_2})}{E(\tau_{\beta_1}) - E(\tau_{\beta_2})}, \tag{20}$$

so that

$$W_{\beta_1,\beta_2}(\rho \to \sigma) = \beta_{\text{eff}}^{-1}[A_{\beta_{\text{eff}}}(\rho) - A_{\beta_{\text{eff}}}(\sigma)]. \tag{21}$$
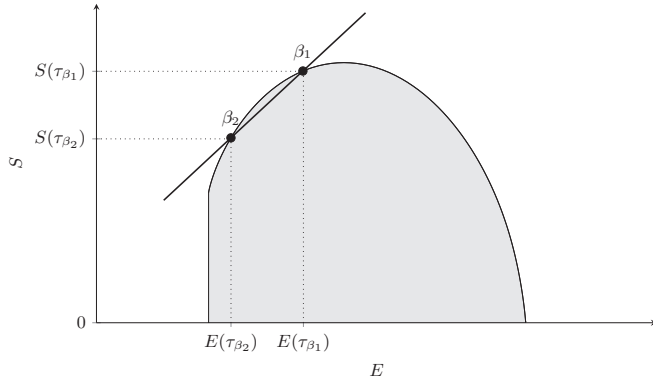
FIG. 4. The visualization of the effective inverse temperature of the reservoir, β_{eff}. The thermal reservoir initially consists of m copies of τ_{β_1}, which are each turned into τ_{β_2} by the state transformation. The value of β_{eff} is given by the slope of the line connecting the two corresponding points in the energy-entropy diagram. When $\beta_2 = \beta_1 + \varepsilon$, for $|\varepsilon| \to 0$, the two points get closer, and the line approaches the tangent to the curve of thermal states. In this case, $\beta_{\text{eff}} = \beta_1 + O(\varepsilon)$, by Remark 3.

In the same way, the equation for heat is equal to the standard one, for the same effective temperature β_{eff}^{-1},

$$Q_{\beta_1,\beta_2}(\rho \to \sigma) = \beta_{\text{eff}}^{-1}[S(\sigma) - S(\rho)]. \qquad (22)$$

This equation can also be seen as a noninfinitesimal generalization of the fundamental thermodynamic relation $dQ = \beta^{-1}dS$. In other words, by the defining Eq. (20), the effective temperature β_{eff}^{-1} can be visualized as a slope in the energy-entropy diagram, as in Fig. 4.

Let us now consider the limiting case of an infinite reservoir, so that the reservoir temperature changes only infinitesimally. In this case, we have $\beta_2 = \beta_1 + \varepsilon$, where $|\varepsilon| \ll 1$. Then, it is straightforward to show that $\beta_{\text{eff}} = \beta_1 + O(\varepsilon)$, and the work and heat we obtain are equal to the standard ones (up to first order in ε), that is,

$$W_{\text{standard}}(\rho \to \sigma) = \beta_1^{-1}[A_{\beta_1}(\rho) - A_{\beta_1}(\sigma)] + O(\varepsilon),$$

$$Q_{\text{standard}}(\rho \to \sigma) = \beta_1^{-1}[S(\sigma) - S(\rho)] + O(\varepsilon).$$

Moreover, we find from Eq. (17) that when we *want* the temperature change to be only $\varepsilon \ll 1$, then the required size of the thermal reservoir per copy of the system S tends to infinity, according to

$$\frac{m}{n} = \frac{S(\sigma) - S(\rho)}{\beta_1 \langle \Delta^2 H \rangle_{\tau_{\beta_1}}} \frac{1}{\varepsilon} + O(1), \qquad (23)$$

where the expectation value in the denominator is the variance of energy in the state τ_{β_1}, or equivalently β_1^{-2} times the heat capacity (at β_1) of a single system.

We close this section with an application of our formalism, relevant for classical and quantum computation, and information processing tasks. We consider the erasure of information, or Landauer's erasure, in the scenario in which the surrounding environment has a finite size [48]. In this scenario, the main system is acting as a memory, and its energy does not change during the transformation, meaning that $E(\rho) = E(\sigma)$. Then we find that when the thermal reservoir has a finite size, the

work required to erase the state ρ and to map it into σ is

$$W_{\beta_1,\beta_2}^{\text{erasure}}(\rho \to \sigma) = \beta_{\text{eff}}^{-1}(\beta_1,\beta_2)[S(\rho) - S(\sigma)], \qquad (24)$$

which converges to the well-known value $\beta_1^{-1}[S(\rho) - S(\sigma)]$ when the size of the reservoir tends to infinity. If the initial state ρ is maximally mixed and the final state σ is pure, we find that the work of erasure is $\beta_{\text{eff}}^{-1}(\beta_1,\beta_2)\log d$.

IX. HEAT ENGINES

We now show how the results of the previous sections can be used in order to analyze the efficiency of heat engines (and refrigerators) utilizing finite-size reservoirs. We do not assume any specific kind of engine consisting of a particular device or using particular mechanisms; instead, we utilize our formalism in order to derive the maximal efficiency of *any* protocol operating on two finite-size reservoirs. As before, our analysis is valid in the limit of many copies and in the case where all systems consist of (approximately) noninteracting microsystems with common Hamiltonian H. For example, we can now see how a heat engine secretly exploits our observation that thermodynamics is a general probabilistic theory (Sec. VI): If we have a system consisting of two subsystems given by thermal states at different temperatures, we start with a macrostate which is a convex combinations of two extreme points of the energy-entropy diagram. Decomposing it in a different way into extreme points, we can therefore extract a certain number of copies of the maximally excited state—which plays the role of extracted work—together with a thermal state at an intermediate temperature.

For heat engines and refrigerators utilizing reservoirs of finite size, we derive explicit expressions for the maximal efficiency depending on two effective temperatures (describing the hot and cold reservoirs, respectively). As per Eqs. (27) and (28), this optimal efficiency with finite-size reservoirs is always lower than the Carnot efficiency.

Our model consists of the same tripartite system as in the previous section, but further specialized to the case where both the initial state ρ and the final state σ are themselves thermal. Hence, the initial state is given by

$$\omega_{\text{in}}^{\text{engine}} = \tau_{\beta_{\text{cold}}}^{\otimes n} \otimes \tau_{\beta_{\text{hot}}}^{\otimes m} \otimes |E_{\min}\rangle\langle E_{\min}|^{\otimes \ell}, \qquad (25)$$

where $\beta_{\text{cold}} > \beta_{\text{hot}}$. The final state, instead, is

$$\omega_{\text{fin}}^{\text{engine}} = \tau_{\beta_{\text{less-cold}}}^{\otimes n} \otimes \tau_{\beta_{\text{less-hot}}}^{\otimes m} \otimes |E_{\max}\rangle\langle E_{\max}|^{\otimes \ell}, \qquad (26)$$

where $\beta_{\text{cold}} > \beta_{\text{less-cold}} > \beta_{\text{less-hot}} > \beta_{\text{hot}}$. The engine uses the hot and cold reservoirs to extract work, but in the meanwhile it degrades these reservoirs, assimilating their temperatures (because these are of finite size).

Since everything that we do is reversible, one can consider both the transformation $\omega_{\text{in}}^{\text{engine}} \to \omega_{\text{fin}}^{\text{engine}}$ (heat engine) as well as the reverse $\omega_{\text{fin}}^{\text{engine}} \to \omega_{\text{in}}^{\text{engine}}$ (refrigerator). We are not concerned with the question of how to realize these transformations; Theorem 1 gives us necessary and sufficient conditions for when they are realizable but does not make any statement about how to implement them, using a "working body" or otherwise. We only know that there exists some unitary acting on the global system together with a small number of ancilla systems which realizes these devices to any
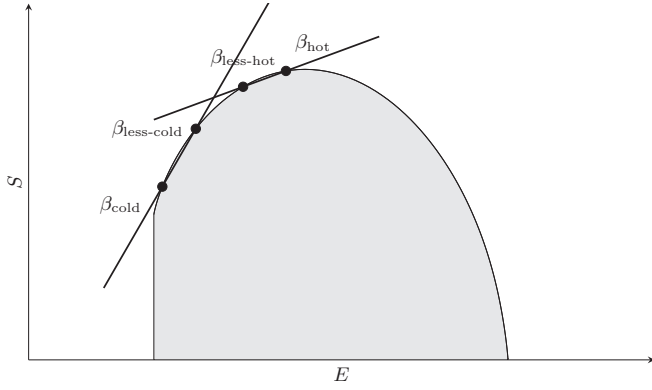
FIG. 5. The two effective temperatures that determine engine efficiency interpreted as slopes in the energy-entropy diagram. For $\beta_{\text{less-cold}}$ very close to β_{cold} and $\beta_{\text{less-hot}}$ very close to β_{hot}, i.e., when the reservoirs are very large compared to the battery size, then the lines approximate tangents and the resulting efficiency approaches the Carnot efficiency.

desired degree of accuracy as $m,n,\ell \to \infty$. And moreover, there is no other device or mechanism that could do better.

In order to evaluate the efficiency of these two devices, we need to evaluate the heat exchanged with the hot reservoir, the work produced or utilized, and the heat exchanged with the cold reservoir. Because of reversibility, these quantities are the same for both devices (at least in absolute value). Using Eqs. (18) and (19), we find the heat exchanged with the hot reservoir Q_{hot}, and the work exchanged W,

$$Q_{\text{hot}} = \beta_{\text{eff}}^{-1}(\beta_{\text{hot}},\beta_{\text{less-hot}})[S(\tau_{\beta_{\text{less-cold}}}) - S(\tau_{\beta_{\text{cold}}})],$$

$$W = [E(\tau_{\beta_{\text{cold}}}) - E(\tau_{\beta_{\text{less-cold}}})]$$
$$- \beta_{\text{eff}}^{-1}(\beta_{\text{hot}},\beta_{\text{less-hot}})[S(\tau_{\beta_{\text{cold}}}) - S(\tau_{\beta_{\text{less-cold}}})],$$

both per copy of the first reservoir system. On the other hand, since the system S is now the cold reservoir, the heat Q_{cold} exchanged with it per copy is equal to

$$Q_{\text{cold}} = E(\tau_{\beta_{\text{less-cold}}}) - E(\tau_{\beta_{\text{cold}}}).$$

We can now evaluate the efficiency of the heat engine, defined as $\eta_{\text{engine}} = \frac{W}{Q_{\text{hot}}}$, and the efficiency of the refrigerator, $\eta_{\text{refrigerator}} = \frac{Q_{\text{cold}}}{W}$. We find that the efficiencies are equal to

$$\eta_{\text{engine}} = 1 - \frac{\beta_{\text{eff}}(\beta_{\text{hot}},\beta_{\text{less-hot}})}{\beta_{\text{eff}}(\beta_{\text{cold}},\beta_{\text{less-cold}})}, \tag{27}$$

$$\eta_{\text{refrigerator}} = \left[\frac{\beta_{\text{eff}}(\beta_{\text{cold}},\beta_{\text{less-cold}})}{\beta_{\text{eff}}(\beta_{\text{hot}},\beta_{\text{less-hot}})} - 1\right]^{-1}, \tag{28}$$

where β_{eff} was defined in Eq. (20). In terms of the interpretation of effective inverse temperatures as slopes in the energy-entropy diagram, we can understand these efficiencies as in Fig. 5: For example, for the heat engine, the quotient of the slopes is always less than the quotient of the two tangent slopes at β_{hot} and β_{cold}, respectively. This implies that $\eta_{\text{engine}} < 1 - \frac{\beta_{\text{hot}}}{\beta_{\text{cold}}}$, and similarly $\eta_{\text{refrigerator}} < (\frac{\beta_{\text{less-cold}}}{\beta_{\text{less-hot}}} - 1)^{-1}$, so that both efficiencies are strictly lower than the Carnot efficiencies. This is due to the fact that the temperature of the two finite-size reservoirs changes during the process. In

the limit where the temperature of the two reservoirs changes only by an infinitesimal amount, both efficiencies approach Carnot's values.

X. OPTIMAL RATES OF CONVERSION AND HOW TO COMPUTE THEM

So far, we have only considered asymptotic equivalence of states, since only energy-preserving unitaries have been allowed. What do we get if we allow in addition that subsystems can be discarded? Building on our previous results, we will now give one possible answer to this question. In all cases, we assume that a state ρ lives on a certain number $n(\rho)$ of microsystems as before, and when we are dealing with two (or more) states ρ and σ, we do not assume $n(\rho) = n(\sigma)$.

Lemma 11 (The asymptotic ordering $\succeq$). For given states ρ and σ, the following are equivalent:

(a) There is a state ϕ such that $\rho \asymp \sigma \otimes \phi$ in the sense of Theorem 1.

(b) There is a state $\sigma' \asymp \rho$ such that $\text{Tr}_{\text{sub}}[\sigma'] = \sigma$ for some subsystem that is not entangled with the rest.

(c) $S(\rho) \geqslant S(\sigma)$ and $A_\beta(\rho) \geqslant A_\beta(\sigma)$ for all $\beta \in (-\infty,+\infty)$.

(d) We have $y(\rho) - y(\sigma) \in \text{Therm}(H)$.

Let us write $\rho \succeq \sigma$ for the ordering relation on states corresponding to these equivalent conditions. In the following, we will investigate this ordering relation a bit further.

Proof. The implication from (a) to (b) is simply by taking $\sigma' := \sigma \otimes \phi$.

Next, we show that (b) implies (c); this follows from the fact that the additive monotones S and A_β respect asymptotic equivalence $\asymp$, and are nonincreasing under tracing out such subsystems. The former is a consequence of Theorem 1, while the latter is a consequence of the no-entanglement assumption in the case of S (the conditional entropy is non-negative) and of the data processing inequality in the case of A_β,

$$A_\beta(\rho) = D\left(\rho \| \tau_\beta^{\otimes n(\rho)}\right) \geqslant D\left(\text{Tr}_{\text{sub}}[\rho] \big\| \text{Tr}_{\text{sub}}\left[\tau_\beta^{\otimes n(\rho)}\right]\right)$$
$$= A_\beta(\text{Tr}_{\text{sub}}[\rho]), \tag{29}$$

where the last equation holds because $\text{Tr}_{\text{sub}}[\tau_\beta]$ is the thermal state on the subsystem.

From (c) to (d), the point $y(\rho) - y(\sigma)$ satisfies all the inequalities of Proposition 10 by assumption, and therefore lies in $\text{Therm}(H)$.

From (d) to (a), the assumption together with Remark 8 guarantees the existence of a state ϕ with $y(\phi) = y(\rho) - y(\sigma)$. Therefore, $y(\rho) = y(\sigma \otimes \phi)$, and then $\rho \asymp \sigma \otimes \phi$ is a consequence of Theorem 1. ∎

The "no entanglement" requirement in condition (b) seems a bit artificial, and it would be interesting to obtain results analogous to the upcoming ones for the ordering relation defined in the analogous way, but where one would be allowed to trace out an *arbitrary* subsystem. We suspect that such a development would require generalizations of Theorems 15 and 1, where instead of characterizing the asymptotic equivalence of states relative to energy-preserving unitaries, one would instead classify the asymptotic *ordering* of states relative to energy-preserving unitaries and discarding

subsystems. We currently do not have such a result and thus use the $\succeq$ relation from Lemma 11.

Definition 12 ([34, Eq. (8.3)]). The *maximal rate* of converting a state ρ into a state σ is given by

$$R_{\max}(\rho \to \sigma) := \sup \left\{ \frac{m}{n} \,\bigg|\, \rho^{\otimes n} \succeq \sigma^{\otimes m} \right\}. \qquad (30)$$

So roughly speaking, we now ask: If we try to convert many copies of ρ into many copies of σ, then how many copies of ρ do we need per copy of σ, where we may discard some additional "junk" states in the process?

Since we already have allowed for sublinear ancillas in the definition of asymptotic equivalence, this notion of maximal rate actually corresponds to the notion of *regularized* maximal rate of Ref. 34, Sec. 8. Building on the methods that we have developed so far, it is not hard to write down a concrete formula for computing maximal rates:

Theorem 13. The maximal rate from ρ to σ can be computed in two ways:

(a) $R_{\max}(\rho \to \sigma)$ is equal to the value of r at which the line in $\mathbb{R}^3$ defined by $r \mapsto y(\rho) - ry(\sigma)$ pinches the boundary of the cone $\text{Therm}(H)$, so that

$$R_{\max}(\rho \to \sigma) = \max\{r \in \mathbb{R}_{\geqslant 0} \mid$$
$$y(\rho) - ry(\sigma) \in \text{Therm}(H)\}. \qquad (31)$$

(b) $R_{\max}(\rho \to \sigma)$ is also equal to the minimal ratio of the value of an additive monotone on ρ versus its value on σ,

$$R_{\max}(\rho \to \sigma) = \min \left\{ \frac{S(\rho)}{S(\sigma)}, \inf_{\beta \in (-\infty, +\infty)} \frac{A_\beta(\rho)}{A_\beta(\sigma)} \right\}, \qquad (32)$$

where the minimization is only over those fractions for which the denominator is nonzero.

Proof. By additivity of y, a rational number $r = \frac{p}{q} \in \mathbb{Q}_{\geqslant 0}$ is an achievable rate if and only if $qy(\rho) - py(\sigma) \in \text{Therm}(H)$, or equivalently $y(\rho) - ry(\sigma) \in \text{Therm}(H)$. This implies (31).

One gets (32) from (31) via Proposition 10, since $y(\rho) - ry(\sigma) \in \text{Therm}(H)$ is equivalent to $S(\rho) \geqslant r S(\sigma)$ together with $A_\beta(\rho) \geqslant r A_\beta(\sigma)$ for all $\beta \in (-\infty, +\infty)$. Hence the condition on r is that it must be less than or equal to $\frac{S(\rho)}{S(\sigma)}$, and also less than or equal to $\frac{A_\beta(\rho)}{A_\beta(\sigma)}$ for every β, for those fractions for which the denominator is nonzero. The largest r that satisfies this is precisely (32). ∎

To understand Eq. (31) intuitively, it may help to normalize the macrostates and phrase the condition in terms of convex combinations in the energy-entropy diagram instead, as per Sec. VI.

What makes the infimum over β in Eq. (32) nontrivial to evaluate is the presence of the partition function term $\log Z_\beta$ in both the numerator and the denominator, due to Eq. (10). Nevertheless, this is a very explicit formula with which one should be able to compute rates in practice. It is an instance of Ref. [34, Theorem 8.24], and the proof is correspondingly similar.

XI. CONCLUSION

Our resource theory for thermodynamics does not make use of an infinite thermal reservoir. Therefore, it is suitable for analyzing state transformations both when the system is decoupled from the environment, e.g., via Eq. (32), and when the system is interacting with a finite reservoir, Eqs. (21) and (22). Moreover, the theory provides a rigorous mathematical explanation (through the Theorem 1) to the fact that, when dealing with macroscopic thermodynamics, we can describe the state of a system with few observables (for instance, energy and entropy). Our approach generalizes the one presented in Ref. [5], where asymptotic state transformations are considered when an infinite reservoir is present.

The results we obtain are valid in a specific regime delineated by several idealized assumptions, such as the assumption that all energy-preserving unitaries are available, the presence of many noninteracting and identical copies of the system, and the constraint of a fixed Hamiltonian for each system. One can think of dropping some of these assumptions, and for example investigate the theory when arbitrary states and interactions are allowed (often called the single-shot regime), or when one has a much more realistic class of operations not requiring such fine-grained control of system and bath [14].

The asymptotic equivalence result presented in Eq. (2) is obtained with the help of a sublinear ancillary system. *A priori*, one might think that this additional system could be used as an unbounded source of work, since we do not require the state of this ancilla to be restored at the end of the process. However, to avoid the possibility of freely modifying the energy of the system by exploiting the ancilla, we constrain the energy spectrum of the latter to be sublinear in the number of copies of the main system.

Recently, resource theories with multiple conserved quantities (even noncommuting ones), have been investigated within the framework of quantum thermodynamics [15,49,50]. However, in these models, emphasis is put on different notions of work, each of them related to a different conserved quantity. Our theory, on the other hand, considers only energy [51]. We expect that our approach can be extended more or less straightforwardly so as to cover multiple commuting conserved quantities as well; generalizing to a treatment of multiple noncommuting conserved quantities may present new challenges.

APPENDIX: ASYMPTOTIC EQUIVALENCE OF STATES UNDER ENERGY-PRESERVING UNITARIES

1. Overview

We show in this section that two states of a quantum system are asymptotically equivalent under energy-preserving unitaries if and only if they have the same entropy and average energy (Theorem 1). Here, we consider two states

asymptotically equivalent when one can turn many identical copies of one state into the same number of copies of the other state, arbitrarily accurately, by applying a global unitary operation which preserves energy. The precise statement is in Theorem 17.

The main difficulty in our proof is in showing that when two states have same energy and entropy, they can be asymptotically mapped one into the other. To prove this implication, we devise a protocol which converts (many copies of) one state into the other, provided that they have the same energy and entropy. We now summarize the protocol, in order to provide a simple and physical idea of its mechanism to the reader. We do this in two cases, one concerning the simpler case of trivial Hamiltonian, and then in general.

When the system has trivial Hamiltonian, we can act on it by means of any unitary operation, and the only assumption we have about the states ρ and σ is that they have the same entropy. Since we work in the asymptotic regime, where we take the tensor product of many copies of these states, we can use the tools developed in Shannon theory [52–54]. In particular, due to the central limit theorem, we can replace the many-copy states $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ with their typical states, Eqs. (A9) and (A10). The use of the typical states highly simplifies the protocol, since in this way we can divide the Hilbert space into a small number of subspaces with common properties. State conversion is achieved in the protocol by mapping the probability distribution of the initial typical state into the one of the final typical state. This is done by introducing an ancillary system with trivial Hamiltonian, whose size is $O(\sqrt{n \log n})$, in the maximally mixed state. This ancilla provides a source of randomness, and we modify the probability distribution of the initial state by applying a global unitary operation on system and ancilla, and tracing out the ancilla. However, a unitary operation can be used only if the transformation is reversible. To assure that this is the case, another ancillary system is introduced, acting as a register, which allows for a dilation to a unitary. Again, the dimension of this second ancillary system is $O(\sqrt{n \log n})$, and the Hamiltonian is trivial. The details of the protocol are in the proof of Theorem 15.

When the system has nontrivial Hamiltonian, we have to reduce the set of allowed unitary operations to the sole energy-preserving ones (the ones that commute with the Hamiltonian). With these operations, we have to devise a protocol which approximately converts many copies of ρ into σ when the two states have same entropy and energy. The protocol which performs this asymptotic transformation is analogous to the one for trivial Hamiltonian. The difference is that in this case we have to add an additional ancillary system with nontrivial Hamiltonian, with which we can exchange both energy and coherence. This ancilla allows us to approximately implement any unitary on the system by applying an energy-preserving unitary on both system and ancilla. Due to the constraints on the energy and entropy of the initial and final state, and to the central limit theorem, we achieve that the size of this additional ancilla is $O(\sqrt{n \log n})$. Moreover, the spectrum of the ancillary Hamiltonian is bounded by $O(n^{\frac{2}{3}})$, so that we modify the amount of energy only by a sublinear amount. The details of the protocol are in the proofs of Lemma 16 and Theorem 17.

It is worth noting that none of the three ancillary systems depends in any way on the state ρ or σ, meaning that the transformation can be performed with the same ancillae for any initial and final state.

2. Asymptotic equivalence of quantum states

Before getting to thermodynamics, it helps to consider an easier case first: two states of a quantum system are many-copies equivalent under unitaries if and only if they have the same entropy. This is what we show first in Theorem 15. Since we can always diagonalize, this is a purely classical problem, and we thus start out with some lemmas for classical information theory. The first one is a simple lemma on randomness extraction which we will apply afterward to approximate the typical set of one distribution by a coarse-graining of the typical set of another. We use the min-entropy, the Hartley entropy, and the Rényi entropy at parameter $-\infty$,

$$H_\infty(p) = -\log \max_x p_x, \tag{A1}$$

$$H_0(p) = \log |\{x \mid p_x > 0\}|, \tag{A2}$$

$$H_{-\infty}(p) = -\log \min_x p_x. \tag{A3}$$

Lemma 14. Let (X,p) and (Y,q) be finite probability spaces. Then there exists a map $f : X \to Y$ such that

$$\|f_*(p) - q\|_1 \leqslant 2^{H_0(q) - H_\infty(p)} \tag{A4}$$

and

$$|f^{-1}(y)| \leqslant 2^{H_{-\infty}(p)} \left(2^{-H_\infty(q)} + 2^{-H_\infty(p)}\right) \tag{A5}$$

for all $y \in Y$.

Here, $f_*(p)$ is the distribution on Y that one obtains by coarse graining p via application of f, that is, by gathering in different sets the elements of X, obtaining a new distribution over a smaller space Y. The $\| \cdot \|_1$ is the total variation distance, i.e., the classical version of the trace distance.

Proof. We choose an arbitrary enumeration of the elements of X as $x_1, \ldots, x_n$, and construct f in piecemeal by defining $f(x_1), \ldots, f(x_n)$ one at a time. At the ith step, we define $f(x_i)$ to be equal to an arbitrary $y \in Y$ whose probability has not yet been completely covered by the p_x that lie in the preimage $f^{-1}(y)$ of the partially defined f, in the sense that

$$q_y > \sum_{x \in f^{-1}(y)} p_x,$$

where the sum is only over those $x \in X$ for which $f(x)$ has already been defined and is equal to y. Finding such a y is always possible since the normalization of p equals the normalization of q. The crucial property of the f thus constructed is that the total probability in a fiber $f^{-1}(y)$ is never significantly larger than q_y,

$$\sum_{x \in f^{-1}(y)} p_x \leqslant q_y + \max_x p_x. \tag{A6}$$

This implies $|f^{-1}(y)| \min_x p_x \leqslant q_y + \max_x p_x$, resulting in (A5). To bound the total variation distance, we also

use (A6),

$$\|f_*(p) - q\|_1 = \sum_y \max\left\{0, \sum_{x \in f^{-1}(y)} p_x - q_y\right\}$$
$$\leqslant \sum_y \max_x p_x = |Y| \max_x p_x.$$

Since we can assume q to have full support without loss of generality, this is the desired inequality (A4). ∎

Turning to quantum information, we use the term "size" of a system to talk about the logarithm of its Hilbert space dimension, i.e., the number of qubits needed to realize it, and write $S(\rho) = -\text{Tr}[\rho \log \rho]$ for the von Neumann entropy of a state ρ.

Theorem 15 (Asymptotic classification of states). For states ρ and σ on any quantum system of dimension d, the following are equivalent:

(a) The states have equal entropy, $S(\rho) = S(\sigma)$.

(b) There exists an ancilla system of size $O(\sqrt{n \log n})$ with state η as well as a unitary U such that

$$\|\text{Tr}_{\text{anc}}[U(\rho^{\otimes n} \otimes \eta)U^\dagger] - \sigma^{\otimes n}\|_1 \xrightarrow{n \to \infty} 0. \quad (A7)$$

(c) There exists an ancilla system of size $o(n)$, with states η and ν as well as unitaries U and V

such that

$$\|\text{Tr}_{\text{anc}}[U(\rho^{\otimes n} \otimes \eta)U^\dagger - V(\sigma^{\otimes n} \otimes \nu)V^\dagger]\|_1 \xrightarrow{n \to \infty} 0. \quad (A8)$$

That is, up to things that happen on an ancilla of sublinear size, two states on a system are many-copies equivalent if and only if they have the same entropy. Condition (b) is a set of requirements on such a many-copies equivalence that we believe to be roughly minimal; in particular, the number of qubits needed to implement the ancilla system grows only barely faster than $O(\sqrt{n})$, and in fact the particular growth rate of $O(\sqrt{n \log n})$ is an arbitrary choice and can be replaced by any function that grows faster than $\sqrt{n}$. Condition (c), in contrast, is a more permissive notion of many-copy equivalence that is still strong enough to imply (a) but it does not provide a different physical intuition than (b).

The superoperators $\text{Tr}_{\text{anc}}[U(\otimes \eta)U^\dagger]$ form channels which are close to unitary in the sense of being implementable with only a sublinear ancilla; these are precisely the channels of subexponential Kraus rank. It may be interesting to study such channels in their own right, and there may be relations to Ref. [55].

Proof. The implication from (b) to (c) is trivial.

Assuming (c), the claim $S(\rho) = S(\sigma)$ can be proven as follows. Let D be the dimension of the ancilla system. Then using the fact that adding or discarding the ancilla cannot change the entropy by more than $\log(D)$ [56], we obtain, writing ε for the left-hand side of (A8),

$$|S(\rho) - S(\sigma)| = \frac{1}{n}|S(\rho^{\otimes n}) - S(\sigma^{\otimes n})| \leqslant \frac{1}{n}|S(\rho^{\otimes n} \otimes \eta) - S(\sigma^{\otimes n} \otimes \nu)| + 2\frac{\log(D)}{n}$$

$$= \frac{1}{n}|S[U^\dagger(\rho^{\otimes n} \otimes \eta)U] - S[V^\dagger(\sigma^{\otimes n} \otimes \nu)V]| + 2\frac{\log(D)}{n}$$

$$\leqslant \frac{1}{n}|S\{\text{Tr}_{\text{anc}}[U^\dagger(\rho^{\otimes n} \otimes \eta)U]\} - S\{\text{Tr}_{\text{anc}}[V^\dagger(\sigma^{\otimes n} \otimes \nu)V]\} + 4\frac{\log(D)}{n}$$

$$\overset{(A8)}{\leqslant} \frac{1}{n}[\log(d^n)\varepsilon + O(1)] + 4\frac{\log(D)}{n} = O(\varepsilon) + O(n^{-1}) + 4\frac{\log(D)}{n},$$

where the last estimate is by Fannes's inequality. Since $\varepsilon \to 0$ as $n \to \infty$ while D grows only subexponentially, it follows that $|S(\rho) - S(\sigma)|$ is smaller than any positive number and therefore $S(\rho) = S(\sigma)$.

To show that (a) implies (b), we can assume by unitary invariance that ρ and σ are diagonal in the same basis, where they are given by $\rho = \text{diag}(p_1, \ldots, p_d)$ and $\sigma = \text{diag}(q_1, \ldots, q_d)$. In other words, we are in a classical situation involving finite probability spaces with distributions $p = (p_1, \ldots, p_d)$ and $q = (q_1, \ldots, q_d)$, and we therefore use classical notation and terminology for the remainder of the proof, and write n_i for the number of times that outcome i occurs upon sampling from $p^{\otimes n}$ or $q^{\otimes n}$. The central limit theorem guarantees that for $p^{\otimes n}$, the set of outcomes that are strongly typical in the sense that

$$n_i \in [(n - \sqrt{n \log n})p_i, (n + \sqrt{n \log n})p_i] \quad (A9)$$

for every $i = 1, \ldots, d$ has a total probability that approaches 1 as $n \to \infty$. Let T_p denote this typical set of outcomes and p_{typ} be the resulting normalized distribution on typical outcomes

that one obtains by conditioning on typicality. Similarly, let T_q be the strongly typical set for $q^{\otimes n}$, corresponding to outcome frequencies n_i restricted by

$$n_i \in [(n - \sqrt{n \log n})q_i, (n + \sqrt{n \log n})q_i], \quad (A10)$$

and q_{typ} be the associated typical distribution. By bounding the lowest and the highest probability of any outcome in this strongly typical set, it is straightforward to show the following inequalities:

$$H_0(p_{\text{typ}}) \geqslant H_\infty(p_{\text{typ}}) \geqslant nS(p)\left(1 - \sqrt{\frac{\log n}{n}}\right),$$

$$H_0(p_{\text{typ}}) \leqslant H_{-\infty}(p_{\text{typ}}) \leqslant nS(p)\left(1 + \sqrt{\frac{\log n}{n}}\right), \quad (A11)$$

where we still write $S(p) = H_1(p)$ for the Shannon entropy, and the second inequality holds for sufficiently large n where the modification due to the conditioning is negligible. The

analogous bounds hold for q_{typ}. Note that the individual probabilities of the typical outcomes may still vary by a factor of up to $2^{2\sqrt{n \log n} \, S(p)}$, so that the typical distributions p_{typ} and q_{typ} may still be far from uniform. The strong typicality inequalities (A9) and (A10) themselves will not be used again; all that we need are the Rényi entropy bounds (A11), and that the probability of typicality approaches 1 as $n \to \infty$.

Now let r_1 be the uniform distribution on $3\sqrt{n \log n} \, S(p)$ many ancilla bits, rounded to the closest integer; in the following, we ignore the irrelevant rounding error. By (A11), this results in the bounds

$$nS(p) + 2\sqrt{n \log n} \, S(p) \leqslant H_\infty(p_{typ} \otimes r_1)$$
$$\leqslant H_{-\infty}(p_{typ} \otimes r_1)$$
$$\leqslant nS(p) + 4\sqrt{n \log n} \, S(p).$$

Hence by Lemma 14, we can find a map $f : T_p \times \{0,1\}^{3\sqrt{n \log n} \, S(p)} \to T_q$ such that

$$\|f_*(p_{typ} \otimes r_1) - q_{typ}\|_1 \leqslant 2^{H_0(q_{typ}) - H_\infty(p_{typ} \otimes r_1)}$$
$$\leqslant 2^{-\sqrt{n \log n} \, S(p)}$$

by (A4), which decays superpolynomially in n. Thanks to (A5), f can be implemented using a register ancilla of dimension at most

$$2^{H_{-\infty}(p_{typ} \otimes r_1)} \big(2^{-H_\infty(q_{typ})} + 2^{-H_\infty(p_{typ} \otimes r_1)} \big)$$
$$\leqslant 2^{5\sqrt{n \log n} \, S(p)} + 2^{2\sqrt{n \log n} \, S(p)} = 2^{O(\sqrt{n \log n})},$$

which is initially taken to carry an arbitrary deterministic distribution r_2 and gets utilized to dilate f to a bijection.

We now put $r := r_1 \otimes r_2$, so that our total ancilla still has size $O(\sqrt{n \log n})$. We take U to be given by the action of f on the system and first ancilla, dilated to a bijection by the second ancilla. Since f_* is contractive, we have

$$\|f_*(p \otimes r_1) - q\|_1 \leqslant \|p_{typ} - p\|_1 + \|q_{typ} - q\|_1$$
$$+ \|f_*(p_{typ} \otimes r_1) - q_{typ}\|_1 \xrightarrow{n \to \infty} 0,$$

since each individual term tends to 0. This establishes (A7) in classical notation. ∎

Getting to thermodynamics, we now also want to take energy preservation into account. To this end, we develop a method to turn every unitary into an energy-preserving unitary, while achieving approximately the same conversion of states. This relies on a protocol modeled after Ref. [5], Appendix E. For finite sets of numbers $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}$, we consider the sumset $\mathcal{A} + \mathcal{B} := \{a + b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$, as studied in additive combinatorics [57], and similarly also the difference set $\mathcal{A} - \mathcal{B} = \{a - b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$. Furthermore,

we write $\|\mathcal{A}\| := \max_{a \in \mathcal{A}} |a|$. And from now on, we also use $\rho \approx_\varepsilon \sigma$ as a shorthand for $\|\rho - \sigma\|_1 \leqslant \varepsilon$.

Lemma 16 (Achieving energy preservation). Let $0 < \delta < 1$ and suppose that $\mathcal{L}, \mathcal{M} \subseteq \mathbb{R}$ are finite sets of numbers such that

$$|\mathcal{M} + \mathcal{L}| \leqslant (1 + \delta)|\mathcal{M}|, \qquad |\mathcal{M} - \mathcal{L}| \leqslant (1 + \delta)|\mathcal{M}|,$$
(A12)

and $\|\mathcal{L}\| \leqslant \|\mathcal{M}\|$. Given a quantum system with Hamiltonian H, suppose that ρ and σ are states supported on energy levels in $\mathcal{L}$, and that there is a unitary U such that $U \rho U^\dagger \approx_\delta \sigma$. Then there is an ancilla system of size $O(\log |\mathcal{M}|)$ with $\|H_{anc}\| \leqslant 4\|\mathcal{M}\|$ and state η as well as an energy-preserving unitary $\tilde{U}$ on the joint system, that is

$$[\tilde{U}, H + H_{anc}] = 0,$$

such that

$$\mathrm{Tr}_{anc}[\tilde{U}(\rho \otimes \eta)\tilde{U}^\dagger] \approx_{4\delta} \sigma.$$

Interestingly, what makes this difficult to prove are the quantum coherences that ρ and σ may have between the energy levels: In the classical case in which neither ρ nor σ has any coherence across energy levels, a unitary can easily be made energy preserving by adding an ancilla in an initial state, which can absorb any energy difference that may arise.

Proof. We do this by distinguishing two cases: First, the case that σ has no coherences across energy levels; second, the case that ρ has no such coherences. In each case, we will use an ancilla of size $O(\log |\mathcal{M}|)$ with $\|H_{anc}\| \leqslant 2\|\mathcal{M}\|$ and obtain a trace distance bound of 2δ. This is sufficient, since in the general case we can choose an arbitrary state τ without energy coherences which has the same spectrum (with multiplicities) as that of ρ or σ, and compose the protocols constructed in the two cases, first from ρ to τ and then from τ to σ. This results in the claimed bounds.

Let the spectral decomposition of the system's Hamiltonian be $H = \sum_{\lambda \in sp(H)} \lambda P_\lambda$, with P_λ the projection onto the corresponding energy eigenspace.

Case 1. σ has no coherences across energy levels, i.e., $P_\lambda \sigma P_\mu = 0$ if $\lambda \neq \mu$.

In this case, let the ancilla space be $\mathcal{H}_{anc} := \mathbb{C}^{|\mathcal{M} - \mathcal{L}|}$ with Hamiltonian given by $H_{anc} = \sum_{h \in \mathcal{M} - \mathcal{L}} h|h\rangle\langle h|$. By (A12) and $\delta < 1$, the ancilla size is indeed $\log |\mathcal{M} - \mathcal{L}| = O(\log |\mathcal{M}|)$ and moreover $\|H_{anc}\| \leqslant 2|\mathcal{M}|$. We take the initial ancilla state to be pure $\eta := |\psi\rangle\langle\psi|$, and given by the Hadamard state

$$|\psi\rangle := |\mathcal{M} - \mathcal{L}|^{-1/2} \sum_{h \in \mathcal{M} - \mathcal{L}} |h\rangle.$$

Furthermore, consider the energy-preserving partial isometry

$$\tilde{V} := \sum_{h \in \mathcal{M}} \sum_{\lambda, \mu \in \mathcal{L}} P_\lambda U P_\mu \otimes |h - \lambda\rangle\langle h - \mu|,$$

Then $\tilde{V}(\rho \otimes \eta)\tilde{V}^\dagger$ evaluates to

$$|\mathcal{M} - \mathcal{L}|^{-1} \sum_{h_1, h_2 \in \mathcal{M}} \sum_{\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathcal{L}} \big(P_{\lambda_1} U P_{\mu_1} \otimes |h_1 - \lambda_1\rangle\langle h_1 - \mu_1| \big)$$

$$\times \left(\rho \otimes \sum_{\ell_1, \ell_2 \in \mathcal{M} - \mathcal{L}} |\ell_1\rangle\langle\ell_2| \right) \left(P_{\mu_2} U^\dagger P_{\lambda_2} \otimes |h_2 - \mu_2\rangle\langle h_2 - \lambda_2| \right)$$

$$= |\mathcal{M} - \mathcal{L}|^{-1} \sum_{\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathcal{L}} P_{\lambda_1} U P_{\mu_1} \rho P_{\mu_2} U^\dagger P_{\lambda_2} \otimes \sum_{h_1, h_2 \in \mathcal{M}} |h_1 - \lambda_1\rangle\langle h_2 - \lambda_2|$$

$$\approx_\delta |\mathcal{M} - \mathcal{L}|^{-1} \sum_{\lambda_1, \lambda_2 \in \mathcal{L}} P_{\lambda_1} \sigma P_{\lambda_2} \otimes \sum_{h_1, h_2 \in \mathcal{M}} |h_1 - \lambda_1\rangle\langle h_2 - \lambda_2|$$

$$= |\mathcal{M} - \mathcal{L}|^{-1} \sum_{\lambda \in \mathcal{L}} P_\lambda \sigma P_\lambda \otimes \sum_{h_1, h_2 \in \mathcal{M}} |h_1 - \lambda\rangle\langle h_2 - \lambda|,$$

where the last step uses the assumption of absence of coherence in σ. The resulting reduced state is therefore

$$\mathrm{Tr}_{\mathrm{anc}}[\tilde{V}(\rho \otimes \eta)\tilde{V}^\dagger] \approx_\delta \sum_{\lambda \in \mathcal{L}} \frac{|\mathcal{M}|}{|\mathcal{M} - \mathcal{L}|} P_\lambda \sigma P_\lambda$$

$$= \frac{|\mathcal{M}|}{|\mathcal{M} - \mathcal{L}|} \sigma \geqslant (1 - \delta)\sigma.$$

So if we take $\tilde{U}$ to be any energy-preserving unitary dilation of V, so that $\tilde{V}$ decomposes into a direct sum of $\tilde{V}$ plus some other arbitrary energy-preserving partial isometry, then the total weight of $\rho \otimes \eta$ on the orthogonal complement of the support of $\tilde{V}$ is at most δ. This shows that

$$\|\mathrm{Tr}_{\mathrm{anc}}[\tilde{U}(\rho \otimes \eta)\tilde{U}] - \sigma\|_1 \leqslant 2\delta,$$

as desired.

Case 2. ρ has no coherences across energy levels, i.e., $P_\lambda \rho P_\mu = 0$ if $\lambda \neq \mu$. It turns out that we can proceed very similarly.

In this case, let the ancilla space be $\mathcal{H}_{\mathrm{anc}} := \mathbb{C}^{|\mathcal{M} + \mathcal{L}|}$ with Hamiltonian given by $H_{\mathrm{anc}} = \sum_{h \in \mathcal{M} + \mathcal{L}} h |h\rangle\langle h|$. By (A12) and $\delta < 1$, the ancilla size is indeed $\log|\mathcal{M} + \mathcal{L}| = O(\log|\mathcal{M}|)$ and moreover $\|H_{\mathrm{anc}}\| \leqslant 2|\mathcal{M}|$. We take the initial ancilla state to be pure $\eta := |\psi\rangle\langle\psi|$, and given by the Hadamard state

$$|\psi\rangle := |\mathcal{M} + \mathcal{L}|^{-1/2} \sum_{h \in \mathcal{M} + \mathcal{L}} |h\rangle.$$

Furthermore, let $\tilde{U}$ to be any energy-preserving dilation of the energy-preserving partial isometry

$$\tilde{V} := \sum_{h \in \mathcal{M}} \sum_{\lambda, \mu \in \mathcal{L}} P_\lambda U P_\mu \otimes |h + \mu\rangle\langle h + \lambda|,$$

so that $\tilde{U}$ decomposes into a direct sum of $\tilde{V}$ plus an arbitrary other partial isometry. Then $\tilde{V}(\rho \otimes \eta)\tilde{V}^\dagger$ evaluates to

$$|\mathcal{M} + \mathcal{L}|^{-1} \sum_{h_1, h_2 \in \mathcal{M}} \sum_{\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathcal{L}} \left(P_{\lambda_1} U P_{\mu_1} \otimes |h_1 + \mu_1\rangle\langle h_1 + \lambda_1| \right) \left(\rho \otimes \sum_{\ell_1, \ell_2 \in \mathcal{M} + \mathcal{L}} |\ell_1\rangle\langle\ell_2| \right)$$

$$\left(P_{\mu_2} U^\dagger P_{\lambda_2} \otimes |h_2 + \lambda_2\rangle\langle h_2 + \mu_2| \right) = |\mathcal{M} + \mathcal{L}|^{-1} \sum_{\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathcal{L}} P_{\lambda_1} U P_{\mu_1} \rho P_{\mu_2} U^\dagger P_{\lambda_2} \otimes \sum_{h_1, h_2 \in \mathcal{M}} |h_1 + \mu_1\rangle\langle h_2 + \mu_2|$$

$$= |\mathcal{M} + \mathcal{L}|^{-1} \sum_{\mu_1, \mu_2 \in \mathcal{L}} U P_{\mu_1} \rho P_{\mu_2} U^\dagger \otimes \sum_{h_1, h_2 \in \mathcal{M}} |h_1 + \mu_1\rangle\langle h_2 + \mu_2|$$

$$= |\mathcal{M} + \mathcal{L}|^{-1} \sum_{\mu \in \mathcal{L}} U P_\mu \rho P_\mu U^\dagger \otimes \sum_{h \in \mathcal{M}} |h + \mu\rangle\langle h + \mu|,$$

where the last step uses the assumption of absence of coherence in ρ. The resulting reduced state is therefore

$$\mathrm{Tr}_{\mathrm{anc}}[\tilde{V}(\rho \otimes \eta)\tilde{V}^\dagger] = \sum_{\mu \in \mathcal{L}} \frac{|\mathcal{M}|}{|\mathcal{M} + \mathcal{L}|} U P_\mu \rho P_\mu U^\dagger$$

$$= \frac{|\mathcal{M}|}{|\mathcal{M} + \mathcal{L}|} U \rho U^\dagger$$

$$\approx_\delta \frac{|\mathcal{M}|}{|\mathcal{M} + \mathcal{L}|} \sigma \geqslant (1 - \delta)\sigma.$$

The claim now follows from the same estimate as in case 1. ∎

We are now sufficiently equipped to approach the proof of the main result. We write $E(\rho) = \text{Tr}[H\rho]$ for the average energy of a state ρ on a system with Hamiltonian H.

Theorem 17 (Asymptotic classification of states in thermodynamics). For states ρ and σ on any quantum system of dimension d with given Hamiltonian H, the following are equivalent:

(a) The states have equal entropy and average energy, $S(\rho) = S(\sigma)$ and $E(\rho) = E(\sigma)$.

(b) There exists an ancilla system of size $O(\sqrt{n \log n})$ whose Hamiltonian H_{anc} satisfies $\|H_{\text{anc}}\| \leqslant O(n^{2/3})$ with state η as well as an energy-preserving unitary U such that

$$\|\text{Tr}_{\text{anc}}[U(\rho^{\otimes n} \otimes \eta)U^\dagger] - \sigma^{\otimes n}\|_1 \xrightarrow{n \to \infty} 0. \quad (A13)$$

(c) There exists an ancilla system of size $o(n)$ whose Hamiltonian H_{anc} satisfies $\|H_{\text{anc}}\| \leqslant o(n)$ with states η and ν as well as energy-preserving unitaries U and V such that

$$\|\text{Tr}_{\text{anc}}[U(\rho^{\otimes n} \otimes \eta)U^\dagger - V(\sigma^{\otimes n} \otimes \nu)V^\dagger]\|_1 \xrightarrow{n \to \infty} 0. \quad (A14)$$

Definition 18. When one (and hence all) of these conditions hold, we say that ρ is asymptotically equivalent to σ, and we write $\rho \asymp \sigma$.

The bound on $\|H_{\text{anc}}\|$ in condition (b) is not tight: Our proof adapts straightforwardly if one replaces the exponent of 2/3 by any other exponent strictly greater than 1/2. We expect that the bound can be reduced even more, down to at least $O(\sqrt{n \log n})$ as in Theorem 15, but proving this will probably require a more fine-grained arithmetical analysis of the energy levels.

Our interpretation of this result is essentially analogous to Theorem 15. The bound on $\|H_{\text{anc}}\|$ is important in that without such a bound, we could transfer an arbitrary amount of energy to or from the ancilla while only modifying the system state marginally (embezzlement). Of course, none of this is specific to the observable under consideration being energy, and the theorem applies likewise to angular momentum or to any other observable. In fact, we expect the analogous theorem to hold for any finite number of commuting observables on the system that are required to be preserved by the unitaries, with very similar proof. The case of noncommuting observables may be more difficult.

Proof. From (c) to (a), equality of entropy follows from Theorem 15. Equality of energy follows from an estimate analogous to the estimate of entropy difference. With $H^{(n)}$ being the n-qudit Hamiltonian and writing ε for the left-hand side of (A14),

$$
\begin{aligned}
|E(\rho) - E(\sigma)| &= \frac{1}{n}|E(\rho^{\otimes n}) - E(\sigma^{\otimes n})| \leqslant \frac{1}{n}|E(\rho^{\otimes n} \otimes \eta) - E(\rho^{\otimes n} \otimes \nu)| + 2\frac{\|H_{\text{anc}}\|}{n} \\
&= \frac{1}{n}|E[U(\rho^{\otimes n} \otimes \eta)U^\dagger] - E[V(\rho^{\otimes n} \otimes \nu)V^\dagger]| + 2\frac{\|H_{\text{anc}}\|}{n} \\
&\leqslant \frac{1}{n}|E\{\text{Tr}_{\text{anc}}[U(\rho^{\otimes n} \otimes \eta)U^\dagger]\} - E\{\text{Tr}_{\text{anc}}[V(\sigma^{\otimes n} \otimes \nu)V^\dagger]\}| + 4\frac{\|H_{\text{anc}}\|}{n} \\
&\leqslant \frac{1}{n}\varepsilon\|H^{(n)}\| + 4\frac{\|H_{\text{anc}}\|}{n}.
\end{aligned}
$$

Since $H^{(n)}$ is additive in n, we have $\|H^{(n)}\| = n\|H\|$, and the first term vanishes as $\varepsilon \to 0$. The second term vanishes as $n \to \infty$ due to the assumption of sublinearity of $\|H_{\text{anc}}\|$. Note that the bound on H_{anc} now plays the role of the bound on entropy change due to the ancilla.

To see that (a) implies (b), we first apply Theorem 15. So for given $\varepsilon > 0$, we have $n \in \mathbb{N}$ together with the other data such that

$$\text{Tr}_{\text{anc}}[U(\rho^{\otimes n} \otimes \eta)U^\dagger] \approx_\varepsilon \sigma^{\otimes n}. \quad (A15)$$

We now need to find another unitary $\tilde{U}$ that achieves something like (A15) while also being energy preserving.

Let the spectral decomposition of the system's Hamiltonian be $H = \sum_{i=1}^\ell E_i P_i$, and let us assume that the Hamiltonian has been shifted such that $E(\rho) = E(\sigma) = 0$ for simplicity. In order to impose strong energy typicality, let us consider the state ρ_{typ} obtained by restricting $\rho^{\otimes n}$ such that a measurement of $P_i^{(n)}$ will result in an outcome in the range $n\text{Tr}[P_i\rho] \pm \sqrt{n \log n}$ with certainty. By taking n to be large enough, we can assume $\rho^{\otimes n} \approx_\varepsilon \rho_{\text{typ}}$ by the central limit theorem. Let $\mathcal{E}_\rho$ denote the set of energy levels of $H^{(n)}$ on this typical subspace,

and let us throw in their negatives and 0 for good measure,

$$\mathcal{L}_\rho := \mathcal{E}_\rho \cup (-\mathcal{E}_\rho) \cup \{0\}.$$

By construction, the set $\mathcal{E}_\rho$ consists of all numbers of the form $\sum_i c_i E_i$, with integer coefficients c_i that satisfy $|c_i - n\text{Tr}[P_i\rho]| \leqslant \sqrt{n \log n}$ for all i. Therefore, every number in $\mathcal{L}_\rho$ is an integer linear combination of any nonzero fixed number in $\mathcal{L}_\rho$ and the single-system energy levels E_i, using coefficients that are $O(\sqrt{n \log n})$. This implies that the k-fold Minkowski sum

$$k\mathcal{L}_\rho = \underbrace{\mathcal{L}_\rho + \cdots + \mathcal{L}_\rho}_{k \text{ times}}$$

also contains only numbers given by some fixed number plus integer linear combinations of the energy levels E_i using coefficients of size $O(k\sqrt{n \log n})$. Therefore, the cardinality $|k\mathcal{L}_\rho|$ is at most polynomial, $O(\text{poly}(nk))$.

With σ_{typ} and $\mathcal{L}_\sigma$ defined in the analogous manner and satisfying the analogous cardinality bound, let us put $\mathcal{L} := \mathcal{L}_\rho \cup \mathcal{L}_\sigma$, which then in particular contains all the energy levels

that are typical for ρ or for σ. We have the bound

$$|k\mathcal{L}| = \left| \bigcup_{j=0}^{k} j\mathcal{L}_\rho + (k-j)\mathcal{L}_\sigma \right|$$

$$\leqslant \sum_{j=0}^{k} |j\mathcal{L}_\rho| \cdot |(k-j)\mathcal{L}_\sigma| = O(\text{poly}(nk)),$$

so let us choose an exponent $\gamma \in \mathbb{N}$ and a coefficient $C > 0$ such that $|k\mathcal{L}| \leqslant C(nk)^\gamma$ for all k; the particular values are not important.

We now aim to apply Lemma 16 using $\mathcal{M} := k\mathcal{L}$. To determine a suitable value of k, we show that if n is sufficiently large, then there is $k \leqslant n^{1/7}$ such that

$$|k\mathcal{L} + \mathcal{L}| \leqslant (1+\varepsilon)|k\mathcal{L}|. \tag{A16}$$

For if this was not the case, then we would have $|(k+1)\mathcal{L}| > (1+\varepsilon)|k\mathcal{L}|$, which yields by induction on k,

$$|k\mathcal{L}| \geqslant (1+\varepsilon)^k |\mathcal{L}|.$$

For $k = n^{1/7}$, we would then be led to conclude

$$(1+\varepsilon)^{n^{1/7}} |\mathcal{L}| \leqslant |n^{1/7}\mathcal{L}| \leqslant C(n^{8/7})^\gamma.$$

Since the left-hand side grows superpolynomially in n while the right-hand side grows only polynomially, this cannot be

the case for all n. It follows that for suitably large n, there is $k \leqslant n^{1/7}$ such that (A16) holds; let us fix such a k. We now equip the existing ancilla in (A15) with the trivial Hamiltonian $H_{\text{anc}} := 0$, so that also $\rho_{\text{typ}} \otimes \eta$ is supported on the energy levels in $\mathcal{L}$. Because $\text{Tr}_{\text{anc}}[U(\rho_{\text{typ}} \otimes \eta)U^\dagger]$ is 3ε close to σ_{typ}, which is also supported on the energy levels in $\mathcal{L}$, it follows that $U(\rho_{\text{typ}} \otimes \eta)U^\dagger$ itself is already 3ε close to being supported on the energy levels in $\mathcal{L}$. Let us write $\hat{\rho} := \rho^{\otimes n} \otimes \eta$ and $\hat{\sigma}$ for the restriction of $U(\rho_{\text{typ}} \otimes \eta)U^\dagger$ to the energy levels in $\mathcal{L}$, so that $\hat{\sigma} \approx_{3\varepsilon} U\hat{\rho}U^\dagger$. By taking $\mathcal{M} := k\mathcal{L}$ in Lemma 16, we can therefore conclude the existence of an additional ancilla system anc' of size $O(\log(|k\mathcal{L}|)) = O(\log n)$ with Hamiltonian bounded by $4\|k\mathcal{L}\| = 4k\|\mathcal{L}\| \leqslant n^{1/7}O(\sqrt{n\log n}) < O(n^{2/3})$ as claimed, with an ancilla state η' and energy-preserving unitary $\tilde{U}$ such that

$$\text{Tr}_{\text{anc}'}[\tilde{U}(\hat{\rho} \otimes \eta')\tilde{U}^\dagger] \approx_{12\varepsilon} \hat{\sigma}.$$

Putting all this together, we therefore have

$$\text{Tr}_{\text{anc,anc}'}[\tilde{U}(\rho^{\otimes n} \otimes \eta \otimes \eta')\tilde{U}^\dagger] = \text{Tr}_{\text{anc,anc}'}[\tilde{U}(\hat{\rho} \otimes \eta')\tilde{U}^\dagger]$$

$$\approx_{12\varepsilon} \text{Tr}_{\text{anc}}[\hat{\sigma}]$$

$$\approx_{3\varepsilon} \text{Tr}_{\text{anc}}[U\hat{\rho}U^\dagger]$$

$$\approx_{\varepsilon} \sigma^{\otimes n},$$

resulting in a total trace distance difference between the left-hand side and the right-hand side of at most 16ε. ∎

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

[2] M. Horodecki and J. Oppenheim, Nat. Commun. **4**, 2059 (2013).

[3] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A **67**, 062104 (2003).

[4] F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, Proc. Natl. Acad. Sci. U.S.A. **112**, 3275 (2015).

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

[6] N. Yunger Halpern and J. M. Renes, Phys. Rev. E **93**, 022126 (2016).

[7] M. P. Woods, N. Ng, and S. Wehner, arXiv:1506.02322 [quant-ph].

[8] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **115**, 210403 (2015).

[9] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. X **5**, 021001 (2015).

[10] H. Wilming, R. Gallego, and J. Eisert, Phys. Rev. E **93**, 042126 (2016).

[11] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern, Phys. Rep. **583**, 1 (2015).

[12] L. Masanes and J. Oppenheim, Nat. Commun. **8**, 14538 (2017).

[13] A. M. Alhambra, L. Masanes, J. Oppenheim, and C. Perry, Phys. Rev. X **6**, 041017 (2016).

[14] C. Perry, P. Ćwikliński, J. Anders, M. Horodecki, and J. Oppenheim, arXiv:1511.06553 [quant-ph].

[15] N. Y. Halpern, P. Faist, J. Oppenheim, and A. Winter, Nat. Commun. **7**, 12051 (2016).

[16] R. Alicki and M. Fannes, Phys. Rev. E **87**, 042123 (2013).

[17] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).

[18] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, New J. Phys. **13**, 053015 (2011).

[19] L. d. Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, Nature (London) **474**, 61 (2011).

[20] J. Åberg, Nat. Commun. **4**, 1925 (2013).

[21] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, Nat. Commun. **6**, 7669 (2015).

[22] P. Skrzypczyk, A. J. Short, and S. Popescu, Nat. Commun. **5**, 4185 (2014).

[23] K. Korzekwa, M. Lostaglio, J. Oppenheim, and D. Jennings, New J. Phys. **18**, 023045 (2016).

[24] V. Narasimhachar and G. Gour, Nat. Commun. **6**, 7689 (2015).

[25] R. Gallego, J. Eisert, and H. Wilming, New J. Phys. **18**, 103017 (2015).

[26] L. J. Schulman and U. V. Vazirani, in *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing*, STOC'99 (ACM, New York, 1999), pp. 322–329.

[27] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B **27**, 1345019 (2012).

[28] B. Coecke, T. Fritz, and R. W. Spekkens, Inf. Comput. **250**, 59 (2016).

[29] L. del Rio, L. Krämer, and R. Renner, arXiv:1511.08818 [quant-ph].

[30] S. Vinjanampathy and J. Anders, Contemp. Phys. **57**, 545 (2015).

[31] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, J. Phys. A **49**, 143001 (2016).

[32] H. Tajima and M. Hayashi, Phys. Rev. E **96**, 012128 (2017).

[33] J. Richens, A. M. Alhambra, and L. Masanes, arXiv:1702.03357 [quant-ph].

[34] T. Fritz, Math. Struct. Comp. Sci. **27**, 850 (2017).

[35] Notice that these quantities are precisely the *asymptotically continuous* ones [41].

[36] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932), Ch. 5.

[37] This terminology was suggested to us by Matteo Smerlak. Without the additive constant, A_β has also been called the "free entropy" in Ref. [50].

[38] N. F. Ramsey, Phys. Rev. **103**, 20 (1956).

[39] G. M. Ziegler, *Lectures on Polytopes*, Graduate Texts in Mathematics Vol. 152 (Springer, Berlin, 1995).

[40] Of course this depends on which observables are considered to be conserved quantities. For us, as indicated by Theorem 17, energy is assumed to be the only observable that is conserved.

[41] B. Synak-Radtke and M. Horodecki, J. Phys. A **39**, L423 (2006).

[42] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Proc. Royal Soc. London A **466**, 3203 (2010).

[43] In the sense that two decompositions do in general not have a common refinement.

[44] J. Barrett, Phys. Rev. A **75**, 032304 (2007).

[45] H. Barnum, J. Barrett, M. Krumm, and M. P. Müller, in *Proceedings of the 12th International Workshop on Quantum Physics and Logic* (EPTCS 195, Oxford, 2015), pp. 43–58.

[46] G. Chiribella and C. M. Scandolo, New J. Phys. **17**, 103027 (2015).

[47] A. E. Allahverdyan, R. Balian, and T. M. Nieuwenhuizen, Europhys. Lett. **67**, 565 (2004).

[48] D. Reeb and M. M. Wolf, New J. Phys. **16**, 103011 (2014).

[49] M. Lostaglio, D. Jennings, and T. Rudolph, New J. Phys. **19**, 043008 (2017).

[50] Y. Guryanova, S. Popescu, A. J. Short, R. Silva, and P. Skrzypczyk, Nat. Commun. **7**, 12049 (2016).

[51] Or any other *single* conserved quantity—if one replaces, e.g., "energy" by "angular momentum" throughout our work, our results are still equally correct and meaningful.

[52] C. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[53] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[54] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).

[55] I. Devetak, A. W. Harrow, and A. Winter, IEEE Trans. Inf. Theory **54**, 4587 (2008).

[56] More precisely, by the fact that the conditional entropy of the ancilla given the system is at most $\log(D)$ in absolute value.

[57] E. Breuillard, B. Green, and T. Tao, in *Erdős Centennial*, Bolyai Society Mathematical Studies Vol. 25 (Springer, Berlin, 2013), pp. 129–151.