Beyond heat baths: Generalized resource theories for small-scale thermodynamics

Nicole Yunger Halpern[1,2,*] and Joseph M. Renes[3,†]

[1]*Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA*
[2]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5*
[3]*Institute for Theoretical Physics, ETH Zürich, Switzerland*

Thermodynamics has recently been extended to small scales with *resource theories* that model heat exchanges. Real physical systems exchange diverse quantities: heat, particles, angular momentum, etc. We generalize thermodynamic resource theories to exchanges of observables other than heat, to baths other than heat baths, and to free energies other than the Helmholtz free energy. These generalizations are illustrated with "grand-potential" theories that model movements of heat and particles. Free operations include unitaries that conserve energy and particle number. From this conservation law and from resource-theory principles, the grand-canonical form of the free states is derived. States are shown to form a quasiorder characterized by free operations, d majorization, the hypothesis-testing entropy, and rescaled Lorenz curves. We calculate the work distillable from—and we bound the work cost of creating—a state. These work quantities can differ but converge to the grand potential in the thermodynamic limit. Extending thermodynamic resource theories beyond heat baths, we open diverse realistic systems to modeling with one-shot statistical mechanics. Prospective applications such as electrochemical batteries are hoped to bridge one-shot theory to experiments.

I. INTRODUCTION

Advances in small-scale experiments and in quantum information have generated interest in "thermodynamics without the thermodynamic limit." Recent experiments involve molecular motors and ratchets [1,2], optical thermal ratchets [3], the unfolding of one DNA or RNA molecule [4–7], and nanoscale walkers [8]. Analyses of these experiments feature thermodynamic concepts such as heat, work, and equilibrium. These concepts are not well-defined outside the thermodynamic limit of $n \to \infty$ particles. Hence, the experimental advances in single-molecule manipulations invite us to extend thermodynamics to small scales.

The resource-theory framework developed in quantum information theory has recently been successfully applied to this problem. Resource theories have been used to calculate how efficiently scarce quantities can be distilled and transformed via cheap, or "free," operations [9]. Perhaps the most famous example is the resource theory of pure bipartite entanglement (which we will call "entanglement theory") [10]. In entanglement theory, agents distill Bell pairs of maximally entangled qubits, usable to simulate quantum channels, from partially entangled states via local operations and classical communications (LOCC). Other resource theories quantify the values of asymmetry [11–13], quantum-computation tools [14], and information [15–17]. Benefits of the resource-theory framework include its operational formulation and the explicit modeling of all resources with physical degrees of freedom.

To an agent with access to a heat bath, nonequilibrium states have value because work can be extracted from them and stored in a battery. Nonequilibrium states' values have been quantified with a family of equivalent resource theories, each associated with an inverse temperature β of the bath [18–23]. We call these resource theories *Helmholtz theories*,

as the central results involve variations on the Helmholtz free energy $F := E - TS$.

Many experiments involve baths other than heat baths, involve interactions other heat exchanges, and are characterized by free energies other than the Helmholtz free energy. The Gibbs free energy $G := E - TS + pV$ describes processes that occur at fixed temperatures and pressures, such as tabletop chemical reactions. The grand potential $\Phi := E - TS - \mu N$ describes heat-and-particle exchanges, and other free energies describe electrochemistry, magnetic fields, mechanical stress and strain, etc. [24,25]. Different types of baths (equivalently, different types of interactions, or different free energies) invite modeling by different families of resource theories. Each family's constituents correspond to different values of the bath's properties. For example, each member of the family of Helmholtz theories corresponds to one value of the inverse temperature β. Altogether, the families describing different baths form an extended family of thermodynamic resource theories amenable to experimental investigation in the present or near future.

We introduce this extended family in this paper, illustrating the formalism with heat-and-particle exchanges. In grand-potential resource theories, free operations conserve energy and particle number. The only states that, if free, prevent such resource theories from being trivial are shown to be grand-canonical ensembles $e^{-\beta(H-\mu N)}/Z$. We derive the grand-canonical ensemble upon establishing rigorously, using the resource-theory formalism, that the free states in Helmholtz theories are canonical ensembles $e^{-\beta H}/Z$. States are shown to form a quasiorder characterized by a variant of majorization called *d majorization*, which is related to binary hypothesis testing. By exploiting the quasiorder, we calculate the work extractable from—and bound the work required to create—one copy of a state $R = (\rho, H, N)$, even by protocols that have a specified probability of failing. The work yield and work cost are shown to differ from each other, in general, unlike in conventional thermodynamics, as observed in Ref. [19]. In the

*nicoleyh@caltech.edu
†renes@phys.ethz.ch

limit as the number of copies of R extracted from or created approaches infinity, the average work yield and work cost approach the difference $\Phi(\rho, H, N) - \Phi(\gamma_\rho, H, N)$ between the state's grand potential and the corresponding equilibrium state's grand potential.

We have structured our results as follows. In the next section, we review the resource-theory framework and define the family of generalized thermodynamic resource theories. We illustrate the family with grand-potential theories. In Sec. III, we deduce the unique form that free states can assume in these theories. In Sec. IV, we interrelate the quasiorder of states, the generalized notion of majorization, and binary hypothesis testing. In Sec. V, we define work in the resource-theory framework and use the quasiorder to determine the work yield and work cost of creating single instances of arbitrary states. In Sec. VI, we show that these one-shot work quantities imply asymptotic results similar to traditional thermodynamics. We conclude by discussing possible applications of our generalized framework to real physical systems.

This work bridges the information-theoretic tool of thermodynamic resource theories to physical reality. We pave the way for physical realizations, with experimental platforms, of entropic predictions about small scales.

II. THERMODYNAMIC RESOURCE THEORIES

First, we introduce the resource-theory framework. We define generalized thermodynamic resource theories, then illustrate them with grand-potential theories.

A. The resource-theory framework

The resource-theory framework models experimental situations in which some physical transformations between quantum states are difficult, while others are easy [9]. In quantum optics, for instance, generating coherent states is easy (e.g., using a laser), while generating number states is difficult. In entanglement theory, classical communication and physical operations on systems possessed by one party each (LOCC) are easy, while quantum operations on systems distributed among multiple parties are impossible. The states that are difficult to create can be regarded as resources since, with free operations, they can simulate difficult operations. Given a maximally entangled state, separated parties restricted to LOCC can implement a quantum channel.

A resource theory is defined by physical operations assumed to be easy, or *free*. Free operations include the creation of *free states*; all other states are resources. This definition specifies an ordering of states: States A and B are ordered as $A \mapsto B$ if free operations can create B from A. The ordering $\mapsto$ is a quasiorder, satisfying reflexivity ($A \mapsto A$) and transitivity ($A \mapsto B$ and $B \mapsto C$ implies $A \mapsto C$). The quasiorder differs from a partial order: Even if $A \mapsto B$ and $B \mapsto A$, A is not necessarily B [26].

Functions of resources that respect the quasiorder, in the sense that $A \mapsto B$ implies $f(A) \geqslant f(B)$, are termed *resource monotones* [17,21,27]. Simple sets of monotones completely characterize the quasiorders in some resource theories, including the resource theories in this paper. Many monotones have operational interpretations [17,21].

We are interested only in resource theories whose quasiorders are nontrivial, in which some transformations between some resources are impossible. When independently specifying a resource theory's free operations and free states, we must prevent free states and free operations from being able, together, to generate arbitrary states.

Having introduced the resource-theory framework, we define generalized thermodynamic resource theories and illustrate them with grand-potential theories.

B. Generalized thermodynamic resource theories

Requiring that free operations conserve particular physical quantities leads to the extended family of thermodynamic resource theories. Which quantities are conserved depends on which physical systems are modeled, as explained in Ref. [28]. The Hamiltonian H is conserved in what we have termed *Helmholtz theories*. Janzing *et al.* first defined Helmholtz theories while investigating the resources required to cool systems, though those authors did not use the term "resource theory" [18]. More recently, Brandão *et al.* studied conversions between resources in the asymptotic limit, as the number n of copies of the converted resource diverges ($n \to \infty$) [20]. Horodecki and Oppenheim extended the analysis of conversions beyond the asymptotic limit [19]. The literature about thermodynamic resource theories has exploded recently. Since the first draft of the present paper was released, for example, coherences [29,30] and correlations have been explored [31], connections have been drawn to fluctuation relations [32–34], and free operations have been generalized [23].

In the *grand-potential* theories focused on in this paper, free operations preserve total energy and total particle number. Free states in thermodynamic resource theories such as grand-potential theories model baths such as heat-and-particle reservoirs. As we see in Sec. III, the free states must be equilibrium states, such as grand-canonical ensembles if H and N are conserved. If nonequilibrium states are free, the resource theory's quasiorder becomes trivial.

We associate different baths (equivalently, different physical quantities that can be exchanged, or, in anticipation of Sec. VI, different free energies) with different families of thermodynamic resource theories. The family of grand-potential theories models heat-and-particle exchanges. The resource theories in each family differ only by the values of the intensive variables that characterize the bath. To specify a grand-potential resource theory, one specifies an inverse temperature β and a chemical potential μ. (For simplicity, we focus on systems that contain particles of only one type. To specify a grand-potential theory that models exchanges of particles of k types, one specifies $\beta, \mu_1, \mu_2, \ldots, \mu_k$.)

Now, let us define the thermodynamic resource theories, their states, and their operations more precisely. Free operations preserve quantities represented, in conventional thermodynamics, by extensive variables. These variables are represented by operators that, with a density operator, define a state. To specify a state R in a grand-potential theory, one specifies a density operator, a Hamiltonian, and a number operator:

$$R = (\rho, H, N). \tag{1}$$

These operators are defined on a quantum state space (Hilbert space) $\mathcal{H}_R$. How a physical system's bath and interactions translate into intensive and extensive variables that define a general family of thermodynamic resource theories is detailed in Ref. [28].

For simplicity, we specialize to states whose operators commute with each other ($[\rho,H] = [\rho,N] = [H,N] = 0$) and have discrete, finite spectra. (Since the initial release of this paper, noncommuting operators have been discussed in Refs. [28,35–37].) However, we do not restrict the forms of H and N further. We denote the dimension of $\mathcal{H}_R$ by d_R. The density operator ρ can be represented by a matrix diagonal relative to the eigenbasis shared by H and N. Such a *quasiclassical* density operator is fully specified by a vector r, which we call the *state vector*, of its eigenvalues. Hence, we also denote the state by $R = (r,H,N)$. The ordering of the elements r_i in r is discussed in Sec. IV.

The composition of the state $R = (\rho,H_R,N_R)$ on $\mathcal{H}_R$ with the state $S = (\sigma,H_S,N_S)$ on $\mathcal{H}_S$ is defined as

$$R + S = (\rho \otimes \sigma, H_R + H_S, N_R + N_S), \qquad (2)$$

wherein $H_R + H_S = H_R \otimes \mathbb{1}_S + \mathbb{1}_R \otimes H_S$ and $N_R + N_S$ is defined similarly.

As detailed in Sec. III, free states have density operators whose the probabilities equal Boltzmann factors. The free states in the grand-potential theory take the form

$$\gamma = e^{-\beta(H - \mu N)}/Z, \qquad (3)$$

wherein β and μ are real numbers and Z is the normalization factor, or partition function. We denote free states by $G = (\gamma,H,N)$ or $G = (g,H,N)$. Each resource $R = (\rho,H,N)$ is associated with an equilibrium state $G_R = (\gamma_R,H,N)$ or $G = (g_R,H,N)$.

Free operations

We call the free operations in thermodynamic resource theories *equilibrating operations*. They are defined in our grand-potential example as follows.

Definition 1: Equilibrating operation. In the grand-potential theory defined by (β,μ), an *equilibrating operation* on a state $R = (\rho,H_R,N_R)$ is any realization of the following three steps:

(a) the drawing of a free state $G = (\gamma,H_G,N_G)$ from the bath;

(b) the performing of a unitary transformation U on $R + G$, wherein $[U,H_R + H_G] = [U,N_R + N_G] = 0$; and

(c) the discarding (tracing out) of any subsystem A associated with its own Hamiltonian and number operator.

The operation is a completely positive trace-preserving linear map of the form

$$R \mapsto R' = (\mathrm{Tr}_A(U[\rho \otimes \gamma]U^\dagger), H_R$$
$$+ H_G - H_A, N_R + N_G - N_A). \qquad (4)$$

Free operations can mix levels whose energies equal each other and whose particle numbers equal each other. We call free operations *equilibrating operations* because (as shown in Sec. IV) free operations monotonically evolve states toward equilibrium states. Equilibrating operations induce on states a quasiorder that we denote by $R \overset{\beta,\mu}{\mapsto} R'$.

In thermodynamic resource theories other than the grand-potential theories we focus on, free unitaries preserve operators associated with other extensive variables. For example, if a system has N_1 particles of species 1 and N_2 particles of species 2, $[U,N_{1_{\mathrm{tot}}}] = [U,N_{2_{\mathrm{tot}}}] = 0$.

Equilibrating operations idealize the operations easily performable by thermodynamic experimentalists. Experimentalists cannot perform all unitaries that preserve energy, particular number, etc. Thermal operations were "coarse grained" to more-realistic operations in Ref. [38]. We expect similar coarse graining to bridge equilibrating operations from idealization to reality.

III. UNIQUE FORM OF FREE STATES

The form of equilibrating operations in Definition 1 implies that only grand-canonical ensembles can be free in grand-potential theories. If any other state is free, the quasiorder breaks down. The breakdown manifests in two ways.

Theorem 1. Consider any grand-potential resource theory in which each pair (H,N) corresponds to exactly one free state $G = (\gamma,H,N)$. If γ does not have the Boltzmann form of Eq. (3), then

(a) some resources R can be generated solely with equilibrating operations: $G \overset{\beta,\mu}{\mapsto} R$; and

(b) equilibrating operations can transform n copies of any state R into m_n copies of any state S, for any m_n and for sufficiently large n: $R^{\otimes n} \overset{\beta,\mu}{\mapsto} S^{\otimes m_n}$.

The proofs appear in Appendix A, but we sketch the main ideas here. First we derive the form of the free states in Helmholtz theories.[1] Then we bootstrap from Helmholtz theories to Theorem 1, which concerns grand-potential theories.

To prove claim (a), we apply the derivation of the forms of the free states in the resource theory of nonuniformity [16], which models closed isolated systems [28]. Consider some energy-and-particle-number eigensubspace S_{E_i,N_j}. The free state γ has some weight on S_{E_i,N_j}. If that weight is distributed nonuniformly across the levels in S_{E_i,N_j}, free operations can redistribute the weight arbitrarily across S_{E_i,N_j}, generating states not defined as free. Claim (b) follows from modifying an argument by Janzing *et al.* [18]. The argument concerns the "effective temperatures" of the states that can be created from given resources in the absence of any bath.

These resource-theory derivations of equilibrium ensembles offer operational alternatives to assumptions such as the ergodic hypothesis. According to the ergodic hypothesis, uniform distributions represent equilibrated isolated systems'

---

[1]In Ref. [21], the canonical form ($e^{-\beta H}/Z,H$) of the free states in Helmholtz theories is argued to follow from [39]. According to [39], only canonical ensembles are completely passive: No work can be extracted from canonical ensembles, even from infinitely many, in the absence of other resources. The work extraction in Ref. [39], however, is not formulated as in Helmholtz resource theories. How to translate the result from [39] into resource theories may not be obvious to all readers. To clarify this subtlety and others, we derive free states' forms directly from the resource-theory framework. Around the time our paper was released, an alternative approach appeared in Ref. [21].

states. Such assumptions have drawn criticism [40,41], lending operational replacements appeal.

IV. QUASIORDER ON STATES

The quasiorder induced by equilibrating operations on quasiclassical states is equivalent to a generalization of majorization. Veinott defined this generalization first, calling it *d majorization* [42]. Ruch and collaborators (who called *d* majorization *the mixing distance*) [43,44] applied *d* majorization to physics, as did Uhlmann and colleagues [45,46]. We dub this quasiorder, in general, thermodynamic resource theories *equimajorization*, because it is *d* majorization relative to equilibrium states.

Definition 2. Let R and S denote states in any grand-potential theory defined by (β,μ). Let g_R and g_S denote the corresponding equilibrium states' state vectors, which contain d_R and d_S elements respectively. R *equimajorizes* S, written as $R \succ_{\beta,\mu} S$, if there exists a $d_S \times d_R$ stochastic matrix M such that

$$Mr = s, \quad Mg_R = g_S, \quad \text{and}$$

$$\sum_{i=1}^{d_S} M_{ij} = 1 \ \forall \ j = 1,2,\ldots,d_R. \tag{5}$$

In the resource theory of nonuniformity, which models closed isolated systems, equilibrium states are microcanonical ensembles: $g_R = \frac{1}{d_R}(1,1,\ldots,1)$ [16,17]. Relative to these uniform states, equimajorization reduces to majorization [27].

Janzing *et al.* established that the quasiorder on quasiclassical resources is equivalent to equimajorization in Helmholtz theories [[18], Theorem 5]. An alternative proof appears in Ref. [19]. In Appendix B, we extend the proof technique in Ref. [18] to grand-canonical theories, obtaining the following result.

Theorem 2. Let R and S denote states in the grand-potential theory defined by (β,μ). There exists an equilibrating operation that maps R to S if and only if R equimajorizes S:

$$R \overset{\beta,\mu}{\mapsto} S \Longleftrightarrow R \succ_{\beta,\mu} S. \tag{6}$$

As mentioned above, sets of resource monotones completely characterize equimajorization and so characterize the existence of equilibrating operations. One such set consists of the *f divergences* [47–49]. Every convex function f corresponds to an f divergence,

$$\phi_f(R) = \sum_{i=1}^{d_R} g_i \, f\left(\frac{r_i}{g_i}\right), \tag{7}$$

wherein g_i denotes the ith element of the state vector g_R of the equilibrium state associated with R. Subsets of the f divergences suffice to characterize equimajorization, as shown below. Various choices of f lead to well-known functions [50]. For example, $f(x) = x \log x$ and $f(x) = -\log x$ lead to the relative entropies $D(r||g_R) = \sum_{i=1}^{d_R} r_j \log(r_i/g_i)$ and $D(g_R||r)$. The function $f(x) = (x^\alpha - 1)/(\alpha - 1)$ leads to the
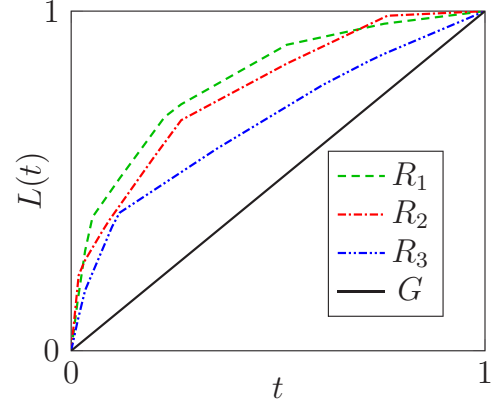

FIG. 1. Rescaled Lorenz curves for three resources (R_1, R_2, R_3) and an equilibrium state (G). The Lorenz curve encodes the quasiorder on states, as equilibrating operations can transform R into R' if and only if $L_R(t) \geqslant L_{R'}(t) \ t \in [0,1]$. Here, $R_1 \overset{\beta,\mu}{\mapsto} R_3$, and $R_2 \overset{\beta,\mu}{\mapsto} R_3$, but R_1 and R_2 are incomparable. The equilibrium state, having the linear Lorenz curve $L_G(t) = t$, is at the bottom of the quasiorder.

Rényi divergences $D_\alpha(r||g_R) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{d_R} r_i^\alpha g_i^{(1-\alpha)}\right)$ for $\alpha \geqslant 0$.[2]

The *Lorenz curve*, introduced by Lorenz in economics [51], encodes another complete set of monotones. Lorenz curves were applied recently to Helmholtz resource theories [19]. In a grand-potential theory, the rescaled Lorenz curve $L_R : [0,1] \to [0,1]$ represents the state R. The curve is the piecewise linear function that connects the points

$$
\begin{aligned}
&(t_k, L_R(t_k)) \\
&= \begin{cases} (0,0) & k = 0, \\ \left(\sum_{j=1}^{k} g_{\pi(j)}, \sum_{j=1}^{k} r_{\pi(j)}\right) & k \in \{1,\ldots,d_R\}, \end{cases}
\end{aligned} \tag{8}
$$

wherein π denotes a permutation such that the sequence $(r_{\pi(j)}/g_{\pi(j)})_j$ is nonincreasing. In accordance with [19,51], we define the rescaled Lorenz curve as a monotonically increasing concave function. (Different conventions appear elsewhere [27].)

Having defined the f divergences and the rescaled Lorenz curve, we state their relationship with equimajorization. Ruch, Schranner, and Seligman first proved this relationship for continuous systems [43], using tools from measure theory. Uhlmann proved the relationship more directly, for discrete systems, which we address [45]. By following Uhlmann, we prove this proposition in Appendix B.

Proposition 1. For any states R and S in the grand-potential resource theory defined by (β,μ), the following are equivalent:

(a) $R \succ_{\beta,\mu} S$;

(b) $L_R(t) \geqslant L_S(t)$ for all $t \in [0,1]$;

(c) $\phi_{f_a}(R) \geqslant \phi_{f_a}(S)$ for every function $f_a(t) = \max\{0, t - a\}$ associated with any $a \in \mathbb{R}$;

---

[2]The foregoing statements hold regardless of the logarithms' bases. Throughout the rest of this paper, statistical mechanics motivates us to specialize to base e.

(d) $\phi_f(R) \geqslant \phi_f(S)$ for all continuous convex functions f. An illustration appears in Fig. 1.

Having characterized L_R in terms of eigenvalues, we explain its relationship with hypothesis testing. The rescaled Lorenz curve is equivalent to the minimal type II error probability, cast as a function of the type I error probability, in an asymmetric hypothesis test. Harremoës noted the relationship between Lorenz curves and hypothesis tests [52]; we establish the relationship more concretely.

An *asymmetric hypothesis test* is used to distinguish whether a given state is ρ or σ. As indicated by our notation, hypothesis testing can be defined in quantum contexts. A test can be thought of as a two-outcome positive operator-valued measure $\{Q, \mathbb{1} - Q\}$. This measure defines a measurement. If the measurement yields the outcome Q, the state is likely ρ. If the measurement yields $\mathbb{1} - Q$, the state is likely σ. A type I error occurs if the state is ρ but $\mathbb{1} - Q$ obtains, so the state seems likely to be σ. A type II error occurs if the state is σ but seems likely to be ρ. The *optimal* test minimizes the type II error probability while preventing the type I error probability from exceeding some tolerance $\epsilon \in [0,1]$.

The optimal type II error probability is

$$b_\epsilon(\rho||\sigma) := \min_{\substack{\text{Tr}[Q\rho] \geqslant 1-\epsilon \\ 0 \leqslant Q \leqslant \mathbb{1}}} \text{Tr}[Q\sigma]. \qquad (9)$$

The condition $\text{Tr}[Q\rho] \geqslant 1 - \epsilon$ is called the *constraint*, and $\text{Tr}[Q\sigma]$ is the *objective function*. Equation (9) defines a semidefinite program, a type of convex optimization, which has a dual form:

$$b_\epsilon(\rho||\sigma) = \max_{\substack{\mu\rho-\sigma \leqslant \tau \\ \mu,\tau \geqslant 0}} \{(1-\epsilon)\mu - \text{Tr}[\tau]\}. \qquad (10)$$

The primal and dual forms' equivalence follows from properties of semidefinite programs [53]. In quasiclassical notation, Q is represented by a matrix, and traces are replaced by sums.

Hypothesis testing can be related to rescaled Lorenz curves as follows. Consider distinguishing between the state vector r in the quasiclassical state $R = (r, H, N)$ and the g_R in the equilibrium state $G_R = (g_R, H, N)$.

Lemma 1. The inverse of $\epsilon \mapsto b_\epsilon(r||g_R)$ is the piecewise linear function that connects the points $(L_R(t_k), 1 - t_k)$, wherein t_k and $L_R(t_k)$ define the rescaled Lorenz curve for R. That is,

$$(t_k, 1 - L_R(t_k)) = (b_\epsilon(r||g_R), \epsilon). \qquad (11)$$

The proof appears in Appendix B.

V. ONE-SHOT WORK YIELD AND COST

Let us quantify the work required to create—and the work extractable from—one copy of a state $R = (r, H, N)$ via protocols that can fail, as realistic protocols can. Upon motivating the calculation, we introduce the hypothesis-testing entropy D_H^ϵ, incorporate a failure probability into equilibrating operations, and define work in thermodynamic resource theories. Finally, we calculate the extractable work and bound the work cost. Proofs appear in Appendix C.

Conventional thermodynamics concerns the average work $\langle W_\text{gain} \rangle$ extractable from—and the average cost $\langle W_\text{cost} \rangle$ of

creating—states by infallible protocols in the asymptotic limit. In the *asymptotic limit*, related to the thermodynamic limit, infinitely many identical copies of R are extracted from or created. $\langle W_\text{gain} \rangle$ and $\langle W_\text{cost} \rangle$ depend on the Shannon entropy S_S, itself an average:

$$S_\text{S}(r) := \sum_i r_i \ln r_i = \langle \ln r_i \rangle_r. \qquad (12)$$

If few copies of a state are extracted from or created, the average cost or yield quantifies the protocol's efficiency poorly. Alternatives to S_S, called *one-shot entropies*, quantify efficiencies in information processing (e.g., [54–58]) and statistical-mechanics (e.g., [19,21,59–62]) problems that involve few systems or trials. In addition to involving finite numbers, realistic protocols have nonzero probabilities of failing to accomplish their purposes. Failure probability has been incorporated into one-shot entropies as a parameter ϵ [53,54].

One alternative to S_S is the *hypothesis-testing entropy* D_H^ϵ. D_H^ϵ is defined in terms of the hypothesis test quantified in Eq. (9). The work extractable from—and the work cost of creating—one copy of a state R is quantified with D_H^ϵ.

Definition 3. The *hypothesis-testing relative entropy* between quantum states ρ and γ is defined as

$$D_\text{H}^\epsilon(\rho||\gamma) := -\ln b_\epsilon(\rho||\gamma) \qquad (13)$$

or, equivalently, by $b_\epsilon(\rho||\gamma) = e^{-D_\text{H}^\epsilon(\rho||\gamma)}$.

Let us incorporate failure probability and work into thermodynamic resource theories.

A faulty operation is defined as a transformation whose output approximates the desired output. Operationally, a state R' approximates a state R if no testing procedure consistent with quantum mechanics can reliably distinguish the states. For simplicity, we focus on approximations R' that differ from R only because of their density operators: If $R = (\rho, H, N)$, then $R' = (\rho', H, N)$. If R' approximates R, we write $R' \approx_\epsilon R$ and say that R' is ϵ *close* to R. (Equivalently, we write $\rho' \approx_\epsilon \rho$ and say that ρ' is ϵ *close* to ρ.) Since density operators' distinguishability is related to the trace distance, we define $R' \approx_\epsilon R$ by $\frac{1}{2}\|\rho' - \rho\|_1 \leqslant \epsilon$.

We define work in terms of the changing of the energy level occupied by a *battery*. In statistical physics and mechanics, work is defined as an integral along a path in real space or in phase space. Quantum states can follow paths along which work integrals cannot easily be calculated [63]. Work on and by quantum systems has been defined more operationally in terms of a "work bit" that has a gap W [19] and in terms of a weight that stores gravitational potential energy [64]. We define work similarly to [64].

Our battery B is any system that has the following qualities: (a) The energies in the range accessed by the agent are finely spaced. (b) The battery occupies an energy eigenstate, being a reliable energy reservoir. By B_E, we denote the battery resource $(|E\rangle\langle E|, H, N)$. We assume $\beta E \gg 1$, for if $\beta E \approx 1$, agents can use the battery's equilibrium state to drive processes that require energy E. Such a use would contradict our physical notion of useful work.

Having defined B, we can define the work extractable from—and the work cost of—a state transformation. If the

battery transitions from B_{E_i} to B_{E_f} while R transforms into S, the transformation outputs the work $E_f - E_i$ (which is negative if the transformation costs work). The ϵ-work value of a resource R is defined as the greatest W for which

$$R + B_E \overset{\beta,\mu}{\mapsto}_\epsilon B_{E+W}. \tag{14}$$

The work cost of ϵ approximately creating R is the least W such that

$$B_{E+W} \overset{\beta,\mu}{\mapsto}_\epsilon R + B_E. \tag{15}$$

Formally,

$$W_{\text{gain}}^\epsilon(R) = \max\{W : R + B_E \overset{\beta,\mu}{\mapsto}_\epsilon B_{E+W}, \beta E \gg 1\}, \tag{16}$$

$$W_{\text{cost}}^\epsilon(R) = \min\{W : B_{E+W} \overset{\beta,\mu}{\mapsto}_\epsilon R + B_E, \beta E \gg 1\}. \tag{17}$$

The simplicity of our battery model facilitates calculations. Realistic features could be incorporated as follows. First, the battery could occupy a mixed state, or a superposition of energy eigenstates, at any stage in either protocol. Second, the system and battery could begin or become entangled. Such entanglement could be analyzed as in Ref. [65]. Frenzel *et al.* point out that a classical field is often assumed to raise and lower a quantum system's energy levels. However, fields are not classical and become entangled with the system. A battery might become entangled similarly.

Having defined work, we state the work value—and bound the work cost—of R.

Theorem 3. The ϵ-work value of a state $R = (r, H, N)$ associated with the free state $G_R = (g_R, H, N)$ is

$$W_{\text{gain}}^\epsilon(R) = \frac{1}{\beta} D_{\text{H}}^\epsilon(r||g_R). \tag{18}$$

The work cost of creating an ϵ approximation to a state R is bounded by

$$\max_{\delta \in (0, 1-\epsilon]} \left[\frac{1}{\beta} D_{\text{H}}^{1-\epsilon-\delta}(r||g_R) - \frac{1}{\beta} \log\left(\frac{1}{\delta}\right) \right]$$
$$\leqslant W_{\text{cost}}^\epsilon(R) \leqslant \frac{1}{\beta} D_{\text{H}}^{1-\epsilon}(r||g_R) - \frac{1}{\beta} \log\left(\frac{1-\epsilon}{\epsilon}\right). \tag{19}$$

A proof appears in Appendix C. Each expression in the theorem contains an entropy $D_{\text{H}}^\epsilon(r||g_R)$ for some error probability ϵ. The factor $\frac{1}{\beta}$ introduces dimensions of energy. Each bound contains a logarithmic correction.

Theorem 3 bounds optimal efficiencies. Thermodynamic optima tend to characterize physically unrealizable processes. Example processes include quasistatic, or infinitely slow, evolutions. Experiments cannot proceed infinitely slowly. What implications can Theorem 3 have for real physical processes? As a process is performed increasingly slowly, its efficiency is expected to approach our predictions. A similar approach has been reported in Ref. [66]. Koski *et al.* erased a bit of information repeatedly. As the erasure's speed dropped, the amount of heat dissipated dropped to near the Landauer limit.

VI. WORK YIELD AND COST OF MANY COPIES OF A RESOURCE

From the previous section's one-shot work quantities, we can recover results reminiscent of traditional thermodynamics and can compare how W_{gain}^ϵ differs from W_{cost}^ϵ as the thermodynamic limit is approached. We denote n copies of $R = (r, H, N)$ by $R^{\otimes n} = (r^{\otimes n}, \sum_{i=1}^n H_i, \sum_{i=1}^n N_i)$. In the asymptotic limit, related to the thermodynamic limit, $n \to \infty$. Also in the limit, we show that $W_{\text{gain}}^\epsilon(R^{\otimes n})$ and $W_{\text{cost}}^\epsilon(R^{\otimes n})$ tend to the difference between the grand potential of R and the grand potential of the associated equilibrium state G_R. As the asymptotic limit is approached, $W_{\text{gain}}^\epsilon(R^{\otimes n})$ and $W_{\text{cost}}^\epsilon(R^{\otimes n})$ differ by terms of order $\sqrt{n}$.

To derive the asymptotic limits of Eq. (18) and inequalities (19), we invoke the Asymptotic Equipartition Property of D_{H}^ϵ [53],

$$\lim_{n \to \infty} \frac{1}{n} D_{\text{H}}^\epsilon(r^{\otimes n}||s^{\otimes n}) = D(r||s) \quad \forall \epsilon \in (0,1), \tag{20}$$

wherein $r = (r_1, r_2, \ldots, r_d)$ and $s = (s_1, s_2, \ldots, s_d)$ denote probability distributions over the same alphabet. We have used the definition

$$D(r||s) := \sum_{i=1}^d r_i \ln\left(\frac{r_i}{s_i}\right) \tag{21}$$

of the relative entropy, defining $0 \ln 0 = 0$ [67].

Applying Eq. (20) to Eq. (18) and to both sides of Inequalities (19) yields

$$\lim_{n \to \infty} \frac{1}{n} W_{\text{gain}}^\epsilon(R^{\otimes n}) = \lim_{n \to \infty} \frac{1}{n} W_{\text{cost}}^\epsilon(R^{\otimes n}) = \frac{1}{\beta} D(r||g_R). \tag{22}$$

In the asymptotic limit, the bounds in Inequalities (19) converge. All strategies of work extraction and state formation, from risky ($\epsilon \approx 1$) to conservative ($\epsilon \approx 0$), become equivalent.

To understand Eq. (22) further, we invoke the definition of the relative entropy:

$$\frac{1}{\beta} D(r||g_R) = \frac{1}{\beta} \sum_i r_i \left(\ln r_i - \ln \frac{e^{-\beta(E_i - \mu n_i)}}{Z} \right) \tag{23}$$

$$= \langle H \rangle_r - T k_{\text{B}} S_{\text{S}}(r) - \mu \langle N \rangle_r + k_{\text{B}} T \ln Z \tag{24}$$

$$= \Phi_{\beta,\mu}(R) - \Phi_{\beta,\mu}(G_R). \tag{25}$$

Recall that $\Phi := E - TS - \mu N$ denotes the grand potential, and $-k_{\text{B}} T \ln Z$ denotes the equilibrium state's free energy, in conventional thermodynamics. Using one-shot information theory, we have recovered the convergence, in the asymptotic limit, of a state's average work cost and average work yield to a difference between free energies.

Equation (22) implies that all resources can be reversibly converted into one another in the asymptotic limit. For any states R and S and for fixed ϵ, there exists an n great enough that $R^{\otimes n} \overset{\beta,\mu}{\mapsto}_\epsilon S^{\otimes m_n}$ for some $m_n \geqslant 1$. To create m_n copies of S from n copies of R, one extracts all the work possible from $R^{\otimes n}$, then constructs $(S^{\otimes m_n})' \approx_\epsilon S^{\otimes m_n}$ from the work. We define the optimal asymptotic conversion rate $\mathsf{R}(R \mapsto S)$ as the asymptotic limit of the supremum of the rates m_n/n achievable by conversion protocols that approximate the desired

output arbitrarily well in the asymptotic limit (protocols for which $\epsilon \rightarrow 0$). This rate is

$$\mathsf{R}(R \mapsto S) = \frac{D(r||g_R)}{D(s||g_S)}. \tag{26}$$

Thus, all nonequilibrium states can be reversibly converted into each other in the asymptotic limit. This result may be surprising. One might have thought that resourcefulness can be "locked" into one form—energy, particle number, or information—preventing an R whose resourcefulness manifests in energy from transforming into an S whose resourcefulness manifests in particle number. Apparently, such locking does not occur. This asymptotic reversible convertibility resembles that in Helmholtz theories [20] and in the nonuniformity theory [15,16]. Asymptotic reversible convertibility in general resource theories is discussed in Ref. [68].

The asymptotic equipartition theorem dictates the leading-order (order-n) behavior of D_{H}^ϵ; a more-refined analysis reveals the next-leading-order terms. Applying techniques from information theory, we show, in Appendix D that the latter terms are of order $\sqrt{n}$:

$$W_{\mathrm{gain}}^\epsilon(R^{\otimes n}) = \frac{1}{\beta}[nD(r||g_R) - O(\sqrt{n})], \quad \text{and} \quad (27)$$

$$W_{\mathrm{cost}}^\epsilon(R^{\otimes n}) = \frac{1}{\beta}[nD(r||g_R) + O(\sqrt{n})]. \tag{28}$$

As one might expect, the work cost $W_{\mathrm{cost}}^\epsilon(R^{\otimes n})$ lies above the asymptotic value, whereas $W_{\mathrm{gain}}^\epsilon(R^{\otimes n})$ lies below. This discrepancy contrasts with conventional thermodynamics, according to which a reversible cycle can extract work from R and use that work to recreate R. Outside the thermodynamic limit, such reversible cycles are impossible. The resource-theory framework refines the second law of thermodynamics, as discussed in Ref. [21].

VII. CONCLUSIONS

We have extended the resource-theory formulation of thermodynamics beyond heat baths. Earlier thermodynamic resource theories model heat exchanges; but many physical systems exchange heat, particles, volume, magnetization, and other observables. We model these exchanges with a set of families of thermodynamic resource theories. Each family corresponds to one free energy, one type of interaction, and one type of bath.

To illustrate mathematical results, we focused on grand-potential theories, whose free operations conserve energy and particle number. We showed, using resource-theory principles, why free states must be grand-canonical ensembles. We characterized the quasiorder on states by an extension of majorization, here termed *equimajorization*. We showed that equimajorization can be formulated in terms of rescaled Lorenz curves and of the optimal error probability in asymmetric hypothesis testing. The hypothesis-testing entropy was shown to be proportional to the amount of work extractable from a state R and to bound the work cost of creating R. In the asymptotic limit as $n \rightarrow \infty$, $W_{\mathrm{gain}}^\epsilon(R^{\otimes n})$ and $W_{\mathrm{cost}}^\epsilon(R^{\otimes n})$ were shown to converge to a difference $\Phi(R) - \Phi(G_R)$ between grand potentials. The convergence rates were shown to differ

on the order of $\sqrt{n}$. In the limit, all states were shown to be reversibly interconvertible.

Opportunities for bringing these resource theories closer to experiments remain. Examples include the finite sizes of heat baths, catalysts (ancillas that facilitate transformations while suffering no or little degradation), limitations on how much of a resource can be exchanged, and the speeds with which transformations can be implemented. In the presence of an infinitely large heat bath and enough work, every resource R can be converted into every other. As the number n of copies R grows large, the size of the required bath scales only superlinearly with n [20]. A finite-sized bath limits the resource hierarchy, possibly spoiling the interconvertibility of all resource states. The effects of the bath's finiteness might be incorporated as in Ref. [69], which concerns the cost of erasing with a small bath (albeit outside the resource-theory framework). Another open question concerns how finite-sized catalysts affect the work cost of resource interconversion [21]. Finally, the optimal efficiencies in the present paper might characterize only quasistatic—infinitely slow—protocols. Realistic protocols proceed at finite rates. Extensions of our results to finite speeds may draw inspiration from [70,71]. As noted in the final reference, finite speeds relate directly to the density of accessible bath levels.

Generalized thermodynamic resource theories open a host of realistic thermodynamic systems to modeling with resource theories. Particular physical platforms call out for modeling: heat-and-particle exchanges, electrochemical batteries, chemical reactions, etc. As thermodynamic potentials other than the Helmholtz free energy F characterize common experiments, generalized thermodynamic resource theories offer opportunities for realizing one-shot statistical mechanics experimentally.

APPENDIX A: DERIVATIONS OF FREE STATES' FORMS

1. Proof of Theorem 1(a)

Below, we prove claims, presented above, about grand-potential theories. We derive the grand-canonical forms of free state vectors; describe the quasiorder; calculate the work $W_{\mathrm{gain}}^\epsilon$ extractable from, and bound the work $W_{\mathrm{cost}}^\epsilon$ required to create,

one copy of a state; and show, via second-order asymptotics, that $W_{\text{gain}}^{\epsilon}$ does not always equal $W_{\text{cost}}^{\epsilon}$.

The state vectors of the free states in grand-potential theories are shown to be grand-canonical ensembles. We first review the derivation, in Ref. [16], of the forms of the free states in the resource theory of nonuniformity, which models closed isolated systems [28]. From the nonuniformity result, we deduce the canonical form of the free states in Helmholtz theories. From the Helmholtz result, we bootstrap to grand-potential theories [see note 1].

Free operations in the nonuniformity theory are called *noisy operations*. Each noisy operation consists of three steps. Any free state u (whose form is to be derived) can be created, any permutation π can be implemented, and any subsystem A can be discarded (marginalized over) [15–17]:

$$r \mapsto \sum_A \pi(r \otimes u). \tag{A1}$$

Resource states are defined here as states that are not free u's (or, in general thermodynamic resource theories, states that are not free G's) that appear explicitly in the definition of free operations. A resource theory is *trivial* if its free operations alone can generate resource states.

As shown in Ref. [16], the free states must be uniform probability distributions, lest the nonuniformity theory be trivial. (Indeed, the quasiorder of resources becomes trivial: From enough copies of any state, free operations can generate any other state.) The free states' form is derived as follows [16]: Suppose that some nonuniform state u_0 is free. By Shannon compressing many copies of u_0 [72], agents can create pure states $(1,0,0,\ldots,0)$ for free. Via noisy operations, agents can create noise for free. Able to generate purity and noise, free operations can generate arbitrary states. Only if all free states are uniform is the nonuniformity theory nontrivial.

General thermodynamic resource theories contain the nonuniformity theory as a special case. In grand-potential theories, for example, free operations can arbitrarily permute levels within each sector $S_{E,N}$ that corresponds to one energy E and one particle number N. Hence, the weight that each free state has on a sector $S_{E,N}$ is distributed uniformly across the levels in $S_{E,N}$. We call this uniformity the *uniform-eigensubspace condition*. The condition is defined in Helmholtz theories as follows.

Definition 4. Let $R = (r,H)$ denote a state in any Helmholtz theory, wherein $r = (r_1,\ldots,r_d)$. R obeys the *uniform-eigensubspace condition* if, for every degenerate eigenvalue E of H, all the r_i associated with E equal each other.

Let us apply the nonuniformity-theory argument to the uniform-eigensubspace condition.

Proposition 2. The free states in each thermodynamic resource theory obey the uniform-eigensubspace condition. If the free states disobeyed the condition, there would exist resources R that equilibrating operations alone could generate: $G \mapsto R$.

Suppose that free states disobeyed the uniform-eigensubspace condition. Each free state's state vector g would have some weight p on each sector S that corresponds to some energy, some particle number, etc. Equilibrating operations could distribute p arbitrarily across the levels in S but could not change the value of p.

The uniform-eigensubspace condition implies the following three lemmas, which complete our derivation of the canonical form of the free states in Helmholtz theories.

Lemma 2. Let H_1 and H_2 denote any Hamiltonians that share an energy gap Δ (and whose spectra are discrete). Let E_1 and $E_1 + \Delta$ denote eigenvalues of H_1, and let E_2 and $E_2 + \Delta$ denote eigenvalues of H_2. Define $G_1 = (g_1,H_1)$ as a Helmholtz-theory state whose weights on E_1 and $E_1 + \Delta$ are $g_1(E_1)$ and $g_1(E_1 + \Delta)$. Define $G_2 = (g_2,H_2)$, $g_2(E_2)$, and $g_2(E_2 + \Delta)$ analogously. If $G_1 + G_2$ satisfies the uniform-eigensubspace condition, the ratio of the weights depends only on the gap:

$$\frac{g_1(E_1 + \Delta)}{g_1(E_1)} = \frac{g_2(E_2 + \Delta)}{g_2(E_2)}. \tag{A2}$$

Proof. The eigenenergy $E_1 + E_2 + \Delta$ of $G_1 + G_2$ has a twofold degeneracy. Since $G_1 + G_2$ satisfies the uniform-eigensubspace condition, the weight of $g_1 \otimes g_2$ on one degenerate level equals the weight on the other:

$$g_1(E_1)g_2(E_2 + \Delta) = g_1(E_1 + \Delta)g_2(E_2). \tag{A3}$$

Since E_1 and E_2 are arbitrary, each ratio of probabilities depends only on Δ; the other details of H_1 and H_2 are irrelevant. ∎

Lemma 3. Let $G = (g,H)$ denote any free Helmholtz-theory state that has weights $g(E)$ and $g(E + \Delta)$ on either side of an energy gap Δ. The ratio of the weights varies exponentially with the gap,

$$\frac{g(E + \Delta)}{g(E)} = e^{-\beta\Delta}, \tag{A4}$$

wherein $\beta \in \mathbb{R}$.

Proof. Consider a state $G = (g,H)$ that has three energies separated by gaps Δ:

$$H = E|1\rangle\langle 1| + (E + \Delta)|2\rangle\langle 2|$$
$$+ (E + 2\Delta)|3\rangle\langle 3| + \sum_{i=4}^{d} E_i|i\rangle\langle i|. \tag{A5}$$

Let us apply Lemma 2 to two copies of G, defining $E_1 = E$ and $E_2 = E + \Delta$. Equation (A2) becomes

$$\frac{g(E + 2\Delta)}{g(E + \Delta)} = \frac{g(E + \Delta)}{g(E)} =: f(\Delta), \tag{A6}$$

which implies

$$[f(\Delta)]^2 = \frac{g(E + 2\Delta)}{g(E + \Delta)}\frac{g(E + \Delta)}{g(E)} = \frac{g(E + 2\Delta)}{g(E)} = f(2\Delta). \tag{A7}$$

This scaling implies that, if f is continuous, it is an exponential,

$$f(\Delta) = e^{-\beta\Delta}, \tag{A8}$$

for some $\beta \in \mathbb{R}$. (The realness of β follows from the probabilities' realness.)

Let G' denote any free state that has the same gap Δ as G. The composition $G + G'$ obeys the uniform-eigensubspace condition. Hence, G' satisfies Eq. (A8), by Lemma 2, even if G' does not have the form in Eq. (A5). ∎

Lemma 4. All the gaps in all the free states in any given Helmholtz theory correspond to the same β.

Proof. Let $G = (g, H)$ and $G' = (g', H')$ denote free states in some Helmholtz theory. Let Δ denote a gap in H; and let $\Delta' = n\Delta$, wherein n denotes a positive integer, denote a gap in H'. We wish to show that g and g' correspond to the same β. The proof will be extended to rational proportionality constants, then to arbitrary constants.

Some free state $G'' = (g'', H'')$ has $p > n$ equally spaced levels $E, E + \Delta, \ldots, E + p\Delta$. For example, G'' might denote a harmonic oscillator. This G'' will serve as a thermometer that interrelates the temperatures of G and G'. Let $E_1 = E$ and $E_2 = E + m\Delta$ for any $m \in \{1, 2, \ldots, p\}$. By an argument like the one used to prove Lemma 3,

$$f(m\Delta) := \frac{g''(E + m\Delta)}{g''(E)} \tag{A9}$$

$$= \left[\frac{g''(E + \Delta)}{g''(E)} \right]^m \tag{A10}$$

$$=: f(\Delta)^m, \tag{A11}$$

wherein $g''(E + k\Delta)$ denotes the weight on level k.

Consider substituting $m = n$ and $\Delta = \frac{1}{n}\Delta'$ into the left-hand side (LHS) of $f(m\Delta) = f(\Delta)^m$:

$$f(\Delta') = f(\Delta)^n. \tag{A12}$$

We have related the ratio of the weights across the gap of G' to the ratio of the weights across the gap of G. That is, we have related the temperature of G' to the temperature of G.

Now, suppose that $\Delta' = \frac{m}{n}\Delta$. Consider substituting $m\Delta = n\Delta'$ into the LHS of $f(m\Delta) = f(\Delta)^m$: $f(n\Delta') = f(\Delta)^m$. This equation's LHS also equals $f(\Delta')^n$, by Eq. (A11). Equating the two expressions for $f(n\Delta')$ yields

$$f(\Delta') = f(\Delta)^{m/n}. \tag{A13}$$

We have related the temperature of G' to the temperature of G, effectively by considering multiple copies of each state.

Finally, suppose that $\Delta' = \alpha\Delta$, wherein α denotes an irrational number. α can be approximated arbitrarily well by a ratio m/n. Arbitrarily many copies of G and G' relate the temperature of G to that of G' via Eq. (A13). ∎

Lemmas 2–4, with the normalization condition, imply that the free states in Helmholtz theories are canonical ensembles. This result will facilitate our proof of Theorem 1 about grand-potential theories.

Theorem 1(a). Consider any grand-potential resource theory in which each pair (H, N) corresponds to exactly one free state $G = (g, H, N)$. If g is not a grand-canonical ensemble, some resources R can be generated solely with equilibrating operations: $G \overset{\beta,\mu}{\mapsto} R$.

Proof. First, we show that each element of g has the form $e^{-\beta(E_i)E_i + \alpha(n_j)n_j}/Z$, wherein $\beta(E_i)$ and $\alpha(n_j)$ denote functions of the energy and particle number. Second, by comparing the grand-potential theory with Helmholtz theories, we show that β and α are constant functions.

Consider the most general state vector associated with H and N. Each element has the form $e^{-\beta(E_i)E_i + \alpha(n_j)n_j + f(E_i, n_j)}/Z$, wherein f denotes some function and Z normalizes the state. Recall that every Helmholtz-theory problem can be decomposed into single-energy lemmas that are equivalent to nonuniformity-theory problems. Likewise, every grand-

potential-theory problem can be decomposed into lemmas that feature just one n_j apiece and that are equivalent to Helmholtz-theory problems. Therefore, the elements of g that correspond to the same n must form a canonical ensemble. (Rather, they would form a canonical ensemble if normalized appropriately.) These g elements could not form a canonical ensemble if f depended on energy nontrivially. Hence, $f(E_i, n_j) = f(n_j)$. By an analogous argument, f cannot depend on n_j. Hence, f is a constant, and each element of g has the form $e^{-\beta(E_i)E_i + \alpha(n_i)n_i}/Z$.

G could feature in a problem in which every number operator is trivial: $N = 0$. Such a problem is equivalent to a problem in a Helmholtz theory. In each Helmholtz theory, all free states share a β that is a constant function of energy. This β must characterize the grand-potential theory. Analogously, all free states in the grand-potential theory share an $\alpha \in \mathbb{R}$. Hence, each element of g has the form $e^{-\beta E_i + \alpha n_j}/Z$. Define $\mu \in \mathbb{R}$ such that $\alpha = -\beta\mu$. ∎

2. Proof sketch of Theorem 1(b)

Here we provide a proof sketch for Theorem 1 (b). For simplicity, we consider a Helmholtz-theory context. We must show that, if the free states do not have the Boltzmann form, the quasiorder on states is trivial: Any state can be created from any other by equilibrating operations. We follow an argument by Janzing *et al.* [18] about the effective temperatures present in many copies of a non-Boltzmann state. These effective temperatures can be used to cool a qubit.

Consider the transformation of $R = (r, H)$ into $S = (s, H)$. It suffices to operate successively on pairs of levels, such as the first and the jth. An equilibrating operation of the following form converts r_j/r_1 into s_j/s_1. Let E_j denote the gap between the two levels. Following [[18], Sec. 3], we suppose that we have access to a free state on three levels, separated by gaps E_j, the probabilities on which are not Boltzmann weighted. As shown by Janzing *et al.*, the product of $n \to \infty$ copies of this free state contains pairs of levels, separated by E_j, characterized by essentially any desired relative probability p_j/p_1. Two levels of the resource-and-free-state composite are degenerate. One degenerate level is the product of the resource's lower level and the free state's upper level; the other degenerate level consists of the reverse. Swapping the degenerate levels is an equilibrating operation. Easy calculation shows that the swap transforms the resource's probabilities into

$$r'_j = r_j + \delta \quad \text{and} \quad r'_1 = r_1 - \delta, \quad \text{wherein}$$
$$\delta = r_1 p_j - r_j p_1. \tag{A14}$$

If $p_j + p_1 = 1$, any relative probability r'_j/r_1 could be reached by appropriate choice of p_j/p_1. Generally, $p_j + p_1$ is very small (exponentially small in n in the Janzing *et al.* example). Hence, the relative weights of the levels in R can be changed by only a tiny amount. Repeating the procedure sufficiently many times, however, yields the desired relative probability s_j/s_1. Unless the free state has Boltzmann weights, therefore, the quasiorder induced by equilibrating operations is trivial.

APPENDIX B: QUASIORDER PROOFS

Let us prove three statements, introduced in Sec. IV, about the quasiorder on states: Theorem 2, Proposition 1, and Lemma 1.

Theorem 2. Let R and S denote states in the grand-potential theory defined by (β,μ). There exists an equilibrating operation that maps R to S if and only if R equimajorizes S:

$$R \overset{\beta,\mu}{\mapsto} S \Longleftrightarrow R \succ_{\beta,\mu} S. \tag{B1}$$

Proof. This proof is adapted from the proof of [[18], Theorem 5], a Helmholtz-theory analog of our Theorem 2.[3] Let $R = (r,H_R,N_R)$ and $S = (s,H_S,N_S)$. By d_R and d_S, we denote the numbers of elements in r and s. By $G_R = (g_R,H_R,N_R)$ and $G_S = (g_S,H_S,N_S)$, we denote the equilibrium states associated with R and S. We begin with the easier part of the proof, showing that the existence of an equilibrating operation implies equimajorization.

Assume that some equilibrating operation maps R to S:

$$R \overset{\beta,\mu}{\mapsto} (\mathcal{E}(R),H_S,N_S) = (S,H_S,N_S). \tag{B2}$$

Let $v = (v_1,\ldots,v_{d_R})$ denote any vector that contains d_R elements. The $d_R \times d_S$ matrix M that implements $\mathcal{E}$ can be defined by

$$Mv = \mathcal{E}(v). \tag{B3}$$

By the definition of $\mathcal{E}$, $Mr = s$. Since equilibrating operations map equilibrium states to equilibrium states, $Mg_R = g_S$. We can see as follows that M is stochastic: If v represents a (normalized) state, $\sum_i v_i = 1$. Equilibrating operations preserve normalization, so

$$1 = \sum_{i=1}^{d_R} [\mathcal{E}(v)]_i = \sum_{i=1}^{d_R} [Mv]_i, \tag{B4}$$

wherein $[w]_i$ denotes the ith element of any vector w. Mapping normalized vectors to normalized vectors, M is stochastic. By Definition 2, $R \succ_{\beta,\mu} S$.

Now, we proceed to the converse claim. Assume that $R \succ_{\beta,\mu} S$. One can prove that some equilibrating operation maps R to S by augmenting three lines in the proof of Theorem 5 by Janzing *et al.* [18]. After outlining the latter proof, we explain how to augment it.

Janzing *et al.* define a particular energy-preserving transformation implemented with a heat bath; consider the limit as the bath's size approaches infinity; and show that, in the limit, the transformation converts the initial state into the equimajorized state. Blending their notation with ours, we denote the initial state by (p,H_p), the final state by $(\tilde{p},H_{\tilde{p}})$, and the associated equilibrium states by (g,H_p) and $(\tilde{g},H_{\tilde{p}})$. The Hamiltonian H_p has l levels, and $H_{\tilde{p}}$ has $\tilde{l}$ levels.

Janzing *et al.* consider the set $\mathcal{S}_n$ of pure eigenstates of $H_p + H_p^n + H_{\tilde{p}}^n + H_{\tilde{p}}$, wherein H^n denotes n copies of H. Each state in $\mathcal{S}_n$ is characterized by two length-$(n+1)$ strings. Each letter in the first (second) string is a number between 1 and

l ($\tilde{l}$) that indicates on which energy level of H_p ($H_{\tilde{p}}$) the state's weight lies. Denote by $u_i \in [0,n+1]$ the number of times that $i \in [1,l]$ appears in the first string and by $v_j \in [0,n+1]$ the number of times that $j \in [1,\tilde{l}]$ appears in the second string. If two states in $\mathcal{S}_n$ correspond to the same pair,

$$u = (u_1,u_2,\ldots,u_l) \quad \text{and} \quad v = (v_1,v_2,\ldots,v_{\tilde{l}}), \tag{B5}$$

the states correspond to the same energy. (u and v are called r and s in Ref. [18].)

A permutation $\pi_n : \mathcal{S}_n \to \mathcal{S}_n$ is defined in terms of the matrix assumed to map p to $\tilde{p}$. Because π_n maps each input to an output that has the same (u,v), π_n conserves energy. π_n is applied to the probability distribution P_n defined by $p \otimes g^{\otimes n} \otimes \tilde{g}^{\otimes n} \otimes \tilde{g}$. A set $\mathcal{T}_n$ of typical (u,v) tuples is defined in terms of P_n and the limit $n \to \infty$. In this limit, Janzing *et al.* show, π_n maps P_n to the distribution defined by $g \otimes g^{\otimes n} \otimes \tilde{g}^{\otimes n} \otimes \tilde{p}$.

To adapt this Helmholtz proof to grand-potential theories, replace p, $\tilde{p}$, g, and $\tilde{g}$ with r, s, g_R, and g_S. If two states in $\mathcal{S}_n$ correspond to the same (u,v), they correspond not only to the same energy, but also to the same particle number. Just as π_n conserves energy, it conserves particle number. The rest of the proof in Ref. [18] shows that, from the equimajorization condition, an equilibrating operation can be constructed. ∎

The proof technique used above extends from grand-potential theories to thermodynamic resource theories in which extensive-variable operators other than H and N commute with each other [28]. Before proceeding to Proposition 1, we establish Lemma 1, for convenience.

Lemma 1. The inverse of $\epsilon \mapsto b_\epsilon(r||g_R)$ is the piecewise linear function that connects the points $(L_R(t_k),1-t_k)$, wherein t_k and $L_R(t_k)$ define the rescaled Lorenz curve for R:

$$(L_R(t_k),1-t_k) = (b_\epsilon(r||g_R),\epsilon). \tag{B6}$$

Proof. Let π denote a permutation such that the sequence $(r_{\pi(k)}/g_{\pi(k)})_k$ is nonincreasing. Let $R_m := \sum_{k=1}^m r_{\pi(k)}$ and $G_m := \sum_{k=1}^m g_{\pi(k)}$ for all $m \in \{1,2,\ldots,d_R\}$, wherein d_R denotes the number of elements in r. For $m = 0$, define $R_0 = G_0 = 0$. The points that define the rescaled Lorenz curve are (G_m,R_m) for $m \in \{0,1,\ldots,d_R\}$. To prove the claim, we first show that $(G_m,1-R_m)$ equals the $(b_\epsilon(r||g_R),\epsilon)$ associated with an optimal hypothesis test for each m. Then, we show that optimal tests interpolate linearly between the points.

We begin with $m = 0$. The optimal test for $\epsilon = 1$ is $Q = 0$; thus, $b_1 = 0$. Hence, $(b_1,1) = (0,1) = (G_0,1-R_0)$. Now, consider the hypothesis test for an $\epsilon_m = 1 - R_m$ for which $m \neq 0$. Define Q_m as a $d_R \times d_R$ matrix that projects onto the m values of k for which $r_{\pi(k)}/g_{\pi(k)}$ is greatest. Operation by Q_m on a vector v preserves the part of the support of v that lies on these m values of k and maps all other elements of v to zero. As $\sum_i [Q_m r]_i = R_m$, Q_m is a feasible measurement element in the primal definition of b_{ϵ_m} [Eq. (9)].[4] Therefore,

$$b_{\epsilon_m}(r||g_R) \leqslant \sum_i [Q_m g_R]_i = G_m. \tag{B7}$$

---

[3]An alternative approach to the Helmholtz-theory analog appears in Ref. [19].

[4]Q is said to be *feasible* if the measurement $\{Q,\mathbb{1}-Q\}$ corresponds to a type I error probability of at most ϵ. The feasible measurement that minimizes the type II error probability is *optimal*.

To show that equality holds, we consider the dual problem in Eq. (10). A feasible pair[5] (μ_m, τ_m) that satisfies the constraint $\mu_m r - g_R \leqslant \tau_m$ is given by defining μ_m such that $r_{\pi(m+1)}/g_{\pi(m+1)} \leqslant 1/\mu_m < r_{\pi(m)}/g_{\pi(m)}$ and $\tau_m = \sum_{k=1}^{m} (\mu_m r_{\pi(k)} - g_{\pi(k)}) e_k e_k^T$, wherein e_k denotes the unit vector that has exactly one nonzero element, which corresponds to the $[\pi(k)]^{\text{th}}$ energy-and-particle-number level, and the superscript T denotes the transpose. Evaluating Eq. (10) shows that the two contributions dependent on μ_m cancel, by $1 - \epsilon_m = R_m$ and by the definition of R_m. Hence, $b_{\epsilon_m}(r||g) \geqslant G_m$. Combining this result with Inequality (B7), shows that

$$b_{\epsilon_m}(r||g_R) = G_m. \tag{B8}$$

Now, consider a type I error for which $\epsilon_m \geqslant \epsilon \geqslant \epsilon_{m+1}$. Set $\lambda \in (0,1)$ such that

$$1 - \epsilon = (1 - \lambda)(1 - \epsilon_m) + \lambda(1 - \epsilon_{m+1}). \tag{B9}$$

Note that $(1-\epsilon) = (1-\epsilon_m) + \lambda r_{\pi(m+1)}$. Since $\epsilon \mapsto b_\epsilon(r||g_R)$ is convex,

$$b_\epsilon(r||g_R) \leqslant (1 - \lambda)b_{\epsilon_m}(r||g_R) + \lambda b_{\epsilon_{m+1}}(r||g_R). \tag{B10}$$

Let us show that Inequality (B10) holds if the inequality is reversed. In the dual problem, if $\mu = g_{\pi(m+1)}/r_{\pi(m+1)}$ and $\tau = \sum_{k=1}^{m} (\mu r_{\pi(k)} - g_{\pi(k)}) e_k e_k^T$, then

$$b_\epsilon(r||g_R) \geqslant \mu[(1 - \epsilon_m) + \lambda r_{\pi(m+1)}] - \mu R_m + G_m \tag{B11}$$

$$= \lambda g_{\pi(m+1)} + G_m \tag{B12}$$

$$= (1 - \lambda)G_m + \lambda G_{m+1} \tag{B13}$$

$$= (1 - \lambda)b_{\epsilon_m}(r||g_R) + \lambda b_{\epsilon_{m+1}}(r||g_R). \tag{B14}$$

The final equality follows from Eq. (B8). Inequalities (B10) and (B14) show that interpolating linearly between $(G_m, 1 - R_m)$ and $(G_{m+1}, 1 - R_{m+1})$ amounts to interpolating linearly between $(b_{\epsilon_m}(r||g_R), \epsilon_m)$ and $(b_{\epsilon_{m+1}}(r||g_R), \epsilon_{m+1})$. ∎

Finally, we give a mostly self-contained proof of Proposition 1.

Proposition 1. For any states R and S in the grand-potential resource theory defined by (β, μ), the following are equivalent.

(a) $R \succ_{\beta,\mu} S$.
(b) $L_R(t) \geqslant L_S(t)$ for all $t \in [0,1]$.
(c) $\phi_{f_a}(R) \geqslant \phi_{f_a}(S)$ for every function $f_a(t) = \max\{0, t - a\}$ and every $a \in \mathbb{R}$.
(d) $\phi_f(R) \geqslant \phi_f(S)$ for all continuous convex functions f.

Proof. We show that $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$. R, S, G_R, and G_S are defined as in the proof of Theorem 2.

$(a) \Rightarrow (b)$ Let M denote the stochastic matrix from the equimajorization condition. Let Q define the optimal test that distinguishes s from g_S with a type I error probability of at most ϵ. To distinguish r from g_R, one can apply M and then measure $\{Q, \mathbb{1} - Q\}$. This test might distinguish between r and g_R suboptimally. Hence, $b_\epsilon(r||g_R) \leqslant b_\epsilon(s||g_S)$. Lemma 1 implies (b).

$(b) \Rightarrow (c)$ The dual formulation of $\epsilon \mapsto b_\epsilon(r||g_R)$ can be written as

$$b_\epsilon(r||g_R) = \max_\mu \left\{ (1 - \epsilon)\mu - \sum_i [\{\mu r - g_R\}_+]_i \right\}. \tag{B15}$$

Hence, $b_{1-\epsilon}(r||g_R)$ is the Legendre transform of

$$\sum_i [\{\mu r - g_R\}_+]_i = \mu \sum_i g_i f_{1/\mu}\left(\frac{r_i}{g_i}\right) = \mu\, \phi_{f_{1/\mu}}(r, g_R), \tag{B16}$$

wherein $f_a(t) = \max\{0, t - a\}$. Since $b_\epsilon(r||g_R) \leqslant b_\epsilon(s||g_S)$, $\phi_{f_{1/\mu}}(r, g_R) \geqslant \phi_{f_{1/\mu}}(s, g_S)$.

$(c) \Rightarrow (d)$ In Ref. [45] (see also [[46], Lemma 1.2.5]), Uhlmann shows that every continuous convex function $f(x)$ can be approximated to arbitrary accuracy by a linear combination, that has positive coefficients, of functions $f_a(x)$. [In Uhlmann's phrasing, a concave $f(x)$ is approximated by positive linear combinations of $-f_a(x)$.] Because $\phi_{f_a}(R) \geqslant \phi_{f_a}(S)$ for all f_a, $\phi_f(R) \geqslant \phi_f(S)$.

$(d) \Rightarrow (a)$ Following [[46], Theorem 1.4.4], we prove the contrapositive. Assume that no stochastic matrix M satisfies $Mr = s$ and $Mg_R = g_S$. Since the set of stochastic matrices is convex and compact (all entries being in the unit interval), the set of vectors $Mr \oplus Mg_R$ is convex and compact. As this set does not contain $s \oplus g_S$, a hyperplane separates $s \oplus g_S$ from $\{Mr \oplus Mg_R\}$ [[73], Theorem 3.5]. That is, a vector $x \oplus y$ satisfies $xs + yg_S > xMr + yMg_R$ for all stochastic matrices M. Taking the maximum over M on the right-hand side (RHS) and denoting by $[g_R]_k$ the kth element of g_R gives

$$xs + yg_S > \max_M \{xMr + yMg_R\} \tag{B17}$$

$$= \max_M \sum_{jk} (x_j M_{jk} r_k + y_j M_{jk}[g_R]_k) \tag{B18}$$

$$= \sum_k \max_j \{x_j r_k + y_j [g_R]_k\}. \tag{B19}$$

Maximizing over the LHS produces

$$\sum_k \max_j \{x_j s_k + y_j [g_S]_k\} > \sum_k \max_j \{x_j r_k + y_j [g_R]_k\}. \tag{B20}$$

However, $f(s,t) = \max_j \{x_j s + y_j t\}$ is a convex function, so $\phi_f(S) \geqslant \phi_f(R)$. The contrapositive implies (a). ∎

APPENDIX C: ONE-SHOT WORK-YIELD AND WORK-COST PROOFS

Equilibrating operations can extract work from one copy of a state $R = (r, H, N)$ and can store the work in a battery. From enough stored work, equilibrating operations can generate R. Each protocol can have a probability $\epsilon \in [0,1]$ of failing to accomplish its purpose. We calculate the maximum work $W_{\text{gain}}^\epsilon(R)$ extractable from—and bound the least work W_{cost}^ϵ required to create—R with error-prone protocols. First we prove a helpful lemma. Anticipating applications of the lemma, we use the notation G and G' associated with free states.

---

[5] (μ, τ) is said to be *feasible* if it satisfies the constraints in Eq. (10).

However, the lemma holds if G denotes an arbitrary quantum state and G' denotes an arbitrary quasiclassical state.

We use the following notation. By e_E, we denote a vector of the eigenvalues of the pure state $|E\rangle\langle E|$. The element associated with energy E is one, and the other elements are zeros. By $[v]_i$, we denote the ith element of any vector v. If r and s denote equal-sized vectors, their scalar product is $r \cdot s := \sum_i r_i s_i$.

Lemma 5. Let $R = (r, H, N)$ denote any state, and let (e_E, H', N') denote any pure state. Let $G = (g, H, N)$ and $G' = (g', H', N')$. The optimal hypothesis test between r and g is related to the optimal test between $r \otimes e_E$ and $g \otimes g'$ by

$$b_\epsilon(r \otimes e_E || g \otimes g') = (g' \cdot e_E) b_\epsilon(r||g). \quad \text{(C1)}$$

Proof. Consider any feasible measurement operator Q in $b_\epsilon(r \otimes e_E || g \otimes g')$. We can find a feasible Q' that gives the same value of the objective function and that has the form $Q' = \sum_i Q_{E_i'} \otimes e_{E_i'} e_{E_i'}^T$. Here, $Q_{E_i'} = (\mathbb{1} \otimes e_{E_i'}) Q (\mathbb{1} \otimes e_{E_i'}^T)$, and the superscript T denotes the transpose. Without loss of generality, we focus on Q' operators that have this form.

The constraint becomes $\sum_i [Q_E r]_i \geqslant 1 - \epsilon$, and the objective function becomes

$$\sum_i [Q'(g \otimes g')]_i = \sum_i \left\{ (g' \cdot e_{E_i'}) \sum_j [Q_{E_i'} g]_j \right\}. \quad \text{(C2)}$$

The minimum follows from setting $Q_{E_i'} = 0$ for all $E_i' \neq E$:

$$b_\epsilon(r \otimes e_E || g \otimes g') = \min \left\{ (g' \cdot e_E) \sum_j [Q_E g]_j \left| \sum_j [Q_E r]_j \right. \right.$$

$$\left. \geqslant 1 - \epsilon, 0 \leqslant Q_E \leqslant \mathbb{1} \right\} \quad \text{(C3)}$$

$$= (g' \cdot e_E) b_\epsilon(r||g). \quad \text{(C4)}$$

∎

This lemma implies another lemma, associated with $\epsilon = 0$, that will facilitate our work-bound proofs.

Lemma 6. Let R denote any state in a grand-potential resource theory defined by β and μ. Let W denote the work extractable from R, and let W' denote the work cost of creating R, with error tolerance $\epsilon = 0$. There exist batteries B such that

$$R + B_E \succ_{\beta,\mu} B_{E+W} \Leftrightarrow R \succ_{\beta,\mu} B_W \quad \text{and} \quad \text{(C5)}$$

$$B_{E+W'} \succ_{\beta,\mu} R + B_E \Leftrightarrow B_{W'} \succ_{\beta,\mu} R. \quad \text{(C6)}$$

Proof. Consider a battery B_{E+W} that consists of two noninteracting parts (e.g., two batteries, B_E and B_W). The total Hamiltonian is the sum of the subsystems' Hamiltonians, and the total number operator is a sum. Suppose that the first Hamiltonian has d_1 eigenvalues and the second has d_2. The joint-system state vector $e_E \otimes e_W$ is an energy-$(E + W)$ eigenstate. The joint system's equilibrium state is the composition of the constituent systems' equilibrium states, whose state vectors we denote by g and g'. That is, $B_{E+W} = B_E + B_W$.

Applying several results, we can prove the equivalences in Eq. (C5):

$$R + B_E \succ_{\beta,\mu} B_W + B_E$$

$$\Leftrightarrow L_{R+B_E}(t) \geqslant L_{B_W+B_E}(t) \quad \forall\, t \in [0,1] \quad \text{(C7)}$$

$$\Leftrightarrow b_\epsilon(r \otimes e_E || g \otimes g') \geqslant b_\epsilon(e_W \otimes e_E || g \otimes g') \quad \text{(C8)}$$

$$\Leftrightarrow b_\epsilon(r||g) \geqslant b_\epsilon(e_W||g) \quad \text{(C9)}$$

$$\Leftrightarrow L_R(t) \geqslant L_{B_W}(t) \quad \forall\, t \in [0,1] \quad \text{(C10)}$$

$$\Leftrightarrow R \succ_{\beta,\mu} B_W. \quad \text{(C11)}$$

The first equivalence follows from Proposition 1; the second, from Lemma 1; the third, from Lemma 5; the fourth, from Lemma 1; and the fifth, from Proposition 1. Similar reasoning underpins the equivalence of Eqs. (C6). ∎

Having simplified the model of work, we calculate W_{gain}^ϵ.

Theorem 3. The ϵ-work value of a state $R = (r, \tilde{H}, N)$ associated with the free state $G_R = (g_R, H, N)$ is

$$W_{\text{gain}}^\epsilon(R) = \frac{1}{\beta} D_{\text{H}}^\epsilon(r||g_R). \quad \text{(C12)}$$

Proof. In the converse part of the proof, we show that the RHS is an upper bound on the extractable work. In the direct part, we construct an equilibrating operation that attains the bound.

For the converse, define an equilibrating operation by $\mathcal{E}(r \otimes e_E) \approx_\epsilon e_E \otimes e_W$. The channel's output is ϵ close, in the l_1 norm, to the desired state:

$$\tfrac{1}{2}|\mathcal{E}(r \otimes e_E) - e_E \otimes e_W|_1 \leqslant \epsilon. \quad \text{(C13)}$$

Since equilibrating operations map equilibrium states to equilibrium states, $\mathcal{E}(g_R \otimes g') = g'$. Using $\mathcal{E}$, we can construct a hypothesis test between $r \otimes e_E$ and $g_R \otimes g'$. The test consists of an application of $\mathcal{E}$ followed by an energy measurement.[6] If the measurement yields $E + W$, we guess that the state is $r \otimes e_E$. Otherwise, we guess $g_R \otimes g'$. By construction, the probability that we correctly guess $r \otimes e_E$ is at least $1 - \epsilon$. The test is feasible for $D_{\text{H}}^\epsilon(r \otimes e_E || g_R \otimes g')$, and

$$e^{-D_{\text{H}}^\epsilon(r \otimes e_E || g_R \otimes g')} \leqslant \mathcal{E}(g_R \otimes g') e_{E+W} \quad \text{(C14)}$$

$$= (g \otimes g') e_{E+W} \quad \text{(C15)}$$

$$= \frac{e^{-\beta(E+W)}}{Z}. \quad \text{(C16)}$$

By Eq. (C1),

$$e^{-D_{\text{H}}^\epsilon(r||g_R)} \leqslant e^{-\beta W}, \quad \text{(C17)}$$

which is equivalent to the upper bound.

For the proof's direct part, we define the state vector

$$\tilde{g}' := \frac{1}{1 - \frac{1}{Z} e^{-\beta(E+W)}} \left[g' - \frac{1}{Z} e^{-\beta(E+W)} e_{E+W} \right]. \quad \text{(C18)}$$

---

[6]The proof does not depend on how or whether measurements are defined in the resource theory. Because the proof is not a protocol for extracting work, the resource-theory agent need not perform the measurement.

Using the optimal measurement Q in $D_{\mathrm{H}}^{\epsilon}(r \otimes e_E || g_R \otimes g')$, we define the operation $\mathcal{E}$ by

$$\mathcal{E}(s \otimes u) = \left\{1 - \sum_i [Q(s \otimes u)]_i\right\} \tilde{g}' \\ + \left\{\sum_i [Q(s \otimes u)]_i\right\} e_{E+W} \qquad (C19)$$

for state vectors s and u. By construction, $\mathcal{E}(r \otimes e_E)$ has the form of the desired output. Since $\mathcal{E}$ is an equilibrating operation, $\mathcal{E}(g_R \otimes g') = g'$. This condition determines the possible values of W and is equivalent to

$$\sum_i [Q(g_R \otimes g')]_i = \frac{e^{-\beta(E+W)}}{Z}. \qquad (C20)$$

This equation is equivalent to Inequality (C16), except for containing an equality. Therefore, free operations can distill at least the work W that satisfies

$$e^{-D_{\mathrm{H}}^{\epsilon}(r||g_R)} = e^{-\beta W}, \qquad (C21)$$

which is equivalent to the lower bound. ∎

Having calculated the work extractable from R, we bound the work cost of creating R [Inequalities (19)].

Theorem 3. The work cost of creating an ϵ approximation to a state R is bounded by

$$\max_{\delta \in (0, 1-\epsilon]} \left[\frac{1}{\beta} D_{\mathrm{H}}^{1-\epsilon-\delta}(r||g_R) - \frac{1}{\beta}\log\left(\frac{1}{\delta}\right)\right] \\ \leqslant W_{\mathrm{cost}}^{\epsilon}(R) \leqslant \frac{1}{\beta} D_{\mathrm{H}}^{1-\epsilon}(r||g_R) - \frac{1}{\beta}\log\left(\frac{1-\epsilon}{\epsilon}\right). \quad (C22)$$

Proof. To derive the lower bound, we suppose that $\mathcal{E}$ is an equilibrating operation that satisfies $\mathcal{E}(e_{E+W}) \approx_{\epsilon} r \otimes e_E$. Using $\mathcal{E}$, we can transform the optimal dual program for e_{E+W} and g' into a feasible dual program for $\mathcal{E}(e_{E+W})$ and $g_R \otimes g'$. This feasible program can be related to the hypothesis-testing entropy of r relative to g_R.

Consider distinguishing between e_{E+W} and g' by hypothesis test. Let 1 denote the state on a one-dimensional space. By Lemma 5,

$$e^{-D_{\mathrm{H}}^{1-\epsilon}(e_{E+W}||g')} = (g' \cdot e_{E+W}) e^{-D_{\mathrm{H}}^{1-\epsilon}(1||1)} \qquad (C23)$$
$$= \epsilon(g' \cdot e_{E+W}). \qquad (C24)$$

The dual formulation of D_H^{ϵ} reads

$$e^{-D_{\mathrm{H}}^{1-\epsilon}(e_{E+W}||g')} = \max_{\substack{\mu \, e_{E+W} - g' \leqslant \tau \\ \mu, \tau \geqslant 0}} \left\{\epsilon\mu - \sum_i \tau_i\right\}, \qquad (C25)$$

wherein τ_i denotes the ith element of τ. Comparing Eqs. (C24) and (C25) shows that $\mu = g' e_{E+W}$ and $\tau = 0$ are the optimal choices in the dual formulation.

Acting on each side of the constraint, $\mu \, e_{E+W} \leqslant g'$, with $\mathcal{E}$ yields $\mu \mathcal{E}(e_{E+W}) \leqslant g_R \otimes g'$. Therefore, $\mu = g' \cdot e_{E+W}$ and $\tau = 0$ are feasible for $D_{\mathrm{H}}^{1-\delta}[\mathcal{E}(e_{E+W})||g_R \otimes g']$,

$$e^{-D_{\mathrm{H}}^{1-\delta}[\mathcal{E}(e_{E+W})||g_R \otimes g']} \geqslant \delta\frac{e^{-\beta(E+W)}}{Z}, \qquad (C26)$$

for all $\delta \in [0,1]$. Since $\frac{1}{2}\|r \otimes e_E - \mathcal{E}(e_{E+W})\|_1 \leqslant \epsilon$,

$$\left|\sum_i [Q\{r \otimes e_E - \mathcal{E}(e_{E+W})\}]_i\right| \leqslant \epsilon \qquad (C27)$$

for every Q. Suppose Q is the optimal choice in $D_{\mathrm{H}}^{1-\eta}(r \otimes e_E || g_R \otimes g')$, such that $\sum_i [Q(r \otimes e_E)]_i = \eta$. For this Q, $\sum_i [Q \, \mathcal{E}(e_{E+W})]_i \geqslant \eta - \epsilon$. Therefore,

$$e^{-D_{\mathrm{H}}^{1-\eta+\epsilon}[\mathcal{E}(e_{E+W})||g_R \otimes g']} \leqslant \sum_i [Q(g_R \otimes g')]_i \qquad (C28)$$

$$= e^{-D_{\mathrm{H}}^{1-\eta}(r \otimes e_E || g_R \otimes g')} \qquad (C29)$$

$$= \frac{e^{-\beta E}}{Z} e^{-D_{\mathrm{H}}^{1-\eta}(r||g_R)}. \qquad (C30)$$

If $\eta = \epsilon + \delta$,

$$\delta e^{-\beta W} \leqslant e^{-D_{\mathrm{H}}^{1-\epsilon-\delta}(r||g_R)}. \qquad (C31)$$

We have lower bounded $W_{\mathrm{cost}}^{\epsilon}(R)$ for every $\delta \in (0, 1 - \epsilon]$. The tightest of these bounds follows from a maximization over δ.

To derive upper bound, we construct an equilibrating operation that maps e_{E+W} to $\tilde{r} \otimes e_E$, wherein $\tilde{r} \approx_{\epsilon} r$, for a suitably chosen value of W. The work system is approximated, whereas the battery is not. The associated work-cost bound may be suboptimal.

By Condition (c) of Proposition 1, such an equilibrating operation exists if and only if

$$K_{\mathrm{in}}(a) \geqslant K_{\mathrm{out}}(a) \qquad (C32)$$

for all $a \in \mathbb{R}$ and for $K_R(a)$ defined as follows. In Proposition 1, the function $f_a(t) := \max\{0, t - a\}$ appears in

$$\phi_{f_a}(R) := \sum_{i=1}^{d_R} g_i f_a\left(\frac{r_i}{g_i}\right) \qquad (C33)$$

$$= \sum_{i=1}^{d_R} g_i \max\left\{0, \frac{r_i}{g_i} - a\right\} \qquad (C34)$$

$$= \sum_{i=1}^{d_R} \max\{0, r_i - g_i a\}. \qquad (C35)$$

To simplify notation, we relabel this sum as $K_R(a)$. Because r and g are normalized, $K_R(a) = (1 - a)$ if $a \leqslant 0$. We can rewrite the LHS of Inequality (C32) as

$$K_{\mathrm{in}}(a) = \left(1 - a\frac{e^{-\beta(E+W)}}{Z}\right)_+. \qquad (C36)$$

This function is linear and satisfies $K_{\mathrm{in}}(0) = 1$ and $K_{\mathrm{in}}(Z e^{\beta(E+W)}) = 0$. We can rewrite the RHS of Inequality (C32) as

$$K_{\mathrm{out}}(a) = \sum_i \left(\tilde{r}_i - g_i a\frac{e^{-\beta E}}{Z}\right)_+. \qquad (C37)$$

Just as for the input state, $K_{\mathrm{out}}(0) = 1$. As a sum of convex functions, $K_{\mathrm{out}}(a)$ is convex. Thus, the condition $K_{\mathrm{in}}(a) \geqslant K_{\mathrm{out}}(a)$ for all $t \in \mathbb{R}$ reduces to

$$K_{\mathrm{out}}(Z e^{\beta(E+W)}) = 0. \qquad (C38)$$

Let us find a value of W for which the transformation is possible. First, we construct a suitable $\tilde{r}$ from the dual form of $D_{\mathrm{H}}^{1-\epsilon}(r||g_R)$. Suppose that μ and τ are the optimal choices, so that

$$e^{-D_{\mathrm{H}}^{1-\epsilon}(r||g_R)} = \epsilon\mu - \sum_i \tau_i \qquad (\text{C39})$$

and $\mu r - g_R \leqslant \tau$. We define $r' = TrT^{\dagger}$ for $T = g_R^{\frac{1}{2}}(g_R + \tau)^{-\frac{1}{2}}$, using the pseudoinverse (the inverse on the support). These definitions satisfy $\mu r' \leqslant g_R$. Let us bound the trace distance

$$\frac{1}{2}\|r - r'\|_1 = \frac{1}{2}\sum_k |r_k - r'_k| \qquad (\text{C40})$$

$$= \frac{1}{2}\sum_k |r_k - g_k r_k(g_k + \tau_k)^{-1}| \qquad (\text{C41})$$

$$= \frac{1}{2}\sum_k r_k |1 - g_k(g_k + \tau_k)^{-1}| \qquad (\text{C42})$$

$$= \frac{1}{2}\sum_k r_k \frac{\tau_k}{g_k + \tau_k} \qquad (\text{C43})$$

$$\leqslant \frac{1}{2}\sum_k \frac{1}{\mu}\tau_k \qquad (\text{C44})$$

$$= \frac{\sum_i \tau_i}{2\mu} \qquad (\text{C45})$$

$$\leqslant \frac{\epsilon}{2}. \qquad (\text{C46})$$

To derive the first inequality, we used the inequality $\mu r \leqslant g_R + \tau$; to derive the second, $\epsilon\mu - \sum_i \tau_i = e^{-D_{\mathrm{H}}^{1-\epsilon}(r||g_R)} \geqslant 0$. Let $\tilde{r} = r'/\sum_i r'_i$. By the triangle inequality,

$$\frac{1}{2}\|r - \tilde{r}\|_1 \leqslant \frac{1}{2}\|r - r'\|_1 + \frac{1}{2}\|\tilde{r} - r'\|_1 \qquad (\text{C47})$$

$$\leqslant \frac{\epsilon}{2} + \frac{1}{2}\left(\frac{1}{\sum_i r'_i} - 1\right)\|r'\|_1 \qquad (\text{C48})$$

$$= \frac{\epsilon}{2} + \frac{1}{2}\left(1 - \sum_i r'_i\right) \qquad (\text{C49})$$

$$\leqslant \epsilon. \qquad (\text{C50})$$

The fourth inequality follows from $\sum_i r_i - \sum_j r'_j \leqslant \|r - r'\|_1$. We have constructed a state $\tilde{r} \approx_{\epsilon} r$.

Moreover, $\mu(\sum_i r'_i)\tilde{r} \leqslant g$. Applying this inequality to (C37) yields

$$K_{\mathrm{out}}(a) \leqslant \sum_i g_i \left(\frac{1}{\mu \sum_i r'_i} - a\frac{e^{-\beta E}}{Z}\right)_+ \qquad (\text{C51})$$

$$= \left(\frac{1}{\mu \sum_i r'_i} - a\frac{e^{-\beta E}}{Z}\right)_+. \qquad (\text{C52})$$

Hence, $K_{\mathrm{out}}(Ze^{\beta(E+W)})$ satisfies

$$K_{\mathrm{out}}(Ze^{\beta(E+W)}) \leqslant \left(\frac{1}{\mu \sum_i r'_i} - e^{\beta W}\right)_+ \qquad (\text{C53})$$

$$\leqslant \left(\frac{1}{\mu(1 - \epsilon)} - e^{\beta W}\right)_+ \qquad (\text{C54})$$

$$\leqslant \left(\frac{\epsilon}{1 - \epsilon}e^{D_{\mathrm{H}}^{1-\epsilon}(r||g)} - e^{\beta W}\right)_+. \qquad (\text{C55})$$

The second inequality follows from $\sum_i r'_i \geqslant 1 - \epsilon$; the third follows from $\epsilon\mu \geqslant \epsilon\mu - \sum_i \tau_i = e^{-D_{\mathrm{H}}^{1-\epsilon}(r||g_R)}$. We can satisfy (C38) by choosing W such that

$$e^{\beta W} = \frac{\epsilon}{1 - \epsilon}e^{D_{\mathrm{H}}^{1-\epsilon}(r||g_R)}. \qquad (\text{C56})$$

Since $W_{\mathrm{cost}}^{\epsilon}(R) \leqslant W$, the upper bound follows directly. ∎

One can show that the upper bound is a nonnegative quantity, using $e^{-D_{\mathrm{H}}^{1-\epsilon}(r||g_R)} \leqslant \epsilon$. (This inequality follows from the choice $Q = \epsilon\mathbb{1}$.) Because $\epsilon \leqslant \frac{\epsilon}{1-\epsilon}$, $e^{-D_{\mathrm{H}}^{1-\epsilon}(r||g_R)} \leqslant \frac{\epsilon}{1-\epsilon}$. The latter implies the upper bound's nonnegativity. The lower bound is nonnegative in all the numerical examples we tested.

APPENDIX D: COMPARISON OF ONE-SHOT WORK YIELD AND WORK COST

We use *second-order asymptotics* to show that $W_{\mathrm{gain}}^{\epsilon}(R^{\otimes n})$ tends to differ from the bounds on $W_{\mathrm{cost}}^{\epsilon}(R^{\otimes n})$ and that the bounds lie arbitrarily close together, as the asymptotic limit is approached. Consider distilling work from, or creating, n copies $R^{\otimes n}$ of $R = (r, H, N)$. The work involved depends on the *normal approximation* to the hypothesis-testing relative entropy [[74], Theorem 5] (see also [75,76]),[7]

$$D_{\mathrm{H}}^{\epsilon}(r^{\otimes n}||g_R^{\otimes n}) = nD(r||g_R)$$
$$+ \sqrt{n}\, s(r||g_R)\, \Phi^{-1}(\epsilon) + O(\log n), \qquad (\text{D1})$$

wherein g_R denotes the state vector of the equilibrium state associated with R, the square-root of the relative-entropy variance is

$$s(r||g_R) := \sqrt{V(r||g_R)}$$
$$= \sqrt{\mathrm{Tr}(r[\log r - \log g_R]^2) - D(r||g_R)^2}, \qquad (\text{D2})$$

and the inverse error function is

$$\Phi^{-1}(\epsilon) := \sup\left\{z \in \mathbb{R} \;\middle|\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{1}{2}t^2}dt \leqslant \epsilon\right\}. \qquad (\text{D3})$$

Equation (D3) admits of the following interpretation. Suppose that, if a hypothesis test is performed on the null-hypothesis state R, the outcome is distributed normally. The probability that a type I error occurs equals $\Phi^{-1}(\epsilon)$. Let us apply Eqs. (D1) to (18) and to inequalities (19).

To characterize the latter expressions' approach toward $D(r||g_R)$, we evaluate the normalized differences between each $W^{\epsilon,\beta}(R^{\otimes n})$ and $\frac{1}{\beta}D(r||g_R)$, assuming n is large. The actual distillable work differs from the asymptotic distillable

---

[7]The quantum version of Eq. (D1) appears in Refs. [74,77], but we have specialized to commuting density operators.

work as

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\left[n\frac{1}{\beta}D(r||g_R)-W_{\mathrm{gain}}^{\epsilon}(R^{\otimes n})\right]=\lim_{n\to\infty}\frac{1}{\beta}\left[-s(R||g_R)\Phi^{-1}(\epsilon)-\frac{O(\log n)}{\sqrt{n}}\right]\tag{D4}$$

$$=\frac{1}{\beta}s(R||g_R)\Phi^{-1}(1-\epsilon).\tag{D5}$$

The final equation follows from $\Phi^{-1}(\epsilon)=-\Phi^{-1}(1-\epsilon)$. If the type I error probability is small ($\epsilon<\frac{1}{2}$), Eq. (D5) is positive because the work distilled at the optimal asymptotic efficiency exceeds the work distilled at any subasymptotic efficiency [i.e., Eq. (D4) is positive].

The lower work-cost bound differs from the asymptotic cost by

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\left[W_{\mathrm{cost}}^{\epsilon}(R^{\otimes n})v-n\frac{1}{\beta}D(r||g_R)\right]\geqslant\lim_{n\to\infty}\frac{1}{\beta}\max_{\delta\in(0,1-\epsilon]}\left[s(r||g_R)\Phi^{-1}(1-\epsilon-\delta)-\frac{\log\frac{1}{\delta}}{\sqrt{n}}\right]$$

$$=\frac{1}{\beta}\max_{\delta\in(0,1-\epsilon]}s(r||g_R)\Phi^{-1}(1-\epsilon-\delta)\tag{D6}$$

$$=\frac{1}{\beta}s(r||g_R)\Phi^{-1}(1-\epsilon).\tag{D7}$$

The first equality holds if δ grows more slowly than $e^{\sqrt{n}}$. The last equality holds since, by the definition and monotonicity of Φ^{-1}, the least possible δ value maximizes $\Phi^{-1}(1-\epsilon-\delta)$. In the limit, this difference arising from the lower bound matches that of the upper bound, as the work cost's upper bound differs from the asymptotic work cost as

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\left[W_{\mathrm{cost}}^{\epsilon}(R^{\otimes n})-n\frac{1}{\beta}D(r||g_R)\right]\leqslant\lim_{n\to\infty}\frac{1}{\beta}\left[s(r||g_R)\Phi^{-1}(1-\epsilon)-\frac{1}{\sqrt{n}}\log\left(\frac{1-\epsilon}{\epsilon}\right)\right]$$

$$=\frac{1}{\beta}s(r||g_R)\Phi^{-1}(1-\epsilon).\tag{D8}$$

The final equality holds if $\frac{1-\epsilon}{\epsilon}$ grows more slowly than $e^{\sqrt{n}}$.

Let us compare these normalized work differences. If n is large, the work-cost bounds exceed the optimal asymptotic work cost $\frac{1}{\beta}D(r||g_R)$ by an amount proportional to $\sqrt{n}$. In contrast, $\frac{1}{\beta}D(r||g_R)$ exceeds the work gain by an amount proportional to $\sqrt{n}$. The work gain and work cost differ, in general, unlike in the asymptotic limit, as in Refs. [19,21]. As creating $R^{\otimes n}$ requires more work than can be extracted from $R^{\otimes n}$, resources degrade.

---

[1] D. Lacoste, A. W. C. Lau, and K. Mallick, Phys. Rev. E **78**, 011915 (2008).

[2] V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh, Nature (London) **445**, 523 (2007).

[3] L. P. Faucheux, L. S. Bourdieu, P. D. Kaplan, and A. J. Libchaber, Phys. Rev. Lett. **74**, 1504 (1995).

[4] C. Bustamante, J. Liphardt, and F. Ritort, Phys. Today **58**, 43 (2005).

[5] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante, Nature (London) **437**, 231 (2005).

[6] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, Science (Washington, DC, US) **296**, 1832 (2002).

[7] A. Alemany and F. Ritort, Europhys. News **41**, 27 (2010).

[8] J. Cheng, S. Sreelatha, R. Hou, A. Efremov, R. Liu, J. R. C. van der Maarel, and Z. Wang, Phys. Rev. Lett. **109**, 238104 (2012).

[9] B. Coecke, T. Fritz, and R. W. Spekkens, arXiv:1409.5531 (2014).

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

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

[12] I. Marvian and R. W. Spekkens, New J. Phys. **15**, 033001 (2013).

[13] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, New J. Phys. **8**, 58 (2006).

[14] V. Veitch, S. A. Hamed Mousavian, D. Gottesman, and J. Emerson, New J. Phys. **16**, 013009 (2014).

[15] M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, Phys. Rev. Lett. **90**, 100402 (2003).

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

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

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

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

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

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

[22] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B **27**, 1345019 (2013).

[23] P. Faist, J. Oppenheim, and R. Renner, New J. Phys. **17**, 043003 (2015).

[24] R. A. Alberty, Pure Appl. Chem. **73**, 1349 (2001).

[25] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley & Sons, New York, 1985).

[26] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic, New York, 1979).

[27] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer, Berlin, 2010).

[28] N. Yunger Halpern, arXiv:1409.7845 (2014).

[29] M. Lostaglio, D. Jennings, and T. Rudolph, Nat. Commun. **6**, 6383 (2015).

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

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

[32] N. Yunger Halpern, A. J. P. Garner, O. C. O. Dahlsten, and V. Vedral, New J. Phys. **17**, 095003 (2015).

[33] S. Salek and K. Wiesner, arXiv:1504.05111 (2015).

[34] Á. M. Alhambra, J. Oppenheim, and C. Perry, arXiv:1504.00020 (2015).

[35] M. Lostaglio, D. Jennings, and T. Rudolph, arXiv:1511.04420 (2015).

[36] Y. Guryanova, S. Popescu, A. J. Short, R. Silva, and P. Skrzypczyk, arXiv:1512.01190 (2015).

[37] N. Yunger Halpern, P. Faist, J. Oppenheim, and A. Winter, arXiv:1512.01189 (2015).

[38] C. Perry, P. Ćwikliński, J. Anders, M. Horodecki, and J. Oppenheim, arXiv:1511.06553 (2015).

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

[40] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).

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

[42] A. F. Veinott, Manage. Sci. **17**, 547 (1971).

[43] E. Ruch, R. Schranner, and T. H. Seligman, J. Chem. Phys. **69**, 386 (1978).

[44] E. Ruch, Theor. Chim. Acta **19**, 225 (1975).

[45] A. Uhlmann, Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R. **27**, 213 (1978).

[46] P. M. Alberti and A. Uhlmann, *Dissipative Motion in State Spaces* (Teubner, Leipzig, 1981).

[47] I. Csiszr, Magyar. Tud. Akad. Mat. Kutato Int. Kozl (Publication of the Mathematical Institute of the Hungarian Academy of Sciences) **8**, 85 (1963).

[48] T. Morimoto, J. Phys. Soc. Jpn. **18**, 328 (1963).

[49] S. M. Ali and S. D. Silvey, J. R. Stat. Soc. Ser. B (Methodological) **28**, 131 (1966).

[50] F. Liese and I. Vajda, IEEE Trans. Inf. Theory **52**, 4394 (2006).

[51] M. O. Lorenz, Publ. Am. Stat. Assoc. **9**, 209 (1905).

[52] P. Harremoës, in *Proceedings of the 2004 International Symposium on Information Theory and its Applications* (Parma, Italy, 2004), pp. 1422–1425.

[53] F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner, in *XVIIth International Congress on Mathematical Physics*, edited by A. Jensen (World Scientific, Singapore, 2013), pp. 134–153.

[54] R. Renner, Security of Quantum Key Distribution, Ph.D. thesis, ETH Zürich, 2005.

[55] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, IEEE Trans. Inf. Theory **57**, 5524 (2011).

[56] J. M. Renes and R. Renner, IEEE Trans. Inf. Theory **57**, 7377 (2011).

[57] J. M. Renes and R. Renner, IEEE Trans. Inf. Theory **58**, 1985 (2012).

[58] L. Wang and R. Renner, Phys. Rev. Lett. **108**, 200501 (2012).

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

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

[61] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, New J. Phys. **17**, 073001 (2015).

[62] O. C. O. Dahlsten, Entropy **15**, 5346 (2013).

[63] G. E. Crooks, J. Stat. Mech.: Theory Exp. (2008) P10023.

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

[65] M. F. Frenzel, D. Jennings, and T. Rudolph, Phys. Rev. E **90**, 052136 (2014).

[66] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin, Proc. Natl. Acad. Sci. USA **111**, 13786 (2014).

[67] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley & Sons, New York, 2012).

[68] G. Gour, I. Marvian, and R. W. Spekkens, Phys. Rev. A **80**, 012307 (2009).

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

[70] C. Browne, A. J. P. Garner, O. C. O. Dahlsten, and V. Vedral, Phys. Rev. Lett. **113**, 100603 (2014).

[71] L. Masanes and J. Oppenheim, arXiv:1412.3828 (2014).

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

[73] J. v. Tiel, *Convex Analysis: An Introductory Text* (Wiley, Chichester, New York, 1984).

[74] K. Li, Ann. Stat. **42**, 171 (2014).

[75] V. Strassen, in *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes Held at Liblice Near Prague, from June 5 to 13, 1962* (Publishing House of the Czechoslovak Academy of Sciences, Prague, 1962), pp. 698–723.

[76] Y. Polyanskiy, H. Poor, and S. Verdu, IEEE Trans. Inf. Theory **56**, 2307 (2010).

[77] M. Tomamichel and M. Hayashi, IEEE Trans. Inf. Theory **59**, 7693 (2013).