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

Nicole Yunger Halpern\textsuperscript{a1,2} and Joseph M. Renes\textsuperscript{b3}

\textsuperscript{1}Institute for Quantum Information and Matter, Caltech, Pasadena, CA 91125, USA
\textsuperscript{2}Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario Canada N2L 2Y5
\textsuperscript{3}Institute for Theoretical Physics, ETH Zürich, Switzerland

(Dated: September 16, 2014)

Abstract

Small-scale heat exchanges have recently been modeled with resource theories intended to extend thermodynamics to the nanoscale and quantum regimes. We generalize these theories to exchanges of quantities 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, \textit{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.

PACS numbers: 05.70.Ce, 89.70.Cf, 05.70.-a, 03.67.-a

Keywords: Resource theory, One-shot, Statistical mechanics, Thermodynamics, Information theory, Nonequilibrium

\textsuperscript{a} E-mail: nicoleyh@caltech.edu
\textsuperscript{b} E-mail: renes@phys.ethz.ch
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 directly 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 $\beta$ of the bath [18–23]. We call these resource theories Helmholtz theories, as the central results concern 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, and other thermodynamic contexts [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 $\beta$. Altogether, the families describing different baths form an extended family of thermodynamic resource theories amenable to experimental investigation in the present or near future.

In this paper, we introduce this extended family of resource theories, 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 [19]. In the limit as the number of copies of \( R \) extracted from or created approaches infinity, the average work yield and work cost approach the difference \( \Phi(\rho, H, N) - \Phi(\gamma, H, N) \) 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 will 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 known from traditional thermodynamics, confirming that work is the useful resource in thermodynamic resource theories. We conclude by discussing possible applications of our generalized framework to realistic systems.

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]. For instance, in quantum optics, generating coherent states is easy (e.g., using a laser), while generating superpositions of 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 amongst 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 \( A \rightarrow B \) if free operations can create \( B \) from \( A \). The ordering \( \rightarrow \) is a quasiorder, satisfying reflexivity (\( A \rightarrow A \)) and transitivity (\( A \rightarrow B \) and \( B \rightarrow C \) implies \( A \rightarrow C \)). The quasiorder differs from a partial order: Even if \( A \rightarrow B \) and \( B \rightarrow A \), \( A \) is not necessarily \( B \) [26].

Functions of resources that respect the quasiorder, in the sense that \( A \rightarrow B \) implies \( f(A) \geq f(B) \), 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 specify-
ing 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 will be explained in [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 they 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].

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 shall see in Section III, the free states must be equilibrium states, such as grand canonical ensembles if \( H \) and \( N \) are conserved, else 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 \( \beta \) and a chemical potential \( \mu \). (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 \( \mathcal{R} \) in a grand-potential theory, one specifies a density operator, a Hamiltonian, and a number operator:

\[
\mathcal{R} = (\rho, H, N). \tag{1}
\]

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 will be detailed in [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. 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 \( \rho \) 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 shall call the *state vector*, of its eigenvalues. Hence we also denote the state by \( \mathcal{R} = (r, H, N) \). 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

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

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

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

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

wherein $\beta$ and $\mu$ 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 $G = (\gamma_R, H, N)$.

1. 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 $(\beta, \mu)$, 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

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

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 $R \overset{\beta,\mu}{\rightarrow} R'$.

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, $[U, N_{1\text{tot}}] = [U, N_{2\text{tot}}] = 0$.

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, else the quasiorder breaks down. The breakdown manifests in two distinct 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 \(\gamma\) does not have the Boltzmann form of Eq. (3), then

(a) some resources \(R\) can be generated solely by equilibrating operations: \(G \xrightarrow{\beta,\mu} R\), and
(b) equilibrating operations can transform one copy of any state \(R\) into one copy of any state \(S\): \(R \xrightarrow{\beta,\mu} S\).

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 \(\gamma\) 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’ states. Such assumptions have drawn criticism [30], 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 [31]. Ruch and collaborators (who called \(d\)-majorization the mixing distance) [32, 33] applied \(d\)-majorization to physics, as did Uhlmann and colleagues [34, 35]. We dub this quasiorder in general thermodynamic resource theories equimajorization, because it is \(d\)-majorization relative to the equilibrium states.

**Definition 2.** Let \(R\) and \(S\) denote states in any grand-potential theory defined by \((\beta,\mu)\). 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

\[
Mr = s, \quad Mg_R = g_S, \quad \text{and} \quad \sum_{i=1}^{d_S} M_{ij} = 1 \quad \forall j = 1, 2, \ldots, d_R.
\]

---

\(^1\) In [21], the canonical form \((e^{-\beta H}/Z,H)\) of the free states in Helmholtz theories is argued to follow from [29]. According to [29], 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 [29], however, is not formulated as in Helmholtz resource theories. How to translate the result from [29] 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.
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 [19]. In Appendix B, we extend the proof technique in [18] to grand-canonical theories, obtaining the following result.

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

\[
R \overset{\beta,\mu}{\rightarrow} S \iff R \succ_{\beta,\mu} S. \tag{6}
\]

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 [36–38]. 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 \( i \)th 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 [39]. For example, \( f(x) = x \log x \) and \( f(x) = -\log x \) lead to the relative entropies \( D(r||g_R) = \sum_{i=1}^{d_R} r_i \log(r_i/g_i) \) and \( D(g_R||r) \). The function \( f(x) = (x^\alpha - 1)/(\alpha - 1) \) leads to the 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 \( \alpha \geq 0 \).

The Lorenz curve, introduced by Lorenz in economics [40], 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 \( L_R : [0,1] \rightarrow [0,1] \) represents the state \( R \). The curve is the piecewise linear function that connects the points

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

wherein \( \pi \) denotes a permutation such that the sequence \( (r_{\pi(j)}/g_{\pi(j)})_j \) is non-increasing. In accordance with [19, 40], 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 proved this relationship for continuous systems [32], using tools from measure theory. Uhlmann proved the relationship more directly, for discrete systems, which we address [34]. By following Uhlmann, we will prove this proposition in Appendix B:

**Proposition 3.** For any states \( R \) and \( S \) in the grand-potential resource theory defined by \((\beta, \mu)\), the following are equivalent:

(a) \( R \succ_{\beta,\mu} S \).
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 \(L_R(t) \geq L_{R'}(t)\) \(t \in [0, 1]\). Here, \(R_1 \xrightarrow{\beta,\mu} R_3\) and \(R_2 \xrightarrow{\beta,\mu} R_3\), 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.

\((b)\) \(L_R(t) \geq L_S(t)\) for all \(t \in [0, 1]\).
\((c)\) \(\phi_{f_a}(R) \geq \phi_{f_a}(S)\) for every function \(f_a(t) = \max\{0, t - a\}\) associated with any \(a \in \mathbb{R}\).
\((d)\) \(\phi_f(R) \geq \phi_f(S)\) for all continuous convex functions \(f\).

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. Harrremoës noted the relationship between Lorenz curves and hypothesis tests \([41]\); we establish the relationship more concretely.

An asymmetric hypothesis test is used to distinguish whether a given state is \(\rho\) or \(\sigma\). 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 measurement (POVM) \(\{Q, \mathbb{1} - Q\}\). If the measurement yields the outcome \(Q\), the state is likely \(\rho\). If \(\mathbb{1} - Q\) obtains, the state is likely \(\sigma\). A Type I error occurs if the state is \(\rho\) but \(\mathbb{1} - Q\) obtains, so the state seems likely to be \(\sigma\). A Type II error occurs if the state is \(\sigma\) but appears likely to be \(\rho\). 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

\[
b_c(\rho||\sigma) := \min_{\text{Tr}[Q\rho] \geq 1 - \epsilon} \frac{\text{Tr}[Q\sigma]}{\text{Tr}[Q]}.
\]

Equation (9) defines a semidefinite program, a type of convex optimization, which has a
The primal and dual forms’ equivalence follows from properties of semidefinite programs [42]. 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 4.** 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,

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

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 $D_{\epsilon H}$, 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*, or 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:

$$S_S(r) := \sum_i r_i \ln r_i = \langle \ln r_i \rangle_r.$$  \hspace{1cm} (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., [43–47]) and statistical-mechanics (e.g., [19, 21, 48–51]) problems that involve few systems or trials. In addition to involving finite numbers of trials, realistic protocols have nonzero probabilities of failing to accomplish their purposes. Failure probability has been incorporated into one-shot entropies as a parameter $\epsilon$ [43, 52].

One alternative to $S_S$ is the hypothesis-testing entropy $D_{\epsilon H}$. $D_{\epsilon H}$ 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 $R$ will be quantified with $D_{\epsilon H}$.

**Definition 3.** The hypothesis-testing relative entropy between quantum states $\rho$ and $\gamma$ is defined as

$$D_{\epsilon H}(\rho||\gamma) := - \ln b_\epsilon(\rho||\gamma)$$  \hspace{1cm} (13)

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


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 \( \rho' \) is \( \epsilon \)-close to \( \rho \). Since density operators’ distinguishability is related to the trace distance, we define \( R' \approx_{\epsilon} R \) by \( \frac{1}{2} \| \rho' - \rho \|_1 \leq \epsilon \).

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 phase-space path traversed by a state. Quantum states can follow paths along which work integrals cannot easily be calculated [53]. 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 [54]. We define work similarly to [54].

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 \( \epsilon \)-work value of a resource \( R \) is defined as the greatest \( W \) for which

\[
R + B_E \xrightarrow{\beta,\mu,\epsilon} B_{E+W}. \tag{14}
\]

The work cost of \( \epsilon \)-approximately creating \( R \) is the least \( W \) such that

\[
B_{E+W} \xrightarrow{\beta,\mu,\epsilon} R + B_E. \tag{15}
\]

Formally,

\[
W_{\text{gain}}^\epsilon (R) = \max \{ W : R + B_E \xrightarrow{\beta,\mu,\epsilon} B_{E+W}, \beta E \gg 1 \} \tag{16}
\]

\[
W_{\text{cost}}^\epsilon (R) = \min \{ W : B_{E+W} \xrightarrow{\beta,\mu,\epsilon} R + B_E, \beta E \gg 1 \} \tag{17}
\]

Having defined work, we state the work value, and bound the work cost, of \( R \).

**Theorem 5.** The \( \epsilon \)-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 \( \epsilon \)-approximation to a state \( R \) is bounded by

\[
\max_{\delta \in [0,1-\epsilon]} \left[ \frac{1}{\beta} D_\text{H}^{1-\epsilon-\delta} (r||g_R) - \frac{1}{\beta} \log \left( \frac{1}{\delta} \right) \right] \leq W_{\text{cost}}^\epsilon (R) \leq \frac{1}{\beta} D_\text{H}^{1-\epsilon} (r||g_R) - \frac{1}{\beta} \log \left( \frac{1-\epsilon}{\epsilon} \right). \tag{19}
\]
VI. WORK YIELD AND COST OF MANY COPIES OF A RESOURCE

From the previous section’s one-shot work quantities, we can recover results from traditional thermodynamics and can compare how \( W^\epsilon_{\text{gain}} \) differs from \( W^\epsilon_{\text{cost}} \) 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 thermodynamic limit, or asymptotic limit, \( n \to \infty \). Also in the limit, \( W^\epsilon_{\text{gain}}(R^\otimes n) \) and \( W^\epsilon_{\text{cost}}(R^\otimes n) \) tend to the difference between the grand potential of \( R \) and the grand potential of the associated equilibrium state \( G_R \). As the thermodynamic limit is approached, \( W^\epsilon_{\text{gain}}(R^\otimes n) \) and \( W^\epsilon_{\text{cost}}(R^\otimes n) \) differ by terms of order \( \sqrt{n} \).

To derive the thermodynamic limits of Eq. (18) and Ineqs. (19), we invoke the Asymptotic Equipartition Property of D\( ^\epsilon_H \) [42]:

\[
\lim_{n \to \infty} \frac{1}{n} D^\epsilon_H(r^\otimes n \| s^\otimes n) = D(r \| s) \quad \forall \epsilon \in (0, 1).
\] (20)

We have used the relative entropy between two probability distributions, \( r = (r_1, r_2, \ldots, r_d) \) and \( s = (s_1, s_2, \ldots, s_d) \), over the same alphabet [55]:

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

wherein \( 0 \ln 0 = 0 \).

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

\[
\lim_{n \to \infty} \frac{1}{n} W^\epsilon_{\text{gain}}(R^\otimes n) = \lim_{n \to \infty} \frac{1}{n} W^\epsilon_{\text{cost}}(R^\otimes n) = \frac{1}{\beta} D(r \| g_R),
\] (22)

wherein \( g_R \) characterizes the equilibrium state associated with \( R \). In the asymptotic limit, the bounds in Ineqs. (19) converge. All strategies of work extraction and state formation, from risky \( (\epsilon \approx 1) \) to conservative \( (\epsilon \approx 0) \), are equivalent in this limit.

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

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

\[
= (H)_r - T \cdot k_B S_S(r) - \mu \langle N \rangle_r + k_B T \ln Z
\] (24)

\[
= \Phi_{\beta, \mu}(R) - \Phi_{\beta, \mu}(G_R).
\] (25)

Recall that \( \Phi := E - TS - \mu N \) denotes the grand potential, and \( -k_B T \ln Z \) 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. As in [19, 21], one-shot information theory implies and extends known thermodynamics.

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 \( \epsilon \), there exists an \( n \) and an \( m_n \) great enough that \( R^\otimes n \xrightarrow{\beta, \mu} S^\otimes m_n \). 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 \( R(R \mapsto S) \) 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 $\epsilon \to 0$). This rate is
\[
R(R \mapsto S) = \frac{D(r\|g_R)}{D(s\|g_S)}.
\]

(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 [56].

The Asymptotic Equipartition Theorem dictates the leading order (order-$n$) behavior of $D'H'$; 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 $\sqrt{n}$:
\[
W_{\text{gain}}(R^{\otimes n}) = \frac{1}{\beta}[nD(r\|g_R) - O(\sqrt{n})], \quad \text{and}
\]
\[
W_{\text{cost}}(R^{\otimes n}) = \frac{1}{\beta}[nD(r\|g_R) + O(\sqrt{n})].
\]

(27)

(28)

As one might expect, the work cost $W_{\text{cost}}(R^{\otimes n})$ lies above the thermodynamic value, whereas $W_{\text{gain}}(R^{\otimes n})$ 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 [21].

VII. CONCLUSIONS

Building on Helmholtz theories’ successes, we have extended the resource-theory formulation of thermodynamics beyond heat baths. We introduced a set of families of thermodynamic resource theories, in which Helmholtz theories form one family. Each family corresponds to one 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 the resource-theory framework, 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 in terms 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 $n \to \infty$, $W_{\text{gain}}(R^{\otimes n})$ and $W_{\text{cost}}(R^{\otimes n})$ 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.

Generalized thermodynamic resource theories open a host of realistic thermodynamic systems to modeling by resource theories. The modeling of heat-and-particle exchanges,
electrochemical batteries, and chemical reactions merit investigation. As thermodynamic potentials other than the Helmholtz free energy \( F \) characterize common experiments, generalized thermodynamic resource theories may offer opportunities for testing one-shot statistical mechanics experimentally.

ACKNOWLEDGEMENTS

NYH is grateful for conversations with Tobias Fritz, Iman Marvian, Markus Müller, Brian Space, and Rob Spekkens. JMR acknowledges helpful conversations with Michael Walter. This work was supported by a Virginia Gilloon Fellowship, an IQIM Fellowship, NSF grant PHY-0803371, the Perimeter Institute for Theoretical Physics, the Swiss National Science Foundation (through the National Centre of Competence in Research Quantum Science and Technology and grant No. 200020-135048), and the European Research Council (grant No. 258932). The Institute for Quantum Information and Matter (IQIM) is an NSF Physics Frontiers Center that receives support from the Gordon and Betty Moore Foundation. Research at the Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. NYH is grateful to Renato Renner for hospitality at ETH Zürich during the development of this paper.

VIII. APPENDICES

Below, we prove claims about grand-potential theories presented above. We derive the grand-canonical forms of free state vectors; describe the quasiorder; calculate the work \( W^\epsilon_{\text{gain}} \) extractable from, and bound the work \( W^\epsilon_{\text{cost}} \) required to create, one copy of a state; and show via second-order asymptotics that \( W^\epsilon_{\text{gain}} \) does not always equal \( W^\epsilon_{\text{cost}} \).

Appendix A  DERIVATIONS OF FREE STATES’ FORMS

A  Proof of Theorem 1(a)

The state vectors of the free states in grand-potential theories are shown to be grand canonical ensembles. We first review the derivation, in [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.\(^1\)

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 \( \pi \) can be implemented, and any subsystem \( A \) can be discarded (marginalized) [15–17]:

\[
r \mapsto \sum_A \pi(r \otimes u).
\]

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 [16], the free states must be uniform probability distributions, lest the nonuniformity theory be trivial (indeed, the quasiorder of resources becomes trivial). The free states’ form is derived as follows [16]: Suppose that some nonuniform state $u_0$ is free. By Shannon compressing many copies of $u_0$, 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 6.** 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 7.** Let $H_1$ and $H_2$ denote any Hamiltonians that share an energy gap $\Delta$ (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, then

$$\frac{g_1(E_1 + \Delta)}{g_1(E_1)} = \frac{g_2(E_2 + \Delta)}{g_2(E_2)}.$$  \hspace{1cm} (29)

That is, the ratio of the weights depends only on the gap.

**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:

$$g_1(E_1)g_2(E_2 + \Delta) = g_1(E_1 + \Delta)g_2(E_2).$$  \hspace{1cm} (30)

Since $E_1$ and $E_2$ are arbitrary, each ratio of probabilities depends only on $\Delta$; the other details of $H_1$ and $H_2$ are irrelevant. \hfill $\square$
Lemma 8. 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 \( \Delta \). The ratio of the weights varies exponentially with the gap:

\[
\frac{g(E + \Delta)}{g(E)} = e^{-\beta \Delta},
\]

wherein \( \beta \in \mathbb{R} \).

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

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

Let us apply Lemma 7 to two copies of \( G \), defining \( E_1 = E \) and \( E_2 = E + \Delta \). Equation (29) becomes

\[
\frac{g(E + 2\Delta)}{g(E + \Delta)} = \frac{g(E + \Delta)}{g(E)} =: f(\Delta),
\]

which implies

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

This scaling implies that, if \( f \) is continuous, it is an exponential:

\[
f(\Delta) = e^{-\beta \Delta}
\]

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

Let \( G' \) denote any free state that has the same gap \( \Delta \) as \( G \). The composition \( G + G' \) obeys the uniform-eigensubspace condition. Hence \( G' \) satisfies Eq. (34), by Lemma 7, even if \( G' \) does not have the form in Eq. (32).

\[\square\]

Lemma 9. All the gaps in all the free states in the same Helmholtz theory have the same \( \beta \).

Proof. Let \( G = (g, H) \) and \( G' = (g', H') \) denote free states in some Helmholtz theory. Let \( \Delta \) 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 \( g' \) have the same \( \beta \). 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 8,

\[
f(m\Delta) := \frac{g''(E + m\Delta)}{g''(E)}
\]

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

\[
=: f(\Delta)^m,
\]

15
wherein \( g''(E + k\Delta) \) denotes the weight on level \( k \).

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

\[
f(\Delta') = f(\Delta)^n.
\]

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. (37). Equating the two expressions for \( f(n\Delta') \) yields

\[
f(\Delta') = f(\Delta)^{m/n}.
\]

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 \( \alpha \) denotes an irrational number. \( \alpha \) 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. (39).

Lemmas 7-9, 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.** 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, then some resources \( R \) can be generated solely by equilibrating operations: \( G \xrightarrow{\beta,\mu} R \).

**Proof.** First, we show that each element of \( g \) has the form \( e^{-\beta(E_i)}e^{\alpha(n_j)n_j}/Z \), 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 will show that \( \beta \) and \( \alpha \) are constant functions.

Consider the most general state vector associated with \( H \) and \( N \). Each element has the form \( e^{-\beta(E_i)}e^{\alpha(n_j)n_j+f(E_i,n_j)}/Z \), 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 \( e^{-\beta(E_i)+\alpha n_j}/Z \).

\( G \) 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 \( \beta \) that is a constant function of energy. This \( \beta \) 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 \( \alpha = -\beta \mu \).
B 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 $S = (s,H)$. To transform the distribution $r$ into $s$, it suffices to successively operate on pairs of levels, such as the first and the $j^{th}$. An equilibrating operation converts $r_j/r_1$ into $s_j/s_1$. The operation has the following form. 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

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

(40)

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, can 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 3, and Lemma 4.

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

$$R \xrightarrow{\beta,\mu} S \iff R \succ_{\beta,\mu} S.$$  

(41)

**Proof.** This proof is adapted from the proof of [18, Theorem 5], a Helmholtz-theory analogue of our Theorem 2. 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.

---

2 An alternative approach to the Helmholtz-theory analogue appears in [19].
Assume that some equilibrating operation maps $R$ to $S$:

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

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

$$Mv = \mathcal{E}(V).$$

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

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

wherein $[w]_i$ denotes the $i^{th}$ 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 $S_n$ of pure eigenstates of $H_p + H_p^n + H_{\tilde{p}} + \tilde{H}_p$, wherein $H^n$ denotes $n$ copies of $H$. Each state in $S_n$ 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 $S_n$ correspond to the same pair

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

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

A permutation $\pi_n : S_n \rightarrow S_n$ is defined in terms of the matrix assumed to map $p$ to $\tilde{p}$. Because $\pi_n$ maps each input to an output that has the same $(u, v)$, $\pi_n$ conserves energy. $\pi_n$ 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 $T_n$ of typical $(u, v)$ tuples is defined in terms of $P_n$ and the limit $n \rightarrow \infty$. In this limit, Janzing et al. show, $\pi_n$ 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 $S_n$ correspond to the same $(u, v)$, they correspond not only to the same energy, but also to the same particle number. Just as $\pi_n$ conserves energy, it conserves particle number. The rest of the proof in [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. Before proceeding to Proposition 3, we establish Lemma 4 for convenience.

**Lemma 4.** 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,

$$(L_R(t_k), 1 - t_k) = (b_\epsilon(r||g_R), \epsilon).$$  \hfill (42)

**Proof.** Let $\pi$ denote a permutation such that the sequence $(r_{\pi(k)}/g_{\pi(k)})_k$ is nonincreasing. Let $\bar{r}_m := \sum_{k=1}^{m} r_{\pi(k)}$ 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 $\bar{r}_0 = G_0 = 0$. The points that define the rescaled Lorenz curve are $(G_m, \bar{r}_m)$ for $m \in \{0, 1, \ldots, d_R\}$. To prove the claim, we first show that $(G_m, 1 - \bar{r}_m)$ 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 $b_{\epsilon_m}$ [Eq. (9)]. \footnote{If $Q$ is said to be feasible if the measurement $\{Q, 1 - Q\}$ corresponds to a Type I error probability of at most $\epsilon$. The feasible measurement that minimizes the Type II error probability is optimal.}

Therefore,

$$b_{\epsilon_m}(r||g_R) \leq \sum_i [Q_m g_R]_i = G_m.$$  \hfill (43)

To show that equality holds, we consider the dual problem in Eq. (10). A feasible pair \footnote{$(\mu, \tau)$ is said to be feasible if they satisfy the constraints in Eq. (10).} $(\mu_m, \tau_m)$ that satisfies the constraint $\mu_m r - g_R \leq \tau_m$ is given by defining $\mu_m$ such that $r_{\pi(m+1)}/g_{\pi(m+1)} \leq 1/\mu_m < r_{\pi(m)}/g_{\pi(m)}$ 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 $\mu_m$ cancel, by $1 - \epsilon_m = R_m$ and by the definition of $R_m$. Hence $b_{\epsilon_m}(r||g) \geq G_m$. Combining this result with Ineq. (43), shows that

$$b_{\epsilon_m}(r||g_R) = G_m.$$  \hfill (44)

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

$$1 - \epsilon = (1 - \lambda)(1 - \epsilon_m) + \lambda(1 - \epsilon_{m+1}).$$  \hfill (45)

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

$$b_\epsilon(r||g_R) \leq (1 - \lambda)b_{\epsilon_m}(r||g_R) + \lambda b_{\epsilon_{m+1}}(r||g_R).$$  \hfill (46)
Let us show that Ineq. (46) 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

\[
\begin{align*}
\phi_e(r||g_R) &\geq \mu[(1 - \epsilon_m) + \lambda r_{\pi(m+1)}] - \mu R_m + G_m \\
&= \lambda g_{\pi(m+1)} + G_m \\
&= (1 - \lambda)G_m + \lambda G_{m+1} \\
&= (1 - \lambda)\phi_{\epsilon_m}(r||g_R) + \lambda \phi_{\epsilon_{m+1}}(r||g_R).
\end{align*}
\]

The final equality follows from Eq. (44). Inequalities (46) and (50) show that interpolating linearly between \((G_m, 1 - R_m)\) and \((G_{m+1}, 1 - R_{m+1})\) amounts to interpolating linearly between \((\phi_{\epsilon_m}(r||g_R), \epsilon_m)\) and \((\phi_{\epsilon_{m+1}}(r||g_R), \epsilon_{m+1})\).

Finally, we give a mostly self-contained proof of Proposition 3.

**Proposition (3).** For any states \( R \) and \( S \) in the grand-potential resource theory defined by \((\beta, \mu)\), the following are equivalent:

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

(b) \( L_R(t) \geq L_S(t) \) for all \( t \in [0, 1] \).

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

(d) \( \phi_f(R) \geq \phi_f(S) \) for all continuous convex functions \( f \).

**Proof.** We will 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. Suppose that \( Q \) defines the optimal test that distinguishes \( s \) from \( g_S \) with a Type I error probability \( \epsilon \). To distinguish \( r \) from \( g_R \), one can apply \( M \) and then measure \( \{Q, 1 - Q\} \). This test might distinguish between \( r \) and \( g_R \) suboptimally. Hence \( b_e(r||g_R) \leq b_e(s||g_S) \). Lemma 4 implies (b).

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

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

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

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

wherein \( f_a(t) = \max\{0, t - a\} \). Since \( b_e(r||g_R) \leq b_e(s||g_S) \), \( \phi_{f_{1/\mu}}(r, g_R) \geq \phi_{f_{1/\mu}}(s, g_S) \).

\((c) \Rightarrow (d)\) In [34] (see also [35, 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 \( \phi_{f_a}(R) \geq \phi_{f_a}(S) \) for all \( f_a \), \( \phi_f(R) \geq \phi_f(S) \).
Lemma 10. Let $R = (r, H, N)$ denote any state; and $(e_E, H', N')$, 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

$$b_e(r \otimes e_E || g \otimes g') = (g' \cdot e_E) b_e(r || g).$$  

Proof. For any feasible measurement operator $Q$ in $b_e(r \otimes e_E || g \otimes g')$, we can find a feasible $Q'$ that gives the same value of the objective function and that has the form

$$Q = Q' \cdot e_E.$$  

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 $W^c_{\text{gain}}(R)$ extractable from, and bound the least work $W^c_{\text{cost}}$ 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. 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 zeroes. By $[v]$, we denote the $i^{th}$ 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 10. Following [35, Theorem 1.4.4], we prove the contrapositive. Assume there is no stochastic matrix $M$ that 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 $\{Mr \oplus Mg_R\}$ [58, Theorem 3.5]. That is, a vector $x \oplus y$ satisfies $x \cdot s + y \cdot g_S > x \cdot Mr + y \cdot M g_R$ for all stochastic matrices $M$. Taking the maximum over $M$ on the right-hand side (RHS) and denoting by $[g_R]k$ the $k^{th}$ element of $g_R$ gives

$$x \cdot s + y \cdot g_S > \max_M \{x \cdot Mr + y \cdot M g_R\}$$  

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

$$= \sum_k \max_j \{x_j r_k + y_j [g_R]k\}. \quad (53)$$  

Maximizing on the LHS produces

$$\sum_k \max_j \{x_j s_k + y_j [g_S]k\} > \sum_k \max_j \{x_j r_k + y_j [g_R]k\}. \quad (54)$$  

But $f(s, t) = \max_j \{x_j s + y_j t\}$ is a convex function, meaning $\phi_f(S) \geq \phi_f(R)$. The contrapositive implies (a).

$\square$
\[ Q' = \sum_i Q_{E'_i} \otimes e_{E'_i}^T_e, \] wherein \( Q_{E'_i} = (\mathbb{1} \otimes e_{E'_i}) Q(\mathbb{1} \otimes e_{E'_i}^T) \) and the superscript \( T \) denotes the transpose. Without loss of generality, focus on \( Q' \) operators that have this form.

The constraint becomes \( \sum_i [Q_E r_i \geq 1 - \epsilon, \text{ and the objective function becomes} \]

\[ \sum_i [Q'(g \otimes g')_i = \sum_i \left\{ (g' \cdot e_{E'_i}) \sum_j [Q_{E'_i} g]_j \right\}. \] (58)

The minimum follows from setting \( Q_{E'_i} = 0 \) for all \( E'_i \neq E' \):

\[ b_e(r \otimes e_E || g \otimes g') = \min\{(g' \cdot e_E) \sum_j [Q_E g]_j \mid \sum_j [Q_E r]_j \geq 1 - \epsilon, 0 \leq Q_E \leq \mathbb{1}\} \] (59)
\[ = (g' \cdot e_E) b_e(r || g). \] (60)
\[ \square \]

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

**Lemma 11.** Let \( R \) denote any state in a grand-potential resource theory defined by \( \beta \) and \( \mu \). 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

\[ R + B_E \succ_{\beta, \mu} B_{E+W} \iff R \succ_{\beta, \mu} B_W \text{ and} \]
\[ B_{E+W'} \succ_{\beta, \mu} R + B_E \iff B_{W'} \succ_{\beta, \mu} R. \] (61) (62)

**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 equivalence of Eqs. (61):

\[ R + B_E \succ_{\beta, \mu} B_W + B_E \iff L_{R+B_E}(t) \geq L_{B_W+B_E}(t) \forall t \in [0,1] \] (63)
\[ \iff b_e(r \otimes e_E || g \otimes g') \geq b_e(e_W \otimes e_E || g \otimes g') \] (64)
\[ \iff b_e(r || g) \geq b_e(e_W || g) \] (65)
\[ \iff L_R(t) \geq L_{B_W}(t) \forall t \in [0,1] \] (66)
\[ \iff R \succ_{\beta, \mu} B_W. \] (67)

The first equivalence follows from Proposition 3; the second, from Lemma 4; the third, from Lemma 10; the fourth, from Lemma 4; and the fifth, from Proposition 3. Similar reasoning justifies the equivalence of Eqs. (62).

\[ \square \]

Having simplified the model of work, we will calculate \( W^\epsilon_{\text{gain}} \).

**Theorem (5).** The \( \epsilon \)-work value of a state \( R = (r, H, N) \) associated with the free state \( G_R = (g_R, H, N) \) is

\[ W^\epsilon_{\text{gain}}(R) = \frac{1}{\beta} D_H^\epsilon(r || g_R). \] (68)

22
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 $E(r \otimes e_E) \approx \epsilon e_E \otimes e_W$. The channel’s output is $\epsilon$-close, in the $l_1$ norm, to the desired state:

$$\frac{1}{2}|\mathcal{E}(r \otimes e_E) - e_E \otimes e_W|_1 \leq \epsilon.$$

(69)

Since equilibrating operations map equilibrium states to equilibrium states, $E(g_R \otimes g') = g'$. Using $E$, we can construct a hypothesis test between $r \otimes e_E$ and $g_R \otimes g'$. The test consists of an application of $E$ followed by an energy measurement. 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 $D_{H}^{\epsilon}(r \otimes e_E || g_R \otimes g')$, and

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

(70)

$$= (g \otimes g') \cdot e_{E+W}$$

(71)

$$= e^{-\beta(E+W)}$$

(72)

By Eq. (57),

$$e^{-D_{H}^{\epsilon}(r || g_R)} \leq e^{-\beta W},$$

(73)

which is equivalent to the upper bound.

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

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

(74)

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

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

(75)

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

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

(76)

This equation is equivalent to Ineq. (72) except contains an equality. Therefore, free operations can distill at least the work $W$ that satisfies

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

(77)

which is equivalent to the lower bound.

---

5 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 does not perform the measurement.
Having calculated the work extractable from $R$, we bound the work cost of creating $R$ [Ineqs. (19)].

**Theorem (5, ctd.).** The work cost of creating an $\epsilon$-approximation to a state $R$ is bounded by

$$
\max_{\delta \in (0, 1-\epsilon)} \left[ \frac{1}{\beta} D^{1-\epsilon-\delta}_H (r \| g_R) - \frac{1}{\beta} \log \left( \frac{1}{\beta} \right) \right] \leq W_{\text{cost}}(R) \leq \frac{1}{\beta} D^{1-\epsilon-\delta}_H (r \| g_R) - \frac{1}{\beta} \log \left( \frac{1-\epsilon}{\epsilon} \right). \quad (78)
$$

**Proof.** To derive the lower bound, we suppose $\mathcal{E}$ is part of 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 10,

$$
e^{-D^{1-\epsilon}_H (e_{E+W} \| g')} = (g' \cdot e_{E+W}) e^{-D^{1-\epsilon}_H (1 \| 1)} = \epsilon (g' \cdot e_{E+W}). \quad (79)$$

The dual formulation of $D^{1-\epsilon}_H$ reads,

$$
e^{-D^{1-\epsilon}_H (e_{E+W} \| g')} = \max_{\mu e_{E+W} - g' \leq \tau} \left\{ \epsilon \mu - \sum_i \tau_i \right\}, \quad (81)$$

wherein $\tau_i$ denotes the $i$th element of $\tau$. Comparing Eqs. (80) and (81) shows that $\mu = g' \cdot e_{E+W}$ and $\tau = 0$ are the optimal choices in the dual formulation.

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

$$
e^{-D^{1-\delta}_H (\mathcal{E}(e_{E+W} \| g_R \otimes g'))} \geq \delta \frac{e^{-\beta (E+W)}}{Z} \quad (82)$$

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

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

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

$$e^{-D^{1-\eta}_H (\mathcal{E}(e_{E+W} \| g_R \otimes g'))} \leq \sum_i [Q (g_R \otimes g')]_i \quad (83)$$

$$= e^{-D^{1-\eta}_H (r \otimes e_E \| g_R \otimes g')} = \frac{e^{-\beta E}}{Z} e^{-D^{1-\delta}_H (r \| g_R)}. \quad (85)$$

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

$$\delta e^{-\beta W} \leq e^{-D^{1-\epsilon-\delta}_H (r \| g_R)}. \quad (86)$$

24
We have lower-bounded $W_{\text{cost}}(R)$ for every $\delta \in (0, 1 - \epsilon]$. The tightest of these bounds follows from a maximization over $\delta$.

To derive upper bound, we construct an equilibrating operation that maps $e_{E+W}$ to $\tilde{r} \otimes e_E$, wherein $\tilde{r} \approx r$, 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 3, such an equilibrating operation exists if and only if

$$K_{\text{in}}(a) \geq K_{\text{out}}(a)$$  \hspace{1cm} (87)

for all $a \in \mathbb{R}$ and for $K_R(a)$ defined as follows. In Proposition 3, the function $f_a(t) := \max\{0, t - a\}$ appears in

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

$$= \sum_{i=1}^{d_R} g_i \max\left\{0, \frac{r_i}{g_i} - a\right\}$$  \hspace{1cm} (88)

$$= \sum_{i=1}^{d_R} \max\{0, r_i - g_ia\}.$$  \hspace{1cm} (89)

To simplify notation, we will relabel this sum as $K_R(a)$. Because $r$ and $g$ are normalized, $K_R(a) = (1 - a)$ if $a \leq 0$. We can rewrite the LHS of Ineq. (87) as

$$K_{\text{in}}(a) = \left(1 - a \frac{e^{-\beta(E+W)}}{Z}\right)_+.$$  \hspace{1cm} (91)

This function is linear and satisfies $K_{\text{in}}(0) = 1$ and $K_{\text{in}}(Ze^{\beta(E+W)}) = 0$. We can rewrite the RHS of Ineq. (87) as

$$K_{\text{out}}(a) = \sum_i \left( \tilde{r}_i - g_i a \frac{e^{-\beta E}}{Z} \right)_+.$$  \hspace{1cm} (92)

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

$$K_{\text{out}}(Ze^{\beta(E+W)}) = 0.$$  \hspace{1cm} (93)

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 $D^{1-\epsilon}_{1|\epsilon}(r||g_R)$. Suppose that $\mu$ and $\tau$ are the optimal choices, so that

$$e^{-D^{1-\epsilon}_{1|\epsilon}(r||g_R)} = \epsilon \mu - \sum_i \tau_i$$  \hspace{1cm} (94)

and $\mu r - g_R \leq \tau$. 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 \( \mu r' \leq g_R \). Let us bound the trace distance

\[
\frac{1}{2} \|r - r'\|_1 = \frac{1}{2} \sum_k |r'_k - r_k| \tag{95}
\]

\[
= \frac{1}{2} \sum_k |r'_k - g_k r_k (g_k + \tau_k)^{-1}| \tag{96}
\]

\[
= \frac{1}{2} \sum_k r_k |1 - g_k (g_k + \tau_k)^{-1}| \tag{97}
\]

\[
= \frac{1}{2} \sum_k r_k \frac{\tau_k}{g_k + \tau_k} \tag{98}
\]

\[
\leq \frac{1}{2} \sum_k \frac{1}{\mu} \tau_k \tag{99}
\]

\[
= \sum_i \frac{\tau_i}{2 \mu} \tag{100}
\]

\[
\leq \frac{\epsilon}{2}. \tag{101}
\]

To derive the first inequality, we have used the inequality \( \mu r \leq g_R + \tau \); to derive the second, \( \epsilon \mu - \sum_i \tau_i = e^{-D_H^{1-\epsilon}(r||g_R)} \geq 0 \). Let \( \tilde{r} = r' / \sum_i r'_i \). By the Triangle Inequality,

\[
\frac{1}{2} \|r - \tilde{r}\|_1 \leq \frac{1}{2} \|r - r'\|_1 + \frac{1}{2} \|\tilde{r} - r'\|_1 \tag{102}
\]

\[
\leq \frac{\epsilon}{2} + \frac{1}{2} \left( \frac{1}{\sum_i r'_i} - 1 \right) \|r'\|_1 \tag{103}
\]

\[
= \frac{\epsilon}{2} + \frac{1}{2} \left( 1 - \sum_i r'_i \right) \tag{104}
\]

\[
\leq \epsilon. \tag{105}
\]

The fourth inequality follows from \( \sum_i r_i - \sum_j r_j' \leq \|r - r'\|_1 \). We have constructed a state \( \tilde{r} \approx \epsilon r \).

Moreover, \( \mu (\sum_i r'_i) \tilde{r} \leq g \). Applying this inequality to (92) yields

\[
K_{\text{out}}(a) \leq \sum_i g_i \left( \frac{1}{\mu \sum_i r'_i} - a \frac{e^{-\beta E}}{Z} \right)_+ \tag{106}
\]

\[
= \left( \frac{1}{\mu \sum_i r'_i} - a \frac{e^{-\beta E}}{Z} \right)_+. \tag{107}
\]

Hence \( K_{\text{out}}(Ze^{\beta(E+W)}) \) satisfies

\[
K_{\text{out}}(Ze^{\beta(E+W)}) \leq \left( \frac{1}{\mu \sum_i r'_i} - e^{\beta W} \right)_+ \tag{108}
\]

\[
\leq \left( \frac{1}{\mu (1 - \epsilon)} - e^{\beta W} \right)_+ \tag{109}
\]

\[
\leq \left( \frac{\epsilon}{1 - \epsilon} e^{D_H^{1-\epsilon}(r||g)} - e^{\beta W} \right)_+. \tag{110}
\]

26
The second inequality follows from $\sum r_i' \geq 1 - \epsilon$; the third, from $\epsilon \mu \geq \epsilon \mu - \sum_i \tau_i = e^{-D_{\mu}^{1-\epsilon}(r||g_R)}$. We can satisfy (93) by choosing $W$ such that

$$e^{\beta W} = \frac{\epsilon}{1 - \epsilon} e^{D_{\mu}^{1-\epsilon}(r||g_R)}.$$  \hspace{1cm} (111)

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

One can show that the upper bound is a nonnegative quantity, using $e^{-D_{\mu}^{1-\epsilon}(r||g_R)} \leq \epsilon$. (This inequality follows from the choice $Q = \epsilon I$.) Because $\epsilon \leq \epsilon_1 - \epsilon$, $e^{-D_{\mu}^{1-\epsilon}(r||g_R)} \leq \epsilon_1 - \epsilon$. The latter implies the upper bound’s nonnegativity. The lower bound is nonnegative in all the numerical examples tested.

Appendix D COMPARISON OF ONE-SHOT WORK YIELD AND WORK COST

We will use second-order asymptotics to show that $W_{\text{gain}}^\epsilon(R^\otimes n)$ tends to differ from the bounds on $W_{\text{cost}}^\epsilon(R^\otimes n)$, and that the bounds lie arbitrarily close together, as the thermodynamic 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 [59, Theorem 5] (see also [60, 61]):

$$D_{H}^{r^\otimes n}(g_{R}^n) = nD(r||g_R) + \sqrt{n} s(r||g_R) \Phi^{-1}(\epsilon) + O(\log n),$$  \hspace{1cm} (112)

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

$$s(r||g_R) := \sqrt{V(r||g_R)} = \sqrt{\text{Tr}(r \log r - \log g_R)^2 - D(r||g_R)^2},$$  \hspace{1cm} (113)

and the inverse error function is

$$\Phi^{-1}(\epsilon) := \sup \left\{ z \in \mathbb{R} \mid \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} dt \leq \epsilon \right\}.$$  \hspace{1cm} (114)

Equation (114) 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 Eq. (112) to Eq. (18) and to Ineqs. (19).

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

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \left[ n \frac{1}{\beta} D(r||g_R) - W_{\text{gain}}^\epsilon((R^\otimes n)) \right] = \lim_{n \to \infty} \frac{1}{\beta} \left[ -s(R||g_R)\Phi^{-1}(\epsilon) - O(\log n) \right] = \frac{1}{\beta} s(R||g_R)\Phi^{-1}(1 - \epsilon).$$  \hspace{1cm} (115)

To evaluate the quantum version of Eq. (112) appears in [59, 62], but we have specialized to commuting density operators.

6 The quantum version of Eq. (112) appears in [59, 62], but we have specialized to commuting density operators.
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. (116) is positive because the work distilled at the optimal asymptotic efficiency exceeds the work distilled at any sub-asymptotic efficiency [i.e., Eq. (115) is positive].

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

$$
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \left[ W_{\text{cost}}(R^\otimes n) v - n \frac{1}{\beta} D(r||g_R) \right] \geq \lim_{n \to \infty} \frac{1}{\beta} \max_{\delta \in [0,1-\epsilon]} \left[ s(r||g_R)\Phi^{-1}(1-\epsilon - \delta) - \frac{\log \frac{1}{\delta}}{\sqrt{n}} \right] \\
= \frac{1}{\beta} \max_{\delta \in [0,1-\epsilon]} s(r||g_R)\Phi^{-1}(1-\epsilon - \delta) \\
= \frac{1}{\beta} s(r||g_R)\Phi^{-1}(1-\epsilon). \quad (117)
$$

The first equality holds if $\delta$ grows more slowly than $e^{\sqrt{n}}$. The last equality holds since, by the definition and monotonicity of $\Phi^{-1}$, the least possible $\delta$-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

$$
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \left[ W_{\text{cost}}^\epsilon(R^\otimes n) - n \frac{1}{\beta} D(r||g_R) \right] \leq \lim_{n \to \infty} \frac{1}{\beta} \left[ s(r||g_R)\Phi^{-1}(1-\epsilon) - \frac{1}{\sqrt{n}} \log \left( \frac{1-\epsilon}{\epsilon} \right) \right] \\
= \frac{1}{\beta} s(r||g_R)\Phi^{-1}(1-\epsilon). \quad (119)
$$

The final line equality 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 $D(r||g_R)$ by an amount proportional to $\sqrt{n}$. In contrast, $D(r||g_R)$ exceeds the work gain by an amount proportional to $\sqrt{n}$. The work gain and work cost differ in general, unlike in the thermodynamic limit, as in [19, 21]. As creating $R^\otimes n$ requires more work than can be extracted from $R^\otimes n$, resources degrade.

REFERENCES

[1] D. Lacoste, A. W. Lau, and K. Mallick, Phys. Rev. E 78, 011915 (2008).
[2] V. Serreli, E. R. Lee, C.-F. and Kay, and D. A. Leigh, Nature 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, Physics Today 58, 43 (2005).
[5] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante, Nature 437, 231 (2005).
[6] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, Science (New York, N.Y.) 296, 1832 (2002).
[7] A. Alemany and F. Ritort, Europhysics 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, (in preparation).
[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 Journal of Physics 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 Journal of Physics 16, 013009 (2014), arXiv:1307.7171 [quant-ph].
[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, ArXiv e-prints (2013), arXiv:1309.6586 [quant-ph].
[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, 1 (2013).
[20] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Physical Review Letters 111, 250404 (2013).
[21] F. G. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, arXiv:1305.5278 (2013).
[22] M. Horodecki and J. Oppenheim, International Journal of Modern Physics B 27, 1345019 (2013), arXiv:1209.2162 [quant-ph].
[23] P. Faist, J. Oppenheim, and R. Renner, ArXiv e-prints (2014), arXiv:1406.3618 [quant-ph].
[24] R. A. Alberty, Pure and Applied Chemistry 73, 1349?380 (2001).
[25] H. B. Callen, Thermodynamics and an Introduction to Thermostatistics, 2nd ed. (John 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, 2010).
[28] N. Yunger Halpern, (in preparation).
[29] W. Pusz and S. Woronowicz, Communications in Mathematical Physics 58, 273 (1978).
[30] E. T. Jaynes, Phys. Rev. 106, 620 (1957).
[31] A. F. Veinott, Manage. Sci. 17, 547 (1971).
[32] E. Ruch, R. Schranner, and T. H. Seligman, The Journal of Chemical Physics 69, 386 (1978).
[33] E. Ruch, Theoretica Chimica Acta 19, 225 (1975).
[34] A. Uhlmann, Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R. 27, 213 (1978).
[35] P. M. Alberti and A. Uhlmann, Dissipative motion in state spaces (Teubner, 1981).
[36] I. Csiszr, Magyar. Tud. Akad. Mat. Kutato Int. Kozl (Publication of the Mathematical Institute of the Hungarian Academy of Sciences) 8, 85 (1963).
[37] T. Morimoto, Journal of the Physical Society of Japan 18, 328 (1963).
[38] S. M. Ali and S. D. Silvey, Journal of the Royal Statistical Society. Series B (Methodological) 28, 131 (1966).
[39] F. Liese and I. Vajda, IEEE Transactions on Information Theory 52, 4394 (2006).
[40] M. O. Lorenz, Publ. Am. Stat. Assoc. 9, 209 (1905).
[41] P. Harremoës, in Proceedings of the 2004 International Symposium on Information Theory and its Applications (Parma, Italy, 2004) pp. 1422–1425.
[42] F. Dupuis, L. Krmer, P. Faist, J. M. Renes, and R. Renner, in XVIIth International Congress on Mathematical Physics, edited by A. Jensen (World Scientific, 2013) pp. 134–153.
[43] R. Renner, Security of Quantum Key Distribution, Ph.D. thesis, PhD Thesis, 2005 (2005).
[44] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, *IEEE Transactions on Information Theory* **57**, 5524 (2011).

[45] J. M. Renes and R. Renner, *IEEE Transactions on Information Theory* **57**, 7377 (2011).

[46] J. M. Renes and R. Renner, *IEEE Transactions on Information Theory* **58**, 1985 (2012).

[47] L. Wang and R. Renner, *Physical Review Letters* **108**, 200501 (2012).

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

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

[50] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, arXiv:1207.0434 (2012).

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

[52] F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner, arXiv:1211.3141 (2012).

[53] G. E. Crooks, *Journal of Statistical Mechanics: Theory and Experiment* **10**, 23 (2008), arXiv:0706.1994 [cond-mat.stat-mech].

[54] P. Skrzypczyk, A. J. Short, and S. Popescu, ArXiv e-prints (2013), arXiv:1302.2811 [quant-ph].

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

[56] G. Gour, I. Marvian, and R. W. Spekkens, *Physical Review A* **80**, 012307 (2009).

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

[58] J. v. Tiel, *Convex analysis: an introductory text* (Wiley, Chichester [West Sussex]; New York, 1984).

[59] K. Li, *The Annals of Statistics* **42**, 171 (2014).

[60] 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.* (Pub. House of the Czechoslovak Academy of Sciences, Prague, 1962) pp. 698–723.

[61] Y. Polyanskiy, H. Poor, and S. Verdu, *IEEE Transactions on Information Theory* **56**, 2307 (2010).

[62] M. Tomamichel and M. Hayashi, *IEEE Transactions on Information Theory* **59**, 7693 (2013).