# Work and reversibility in quantum thermodynamics

Álvaro M. Alhambra,[1] Stephanie Wehner,[2,3] Mark M. Wilde,[4] and Mischa P. Woods[1,2,3]

[1]*University College London, Department of Physics & Astronomy, London WC1E 6BT, United Kingdom*
[2]*QuTech, Delft University of Technology, Lorentzweg 1, 2611 CJ Delft, Netherlands*
[3]*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*
[4]*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy,
Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA*
(Dated: February 21, 2017)

It is a central question in quantum thermodynamics to determine how much work can be gained by a process that transforms an initial state $\rho$ to a final state $\sigma$. For example, we might ask how much work can be obtained by thermalizing $\rho$ to a thermal state $\sigma$ at temperature $T$ of an ambient heat bath. Here, we show that, for different sets of resource-theoretic thermodynamic operations involving catalysis, the amount of work is characterized by how reversible the process is. More specifically, the amount of work to be gained depends on how well we can return the state $\sigma$ to its original form $\rho$ *without* investing any work. We proceed to exhibit an explicit reversal operation in terms of the Petz recovery channel coming from quantum information theory. Our result establishes a quantitative link between the reversibility of thermodynamical processes and the corresponding work gain.

## I. INTRODUCTION

Quantum thermodynamics is experiencing a renaissance in which ideas from quantum information theory enable us to understand thermodynamics for even the smallest quantum systems. Our inability to apply statistical methods to a small number of particles and the presence of quantum coherences make this a challenging undertaking. Yet, we are now indeed able to construct very small quantum devices allowing us to probe such regimes experimentally [1–3]. Theoretical results studying the efficiency of small thermal machines [4–8], catalysis [9–11], work extraction [12–19], and the second laws of quantum thermodynamics [20, 21] have furthermore led to the satisfying conclusion that the usual laws of thermodynamics as we know them can be derived from the laws of quantum mechanics in an appropriate limit.

Here we will be concerned with the fundamental problem of how much work is gained (or invested) by the transformation of a state $\rho_S$ to a state $\sigma_S$ of some system $S$ in the presence of a thermal bath. The second laws [20] provide general bounds on the amount of work, which are tight if $\rho_S$ is diagonal in the energy eigenbasis of the system. Special instances of this problem have drawn particular attention, such as *gaining* the maximum amount of work from $\rho_S$ by thermalizing it to the temperature of the surrounding bath [12], extracting work from correlations among different subsystems when $\rho_S$ is a multipartite state (see, e.g. [22]), as well as the case when $\sigma_S$ results from a measurement on $\rho_S$ [23–25]. When thinking about *investing* work, one of the most well studied instances is Landauer's erasure [26], which is concerned with the amount of energy necessary to take an arbitrary state $\rho_S$ to a pure state $\sigma_S$.

We adopt the resource theory approach of [12, 27, 28], which has the appealing feature of explictly accounting for all energy flows. Let us first establish what we mean by gaining work. In the macroscopic world, work is often illustrated as raising a weight by a certain amount. An analogue in the quantum regime corresponds to raising a system from a state of low energy to an excited state. We will call a system that is used to store energy a *battery*. The simplest model of such a battery is given by the 'wit' [12], a two-level system that is normalized such that the ground state $|0\rangle$ has energy '0', and the excited state $|1\rangle$ has energy $W$, corresponding to the amount of work $W$ we wish to store. That is, the Hamiltonian of the battery is given by $H_W = W|1\rangle\langle 1|$. Note that raising the system from the ground state to the excited state not only changes its energy by $W$, but this energy transfer takes place in a fully ordered form: we *know* the final state and can thus, in principle, transfer all of the energy onto a third system. This is analogous to the macroscopic notion of lifting a weight on a string and thus does not include any contributions from heat. Clearly, in a specific physical system, it can be difficult to realize an energy gap of precisely the amount of work $W$ that we wish to extract. However, one can imagine this two-level system as being part of a quasi-continuum battery from which we pick two levels with a suitable energy gap [29]. We refer to [29] for a discussion on (approximate) work in the microregime. When gaining work $W$ by transforming $\rho_S$ to $\sigma_S$, we are thus implementing the process

$$\rho_S \otimes |0\rangle\langle 0|_W \to \sigma_S \otimes |1\rangle\langle 1|_W . \tag{1}$$

The situation of investing work can be described similarly, except that instead of raising the system to the excited state, we draw energy from the battery by lowering it to the ground state. When investing work $W$ in order to enable

a transformation, we are thus implementing the process

$$\rho_S \otimes |1\rangle\langle 1|_W \to \sigma_S \otimes |0\rangle\langle 0|_W \ . \tag{2}$$

Within the resource theory approach to quantum thermodynamics, there are various different classes of thermodynamic operations. Most of the classes we will study in this paper have in common that, given a particular fixed temperature $T$, we may access a bath described by a Hamiltonian $H_B$ and thermal state $\hat{\tau}_B = \exp(-\beta H_B)/Z_B$, where $\beta = 1/(kT)$ is the inverse temperature, $k$ is the Boltzmann constant, and $Z_B$ is the partition function. To help us, we may also make use of catalysts in states $\eta_{C_i}$ with Hamiltonian $H_C = \oplus_{i=1}^n H_{C_i}$. Let $H_S$ be the Hamiltonian associated with the system $S$ and let $U$ denote a unitary that acts on the system $S$, the battery $W$, the catalysts $C = \otimes_{i=1}^n C_i$, and the bath $B$. The only unitary transformations $U$ that are allowed are those which conserve total energy. That is, the allowed unitaries are such that $[U, H] = 0$, where $H = H_S + H_W + H_C + H_B$ is the total Hamiltonian. The transformation $\mathcal{T}$ performing the mapping $\mathcal{T}(\rho_S \otimes |0\rangle\langle 0|_W) = \sigma_S \otimes |1\rangle\langle 1|_W$ then takes the following form

$$\mathcal{T}(\omega_{SW}) = \text{Tr}_{CB}[U(\omega_{SW} \otimes \eta_{C_1} \ldots \eta_{C_n} \otimes \hat{\tau}_B)U^\dagger] \tag{3}$$

for some input state $\omega_{SW}$ of the system and the battery. The resource theoretic operations we study in the paper, all differ in the nature of the physical restrictions imposed on the catalysts. For example, in the class of so-called catalytic thermal operations [20] we only allow for a single catalyst, and, in the regime of exact catalysis, we furthermore require that the catalyst is conserved exactly. That is, in such a case the output state $\omega_{SW}^{\text{out}}$ on $SW$ and the output state $\eta_C^{\text{out}}$ on $C$ satisfies

$$\text{Tr}_B[U(\omega_{SW} \otimes \eta_C \otimes \hat{\tau}_B)U^\dagger] = \omega_{SW}^{\text{out}} \otimes \eta_C^{\text{out}} \tag{4}$$

with $\eta_C = \eta_C^{\text{out}}$. Whenever the catalyst is absent, the allowed transformations are a subset that is simply called *thermal operations* [12]. The other classes of resource thermodynamic operations will be discussed in Sections V and VI. Given that $U$ conserves total energy, it is clear that this framework accounts for all energy flows, making it particularly appealing for studying quantum thermodynamics.

How much work could we gain by transforming $\rho_S$ to $\sigma_S$ using a bath of temperature $T$? It is clear from the above, that this question can be answered by asking about the largest value of $W_{\text{gain}}(\rho_S \to \sigma_S) = W$ that can be achieved by a thermodynamical operation belonging to the particular class in question, e.g. catalytic thermal operations, in the transition of Eq. (1). The standard second law tells us that this transformation is possible only if

$$F(\rho_S \otimes |0\rangle\langle 0|_W) \geq F(\sigma_S \otimes |1\rangle\langle 1|_W) \ , \tag{5}$$

where $F(\omega_{SW}) = \text{Tr}[H_{SW}\omega_{SW}] - kT\,S(\omega_{SW})$ is the Helmholtz free energy with $H_{SW} = H_S + H_W$ and $S(\omega_{SW}) = -\text{Tr}[\omega_{SW} \log \omega_{SW}]$ [1]. When $\rho_S$ is diagonal in the energy eigenbasis, then the standard second law does in fact give necessary and sufficient conditions for the desired transformation to exist whenever the systems are extremely large [28] or if we allow for a slightly inexact catalysis [20]. Specifically, if an arbitrary catalyst can be used, what we mean by this is that the "error per particle" in the output catalyst is bounded as $\|\eta_C - \eta_C^{\text{out}}\|_1 \leq \varepsilon/\log d_C$, where $d_C$ is the dimension of the catalyst and $\varepsilon > 0$ is some tolerance [20]. Similarly, if inexact catalysis takes on the form of allowing small correlations in the output catalyst, only the standard free energy is relevant [10]. This is the case of correlating-catalytic thermal operations, which we will study in Section V. While this regime of inexact catalysis is—due to an accumulation of errors—highly undesirable when analyzing cyclic or quasi-static processes as is the case in thermal machines [29], it is an otherwise very well motivated regime in physical implementations where an ever so slight error is essentially unavoidable. In the present Section this regime will be our primary focus.

Using the fact that $\text{Tr}[H_W|0\rangle\langle 0|_W] = 0$ and $\text{Tr}[H_W|1\rangle\langle 1|_W] = W_{\text{gain}}(\rho_S \to \sigma_S)$, it is easy to use (5) to obtain the following upper bound on the amount of work that we can hope to obtain

$$W_{\text{gain}}(\rho_S \to \sigma_S) \leq F(\rho_S) - F(\sigma_S) \ . \tag{6}$$

If $\rho_S$ is diagonal in the energy eigenbasis, then (6) is *tight*. It will be convenient to note [30] that the free energy can also be expressed in terms of the quantum relative entropy

$$D(\rho_S \| \tau_S) := \text{Tr}[\rho_S \log \rho_S] - \text{Tr}[\rho_S \log \tau_S], \tag{7}$$

defined as above when $\text{supp}(\rho_S) \subseteq \text{supp}(\tau_S)$ and equal to $+\infty$ otherwise. Note that here and throughout, we consider that the operator logarithm is evaluated only on the support of its argument. Specifically, $F(\rho_S) = kT[D(\rho_S \| \tau_S) -$

---

[1] All logarithms in this paper are base $e$.

$\log Z_S$], where $\tau_S = \exp(-\beta H_S)/Z_S$ is the thermal state of the system at the temperature $T$ of the ambient bath. Since we do not change the Hamiltonian of the system, we can hence express the amount of work in regimes where the standard free energy is relevant, as

$$W_{\text{gain}}(\rho_S \to \sigma_S) = kT\Delta , \tag{8}$$

where the difference $\Delta$ of relative entropies will play a very special role

$$\Delta = D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) . \tag{9}$$

Note that when $\Delta \geq 0$, the standard second law allows for the transformation $\rho_S \to \sigma_S$.

How about the case in which we need to *invest* work? In regimes where only the standard free energy is relevant to dictate the transition (2), a similar calculation yields that $W_{\text{inv}}(\rho_S \to \sigma_S) = kT(D(\sigma_S\|\tau_S) - D(\rho_S\|\tau_S))$. This means that in such regimes $W_{\text{inv}}(\rho_S \to \sigma_S) = -W_{\text{gain}}(\rho_S \to \sigma_S) = W_{\text{gain}}(\sigma_S \to \rho_S)$, i.e., the amount of energy that we need to invest to transform $\rho_S$ to $\sigma_S$ is precisely equal to the amount of work that we can gain by transforming $\sigma_S$ back to $\rho_S$. In the standard free energy regime, we thus see that we do not need to treat the amount of work invested as a separate case, but rather it can be understood fully in terms of the transformation of $\sigma_S$ back to $\rho_S$ in which work can be gained. It is useful to note that for systems $S$ that are truly small [12], or when we are interested in the case of *exact* catalysis, then this is not the case. In these situations, the standard second law needs to be augmented with more refined conditions [20] that lead to differences. With some abuse of terminology, we will refer to this as the *nano* regime. In place of just one free energy, the nano regime requires that a family of free energies $F_\alpha$ satisfies

$$F_\alpha(\rho_S) \geq F_\alpha(\sigma_S) , \tag{10}$$

for all $\alpha \geq 0$. These generalized free energies can be expressed in terms of the $\alpha$-Rényi divergences as

$$F_\alpha(\rho_S) = kT[D_\alpha(\rho_S\|\tau_S) - \log Z_S] , \tag{11}$$

where the general definition of $D_\alpha$ [2] takes on a simplified form if $\rho_S$ is diagonal in the energy eigenbasis. More precisely,

$$D_\alpha(\rho_S\|\tau_S) = \frac{1}{\alpha - 1} \log \sum_j \rho_j^\alpha \tau_j^{1-\alpha} , \tag{12}$$

where $\rho_j$ and $\tau_j$ are the eigenvalues of $\rho_S$ and $\tau_S$ respectively. The standard free energy is a member of this family for $\alpha \to 1$. A short calculation [20] yields that in this regime

$$W_{\text{gain}}^{\text{nano}}(\rho_S \to \sigma_S) \leq \inf_{\alpha \geq 0} kT \left( D_\alpha(\rho_S\|\tau_S) - D_\alpha(\sigma_S\|\tau_S) \right) , \tag{13}$$

$$W_{\text{inv}}^{\text{nano}}(\rho_S \to \sigma_S) \geq \sup_{\alpha \geq 0} kT \left( D_\alpha(\sigma_S\|\tau_S) - D_\alpha(\rho_S\|\tau_S) \right)$$

$$\geq kT \left( D(\sigma_S\|\tau_S) - D(\rho_S\|\tau_S) \right) , \tag{14}$$

where (the first) inequalities are again attained if $\rho_S$ is diagonal in the energy eigenbasis. In the nano-regime, it is thus possible that $W_{\text{gain}}^{\text{nano}} \neq W_{\text{inv}}^{\text{nano}}$.

## II. RESULT

**Gaining work.** Our main result is to prove that the amount of work $W_{\text{gain}}(\rho_S \to \sigma_S) \geq 0$ gained when transforming $\rho_S$ to $\sigma_S$ with a particular class of thermodynamic operations, can be characterized by how well we can recover the state $\rho_S$ from $\sigma_S$ using a thermodynamic operation of the same class which requires *no* work at all. This result applies to all situations in which $W$ is characterized by the standard free energy (see above). It forms a powerful link between the reversibility of the initial process and the amount of work drawn from it. Loosely speaking, we will see that if little work is obtained when transforming $\rho_S$ to $\sigma_S$ with a particular class of thermodynamic operation, then there

---

[2] For arbitrary states $\rho_S$, we have for $0 \leq \alpha < 1/2$ that $D_\alpha(\rho_S\|\tau_S) = \frac{1}{\alpha-1} \log \text{Tr}[\rho_S^\alpha \tau_S^{1-\alpha}]$ [31] and for $\alpha \geq 1/2$, $D_\alpha(\rho_S\|\tau_S) = \frac{1}{\alpha-1} \log \left[ \text{Tr} \left( \tau_S^{(1-\alpha)/2\alpha} \rho_S \tau_S^{(1-\alpha)/2\alpha} \right)^\alpha \right]$ [32, 33].

exists a thermodynamic operation of the same class that can recover $\rho_S$ from $\sigma_S$ quite well. Or stated differently, if this thermodynamic operation performs badly at recovering $\rho_S$, then the amount of work can be large. More precisely

$$W_{\text{gain}}(\rho_S \to \sigma_S) \geq -kT \log F(\rho_S, \mathcal{R}_{\rho \to \sigma}(\sigma_S)) \,, \tag{15}$$

where $F(\rho, \sigma) = \text{Tr}[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}]^2$ is the fidelity, and $\mathcal{R}_{\rho \to \sigma}$ is a reversal operation using a bath at temperature $T$, and (potentially) catalysts that takes $\sigma_S$ close to the original state $\rho_S$. $W_{\text{gain}}(\rho_S \to \sigma_S)$ is thus related to the reversal operation

$$\rho_S \overset{W_{\text{gain}}}{\underset{\mathcal{R}}{\rightleftarrows}} \sigma_S \,. \tag{16}$$

This reversal operation $\mathcal{R}_{\rho \to \sigma}$ is generated by using the same bath (and possibly catalysts too) as in the original, but reversing the global dynamics. That is, if the original operation is generated by a global unitary evolution $U(\cdot)U^\dagger$, the reversal operation is generated by $U^\dagger(\cdot)U$. Hence, no work is required to perform it. Below we also show that the reversal operation coincides with the *Petz recovery map* [34–37] from quantum information theory.

**Investing work.** As outlined above, in regimes where only the standard free energy is relevant, the amount of work $W_{\text{inv}}(\rho_S \to \sigma_S) \geq 0$ we need to invest to transform $\rho_S$ to $\sigma_S$ can be understood fully in terms of the work $W_{\text{gain}}(\sigma_S \to \rho_S)$ gained by the inverse process. This is not true in the nano regime. However, it is a nice application of our analysis to show that the reversal operation from the transformation $\sigma_S \to \rho_S$ that allows us to *gain* work, nevertheless allows us to make statements about $W_{\text{inv}}^{\text{nano}}$. Specifically,

$$W_{\text{inv}}^{\text{nano}}(\rho_S \to \sigma_S) \geq -kT \log F(\sigma_S, \mathcal{R}_{\sigma \to \rho}(\rho_S)) \,, \tag{17}$$

where $\mathcal{R}_{\sigma \to \rho}$ is the reversal operation constructed from the process that transforms $\sigma_S$ to $\rho_S$—from which work might be gained—while depositing such energy into the bath. That is, the reversal operation obtained from

$$\rho_S \overset{W_{\text{gain}}}{\underset{\mathcal{R}}{\leftrightharpoons}} \sigma_S \tag{18}$$

can be used to transform $\rho_S$ to $\sigma_S$ at *no energy cost* to the battery, and allows us to make statements about $W_{\text{inv}}$:

$$\rho_S \overset{W_{\text{inv}}}{\underset{\mathcal{R}}{\rightrightarrows}} \sigma_S \,. \tag{19}$$

Needless to say, since for example the erasure of a thermal state $\rho_S = \tau_S$ to a pure state $\sigma_S$ costs a significant amount of work, the operation $\mathcal{R}_{\sigma \to \rho}$ cannot always achieve a very high fidelity. This is indeed, precisely what we see here since if the fidelity is very small, then the amount of work that we need to invest is large. We illustrate this more subtle application of our result in Section IV by means of a simple example of a harmonic oscillator bath. We emphasize that the bound on $W_{\text{inv}}^{\text{nano}}$ is valid for all the classes thermodynamic operations we study in this paper.

## III. PROOF FOR THERMAL OPERATIONS (NO CATALYSTS)

We now give details of our main result for the simpler case of thermal operations (TO) without catalysts. Section V contains details of other, more general sets of operations.

Let us first suppose that we can draw a positive amount of work by transforming $\rho_S$ to $\sigma_S$, so that $\Delta > 0$. Note that in regimes dictated by the standard free energy, $\Delta > 0$ implies that there exists a different thermal operation taking $\rho_S$ to $\sigma_S$ without drawing any work at all [20, 21]—in this case the additional energy can be deposited into the bath. Let $V$ be the energy-conserving unitary that realizes this latter thermal operation, and let $(\hat{\tau}_B, H_B)$ be the thermal state and Hamiltonian of the bath, such that $\sigma_S = \text{Tr}_B[V(\rho_S \otimes \hat{\tau}_B)V^\dagger]$. Note that $V$ acts on systems $S$ and $B$ and $[V, H_S + H_B] = 0$. We have the following theorem:

**Theorem 1.** *Let $\mathcal{T}$ be a thermal operation given by*

$$\mathcal{T}(\cdot)_S = \text{Tr}_B[V((\cdot)_S \otimes \hat{\tau}_B)V^\dagger], \tag{20}$$

*where $V$ and $\hat{\tau}_B$ are defined above. Then it obeys the inequality*

$$D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S) \geq D(\rho_S \| \mathcal{R}(\sigma_S)) \geq -\log(F(\rho_S, \mathcal{R}(\sigma_S))), \tag{21}$$

*where $\mathcal{R}(\cdot)$ is a recovery map, which is another thermal operation given by*

$$\mathcal{R}(\cdot) = \text{Tr}_B[V^\dagger((\cdot)_S \otimes \hat{\tau}_B)V]. \tag{22}$$

*Proof.* Our proof is divided into two main steps.

**Step 1: Rewriting the relative entropy difference.**

Our first step will be to rewrite $\Delta = D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S)$ as an *equality* involving the operation $V$. Observe that

$$D(\rho_S\|\tau_S) = D(\rho_S \otimes \hat{\tau}_B \| \tau_S \otimes \hat{\tau}_B) \tag{23}$$

$$= D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \| V(\tau_S \otimes \hat{\tau}_B)V^\dagger) \tag{24}$$

$$= D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \| \tau_S \otimes \hat{\tau}_B), \tag{25}$$

where we have the used the facts that the relative entropy is invariant with respect to tensoring an ancilla state or applying a unitary, and $V$ is an energy-conserving unitary so that $V(\tau_S \otimes \hat{\tau}_B)V^\dagger = \tau_S \otimes \hat{\tau}_B$.

For density operators $\eta_{CD}$ and $\theta_{CD}$ such that $\mathrm{supp}(\eta_{CD}) \subseteq \mathrm{supp}(\theta_{CD})$, it is possible to write

$$D(\eta_{CD}\|\theta_{CD}) - D(\eta_D\|\theta_D) = \mathrm{Tr}(\eta_{CD}[\log \eta_{CD} - \log \theta_{CD} - \log I_C \otimes \eta_D + \log I_C \otimes \theta_D]) . \tag{26}$$

Using these two facts, we can rewrite $\Delta$ as follows:

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) = \mathrm{Tr}\left(V(\rho_S \otimes \hat{\tau}_B)V^\dagger[\log V(\rho_S \otimes \hat{\tau}_B^\epsilon)V^\dagger - \log \tau_S \otimes \hat{\tau}_B - \log \sigma_S \otimes I_B + \log \tau_S \otimes I_B]\right). \tag{27}$$

We can simplify the operator consisting of the last three terms on the right above as

$$-\log \tau_S \otimes \hat{\tau}_B - \log \sigma_S \otimes I_B + \log \tau_S \otimes I_B = -\log I_S \otimes \hat{\tau}_B - \log \sigma_S \otimes I_B \tag{28}$$

$$= -\log \sigma_S \otimes \hat{\tau}_B, \tag{29}$$

and thus conclude that

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) = D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \| \sigma_S \otimes \hat{\tau}_B). \tag{30}$$

Hence we have that the right-hand side is equal to

$$D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \| \sigma_S \otimes \hat{\tau}_B) = D(\rho_S \otimes \hat{\tau}_B \| V^\dagger(\sigma_S \otimes \hat{\tau}_B)V). \tag{31}$$

Putting everything together, we see that

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) = D(\rho_S \otimes \hat{\tau}_B \| V^\dagger(\sigma_S \otimes \hat{\tau}_B)V). \tag{32}$$

Thus, the quantity $\Delta$ related to the work gain in (8) is exactly equal to the "relative entropy distance" between the original state $\rho_S \otimes \hat{\tau}_B$ and the state resulting from the following thermal operation:

$$\sigma_S \to V^\dagger(\sigma_S \otimes \hat{\tau}_B)V, \tag{33}$$

which consists of adjoining $\sigma_S$ with a thermal state $\hat{\tau}_B$ and performing the inverse of the unitary $V$. Note that this statement is non-trivial, since $\sigma_S \otimes \hat{\tau}_B \neq V(\sigma_S \otimes \hat{\tau}_B)V^\dagger$. The forward operation $V$ can create correlations between the system and the bath, whereas $V^\dagger$ is applied to a fresh and entirely uncorrelated bath, making it a thermal operation.

**Step 2: A lower bound using the recovery map.** Due to the fact that the quantum relative entropy can never increase under the action of a partial trace [38, 39], we can conclude from (32) that the following inequality holds

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) \geq D(\rho_S \| \mathcal{R}_{\rho \to \sigma}(\sigma_S)) , \tag{34}$$

where

$$\mathcal{R}_{\rho \to \sigma}(\sigma_S) = \mathrm{Tr}_B[V^\dagger(\sigma_S \otimes \hat{\tau}_B)V] . \tag{35}$$

Note that this operation is a thermal operation, and requires no work. Our claim (15) now follows from the relation between the $\alpha$-Rényi divergences [32]

$$D(\rho\|\sigma) \geq D_{\frac{1}{2}}(\rho\|\sigma) = -\log F(\rho, \sigma) . \tag{36}$$

This concludes the proof. $\qquad\square$

**Investing work:** Our result for the case where we *invest* rather than *gain* work follows from the very same analysis as above by exchanging the roles of $\rho_S$ and $\sigma_S$: if $D(\sigma_S\|\tau_S) - D(\rho_S\|\tau_S) \geq 0$, then we could gain work by transforming $\sigma_S$ and $\rho_S$. The above argument yields a recovery map $\mathcal{R}_{\sigma \to \rho}$ that lower bounds the relative entropy difference. Our claim of (17) now follows from (14).

## A.  Remark: Petz recovery map

We remark that $\mathcal{R}$ is actually a special quantum map, called the Petz recovery map [34–37]. For a general quantum map $\mathcal{N}$ and a given density operator $\theta$, this recovery map is defined as

$$\tilde{\mathcal{N}}(\cdot) = \theta^{1/2}\mathcal{N}^{\dagger}[\mathcal{N}(\theta)^{-1/2}(\cdot)\mathcal{N}(\theta)^{-1/2}]\theta^{1/2}, \tag{37}$$

where $\mathcal{N}^{\dagger}$ is the adjoint of the channel $\mathcal{N}$ [40]. As a consequence, we can conclude that the main conjecture from [41] holds for the special case of thermal operations. We show this in the following lemma.

**Lemma 1.** *The map $\mathcal{R}(\cdot)$ in (35) is the Petz recovery map of the original thermal operation, provided we choose the $\theta$ in (37) to be the thermal state $\tau_S$.*

*Proof.* Consider that for two density operators $\eta$ and $\theta$ and a quantum channel $\mathcal{N}$, we can associate the relative entropy difference $D(\eta\|\theta) - D(\mathcal{N}(\eta)\|\mathcal{N}(\theta))$ and the Petz recovery channel:

$$\tilde{\mathcal{N}}(\cdot) = \theta^{1/2}\mathcal{N}^{\dagger}[\mathcal{N}(\theta)^{-1/2}(\cdot)\mathcal{N}(\theta)^{-1/2}]\theta^{1/2}. \tag{38}$$

For our case, we have that

$$\eta = \rho_S, \qquad \theta = \tau_S, \qquad \mathcal{N}(\cdot) = \mathrm{Tr}_B[V((\cdot)_S \otimes \hat{\tau}_B)V^{\dagger}], \tag{39}$$

which implies that $\mathcal{N}(\theta) = \tau_S$. Using the definition of the adjoint, one can show that

$$\mathcal{N}^{\dagger}(\cdot) = \mathrm{Tr}_B\left[\hat{\tau}_B^{1/2}V^{\dagger}[(\cdot)_S \otimes I_B]V\hat{\tau}_B^{1/2}\right], \tag{40}$$

which implies for our case that the Petz recovery channel takes the following form:

$$\tilde{\mathcal{N}}(\cdot) = \tau_S^{1/2}\mathrm{Tr}_B\left[\hat{\tau}_B^{1/2}V^{\dagger}[\tau_S^{-1/2}(\cdot)_S\tau_S^{-1/2} \otimes I_B]V\hat{\tau}_B^{1/2}\right]\tau_S^{1/2}. \tag{41}$$

We can rewrite this as follows:

$$\mathrm{Tr}_B\left[(\tau_S^{\frac{1}{2}} \otimes \hat{\tau}_B^{\frac{1}{2}})V^{\dagger}[\tau_S^{-\frac{1}{2}}(\cdot)_S\tau_S^{-\frac{1}{2}} \otimes I_B]V(\tau_S^{\frac{1}{2}} \otimes \hat{\tau}_B^{\frac{1}{2}})\right] = \mathrm{Tr}_B\left[(\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}}V^{\dagger}[\tau_S^{-\frac{1}{2}}(\cdot)_S\tau_S^{-\frac{1}{2}} \otimes I_B]V(\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}}\right] \tag{42}$$

$$= \mathrm{Tr}_B\left[V^{\dagger}(\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}}\tau_S^{-\frac{1}{2}}(\cdot)_S\tau_S^{-\frac{1}{2}} \otimes I_B](\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}}V\right] \tag{43}$$

$$= \mathrm{Tr}_B[V^{\dagger}((\cdot)_S \otimes \hat{\tau}_B)V], \tag{44}$$

where we have used that $[V, \tau_S \otimes \hat{\tau}_B] = 0$. $\square$

## IV.  EXAMPLE FOR THERMAL OPERATIONS

Let us illustrate the reversal operation $\mathcal{R}_{\sigma\to\rho}$ by means of a simple example. Let $S$ be a two-level system, with Hamiltonian $H_S = E_S|1\rangle\langle1|$. Suppose that $S$ is in the mixed state $\rho_S = p_0|0\rangle\langle0|_S + p_1|1\rangle\langle1|_S$ with $p_0 \in [1-e^{-\beta E_S}, 1]$, and we invest work $W_{\mathrm{inv}}^{\mathrm{nano}}(\rho_S \to \sigma_S)$ to bring the system to the ground state $\sigma_S = |0\rangle\langle0|_S$. That is, we are performing Landauer erasure. Recall that the reversal operation associated with the lower bound for the work invested (17) is determined by the operation that takes $\sigma_S = |0\rangle\langle0|_S$ to $\rho_S$ without drawing any work, but instead dumping the resulting energy into the bath.

For our simple example, consider a bath comprised of a harmonic oscillator $H_B = \sum_{n=0}^{\infty} E_n|n\rangle\langle n|_B$ where $E_n = n\hbar\omega$ [3]. Note that, for each $n$, the gap between $n$ and $n+1$ is constant: $G = E_{n+1} - E_n = \hbar\omega$. To illustrate, let us consider the energy gap of the system to be equal to $E_S = \hbar\omega$—an example in which $E_S$ is a multiple of $\hbar\omega$ is analogous.

––––––––

[3] We could have also written $E_n = (2n+1)\frac{\hbar}{2}\omega$, which is the same after re-normalizing. For notational convenience we have subtracted the constant $\frac{\hbar}{2}\omega$.

### 1. Transforming $\sigma_S$ to $\rho_S$

Our first goal is to find the explicit operation that takes $\sigma_S$ to $\rho_S$, that has the effect of mixing the ground state of the system. We will see that no catalyst is required for this transformation, that is $V = U \otimes \mathbb{1}_C$ where $U$ acts on the system and bath. Note that since $U$ conserves energy, $U$ is block diagonal in the energy eigenbasis belonging to different energies. More precisely, if the total Hamiltonian $H = H_S + H_B$ is block diagonal $H = \bigoplus_n E_n \Pi_{E_n}$ where $\Pi_{E_n}$ is the projector onto the subspace of energy $E_n = n\hbar\omega$ spanned by $|0\rangle_S |0\rangle_B$ for $n = 0$ and $\{|0\rangle_S |n\rangle_B, |1\rangle_S |n-1\rangle_B\}$ for $n = 1, 2, 3, \ldots$, then $U = \bigoplus_n U_{E_n}$ where $U_{E_n}$ is a unitary acting only on the subspace of energy $E_n$. That is, $\Pi_{E_n} U_{E_n} \Pi_{E_n} = U_{E_n}$.

Consider the unitary transformations $U_{E_n}$ defined by the following action:

$$U_{E_0} |0\rangle_S |0\rangle_B = |0\rangle_S |0\rangle_B =: |\Psi_{E_0}\rangle \ , \tag{45}$$

$$U_{E_n} |0\rangle_S |n\rangle_B = \sqrt{b} |0\rangle_S |n\rangle_B + \sqrt{1-b} |1\rangle_S |n-1\rangle_B =: |\Psi_{E_n}\rangle \quad \text{for } n = 1, 2, 3, \ldots \ , \tag{46}$$

$$U_{E_n} |1\rangle_S |n-1\rangle_B = \sqrt{1-b} |0\rangle_S |n\rangle_B - \sqrt{b} |1\rangle_S |n-1\rangle_B =: |\Psi_{E_n}^\perp\rangle \quad \text{for } n = 1, 2, 3, \ldots \ , \tag{47}$$

where $0 \leq b \leq 1$ is a parameter that will be chosen in accordance with the desired target state $\rho_S$ below. It will be useful to observe that in the subspace $\{|0\rangle_S |n\rangle_B, |1\rangle_S |n-1\rangle_B\}$, the unitary $U_{E_n}$ can be written as

$$U_{E_n} = \begin{pmatrix} \sqrt{b} & \sqrt{1-b} \\ \sqrt{1-b} & -\sqrt{b} \end{pmatrix} \ , \tag{48}$$

which makes it easy to see that $U = U^\dagger$ is Hermitian. Note that the states are normalized and $\langle\Psi_{E_n}|\Psi_{E_n}^\perp\rangle = 0$ for $n = 1, 2, 3, \ldots$. The bath thermal state is

$$\hat{\tau}_B = \frac{1}{Z_B} \sum_{n=0}^{\infty} e^{-nE_S\beta} |n\rangle\langle n|_B \ , \tag{49}$$

where $Z_B = \sum_{n=0}^{\infty} e^{-nE_S\beta} = 1/(1 - e^{-E_S\beta})$ is the partition function of the bath, and we have used the fact that $E_n = n\hbar\omega = nE_S$. The unitary thus transforms the overall state as

$$U(|0\rangle\langle 0|_S \otimes \hat{\tau}_B)U^\dagger = \frac{1}{Z_B} \sum_{n=0}^{\infty} e^{-nE_S\beta} U \left(|0\rangle_S |n\rangle_B \ _S\langle 0| \ _B\langle n|\right) U^\dagger \tag{50}$$

$$= \frac{1}{Z_B} \sum_{n=1}^{\infty} e^{-nE_S\beta} U \left(|0\rangle_S |n\rangle_B \ _S\langle 0| \ _B\langle n|\right) U^\dagger + \frac{1}{Z_B} U \left(|0\rangle_S |0\rangle_B \ _S\langle 0| \ _B\langle 0|\right) U^\dagger \tag{51}$$

$$= \frac{1}{Z_B} \sum_{n=1}^{\infty} e^{-nE_S\beta} |\Psi_{E_n}\rangle\langle\Psi_{E_n}| + \frac{1}{Z_B} |0\rangle\langle 0|_S \otimes |0\rangle\langle 0|_B \tag{52}$$

$$=: \rho_{SB}^0 \ . \tag{53}$$

By linearity of the partial trace operation, we have that

$$\text{Tr}_B(\rho_{SB}^0) = \frac{Z_B - 1}{Z_B} \left(b|0\rangle\langle 0|_S + (1-b)|1\rangle\langle 1|_S\right) + \frac{1}{Z_B} |0\rangle\langle 0|_S \tag{54}$$

$$= p_0 |0\rangle\langle 0|_S + p_1 |1\rangle\langle 1|_S \ , \tag{55}$$

where

$$p_0 = \frac{1}{Z_B} \left((Z_B - 1)b + 1\right) \ , \tag{56}$$

$$p_1 = 1 - p_0 \ . \tag{57}$$

Note that since $0 \leq b \leq 1$, $p_0 \in [1/Z_B, 1] = [1 - e^{-E_S\beta}, 1]$. Solving (56) for $b$ gives

$$b = \frac{p_0 Z_B - 1}{Z_B - 1} \ . \tag{58}$$

## 2. The reversal operation

Let us now construct the reversal map $\mathcal{R}_{\sigma \to \rho}$. Since no catalyst is needed, we can write the reversal map as

$$\mathcal{R}_{\sigma \to \rho}(\rho_S) = \mathrm{Tr}_B \left[ U^\dagger (\rho_S \otimes \hat{\tau}_B) U \right] = \mathrm{Tr}_B \left[ U(\rho_S \otimes \hat{\tau}_B) U^\dagger \right] , \tag{59}$$

where we have used the fact that $U = U^\dagger$. To evaluate the reversal map for arbitrary $\rho_S$, let us first note that by a calculation similar to the above

$$U(|1\rangle\langle 1|_S \otimes \hat{\tau}_B) U^\dagger = \frac{1}{Z_B} \sum_{n=0}^\infty e^{-nE_S\beta} |\Psi_{E_{n+1}}^\perp\rangle\langle\Psi_{E_{n+1}}^\perp| =: \rho_{SB}^1 . \tag{60}$$

Using the linearity of the partial trace, we furthermore observe that

$$\mathrm{Tr}_B \left[ \rho_{SB}^1 \right] = (1 - b)|0\rangle\langle 0|_S + b|1\rangle\langle 1|_S . \tag{61}$$

Using (53) and (60) together with (55) and (61), we then have

$$\mathcal{R}_{\sigma \to \rho}(\rho_S) = \mathrm{Tr}_B \left[ U(\rho_S \otimes \hat{\tau}_B) U^\dagger \right] \tag{62}$$
$$= p_0 \mathrm{Tr}_B \left[ \rho_{SB}^0 \right] + p_1 \mathrm{Tr}_B \left[ \rho_{SB}^1 \right] \tag{63}$$
$$= p_0 \left( p_0 |0\rangle\langle 0|_S + p_1 |1\rangle\langle 1|_S \right) + p_1 \left( (1 - b)|0\rangle\langle 0|_S + b|1\rangle\langle 1|_S \right) \tag{64}$$
$$= P_0^{\mathcal{R}} |0\rangle\langle 0|_S + P_1^{\mathcal{R}} |1\rangle\langle 1|_S , \tag{65}$$

with

$$P_1^{\mathcal{R}} := 1 - P_0^{\mathcal{R}} , \tag{66}$$

$$P_0^{\mathcal{R}} := (p_0)^2 + (p_1)^2 \frac{Z_B}{Z_B - 1} = (p_0)^2 + (1 - p_0)^2 e^{E_S\beta} , \tag{67}$$

where we have used the fact that $p_0 + p_1 = 1$ and $Z_B = 1/(1 - e^{-E_S\beta})$. We can now compute the lower bound for $W_{\mathrm{inv}}^{\mathrm{nano}}(\rho \to \sigma)$. We find

$$W_{\mathrm{inv}}^{\mathrm{nano}}(\rho \to \sigma) \geq -kT \log F(\sigma_S, \mathcal{R}_{\sigma \to \rho}(\rho_S)) \tag{68}$$
$$= -kT \log P_0^{\mathcal{R}} . \tag{69}$$

Plugging in (66) followed by (58) into (68) we find

$$W_{\mathrm{inv}}^{\mathrm{nano}}(\rho \to \sigma) \geq -kT \log \left[ (p_0)^2 + (1 - p_0)^2 e^{E_S\beta} \right] , \tag{70}$$

where we recall $p_0 \in [1/Z_B, 1] = [1 - e^{-E_S\beta}, 1]$.

## 3. Three special cases

We examine three special cases of (70):

1) Consider $p_0 = 1$. In this case we want to form the state $|0\rangle\langle 0|_S$ from the state $|0\rangle\langle 0|_S$. The work invested must clearly be zero in this case. The RHS of (70) is also zero, and hence the bound (17) is tight for this case.

2) Consider $p_0 = 1/Z_S = 1/(1 + e^{-E_S\beta})$. That is, we want to form a pure state from the thermal state $\rho_S = \tau_S$. In this case, the RHS of (70) simplifies to $-kT \log F(\sigma_S, \mathcal{R}_{\sigma \to \rho}(\rho_S)) = kT \log Z_S$. By direct calculation using the 2nd laws (using Eqs. (13)–(14)) we find $W_{\mathrm{gain}}^{\mathrm{nano}} = W_{\mathrm{inv}}^{\mathrm{nano}} = (\log Z_S)/\beta$ and thus the bound is also tight for this case.

3) Consider $p_0 = 1/Z_B$. That is, we want to form a pure state from the state whose ground state population is the same as the ground state population of the harmonic oscillator bath. In this case, (70) reduces to $-kT \log F(\sigma_S, \mathcal{R}_{\sigma \to \rho}(\rho_S)) = -kT \log[1 + e^{-2E_S\beta} - e^{-E_S\beta}]$.
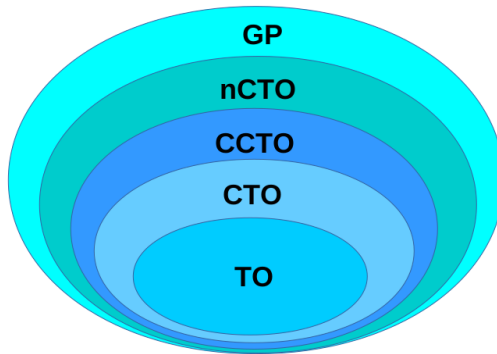
FIG. 1. We show how the different classes of resource theoretic thermodynamic operations are included within each other. The most general ones are the maps that preserve the Gibbs state of a system, known as Gibbs Preserving maps (GP), followed by those in which we allow for an arbitrary number of catalysts (CCTO, and the slightly wider darker region where we allow the catalysts to correlate with the system, nCTO), followed by catalytic thermal operations (CTO) with a single catalyst, and finally the simplest one, thermal operations (TO). If one has access to the full set of CCTO, then the standard second law determines when transitions are possible, and hence Eq. (8) holds. In the main section, we show that the recovery map of an operation within one of the classes is inside the same class, and thus if one can implement all the operations of a specific class, they can also implement the corresponding recovery operation without the need for additional resources

.

## V.   EXTENDING TO MORE GENERAL OPERATIONS INVOLVING CATALYSTS

In this section, we show how the results of Section III hold for increasingly general sets of resource theoretic thermodynamic operations. In particular, we give proofs for classes of thermodynamic operations consisting of:

1. Catalytic thermal operations (CTO), with a single catalyst in a state that is left unchanged at the end, as introduced in [20].

2. The regime of approximate catalytic thermal operations with various catalysts, where their local states do not change, but can become correlated [10]. We refer to this class as "correlating catalytic thermal operations"(CCTO).

3. Operations with various catalysts where they also become correlated with the system, which is a potentially physically relevant extension that has not been previously defined. We will call this more general class *n-catalytic thermal operations* (nCTO).

4. The most general kind of maps that can be taken as free operations of a thermodynamic resource theory [42]: the set of maps that leave the thermal state invariant, called Gibbs-preserving maps (GP). A resource theory defined by these free operations fits the requirements of [43].

We show how each set is included in the others in the diagram of Fig 1.

For 3), it is known [10] that even in the nano regime the free energy is recovered as the unique measure for thermodynamical processes. This case is hence of particular importance in this context, as the main results of the present work are written in terms of the free energy: if one has access to the full set of such operations, then we are in a regime in which Eq. (8) holds.

First we prove the following lemma, which highlights a particular feature shared by most of maps for which the main results can be shown. For the last, more general, set of operations, Gibbs preserving maps, we give a separate proof in Section VI, as the result differs in form slightly.

**Lemma 2.** *Let $\mathcal{T}(\cdot)$ be a quantum channel with a full-rank steady state $\tau_S = \mathcal{T}(\tau_S)$ and a Stinespring dilation*

$$\mathcal{T}(\cdot)_S = \mathrm{Tr}_E[U((\cdot)_S \otimes \rho_E)U^\dagger], \tag{71}$$

*for some unitary $U$ and an environment state $\rho_E$, such that*

$$U(\tau_S \otimes \rho_E)U^\dagger = \tau_S \otimes \rho_E. \tag{72}$$

*That is, at the fixed point no correlations with the environment are created. It then holds that for an arbitrary initial state $\rho_S$, and $\sigma_S = \mathcal{T}(\rho_S)$*

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) \geq D(\rho_S\|\mathcal{R}(\sigma_S)), \tag{73}$$

*where $\mathcal{R}(\cdot)$ is the Petz recovery map, given by*

$$\mathcal{R}(\cdot) = \mathrm{Tr}_E[U^\dagger((\cdot)_S \otimes \rho_E)U]. \tag{74}$$

*Proof.* Our proof follows the steps of the particular case of thermal operations shown previously. We first write

$$D(\rho_S\|\tau_S) = D(\rho_S \otimes \rho_E\|\tau_S \otimes \rho_E) \tag{75}$$

$$= D(U(\rho_S \otimes \rho_E)U^\dagger\|U(\tau_S \otimes \rho_E)U^\dagger) \tag{76}$$

$$= D(U(\rho_S \otimes \rho_E)U^\dagger\|\tau_S \otimes \rho_E), \tag{77}$$

where we have used the main assumption of the lemma in (72) and the facts that the relative entropy is invariant with respect to tensoring an ancilla state or applying a unitary.

Now we recall the identity of (26) from the proof of Theorem 1:

$$D(\eta_{CD}\|\theta_{CD}) - D(\eta_D\|\theta_D) = \mathrm{Tr}(\eta_{CD}[\log \eta_{CD} - \log \theta_{CD} - \log I_C \otimes \eta_D + \log I_C \otimes \theta_D]), \tag{78}$$

where $\mathrm{supp}(\eta_{CD}) \subseteq \mathrm{supp}(\theta_{CD})$. We use it together with (75) to write

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) = \mathrm{Tr}(U(\rho_S \otimes \rho_E)U^\dagger[\log U(\rho_S \otimes \rho_E)U^\dagger - \log \tau_S \otimes \rho_E - \log \sigma_S \otimes I_E + \log \tau_S \otimes I_E]). \tag{79}$$

The last three terms on the right-hand side above can be simplified significantly

$$- \log \tau_S \otimes \rho_E - \log \sigma_S \otimes I_E + \log \tau_S \otimes I_E = - \log I_S \otimes \rho_E - \log \sigma_S \otimes I_E \tag{80}$$

$$= - \log \sigma_S \otimes \rho_E, \tag{81}$$

which leads to

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) = D(U(\rho_S \otimes \rho_E)U^\dagger\|\sigma_S \otimes \rho_E) \tag{82}$$

We also have that

$$D(U(\rho_S \otimes \rho_E)U^\dagger\|\sigma_S \otimes \rho_E) = D(\rho_S \otimes \rho_E\|U^\dagger(\sigma_S \otimes \rho_E)U). \tag{83}$$

Putting everything together, we see that

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) = D(\rho_S \otimes \rho_E\|U^\dagger(\sigma_S \otimes \rho_E)U) \tag{84}$$

$$\geq D(\rho_S\|\mathcal{R}(\sigma_S)). \tag{85}$$

What is left is to show that $\mathcal{R}$ is indeed the Petz recovery map. Again this follows through the same reasoning as it does for thermal operations.

The adjoint of the map $\mathcal{T}(\cdot)$ is as follows

$$(\cdot)_S \to \mathrm{Tr}_E\left[\rho_E^{1/2}U^\dagger((\cdot)_S \otimes I_E)U\rho_E^{1/2}\right], \tag{86}$$

by definition the Petz recovery channel is given as

$$(\cdot)_S \to \tau_S^{1/2}\mathrm{Tr}_E\left[\rho_E^{1/2}U^\dagger\left(\tau_S^{-1/2}(\cdot)_S\tau_S^{-1/2} \otimes I_E\right)U\rho_E^{1/2}\right]\tau_S^{1/2}. \tag{87}$$

By a series of steps similar to those shown previously, we have that

$$\tau_S^{1/2}\mathrm{Tr}_E\left[\rho_E^{1/2}U^\dagger\left(\tau_S^{-1/2}(\cdot)_S\tau_S^{-1/2} \otimes I_E\right)U\rho_E^{1/2}\right]\tau_S^{1/2}$$

$$= \mathrm{Tr}_E\left[(\tau_S \otimes \rho_E)^{1/2}U^\dagger\left(\tau_S^{-1/2}(\cdot)_S\tau_S^{-1/2} \otimes I_E\right)U(\tau_S \otimes \rho_E)^{1/2}\right] \tag{88}$$

$$= \mathrm{Tr}_E\left[U^\dagger(\tau_S \otimes \rho_E)^{1/2}\left(\tau_S^{-1/2}(\cdot)_S\tau_S^{-1/2} \otimes I_E\right)(\tau_S \otimes \rho_E)^{1/2}U\right] \tag{89}$$

$$= \mathrm{Tr}_E\left[U^\dagger((\cdot)_S \otimes \rho_E)U\right] \tag{90}$$

$$= \mathcal{R}(\cdot), \tag{91}$$

We note that the matrix $(\tau_S \otimes \rho_E)^{1/2}$ is diagonal in the same basis as $\tau_S \otimes \rho_E$, and both have the same degeneracies. Hence, if the latter commutes with $U$, so does the former. This allows us to go from the second to the third line by using the assumption of (72). $\square$

This lemma implies that for any quantum channel that has a dilation satisfying the condition in (72), we can show a generalization of (34). In the next lemma, we develop a further relevant set of maps for which the condition in (72) holds. We call these *n-catalytic thermal operations* (nCTO), which are defined through an energy-conserving unitary $V$ acting on the system $S$, a set $C$ of $n$ catalysts $C = \bigotimes_{i=1}^{n} C_i$ with initial states $\eta_{C_i}$, and the bath $B$, such that

$$\text{Tr}_B\left[V\left(\rho_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger\right] = \sigma_{SC}\,, \tag{92}$$

where $\text{Tr}_C[\sigma_{SC}] = \sigma_S$, and we impose the constraint $\text{Tr}_{S\,C\backslash C_l}[\sigma_{SC}] = \eta_{C_i}\,\forall i$, meaning that the local states of the catalysts return to their original form. The unitary conserves the energy of bath, system, and all the catalysts, so that $[V, H_S + H_B + \sum_{i=1}^{n} H_{C_i}] = 0$, but correlations between the different parts may be allowed in the final state. We can then show the following:

**Lemma 3.** *For a given n-catalytic thermal operation (nCTO) of the form in* (92)*, if the input state $\rho_S$ is the thermal state $\tau_S$ (the fixed point), then*

$$V\left(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger = \tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\,. \tag{93}$$

*Proof.* Let $\hat{\rho}_B = \text{Tr}_{SC}[V\left(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger]$ denote the local state of the bath after the transformation, and denote the total Hamiltonian as $H = H_S + H_B + \sum_{i=1}^{n} H_{C_i}$, the sum of all the local ones. Conservation of energy before and after the operation corresponds to the following:

$$\text{Tr}[HV\left(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger] = \text{Tr}[H(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i})] = \text{Tr}[H_S \tau_S] + \text{Tr}[H_B \hat{\tau}_B] + \sum_{i=1}^{n} \text{Tr}[H_{C_i} \eta_{C_i}]. \tag{94}$$

Also, the total average energy is the sum of the local energies

$$\text{Tr}[HV\left(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger] = \text{Tr}[H_S \tau_S] + \text{Tr}[H_B \hat{\rho}_B] + \sum_{i=1}^{n} \text{Tr}[H_{C_i} \eta_{C_i}], \tag{95}$$

and hence $\text{Tr}[H_B \hat{\rho}_B] = \text{Tr}[H_B \hat{\tau}_B]$. For a given fixed energy, the thermal state is the unique state with the highest entropy [44, Theorem 1.3], and hence $S(\hat{\tau}_B) \geq S(\hat{\rho}_B)$.

Now we consider the entropy $S(\rho) = -\text{Tr}[\rho \log \rho]$ before and after the transformation. Since the joint operation is a unitary, we have from unitary invariance and subadditivity of quantum entropy that

$$S(\tau_S) + S(\hat{\tau}_B) + \sum_{i=1}^{n} S(\eta_{C_i}) = S(V\left(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger) \tag{96}$$

$$\leq S(\tau_S) + S(\hat{\rho}_B) + \sum_{i=1}^{n} S(\eta_{C_i}), \tag{97}$$

and hence that $S(\hat{\tau}_B) \leq S(\hat{\rho}_B)$. Given our conclusion in (94) regarding conservation of energy, it must then be the case that $S(\hat{\tau}_B) = S(\hat{\rho}_B)$, and thus that the inequality in (97) is an equality. Saturation of subadditivity of entropy in (97) occurs if and only if

$$V\left(\tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}\right)V^\dagger = \tau_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}, \tag{98}$$

completing the proof. □

Putting together Lemmas 2 and 3 and taking the environment state $\rho_E$ from Lemma 2 to be the state of the bath and the set of catalysts (i.e., $\rho_E \equiv \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i}$), we arrive at the following conclusion:

**Theorem 2.** *Let $\mathcal{T}$ be an nCTO of the form in* (92) *and given by*

$$\mathcal{T}(\cdot) = \text{Tr}_{BC}[U((\cdot)_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i})U^\dagger]. \tag{99}$$

*Then it obeys the inequality*

$$D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) \geq D(\rho_S\|\mathcal{R}(\sigma_S)), \tag{100}$$

*where $\mathcal{R}(\cdot)$ is the Petz recovery map, given by*

$$\mathcal{R}(\cdot) = \text{Tr}_{BC}[U^\dagger((\cdot)_S \otimes \hat{\tau}_B \otimes_{i=1}^{n} \eta_{C_i})U], \tag{101}$$

*which is a thermodynamic operation that belongs to the same class as the original operation $\mathcal{T}$ (so for instance, the recovery map of a CTO is also a CTO), but with the global reversible dynamics generated by inverse unitary $U^\dagger$ instead.*

There are three relevant subsets of $n$-catalytic thermal operations to which Theorem 2 applies:

1. Catalytic thermal operations, as defined in [20], where we have a single catalyst that does not become correlated with the system.

2. Catalytic thermal operations with multiple catalysts that may become correlated between themselves, as defined in [10].

3. Operations where the catalysts and the system are allowed to correlate. This set, although slightly more general and possibly powerful, has not been studied in detail up to date.

As already pointed above, the second kind is of particular interest because it can be shown that if one uses three or more catalysts in a process in such a way that they correlate between themselves, while their local states remain unchanged, the criteria for when thermodynamical transitions are possible reduces to the standard free energy decreasing [10], as in (5).

We have shown that in this context, the Petz recovery map can be understood as the map generated by the inverse of a global dynamics, which can be understood as a *time reversal* of the original map. This fact has been noted in quantum thermodynamics in the context of fluctuation theorems [45–47].

## VI. GIBBS PRESERVING MAPS

The only set of maps to which the conditions of Lemma 2 do not apply is the most general set of Gibbs preserving maps [42], and we hence need a different method. To prove a bound of the form of (21), we use the following general result for quantum maps:

**Theorem 3** (From [48]). *Let $\mathcal{N}(\cdot)$ be a quantum channel, and let $\eta$ and $\theta$ be quantum states. We have that*

$$D(\eta\|\theta) - D(\mathcal{N}(\eta)\|\mathcal{N}(\theta)) \geq -2\log F(\eta, \int_{\mathbb{R}} dt \; p(t) \; \tilde{\mathcal{N}}_t(\mathcal{N}(\eta))), \tag{102}$$

*where $F(\rho, \sigma) = \mathrm{Tr}[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}]$ is the quantum fidelity, the map $\tilde{\mathcal{N}}_t$ is the following* rotated *recovery map*

$$\tilde{\mathcal{N}}_t(\cdot) = \theta^{it/2}\tilde{\mathcal{N}}(\mathcal{N}(\theta)^{-it/2}(\cdot)\mathcal{N}(\theta)^{it/2})\theta^{-it/2}, \tag{103}$$

*with $\tilde{\mathcal{N}}$ defined as in (38) and $p(t) = \frac{\pi}{2}(\cosh(\pi t) + 1)^{-1}$ is a probability density function.*

It can be seen by inspection that if we take the map $\mathcal{N}$ to be Gibbs-preserving so that $\tau_S = \mathcal{N}(\tau_S)$ and if we set $\theta = \tau_S$, then the rotated recovery map is Gibbs-preserving as well, namely $\tilde{\mathcal{N}}_t(\tau_S) = \tau_S$. More explicitly, the bound on $\Delta$ is as follows:

$$\Delta = D(\rho_S\|\tau_S) - D(\sigma_S\|\tau_S) \geq -2\log F(\rho_S, \int_{\mathbb{R}} dt \; p(t)\tilde{\mathcal{N}}_t(\sigma_S)), \tag{104}$$

where instead of having the relative entropy, we have the fidelity in the lower bound for the decrease of free energy.

## VII. CONCLUSION

We have shown that the amount of work drawn from a process is directly linked to the reversibility of the process. Specifically, we see that there are some regimes where if the amount of work drawn is small, then there exists a recovery operation that approximately restores the system to its initial state at no work cost at all. It is interesting to note that our result also makes a statement about the standard second law whenever $\rho_S$ can be transformed to $\sigma_S$. When $\Delta$ is small, then the transformation $\rho_S$ to $\sigma_S$ can be reversed very well by the map $\mathcal{R}_{\rho\to\sigma}$.

Our main result applies to all regimes where the standard free energy is relevant, and it is a very interesting open question to extend our result to regimes in which we require the full set of second laws [20]. What makes this question challenging is that (26) does not carry over to the regime of $D_\alpha$ for $\alpha \neq 1$, and indeed recent work [49] suggests that other quantities naturally generalize the *difference* of relative entropies—and this generalization does not always result in the difference of $\alpha$-Rényi relative entropies. It hence forms a more fundamental challenge to understand whether the difference of such $\alpha$-relative entropies, or the quantities suggested in [49] should be our starting point. However, the quantities in [49] would require a proof of a new set of second laws.

We have applied our analysis to the case of investing work, which in the regime where only the standard free energy is relevant can be characterized fully by how much work can be *gained* by the inverse process. This relation to the inverse process is not true in the nano-regime where all the refined second laws of [20] become relevant. Nevertheless, we have shown that the reversal operation of said inverse process can indeed be used to understand the amount of work that needs to be invested, adding another piece to the growing puzzle that is quantum thermodynamics.

Since this initial work, there has been a series of recent results consisting in giving lower bounds to the decrease of relative entropy in different cases of interest, covering a number of different branches within quantum information theory, such as [50–54], which shows the importance of the concept of recoverability and of recovery maps.

## VIII.   ACKNOWLEDGMENTS

[1] Jonne V. Koski, Ville F. Maisi, Jukka P. Pekola, and Dmitri V. Averin. Experimental realization of a Szilard engine with a single electron. *Proceedings of the National Academy of Sciences*, 111(38):13786–13789, 2014. arXiv:1402.5907.

[2] Ignacio A. Martínez, Édgar Roldán, Luis Dinis, Dmitri Petrov, and Raúl A. Rica. Adiabatic processes realized with a trapped Brownian particle. *Physical Review Letters*, 114(12):120601, March 2015. arXiv:1409.7578.

[3] Ignacio A. Martínez, Édgar Roldán, Luis Dinis, Dmitri Petrov, Juan M. R. Parrondo, and Raúl A. Rica. Brownian Carnot engine. *Nature Physics*, 12(1):67–70, 2016. arXiv:1412.1282.

[4] Hiroyasu Tajima and Masahito Hayashi. Refined Carnot's theorem: Asymptotics of thermodynamics with finite-size heat baths. 2014. arXiv:1405.6457.

[5] Paul Skrzypczyk, Nicolas Brunner, Noah Linden, and Sandu Popescu. The smallest refrigerators can reach maximal efficiency. *Journal of Physics A: Mathematical and Theoretical*, 44(49):492002–492008, December 2011. arXiv:1009.0865.

[6] Nicolas Brunner, Marcus Huber, Noah Linden, Sandu Popescu, Ralph Silva, and Paul Skrzypczyk. Entanglement enhances cooling in microscopic quantum refrigerators. *Physical Review E*, 89(3):032115, March 2014. arXiv:1305.6009.

[7] Kay Brandner, Michael Bauer, Michael T Schmid, and Udo Seifert. Coherence-enhanced efficiency of feedback-driven quantum engines. *New Journal of Physics*, 17(6):065006, June 2015. arXiv:1503.04865.

[8] Mark T. Mitchison, Mischa P. Woods, Javier Prior, and Marcus Huber. Coherence-assisted single-shot cooling by quantum absorption refrigerators. *New Journal of Physics*, 17(11):115013, 2015. arXiv:1504.01593.

[9] Nelly Ng, Laura Mančinska, Cristina Cirstoiu, Jens Eisert, and Stephanie Wehner. Limits to catalysis in quantum thermodynamics. *New Journal of Physics*, 17(8):085004, 2015. arXiv:1405.3039.

[10] Matteo Lostaglio, Markus P. Mueller, and Michele Pastena. Extracting work from absence of correlations. *Physical Review Letters*, 115(15):150402, 2015. arXiv:1409.3258.

[11] Johan Åberg. Catalytic coherence. *Physical Review Letters*, 113(15):150402, October 2014. arXiv:1304.1060.

[12] Michal Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nano thermodynamics. *Nature Communications*, 4:2059, June 2013. arXiv:1111.3834.

[13] Johan Åberg. Truly work-like work extraction via a single-shot analysis. *Nature Communications*, 4:1925, June 2013.

[14] Oscar Dahlsten, Renato Renner, Elisabeth Rieper, and Vlatko Vedral. Inadequacy of von Neumann entropy for characterizing extractable work. *New Journal of Physics*, 13(5):053015, May 2011. arXiv:0908.0424.

[15] Rodrigo Gallego, Jens Eisert, and Henrik Wilming. Thermodynamic work from operational principles. *New Journal of Physics*, 18(10):103017, October 2016. arXiv:1504.05056.

[16] Jochen Gemmer and Janet Anders. From single-shot towards general work extraction in a quantum thermodynamic framework. *New Journal of Physics*, 17(8):085006, 2015. arXiv:1504.05061.

[17] Marlan O. Scully, M. Suhail Zubairy, Girish S. Agarwal, and Herbert Walther. Extracting work from a single heat bath via vanishing quantum coherence. *Science*, 299(5608):862–864, February 2003.

[18] Felix Binder, Sai Vinjanampathy, Kavan Modi, and John Goold. Quantum thermodynamics of general quantum processes. *Physical Review E*, 91(3):032119, March 2015. arXiv:1406.2801.

[19] Sina Salek and Karoline Wiesner. Fluctuations in single-shot $\varepsilon$-deterministic work extraction. 2015. arXiv:1504.05111.

[20] Fernando G. S. L. Brandão, Michal Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. The second laws of quantum thermodynamics. *Proceedings of the National Academy of Sciences*, 112(11):3275–3279, March 2015. arXiv:1305.5278.

[21] Matteo Lostaglio, Kamil Korzekwa, David Jennings, and Terry Rudolph. Quantum coherence, time-translation symmetry, and thermodynamics. *Physical Review X*, 5(2):021001, April 2015. arXiv:1410.4572.

[22] Marti Perarnau-Llobet, Karen V. Hovhannisyan, Marcus Huber, Paul Skrzypczyk, Nicolas Brunner, and Antonio Acin. Extractable work from correlations. *Physical Review X*, 5(4):041011, 2015. arXiv:1407.7765.

[23] Philipp Kammerlander and Janet Anders. Coherence and measurement in quantum thermodynamics. *Scientific Reports*, 6:22174, February 2016. arXiv:1502.02673.

[24] Takahiro Sagawa and Masahito Ueda. Second law of thermodynamics with discrete quantum feedback control. *Physical Review Letters*, 100(8):080403, February 2008. arXiv:0710.0956.

[25] Kurt Jacobs. Second law of thermodynamics and quantum feedback control: Maxwell's demon with weak measurements. *Physical Review A*, 80(1):012322, July 2009. arXiv:0906.4146.

[26] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 1961.

[27] Dominik Janzing, Pawel Wocjan, Robert Zeier, Rubino Geiss, and Thomas Beth. The thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the second law. *International Journal of Theoretical Physics*, 39(12):2217–2753, 2000. arXiv:quant-ph/0002048.

[28] Fernando G. S. L. Brandão, Michal Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens. Resource theory of quantum states out of thermal equilibrium. *Physical Review Letters*, 111(25):250404, December 2013. arXiv:1111.3882.

[29] Mischa Woods, Nelly Ng, and Stephanie Wehner. The maximum efficiency of nano heat engines depends on more than temperature. June 2015. arXiv:1506.02322.

[30] Matthew J. Donald. Free energy and the relative entropy. *Journal of Statistical Physics*, 49(1-2):81–87, October 1987.

[31] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports in Mathematical Physics*, 23(1):57–65, February 1986.

[32] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: a new definition and some properties. *Journal of Mathematical Physics*, 54(12):122203, December 2013. arXiv:1306.3142.

[33] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, October 2014. arXiv:1306.1586.

[34] Dénes Petz. Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Communications in Mathematical Physics*, 105(1):123–131, 1986.

[35] Dénes Petz. Sufficiency of channels over von Neumann algebras. *The Quarterly Journal of Mathematics*, 39(1):97–108, 1988.

[36] Howard Barnum and Emanuel Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43(5):2097–2106, May 2002. arXiv:quant-ph/0004088.

[37] Patrick Hayden, Richard Jozsa, Denes Petz, and Andreas Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Communications in Mathematical Physics*, 246(2):359–374, April 2004. arXiv:quant-ph/0304007.

[38] Goran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, 1975.

[39] Armin Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics*, 54(1):21–32, 1977.

[40] John Watrous. Theory of quantum information lecture notes. https://cs.uwaterloo.ca/~watrous/LectureNotes.html, Fall 2011.

[41] Andreas Winter and Ke Li. Squashed entanglement, k-extendibility, quantum Markov chains, and recovery maps. October 2014. arXiv:1410.4184.

[42] Philippe Faist, Jonathan Oppenheim, and Renato Renner. Gibbs-preserving maps outperform thermal operations in the quantum regime. *New Journal of Physics*, 17(4):043003, 2015. arXiv:1406.3618.

[43] Fernando G. S. L. Brandão and Gilad Gour. Reversible framework for quantum resource theories. *Physical Review Letters*, 115(7):070503, 2015. arXiv:1502.03149.

[44] Eric A. Carlen. Trace inequalities and quantum entropy: An introductory course. *Contemporary Mathematics*, 529:73–140, 2010.

[45] Gavin E Crooks. Quantum operation time reversal. *Physical Review A*, 77(3):034101, 2008.

[46] Johan Aberg. Fully quantum fluctuation theorems. *arXiv preprint arXiv:1601.01302*, 2016.

[47] Álvaro M. Alhambra, Lluis Masanes, Jonathan Oppenheim, and Christopher Perry. Fluctuating work: From quantum thermodynamical identities to a second law equality. *Phys. Rev. X*, 6:041017, Oct 2016.

[48] Marius Junge, Renato Renner, David Sutter, Mark M. Wilde, and Andreas Winter. Universal recovery from a decrease of quantum relative entropy. 2015. arXiv:1509.07127.

[49] Kaushik P. Seshadreesan, Mario Berta, and Mark M. Wilde. Rényi squashed entanglement, discord, and relative entropy differences. *Journal of Physics A: Mathematical and Theoretical*, 48(39):395303, 2015. arXiv:1410.1443.

[50] Mario Berta, Stephanie Wehner, and Mark M Wilde. Entropic uncertainty and measurement reversibility. *New Journal of Physics*, 18(7):073004, 2016.

[51] Francesco Buscemi, Siddhartha Das, and Mark M Wilde. Approximate reversibility in the context of entropy gain, information gain, and complete positivity. *Physical Review A*, 93(6):062314, 2016.

[52] Álvaro M Alhambra and Mischa P Woods. Dynamical maps, quantum detailed balance and petz recovery map. *arXiv preprint arXiv:1609.07496*, 2016.

[53] Marius Lemm and Mark M Wilde. Information-theoretic limitations on approximate quantum cloning and broadcasting. *arXiv preprint arXiv:1608.07569*, 2016.

[54] Iman Marvian and Seth Lloyd. From clocks to cloners: Catalytic transformations under covariant operations and recoverability. *arXiv preprint arXiv:1608.07325*, 2016.