

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Work and reversibility in quantum thermodynamics

Álvaro M. Alhambra, Stephanie Wehner, Mark M. Wilde, and Mischa P. Woods

Phys. Rev. A **97**, 062114 — Published 15 June 2018

DOI: [10.1103/PhysRevA.97.062114](https://doi.org/10.1103/PhysRevA.97.062114)

# Work and Reversibility in Quantum Thermodynamics

Álvaro M. Alhambra,<sup>1</sup> Stephanie Wehner,<sup>2,3</sup> Mark M. Wilde,<sup>4</sup> and Mischa P. Woods<sup>1,2,3</sup>

<sup>1</sup>*University College London, Department of Physics & Astronomy, London WC1E 6BT, United Kingdom*

<sup>2</sup>*QuTech, Delft University of Technology, Lorentzweg 1, 2611 CJ Delft, Netherlands*

<sup>3</sup>*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*

<sup>4</sup>*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA*

It is a central question in quantum thermodynamics to determine how irreversible is a process that transforms an initial state  $\rho$  to a final state  $\sigma$ , and whether such irreversibility can be thought of as a useful resource. 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, the amount of entropy produced along a transition is characterized by how reversible the process is. More specifically, this entropy production depends on how well we can return the state  $\sigma$  to its original form  $\rho$  without investing any work. At the same time, the entropy production can be linked to the work that can be extracted along a given transition, and we explore the consequences that this fact has for our results. We also 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 [4–9], catalysis [10–12], work extraction [13–20], and the second laws of quantum thermodynamics [21, 22] 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 are concerned with the fundamental problem of how irreversible is the transformation of a state  $\rho_S$  to a state  $\sigma_S$  of some system  $S$  in the presence of a thermal bath, and how that irreversibility is related to the work that can be extracted through the same transformation. In this regard, the second laws [21] provide general constraints on these transitions, which are necessary and sufficient 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 [13], extracting work from correlations among different subsystems when  $\rho_S$  is a multipartite state (see, e.g. [23]), as well as the case when  $\sigma_S$  results from a measurement on  $\rho_S$  [24–26]. When thinking about *investing* work, one of the most well studied instances is Landauer’s erasure [27], 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 [13, 28, 29], which has the appealing feature of explicitly accounting for all energy flows. We will focus on the quantitative features of the irreversibility of thermodynamical processes that take an initial state  $\rho_S$  to a final state  $\sigma_S$ . In particular, we here show that a key quantity, namely the decrease of free energy or *entropy production* is related to how well a particular thermodynamical process can be reversed. In turn, this quantity is directly related to how much work can be extracted in the transition  $\rho_S \rightarrow \sigma_S$ .

## II. PRELIMINARIES

Let us now describe a prominent class of processes that we will be dealing with, known in the resource theory approach as *thermal operations* [21]. 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<sup>1</sup> and  $Z_B$  is the partition function. Let  $H_S$  be the Hamiltonian associated with the system  $S$  and let  $U$  denote a unitary that acts on the system  $S$ , a battery system  $W$ , and the bath  $B$ . The only unitary transformations  $U$  that are allowed are

---

<sup>1</sup> Here,  $k$  is the Boltzmann constant.

those that conserve total energy. That is, the allowed unitaries are such that  $[U, H] = 0$ , where  $H = H_S + H_W + H_B$  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

$$\mathcal{T}(\eta_{SW}) = \text{Tr}_B[U(\eta_{SW} \otimes \hat{\tau}_B)U^\dagger] \quad (1)$$

for some input state  $\eta_{SW}$  of the system and the battery. Other classes of thermodynamic operations are discussed in Sections VI and VII. 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.

We shall focus on the following quantity, known as entropy production

$$F(\rho_S) - F(\sigma_S), \quad (2)$$

where  $F(\omega_S) = \text{Tr}[H_S \omega_S] - kT S(\omega_S)$  is the Helmholtz free energy and the von Neumann entropy is defined as  $S(\omega_S) = -\text{Tr}[\omega_S \log \omega_S]$ .<sup>2</sup> This entropy production is always non-negative under the action of thermal operations [29]. In the absence of work extraction or expenditure, the change of energy in the system is equal to the negated change of energy in the bath

$$\text{Tr}[H_S \sigma_S] - \text{Tr}[H_S \rho_S] = -Q, \quad (3)$$

where  $Q$  denotes heat. This is due to energy conservation. In the limit of an infinite heat bath, we have that  $Q = \beta \delta S_B$  (the heat and the change of entropy of the bath are proportional). Thus, in that case, we can understand the quantity in (2) as the sum of the change of entropy of the system and bath separately, which is always positive.

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

$$F(\rho_S \otimes |0\rangle\langle 0|_W) \geq F(\sigma_S \otimes |1\rangle\langle 1|_W), \quad (4)$$

where  $H_{SW} = H_S + H_W$  and  $S(\omega_{SW}) = -\text{Tr}[\omega_{SW} \log \omega_{SW}]$ . Using the fact that  $\text{Tr}[H_W |0\rangle\langle 0|_W] = 0$  and  $\text{Tr}[H_W |1\rangle\langle 1|_W] = W_{\text{gain}}(\rho_S \rightarrow \sigma_S)$ , we can use (4) to obtain the following upper bound on the amount of work that we can hope to obtain

$$W_{\text{gain}}(\rho_S \rightarrow \sigma_S) \leq F(\rho_S) - F(\sigma_S). \quad (5)$$

That is, the entropy production upper bounds the deterministic work that can be extracted along a transition  $\rho_S \rightarrow \sigma_S$ .

In regimes in which the second law gives necessary and sufficient conditions for particular transitions  $\rho_S \rightarrow \sigma_S$  to be possible, it follows that (4) can be saturated for any states  $\rho_S$  and  $\sigma_S$ , in which case we have a very tight relation between entropy production and work. An example of a regime where (5) gives necessary and sufficient conditions occurs if we consider a non-deterministic work paradigm and allow the amount of work to fluctuate arbitrarily, in a transition in which the states are both diagonal in the energy eigenbasis and work is characterised by the mean value of the battery only [30]. Other examples are those in which the systems are extremely large [29] or if we allow for a slightly inexact catalysis [21]. 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  $\|\eta_C^{\text{in}} - \eta_C^{\text{out}}\|_1 \leq \varepsilon / \log d_C$ , where  $d_C$  is the dimension of the catalyst and  $\varepsilon > 0$  is some tolerance [21]. Similarly, if inexact catalysis takes on the form of allowing small correlations in the output catalyst, only the standard free energy is relevant [11]. A small caveat is that this regime is only achieved in transitions in which both  $\rho_S$  and  $\sigma_S$  are diagonal in the energy eigenbasis [31].

It is convenient to note [32] that the free energy can also be expressed in terms of the quantum relative entropy. Specifically,  $F(\rho_S) = kT[D(\rho_S \|\tau_S) - \log Z_S]$ , where  $\tau_S = \exp(-\beta H_S)/Z_S$  is the thermal state of the system at the temperature  $T$  of the ambient bath. The relative entropy is defined as [33]

$$D(\rho \|\tau) := \text{Tr}[\rho \log \rho] - \text{Tr}[\rho \log \tau], \quad (6)$$

when  $\text{supp}(\rho) \subseteq \text{supp}(\tau)$  and equal to  $+\infty$  otherwise. Since we do not change the Hamiltonian of the system, we can hence express the amount of work in regimes in which the standard free energy is relevant as

$$W_{\text{gain}}(\rho_S \rightarrow \sigma_S) = kT \Delta, \quad (7)$$

---

<sup>2</sup> All logarithms in this paper are base  $e$ . Furthermore, here and throughout, we take the convention that the operator logarithm is evaluated only on the support of its argument.

where we define the difference  $\Delta$  of relative entropies, which plays a special role as it is proportional to the entropy production

$$\Delta \equiv D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S) = \beta F(\rho_S) - \beta F(\sigma_S). \quad (8)$$

We also define a related quantity, which is the work that needs to be invested in doing the opposite transition, as  $W_{\text{inv}}(\sigma_S \rightarrow \rho_S)$ . This is how much work is needed to go deterministically from  $\sigma_S$  to  $\rho_S$ . In the nano-regime, it is possible that  $W_{\text{gain}} \neq W_{\text{inv}}$ . In fact, in general, we have the following relation:

$$W_{\text{gain}}(\rho_S \rightarrow \sigma_S) \leq F(\rho_S) - F(\sigma_S) \leq W_{\text{inv}}(\sigma_S \rightarrow \rho_S). \quad (9)$$

This also means that in the regimes in which the free energy gives necessary and sufficient conditions,  $W_{\text{gain}}(\rho_S \rightarrow \sigma_S) = W_{\text{inv}}(\sigma_S \rightarrow \rho_S)$ ; 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  $\rho_S$ . Thus in this “standard” free energy regime governed by the Helmholtz free energy  $F(\rho_S)$ , we see that we do not need to treat the amount of work gained 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 needs to be spent.

It is useful to note that for systems  $S$  that are truly small [13], or when we are interested in the case of *exact* catalysis, this is not the case in general. In these situations, the standard second law needs to be augmented with more refined conditions [21] that lead to differences. With some abuse of terminology, we 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

$$F_\alpha(\rho_S) \geq F_\alpha(\sigma_S), \quad (10)$$

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

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

where the general definition of  $D_\alpha$  <sup>3</sup> takes on a simplified form if  $\rho_S$  is diagonal in the energy eigenbasis. More precisely,

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

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  $\alpha \rightarrow 1$ . A short calculation [21] yields that in this regime

$$W_{\text{gain}}(\rho_S \rightarrow \sigma_S) \leq \inf_{\alpha \geq 0} kT [D_\alpha(\rho_S \| \tau_S) - D_\alpha(\sigma_S \| \tau_S)], \quad (13)$$

$$\begin{aligned} W_{\text{inv}}(\sigma_S \rightarrow \rho_S) &\geq \sup_{\alpha \geq 0} kT [D_\alpha(\rho_S \| \tau_S) - D_\alpha(\sigma_S \| \tau_S)] \\ &\geq kT [D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S)], \end{aligned} \quad (14)$$

where (the first) inequalities are again attained if  $\rho_S$  is diagonal in the energy eigenbasis.

### III. RESULT

Our main result is the following relation between entropy production along a change of state and how well a particular change can be undone. It takes the following form:

$$F(\rho_S) - F(\sigma_S) \geq kT D(\rho_S \| \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)), \quad (15)$$

where  $\mathcal{R}_{\sigma \rightarrow \rho}$  is a thermal reversal operation using a bath at temperature  $T$  that, when  $F(\rho_S) - F(\sigma_S)$  is small, takes  $\sigma_S$  close to the original state  $\rho_S$ . If  $\rho_S \rightarrow \sigma_S$  through a map of the form of (1), this reversed channel is defined as

$$\mathcal{R}_{\sigma \rightarrow \rho}(\cdot) = \text{Tr}_B[U^\dagger((\cdot) \otimes \tau_B)U]. \quad (16)$$

---

<sup>3</sup> For arbitrary states  $\rho_S$ , we have for  $0 \leq \alpha < 1/2$  that  $D_\alpha(\rho_S \| \tau_S) = \frac{1}{\alpha - 1} \log \text{Tr}[\rho_S^\alpha \tau_S^{1-\alpha}]$  [34] and for  $\alpha \geq 1/2$ ,  $D_\alpha(\rho_S \| \tau_S) = \frac{1}{\alpha - 1} \log \left[ \text{Tr} \left( \tau_S^{(1-\alpha)/(2\alpha)} \rho_S \tau_S^{(1-\alpha)/(2\alpha)} \right)^\alpha \right]$  [35, 36].

That is, the global unitary is reversed after using a new copy of the thermal bath state. This way, in the reversal operation we are ignoring both the correlations with the bath, and its change of state. We now relate the inequality in (15) to the work relative to the transition  $\rho_S \rightarrow \sigma_S$ .

**Investing work.** As outlined above, in the general regime in which not only the standard free energy is relevant, the amount of work  $W_{\text{inv}}(\sigma_S \rightarrow \rho_S) \geq 0$  we need to invest to transform  $\sigma_S$  to  $\rho_S$  is larger than the entropy production. Thus (15) together with (14) guarantee that

$$W_{\text{inv}}(\sigma_S \rightarrow \rho_S) \geq kT D(\rho_S \| \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)), \quad (17)$$

where  $\mathcal{R}_{\rho \rightarrow \sigma}$  is again the reversal operation. This says, for instance, that if not very much work needs to be spent in restoring  $\rho_S$  from  $\sigma_S$ , then a particular thermal operation not involving any work would also recover  $\rho_S$  from  $\sigma_S$  well, as measured by the relative entropy distance.

This may not always be the case as for example the erasure of a thermal state  $\sigma_S = \tau_S$  to a pure state  $\rho_S$  costs a significant amount of work. There, the operation  $\mathcal{R}_{\rho \rightarrow \sigma}$  will not change the thermal state of the system, effectively not recovering at all. Indeed this inequality also says that if the relative entropy is large, then the amount of work that we need to invest is large too. We illustrate this application of our result in Section V by means of a simple example of a harmonic oscillator bath.

**Gaining work.** Let us focus on particular situations in which  $W$  is characterized by the standard free energy (for physical examples of this regime see paragraph after that containing (5) in Section II). There, we have that  $kT\Delta = W_{\text{gain}}(\rho_S \rightarrow \sigma_S)$ . In that case (15) states that the amount of work  $W_{\text{gain}}(\rho_S \rightarrow \sigma_S) \geq 0$  gained when transforming  $\rho_S$  to  $\sigma_S$  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. More precisely

$$W_{\text{gain}}(\rho_S \rightarrow \sigma_S) \geq kT D(\rho_S \| \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)). \quad (18)$$

In this particular case a link is established between the reversibility of some transition and the amount of work that could be drawn from it. Loosely speaking, if little work can be obtained when transforming  $\rho_S$  to  $\sigma_S$  with a thermodynamic operation, then there 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 that can be obtained in the transition  $\rho_S \rightarrow \sigma_S$  can be large.

#### IV. PROOF FOR THERMAL OPERATIONS

We now give details of our main result, which applies to the set of thermal operations (TO) without catalysts. Section VI 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 [21, 22]—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*

$$\mathcal{T}(\cdot)_S = \text{Tr}_B[V((\cdot)_S \otimes \hat{\tau}_B)V^\dagger], \quad (19)$$

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

$$D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S) \geq D(\rho_S \| \mathcal{R}(\sigma_S)), \quad (20)$$

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

$$\mathcal{R}(\cdot) = \text{Tr}_B[V^\dagger((\cdot)_S \otimes \hat{\tau}_B)V]. \quad (21)$$

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

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

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

$$D(\rho_S \| \tau_S) = D(\rho_S \otimes \hat{\tau}_B \| \tau_S \otimes \hat{\tau}_B) \quad (22)$$

$$= D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \| V(\tau_S \otimes \hat{\tau}_B)V^\dagger) \quad (23)$$

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

where we have 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  $\text{supp}(\eta_{CD}) \subseteq \text{supp}(\theta_{CD})$ , it is possible to write

$$D(\eta_{CD} \parallel \theta_{CD}) - D(\eta_D \parallel \theta_D) = \text{Tr}(\eta_{CD} [\log \eta_{CD} - \log \theta_{CD} - \log I_C \otimes \eta_D + \log I_C \otimes \theta_D]). \quad (25)$$

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

$$D(\rho_S \parallel \tau_S) - D(\sigma_S \parallel \tau_S) = \text{Tr} \left( V(\rho_S \otimes \hat{\tau}_B)V^\dagger [\log V(\rho_S \otimes \hat{\tau}_B)V^\dagger - \log \tau_S \otimes \hat{\tau}_B - \log \sigma_S \otimes I_B + \log \tau_S \otimes I_B] \right). \quad (26)$$

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

$$-\log \tau_S \otimes \hat{\tau}_B - \log \sigma_S \otimes I_B + \log \tau_S \otimes I_B = -\log I_S \otimes \hat{\tau}_B - \log \sigma_S \otimes I_B \quad (27)$$

$$= -\log \sigma_S \otimes \hat{\tau}_B, \quad (28)$$

and thus conclude that

$$D(\rho_S \parallel \tau_S) - D(\sigma_S \parallel \tau_S) = D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \parallel \sigma_S \otimes \hat{\tau}_B). \quad (29)$$

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

$$D(V(\rho_S \otimes \hat{\tau}_B)V^\dagger \parallel \sigma_S \otimes \hat{\tau}_B) = D(\rho_S \otimes \hat{\tau}_B \parallel V^\dagger(\sigma_S \otimes \hat{\tau}_B)V). \quad (30)$$

Putting everything together, we see that

$$D(\rho_S \parallel \tau_S) - D(\sigma_S \parallel \tau_S) = D(\rho_S \otimes \hat{\tau}_B \parallel V^\dagger(\sigma_S \otimes \hat{\tau}_B)V). \quad (31)$$

Thus, the quantity  $\Delta$  related to the work gain in (7) 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:

$$\sigma_S \rightarrow V^\dagger(\sigma_S \otimes \hat{\tau}_B)V, \quad (32)$$

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 unitary 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 [37, 38], we can conclude from (31) that the following inequality holds

$$D(\rho_S \parallel \tau_S) - D(\sigma_S \parallel \tau_S) \geq D(\rho_S \parallel \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)), \quad (33)$$

where

$$\mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S) = \text{Tr}_B[V^\dagger(\sigma_S \otimes \hat{\tau}_B)V]. \quad (34)$$

This concludes the proof. Note that this operation is a thermal operation, and requires no work.  $\square$

#### A. Remark: Petz recovery map

We remark that  $\mathcal{R}$  is actually a special quantum map, called the Petz recovery map [39–42]. For a general quantum channel  $\mathcal{N}$  and a given density operator  $\theta$ , this recovery map is defined as

$$\tilde{\mathcal{N}}(\cdot) = \theta^{1/2} \mathcal{N}^\dagger[\mathcal{N}(\theta)^{-1/2}(\cdot)\mathcal{N}(\theta)^{-1/2}] \theta^{1/2}, \quad (35)$$

where  $\mathcal{N}^\dagger$  is the adjoint of the channel  $\mathcal{N}$  [43]. As a consequence, we can conclude that the main conjecture from [44] holds for the special case of thermal operations. We show this in the following lemma.

**Lemma 1.** *The map  $\mathcal{R}(\cdot)$  in (34) is the Petz recovery map of the original thermal operation, provided we choose the state  $\theta$  in (35) to be the thermal state  $\tau_S$ .*

*Proof.* Consider that to 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 of Eq. (35) above. For our case, we have that

$$\eta = \rho_S, \quad \theta = \tau_S, \quad \mathcal{N}(\cdot) = \text{Tr}_B[V((\cdot)_S \otimes \hat{\tau}_B)V^\dagger], \quad (36)$$

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

$$\mathcal{N}^\dagger(\cdot) = \text{Tr}_B \left[ \hat{\tau}_B^{1/2} V^\dagger [(\cdot)_S \otimes I_B] V \hat{\tau}_B^{1/2} \right], \quad (37)$$

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

$$\tilde{\mathcal{N}}(\cdot) = \tau_S^{1/2} \text{Tr}_B \left[ \hat{\tau}_B^{1/2} V^\dagger [\tau_S^{-1/2}(\cdot)_S \tau_S^{-1/2} \otimes I_B] V \hat{\tau}_B^{1/2} \right] \tau_S^{1/2}. \quad (38)$$

We can rewrite this as follows:

$$\text{Tr}_B \left[ (\tau_S^{\frac{1}{2}} \otimes \hat{\tau}_B^{\frac{1}{2}}) V^\dagger [\tau_S^{-\frac{1}{2}}(\cdot)_S \tau_S^{-\frac{1}{2}} \otimes I_B] V (\tau_S^{\frac{1}{2}} \otimes \hat{\tau}_B^{\frac{1}{2}}) \right] = \text{Tr}_B \left[ (\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}} V^\dagger [\tau_S^{-\frac{1}{2}}(\cdot)_S \tau_S^{-\frac{1}{2}} \otimes I_B] V (\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}} \right] \quad (39)$$

$$= \text{Tr}_B \left[ V^\dagger (\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}} \tau_S^{-\frac{1}{2}}(\cdot)_S \tau_S^{-\frac{1}{2}} \otimes I_B (\tau_S \otimes \hat{\tau}_B)^{\frac{1}{2}} V \right] \quad (40)$$

$$= \text{Tr}_B[V^\dagger((\cdot)_S \otimes \hat{\tau}_B)V], \quad (41)$$

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

## V. EXAMPLE FOR THERMAL OPERATIONS

Let us illustrate the reversal operation  $\mathcal{R}_{\sigma \rightarrow \rho}$  by means of a simple example. Let  $S$  be a two-level system, with Hamiltonian  $H_S = E_S|1\rangle\langle 1|$ . Let us take  $\rho_S = |0\rangle\langle 0|_S$  and  $\sigma_S = p_0|0\rangle\langle 0|_S + p_1|1\rangle\langle 1|_S$  with  $p_0 \in [1 - e^{-\beta E_S}, 1]$ . When  $p_0 = 1/2$ , the opposite operation  $\sigma_S \rightarrow \rho_S$  corresponds to Landauer erasure. Recall that the reversal operation associated with the lower bound for the work in (17) is determined by the operation that takes  $\rho_S = |0\rangle\langle 0|_S$  to  $\sigma_S$  without drawing any work.

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  $E_n = n\hbar\omega$ .<sup>4</sup> 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.

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

Our first goal is to find the explicit operation that takes  $\rho_S$  to  $\sigma_S$ , which has the effect of mixing the ground state of the system. 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  $n=1, 2, 3, \dots$ , 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:

$$U_{E_0} |0\rangle_S |0\rangle_B = |0\rangle_S |0\rangle_B =: |\Psi_{E_0}\rangle, \quad (42)$$

$$U_{E_n} |0\rangle_S |n\rangle_B = \sqrt{b} |0\rangle_S |n\rangle_B + \sqrt{1-b} |1\rangle_S |n-1\rangle_B =: |\Psi_{E_n}\rangle \quad \text{for } n=1, 2, 3, \dots, \quad (43)$$

$$U_{E_n} |1\rangle_S |n-1\rangle_B = \sqrt{1-b} |0\rangle_S |n\rangle_B - \sqrt{b} |1\rangle_S |n-1\rangle_B =: |\Psi_{E_n}^\perp\rangle \quad \text{for } n=1, 2, 3, \dots, \quad (44)$$

where  $0 \leq b \leq 1$  is a parameter that will be chosen in accordance with the desired target state  $\rho_S$  below. It is 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

$$U_{E_n} = \begin{pmatrix} \sqrt{b} & \sqrt{1-b} \\ \sqrt{1-b} & -\sqrt{b} \end{pmatrix}, \quad (45)$$

---

<sup>4</sup> 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  $\frac{\hbar}{2}\omega$ .



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  $n = 1, 2, 3, \dots$ . The bath thermal state is

$$\hat{\tau}_B = \frac{1}{Z_B} \sum_{n=0}^{\infty} e^{-nE_S\beta} |n\rangle\langle n|_B, \quad (46)$$

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

$$U(|0\rangle\langle 0|_S \otimes \hat{\tau}_B)U^\dagger = \frac{1}{Z_B} \sum_{n=0}^{\infty} e^{-nE_S\beta} U(|0\rangle_S |n\rangle_B {}_S\langle 0|_B \langle n|)U^\dagger \quad (47)$$

$$= \frac{1}{Z_B} \sum_{n=1}^{\infty} e^{-nE_S\beta} U(|0\rangle_S |n\rangle_B {}_S\langle 0|_B \langle n|)U^\dagger + \frac{1}{Z_B} U(|0\rangle_S |0\rangle_B {}_S\langle 0|_B \langle 0|)U^\dagger \quad (48)$$

$$= \frac{1}{Z_B} \sum_{n=1}^{\infty} e^{-nE_S\beta} |\Psi_{E_n}\rangle\langle \Psi_{E_n}| + \frac{1}{Z_B} |0\rangle\langle 0|_S \otimes |0\rangle\langle 0|_B \quad (49)$$

$$=: \sigma_{SB}^0. \quad (50)$$

By linearity of the partial trace operation, we have that

$$\text{Tr}_B(\sigma_{SB}^0) = \frac{Z_B - 1}{Z_B} (b|0\rangle\langle 0|_S + (1-b)|1\rangle\langle 1|_S) + \frac{1}{Z_B} |0\rangle\langle 0|_S \quad (51)$$

$$= p_0|0\rangle\langle 0|_S + p_1|1\rangle\langle 1|_S, \quad (52)$$

where

$$p_0 = \frac{1}{Z_B} ((Z_B - 1)b + 1), \quad (53)$$

$$p_1 = 1 - p_0. \quad (54)$$

Note that since  $0 \leq b \leq 1$ ,  $p_0 \in [1/Z_B, 1] = [1 - e^{-E_S\beta}, 1]$ . Solving (53) for  $b$  gives

$$b = \frac{p_0 Z_B - 1}{Z_B - 1}. \quad (55)$$

## 2. The reversal operation

Let us now construct the reversal map  $\mathcal{R}_{\sigma \rightarrow \rho}$ . This map can be written as

$$\mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S) = \text{Tr}_B [U^\dagger(\sigma_S \otimes \hat{\tau}_B)U] = \text{Tr}_B [U(\sigma_S \otimes \hat{\tau}_B)U^\dagger], \quad (56)$$

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

$$U(|1\rangle\langle 1|_S \otimes \hat{\tau}_B)U^\dagger = \frac{1}{Z_B} \sum_{n=0}^{\infty} e^{-nE_S\beta} |\Psi_{E_{n+1}}^\perp\rangle\langle \Psi_{E_{n+1}}^\perp| =: \sigma_{SB}^1. \quad (57)$$

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

$$\text{Tr}_B [\sigma_{SB}^1] = (1-b)|0\rangle\langle 0|_S + b|1\rangle\langle 1|_S. \quad (58)$$

Using (50) and (57) together with (52) and (58), we then have

$$\mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S) = \text{Tr}_B [U(\sigma_S \otimes \hat{\tau}_B)U^\dagger] \quad (59)$$

$$= p_0 \text{Tr}_B [\sigma_{SB}^0] + p_1 \text{Tr}_B [\sigma_{SB}^1] \quad (60)$$

$$= p_0 (p_0|0\rangle\langle 0|_S + p_1|1\rangle\langle 1|_S) + p_1 ((1-b)|0\rangle\langle 0|_S + b|1\rangle\langle 1|_S) \quad (61)$$

$$= P_0^{\mathcal{R}}|0\rangle\langle 0|_S + P_1^{\mathcal{R}}|1\rangle\langle 1|_S, \quad (62)$$



with

$$P_1^{\mathcal{R}} := 1 - P_0^{\mathcal{R}}, \quad (63)$$

$$P_0^{\mathcal{R}} := (p_0)^2 + (p_1)^2 \frac{Z_B}{Z_B - 1} = (p_0)^2 + (1 - p_0)^2 e^{E_S \beta}, \quad (64)$$

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  $W_{\text{inv}}(\sigma_S \rightarrow \rho_S)$ . We find

$$W_{\text{inv}}(\sigma_S \rightarrow \rho_S) \geq kTD(\rho_S \| \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)) \quad (65)$$

$$= -kT \log P_0^{\mathcal{R}}. \quad (66)$$

Plugging in (64) into (65) we find

$$W_{\text{inv}}(\sigma_S \rightarrow \rho_S) \geq -kT \log [(p_0)^2 + (1 - p_0)^2 e^{E_S \beta}], \quad (67)$$

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 (67):

- 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 (67) 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 “recover” from a thermal state  $\sigma_S = \tau_S$ , as to get close to a ground state. In this case, the RHS of (67) simplifies to  $kTD(\rho_S \| \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)) = kT \log Z_S$ . By direct calculation using the 2nd laws (using (13)–(14)) we find  $W_{\text{gain}}^{\text{nano}} = W_{\text{inv}}^{\text{nano}} = (\log Z_S)/\beta$  and thus the bound is also tight for this case.
- 3) Consider  $p_0 = 1/Z_B$ . That is, we want the recovery map to approach to 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, (67) reduces to  $kTD(\rho_S \| \mathcal{R}_{\sigma \rightarrow \rho}(\sigma_S)) = -kT \log[1 + e^{-2E_S \beta} - e^{-E_S \beta}]$ .

## VI. EXTENDING TO MORE GENERAL OPERATIONS INVOLVING CATALYSTS

We now prove the following lemma, which highlights the condition that a given map has to obey for the proof of Section IV to still hold.

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

$$\mathcal{T}(\cdot)_S = \text{Tr}_E[U((\cdot)_S \otimes \rho_E)U^\dagger], \quad (68)$$

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

$$U(\tau_S \otimes \rho_E)U^\dagger = \tau_S \otimes \rho'_E. \quad (69)$$

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

$$D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S) \geq D(\rho_S \| \mathcal{R}(\sigma_S)), \quad (70)$$

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

$$\mathcal{R}(\cdot) = \text{Tr}_E[U^\dagger((\cdot)_S \otimes \rho'_E)U]. \quad (71)$$

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

$$D(\rho_S \|\tau_S) = D(\rho_S \otimes \rho_E \|\tau_S \otimes \rho_E) \quad (72)$$

$$= D(U(\rho_S \otimes \rho_E)U^\dagger \| U(\tau_S \otimes \rho_E)U^\dagger) \quad (73)$$

$$= D(U(\rho_S \otimes \rho_E)U^\dagger \|\tau_S \otimes \rho'_E), \quad (74)$$

where we have used the main assumption of the lemma from (69) 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 (25) from the proof of Theorem 1:

$$D(\eta_{CD} \|\theta_{CD}) - D(\eta_D \|\theta_D) = \text{Tr}(\eta_{CD} [\log \eta_{CD} - \log \theta_{CD} - \log I_C \otimes \eta_D + \log I_C \otimes \theta_D]), \quad (75)$$

where  $\text{supp}(\eta_{CD}) \subseteq \text{supp}(\theta_{CD})$ . We use it together with (72) to write

$$D(\rho_S \|\tau_S) - D(\sigma_S \|\tau_S) = \text{Tr}(U(\rho_S \otimes \rho_E)U^\dagger [\log U(\rho_S \otimes \rho_E)U^\dagger - \log \tau_S \otimes \rho'_E - \log \sigma_S \otimes I_E + \log \tau_S \otimes I_E]). \quad (76)$$

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

$$-\log \tau_S \otimes \rho'_E - \log \sigma_S \otimes I_E + \log \tau_S \otimes I_E = -\log I_S \otimes \rho'_E - \log \sigma_S \otimes I_E \quad (77)$$

$$= -\log \sigma_S \otimes \rho'_E, \quad (78)$$

which leads to

$$D(\rho_S \|\tau_S) - D(\sigma_S \|\tau_S) = D(U(\rho_S \otimes \rho_E)U^\dagger \|\sigma_S \otimes \rho'_E). \quad (79)$$

We also have that

$$D(U(\rho_S \otimes \rho_E)U^\dagger \|\sigma_S \otimes \rho'_E) = D(\rho_S \otimes \rho_E \| U^\dagger(\sigma_S \otimes \rho'_E)U). \quad (80)$$

Putting everything together, we see that

$$D(\rho_S \|\tau_S) - D(\sigma_S \|\tau_S) = D(\rho_S \otimes \rho_E \| U^\dagger(\sigma_S \otimes \rho'_E)U) \quad (81)$$

$$\geq D(\rho_S \|\mathcal{R}'(\sigma_S)). \quad (82)$$

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

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

$$(\cdot)_S \rightarrow \text{Tr}_E \left[ \rho_E^{1/2} U^\dagger ((\cdot)_S \otimes I_E) U \rho_E^{1/2} \right], \quad (83)$$

and, by definition the Petz recovery channel is given as

$$(\cdot)_S \rightarrow \tau_S^{1/2} \text{Tr}_E \left[ \rho_E^{1/2} U^\dagger \left( \tau_S^{-1/2} (\cdot)_S \tau_S^{-1/2} \otimes I_E \right) U \rho_E^{1/2} \right] \tau_S^{1/2}. \quad (84)$$

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

$$\begin{aligned} & \tau_S^{1/2} \text{Tr}_E \left[ \rho_E^{1/2} U^\dagger \left( \tau_S^{-1/2} (\cdot)_S \tau_S^{-1/2} \otimes I_E \right) U \rho_E^{1/2} \right] \tau_S^{1/2} \\ &= \text{Tr}_E \left[ (\tau_S \otimes \rho_E)^{1/2} U^\dagger \left( \tau_S^{-1/2} (\cdot)_S \tau_S^{-1/2} \otimes I_E \right) U (\tau_S \otimes \rho_E)^{1/2} \right] \end{aligned} \quad (85)$$

$$= \text{Tr}_E \left[ U^\dagger (\tau_S \otimes \rho'_E)^{1/2} \left( \tau_S^{-1/2} (\cdot)_S \tau_S^{-1/2} \otimes I_E \right) (\tau_S \otimes \rho'_E)^{1/2} U \right] \quad (86)$$

$$= \text{Tr}_E \left[ U^\dagger ((\cdot)_S \otimes \rho'_E) U \right] \quad (87)$$

$$= \mathcal{R}(\cdot), \quad (88)$$

We note that from (69), multiplying by  $U$  and  $U^\dagger$  and taking the square root on both sides of the equation allows us to conclude  $(\tau_S \otimes \rho_E)^{1/2} U^\dagger = U^\dagger (\tau_S \otimes \rho'_E)^{1/2}$  and  $U(\tau_S \otimes \rho_E)^{1/2} = (\tau_S \otimes \rho'_E)^{1/2} U$ . These equalities allow us to go from the second to the third line by using the assumption of (69).  $\square$

This lemma implies that for any quantum channel that has a dilation satisfying the condition in (69), we arrive at an inequality like that in (33). In the next lemma, we define a further set of maps for which the condition in (69) holds. We say there is an *isentropic catalytic thermal operation* (ICTO) from  $\rho_S$  to  $\sigma_S$ , if there exists an energy-conserving unitary  $V$  acting on the system  $S$ , the bath  $B$ , and a set of  $n$  *isentropic catalysts*  $\otimes_{i=1}^n \eta_{C_i} =: \eta_C$  on  $C = \otimes_{i=1}^n C_i$  with initial states  $\eta_{C_i}$ , such that

$$\text{Tr}_B [V (\rho_S \otimes \hat{\tau}_B \otimes \eta_C) V^\dagger] = \sigma_{SC}, \quad (89)$$

where  $\text{Tr}_C [\sigma_{SC}] = \sigma_S$ . The unitary  $V$  conserves the energy of the bath, the system, and all the catalysts, so that  $[V, H_S + H_B + H_C] = 0$ , where  $H_C := \sum_{i=1}^n H_{C_i}$  are the Hamiltonians of the catalysts. This said, correlations between the different catalysts  $\eta_{C_i}$  are allowed in the final state. For every ICTO, we can define an associated channel  $\mathcal{H}_S \rightarrow \mathcal{H}_S$ :

$$\mathcal{T}_{ICTO}(\cdot) := \text{Tr}_{BC} [\sigma'_{BC}(\cdot)], \quad \sigma'_{BC}(\cdot) := V ((\cdot) \otimes \hat{\tau}_B \otimes \eta_C) V^\dagger. \quad (90)$$

The isentropic catalysts are required to satisfy the following:

- 1)  $S(\text{Tr}_{S \setminus C_i} [\sigma_{SC}]) = S(\eta_{C_i}) \forall i$ , meaning that the local states of the catalysts return to states of equal entropy to the initial states.
- 2) When the input to the channel in (90) is the thermal state  $\tau_S$  of the system, the entropy and mean energy of the catalysts are non-increasing and non-decreasing respectively:  $S(\eta_C) \geq S(\sigma'_C)$  and  $\text{Tr}[H_C \eta_C] \leq \text{Tr}[H_C \sigma'_C]$ .

Condition 1) guarantees that the catalysts are not degraded in the sense of an entropy change, while (as will become evident in the following lemma) condition 2) guarantees that the channel  $\mathcal{T}_{ICTO}(\cdot)$  is Gibbs preserving, which is a physically relevant condition for a channel resulting from a thermodynamic process. This new class of operations is between TO and Gibbs preserving maps.

These conditions are different from the ones that apply to catalytic thermal operations as usually defined in the literature [21]. For those, a stronger version of condition 1) holds, but condition 2) does not necessarily hold. However, since condition 2) is only required to hold for the von Neumann entropy and mean energy, rather than requiring exact catalysis, it is feasible (given what is known about work embezzlement with inexact catalysts [10]) that one can always construct a catalyst large enough, such that for every catalytic thermal operation transforming  $\rho \rightarrow \sigma$ , there exists another catalytic thermal operation also transforming  $\rho \rightarrow \sigma$  (possibly with a larger catalyst) such that condition 2) is satisfied. If such a family of catalytic thermal operations exists, it would be very satisfying since via the following lemma it would mean that there is a subset of catalytic thermal operations which allow for all possible transformations as the full set, yet with the additional physically relevant property of belonging to the class of Gibbs preserving maps.

We now show that given the assumptions above, the operations defined as such obey the conditions of Theorem 2.

**Lemma 3.** *For every ICTO channel as defined in (90) the following hold:*

- 1) *The channel is Gibbs preserving:  $\mathcal{T}_{ICTO}(\tau_S) = \tau_S$ .*
- 2) *The isentropic catalysts do not become correlated with the bath or the system when the input to the channel is thermal:*

$$V (\tau_S \otimes \hat{\tau}_B \otimes \eta_C) V^\dagger = \tau_S \otimes \hat{\tau}_B \otimes \sigma'_C. \quad (91)$$

*Proof.* Let  $\hat{\rho}_{SB} = \text{Tr}_C [V (\tau_S \otimes \hat{\tau}_B \otimes \eta_C) V^\dagger]$  denote the local state of the system and 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:

$$\text{Tr}[HV (\tau_S \otimes \hat{\tau}_B \otimes \eta_C) V^\dagger] = \text{Tr}[H(\tau_S \otimes \hat{\tau}_B \otimes \eta_C)] = \quad (92)$$

$$\text{Tr}[(H_S + H_B)(\tau_S \otimes \tau_B)] + \text{Tr}[H_C \eta_C] \leq \text{Tr}[(H_S + H_B)(\tau_S \otimes \tau_B)] + \text{Tr}[H_C \sigma'_C]. \quad (93)$$

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

$$\text{Tr}[HV (\tau_S \otimes \hat{\tau}_B \otimes \eta_C) V^\dagger] = \text{Tr}[(H_S + H_B)\hat{\rho}_{SB}] + \text{Tr}[H_C \sigma'_C], \quad (94)$$

and hence  $\text{Tr}[(H_S + H_B)\hat{\rho}_{SB}] \leq \text{Tr}[(H_S + H_B)(\tau_S \otimes \hat{\tau}_B)]$ . Let  $T'$  be the temperature of the Gibbs state  $\tau'_{SB}$  such that  $\text{Tr}[(H_S + H_B)\hat{\rho}_{SB}] = \text{Tr}[(H_S + H_B)\tau'_{SB}]$ <sup>5</sup>. For a given fixed energy, the thermal state is the unique state with the highest

<sup>5</sup> Note that such a  $T' \geq 0$  always exists since  $T' = 0$  is the ground state, the Gibbs state mean energy is monotonically increasing with its temperature, and the mean energy of  $\hat{\rho}_{SB}$  is upper bounded by a thermal state of the same Hamiltonian

entropy [45, Theorem 1.3], and hence  $S(\tau'_{SB}) \geq S(\hat{\rho}_{SB})$ . Yet, since,  $\text{Tr}[(H_S + H_B)\tau'_{SB}] \leq \text{Tr}[(H_S + H_B)(\tau_S \otimes \hat{\tau}_B)]$ , it follows that the temperature  $T$  of state  $\tau_S \otimes \hat{\tau}_B$  satisfies  $T \geq T'$ , and thus by direct calculation  $S(\tau_S \otimes \hat{\tau}_B) \geq S(\tau'_{SB})$ . So we conclude that

$$S(\tau_S \otimes \hat{\tau}_B) \geq S(\hat{\rho}_{SB}). \quad (95)$$

Now we consider the entropy before and after the transformation. Since the joint operation is a unitary, we have from unitary invariance and sub-additivity of quantum entropy that

$$S(\tau_S \otimes \hat{\tau}_B) + S(\eta_C) = S(V(\tau_S \otimes \hat{\tau}_B \otimes \eta_C)V^\dagger) \quad (96)$$

$$= S(\hat{\rho}_{SBC}) \quad (97)$$

$$\leq S(\hat{\rho}_{SB}) + S(\sigma'_C) \quad (98)$$

$$\leq S(\hat{\rho}_{SB}) + S(\eta_C), \quad (99)$$

where we have defined  $\hat{\rho}_{SBC} := V(\tau_S \otimes \hat{\tau}_B \otimes \eta_C)V^\dagger$ . Hence  $S(\tau_S \otimes \hat{\tau}_B) \leq S(\hat{\rho}_{SB})$ . Given our conclusion in (95) regarding conservation of energy, it must then be the case that  $S(\tau_S \otimes \hat{\tau}_B) = S(\hat{\rho}_{SB})$ . At fixed von Neumann entropy, the thermal (Gibbs) state minimises the mean energy, and thus  $S(\tau_S \otimes \hat{\tau}_B) = S(\hat{\rho}_{SB})$  implies  $\text{Tr}[(H_S + H_B)\hat{\rho}_{SB}] \geq \text{Tr}[(H_S + H_B)(\tau_S \otimes \hat{\tau}_B)]$ . Since previously we concluded  $\text{Tr}[(H_S + H_B)\hat{\rho}_{SB}] \leq \text{Tr}[(H_S + H_B)(\tau_S \otimes \hat{\tau}_B)]$ , we then have that  $\text{Tr}[(H_S + H_B)\hat{\rho}_{SB}] = \text{Tr}[(H_S + H_B)(\tau_S \otimes \hat{\tau}_B)]$ . Thus the last equality together with  $S(\tau_S \otimes \hat{\tau}_B) = S(\hat{\rho}_{SB})$  implies

$$\tau_S \otimes \hat{\tau}_B = \hat{\rho}_{SB} = \hat{\rho}_S \otimes \hat{\rho}_B. \quad (100)$$

Using (100), (99), and noting that by definition  $\hat{\rho}_C = \sigma'_C$ , we conclude

$$S(\hat{\rho}_{SBC}) \leq S(\hat{\rho}_S \otimes \hat{\rho}_B \otimes \sigma'_C) = S(\hat{\rho}_S \otimes \hat{\rho}_B \otimes \hat{\rho}_C) \leq S(\hat{\rho}_S \otimes \hat{\rho}_B \otimes \eta_C) = S(\hat{\rho}_{SBC}) \quad (101)$$

Hence  $S(\hat{\rho}_S \otimes \hat{\rho}_B \otimes \hat{\rho}_C) = S(\hat{\rho}_{SBC})$ , which is true iff  $\hat{\rho}_{SBC} = \hat{\rho}_S \otimes \hat{\rho}_B \otimes \hat{\rho}_C$ . Writing this in terms of  $\tau_S$ ,  $\hat{\tau}_B$ ,  $\sigma'_C$  and  $V(\tau_S \otimes \hat{\tau}_B \otimes \eta_C)V^\dagger$  gives us (91), completing the proof.  $\square$

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.,  $\rho_E \equiv \hat{\tau}_B \otimes \eta_C$ ), we arrive at the following conclusion:

**Theorem 2.** *Let  $\mathcal{T}_{ICTO}(\cdot)$  be an ICTO channel of the form in (90) given by*

$$\mathcal{T}_{ICTO}(\cdot) = \text{Tr}_{BC}[U((\cdot)_S \otimes \hat{\tau}_B \otimes \eta_C)U^\dagger]. \quad (102)$$

*Then it obeys the inequality*

$$D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S) \geq D(\rho_S \| \mathcal{R}(\sigma_S)), \quad (103)$$

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

$$\mathcal{R}(\cdot) = \text{Tr}_{BC}[U^\dagger((\cdot)_S \otimes \hat{\tau}_B \otimes \sigma'_C)U]. \quad (104)$$

*The Petz recovery map preserves the Gibbs state  $\mathcal{R}(\tau_S) = \tau_S$ .*

*Proof.* (103) and (104) are a direct consequence of Lemmas 2 and 3. The fact that the Petz recovery map preserves the thermal state follows from Lemma 3 by inspection.  $\square$

## VII. GIBBS PRESERVING MAPS

A general set of maps to which the conditions of Lemma 2 do not apply is that of Gibbs preserving maps [46], and we hence need a different method to prove an analogous result. To prove a bound in (20), we use the following general result for quantum maps from [47]:

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

$$D(\eta \| \theta) - D(\mathcal{N}(\eta) \| \mathcal{N}(\theta)) \geq - \int_{\mathbb{R}} dt p(t) \log F(\eta, \tilde{\mathcal{N}}_t(\mathcal{N}(\eta))), \quad (105)$$

*where  $F(\rho, \sigma) = (\text{Tr}[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}])^2$  is the quantum fidelity, the map  $\tilde{\mathcal{N}}_t$  is the following rotated recovery map*

$$\tilde{\mathcal{N}}_t(\cdot) = \theta^{it/2} \tilde{\mathcal{N}}(\mathcal{N}(\theta)^{-it/2}(\cdot)\mathcal{N}(\theta)^{it/2})\theta^{-it/2}, \quad (106)$$

*with  $\tilde{\mathcal{N}}$  defined as in (35) and  $p(t) = \frac{\pi}{2}(\cosh(\pi t) + 1)^{-1}$  is a probability density function.*

In the same way as in Theorem 2, 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:

$$\Delta = D(\rho_S \| \tau_S) - D(\sigma_S \| \tau_S) \geq - \int_{\mathbb{R}} dt \, p(t) \log F(\rho_S, \tilde{\mathcal{N}}_t(\sigma_S)), \quad (107)$$

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

## VIII. CONCLUSION

We have shown how the amount of entropy produced along a thermal process that takes  $\rho_S$  to  $\sigma_S$  is directly linked to the reversibility of the process. Specifically, we see that if this quantity is small, then there exists a recovery operation that approximately restores the system to its initial state at no work cost at all. For thermal operations, this map is another operation in the same set and can be taken to be  $\mathcal{R}_{\sigma \rightarrow \rho}$ .

Our main result applies to the decrease of standard free energy, and it is a very interesting open question to extend our result to regimes in which we require the full set of second laws [21]. What makes this question challenging is that (25) does not carry over to the regime of  $D_\alpha$  for  $\alpha \neq 1$ , and indeed recent work [48] 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 [48] should be our starting point. However, the quantities in [48] 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 [21] 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 have 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 [49–53], which shows the importance of the concept of recoverability and of recovery maps.

## IX. ACKNOWLEDGMENTS

We thank Jonathan Oppenheim, Christopher Perry, and Renato Renner for insightful discussions. MMW is grateful to SW and her group for hospitality during a research visit to QuTech, and acknowledges support from startup funds from the Department of Physics and Astronomy at LSU, the NSF under Award No. CCF-1350397, and the DARPA Quiness Program through US Army Research Office award W31P4Q-12-1-0019. MPW and SW acknowledge support from MOE Tier 3A Grant MOE2012-T3-1-009 and STW, Netherlands.

- 
- [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] Mischa Woods, Nelly Ng, and Stephanie Wehner. The maximum efficiency of nano heat engines depends on more than temperature. June 2015. arXiv:1506.02322.
  - [5] Hiroyasu Tajima and Masahito Hayashi. Finite-size effect on optimal efficiency of heat engines. *Physical Review E*, 96(1):012128, July 2017. arXiv:1405.6457.
  - [6] 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.
  - [7] 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.
  - [8] 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.

- [9] 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.
- [10] 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.
- [11] Matteo Lostaglio, Markus P. Mueller, and Michele Pastena. Extracting work from absence of correlations. *Physical Review Letters*, 115(15):150402, 2015. arXiv:1409.3258.
- [12] Johan Åberg. Catalytic coherence. *Physical Review Letters*, 113(15):150402, October 2014. arXiv:1304.1060.
- [13] Michal Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nano thermodynamics. *Nature Communications*, 4:2059, June 2013. arXiv:1111.3834.
- [14] Johan Åberg. Truly work-like work extraction via a single-shot analysis. *Nature Communications*, 4:1925, June 2013. arXiv:1110.6121.
- [15] 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.
- [16] Rodrigo Gallego, Jens Eisert, and Henrik Wilming. Thermodynamic work from operational principles. *New Journal of Physics*, 18(10):103017, October 2016. arXiv:1504.05056.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] Sina Salek and Karoline Wiesner. Fluctuations in single-shot  $\epsilon$ -deterministic work extraction. 2015. arXiv:1504.05111.
- [21] 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.
- [22] 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.
- [23] Marti Perarnau-Llobet, Karen V. Hovhannisyan, Marcus Huber, Paul Skrzypczyk, Nicolas Brunner, and Antonio Acín. Extractable work from correlations. *Physical Review X*, 5(4):041011, 2015. arXiv:1407.7765.
- [24] Philipp Kammerlander and Janet Anders. Coherence and measurement in quantum thermodynamics. *Scientific Reports*, 6:22174, February 2016. arXiv:1502.02673.
- [25] 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.
- [26] 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.
- [27] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 1961.
- [28] 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.
- [29] 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.
- [30] Paul Skrzypczyk, Anthony J Short, and Sandu Popescu. Work extraction and thermodynamics for individual quantum systems. *Nature Communications*, 5, 2014. arXiv:1307.1558.
- [31] Matteo Lostaglio, David Jennings, and Terry Rudolph. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nature Communications*, 6, 03 2015. arXiv:1405.2188.
- [32] Matthew J. Donald. Free energy and the relative entropy. *Journal of Statistical Physics*, 49(1-2):81–87, October 1987.
- [33] Hisaharu Umegaki. Conditional expectation in an operator algebra. In *Kodai Mathematical Seminar Reports*, volume 14, pages 59–85. Department of Mathematics, Tokyo Institute of Technology, 1962.
- [34] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports in Mathematical Physics*, 23(1):57–65, February 1986.
- [35] 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.
- [36] 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.
- [37] Goran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, 1975.
- [38] Armin Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics*, 54(1):21–32, 1977.
- [39] Dénes Petz. Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Communications in Mathematical Physics*, 105(1):123–131, 1986.
- [40] Dénes Petz. Sufficiency of channels over von Neumann algebras. *The Quarterly Journal of Mathematics*, 39(1):97–108, 1988.



- [41] 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.
- [42] 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.
- [43] John Watrous. Theory of quantum information lecture notes. <https://cs.uwaterloo.ca/~watrous/LectureNotes.html>, Fall 2011.
- [44] Andreas Winter and Ke Li. Squashed entanglement, k-extendibility, quantum Markov chains, and recovery maps. October 2014. arXiv:1410.4184.
- [45] Eric A. Carlen. Trace inequalities and quantum entropy: An introductory course. *Contemporary Mathematics*, 529:73–140, 2010.
- [46] 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.
- [47] 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.
- [48] 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.
- [49] Mario Berta, Stephanie Wehner, and Mark M. Wilde. Entropic uncertainty and measurement reversibility. *New Journal of Physics*, 18(7):073004, 2016. arXiv:1511.00267.
- [50] 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. arXiv:1601.01207.
- [51] Álvaro M. Alhambra and Mischa P. Woods. Dynamical maps, quantum detailed balance, and the petz recovery map. *Physical Review A*, 96(2):022118, 2017. arXiv:1609.07496.
- [52] Marius Lemm and Mark M. Wilde. Information-theoretic limitations on approximate quantum cloning and broadcasting. *Physical Review A*, 96(1):012304, 2017. arXiv:1608.07569.
- [53] Iman Marvian and Seth Lloyd. From clocks to cloners: Catalytic transformations under covariant operations and recoverability. *arXiv preprint arXiv:1608.07325*, 2016.