Thermodynamically optimal protocols for dual-purpose qubit operations

Joe Dunlop ©,[1,*] Federico Cerisola ©,[1,2] Jorge Tabanera-Bravo ©,[3,4] and Janet Anders ©[1,5]

[1]*Physics and Astronomy, University of Exeter, Exeter EX4 4QL, England, United Kingdom*
[2]*Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, England, United Kingdom*
[3]*Departamento Estructura de la Materia, Física Térmica y Electrónica*
and GISC, Universidad Complutense de Madrid, 28040 Madrid, Spain
[4]*Mathematical Biophysics Group, Max Planck Institute for Multidisciplinary Sciences, Göttingen 37077, Germany*
[5]*Institut für Physik, Potsdam University, 14476 Potsdam, Germany*

Information processing, quantum or classical, relies on channels transforming multiple input states to different corresponding outputs. Previous research has established bounds on the thermodynamic resources required for such operations, but no protocols have been specified for their optimal implementation. For the insightful case of qubits, we here develop explicit protocols to transform two states in an energetically optimal manner. We first prove conditions on the feasibility of carrying out such transformations at all, and then quantify the achievable work extraction. Our results uncover a fundamental incompatibility between the thermodynamic ideal of slow, quasistatic processes and the information-theoretic requirement to preserve distinguishability between different possible output states.

I. INTRODUCTION

In introducing entropy as an information measure, Shannon stressed an essential property of communication channels: that the message to be transmitted must belong to a set of possible messages, and that "the system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design" [1]. In general, information processing involves *multipurpose operations* mapping each of a collection of possible input signals to their corresponding outputs via some standard protocol (Fig. 1) [2]. Signals must be encoded in the states of a physical system, with stored information manifested as entropy of the medium, and all operations on them must obey thermodynamic laws—in particular the Clausius inequality, applied by Landauer to lower bound the energetic cost of erasure in the presence of a thermal environment [3–6].

Quantum information theory and the development of quantum technologies [7] have driven a surge of interest in basic limits on thermodynamic performance [8–10], and the role of novel quantum resources such as coherence [11–17] and entanglement [18–24]. Beyond foundational interest, energy consumption and heat generation may be important factors in the viability of future generations of hardware [25,26].

The energetic requirements for transforming a given quantum state to another have been studied in some detail. If the entropy of the output state differs from the input, the transformation must involve interaction with a heat bath. Nonetheless, in the absence of additional constraints (e.g., imperfect control or finite duration [27–29]), there is no fundamental barrier to these processes being carried out reversibly, provided the protocol is optimized for the particular input state present. Indeed, concrete protocols have been developed which in theory saturate the reversible limit, where the work extraction is equal to the reduction of the system's free energy [11,30,31].

On the other hand, relatively little attention has been given to the thermodynamics of multipurpose quantum operations, which map each of a collection of possible input states ρ_n to respective outputs η_n, the prototypical example being a single-qubit operation with two input states (Fig. 1). While the set of inputs is fixed, it is not known *which* of the inputs is present in a given instance [2]. The task is now to find a multipurpose operation $\mathcal{M}$ that maps *every* possible input state to its target output state. Constructing a one-size-fits-all protocol presents a more constrained problem: the optimal strategy for transforming ρ_1 to η_1 might fail to transform ρ_2 to η_2.

It has been shown by Bedingham and Maroney that in general, multipurpose operations necessarily incur an irreversible energy cost which is dissipated to the thermal environment [32], in addition to the reversible work implied by the Landauer bound. Chiribella *et al.* have explored the nonequilibrium cost of carrying out multipurpose operations *approximately*, using a resource-theoretic approach to lower bound the number of pure qubit states consumed for a given accuracy of outputs [33].

However, key questions remain open: there is no prescription for *how* an operation with two or more inputs might be carried out in a thermodynamically optimal way. Moreover, no upper bound has been placed on the energy cost.

---
*Contact author: he/him/his; j.dunlop@exeter.ac.uk
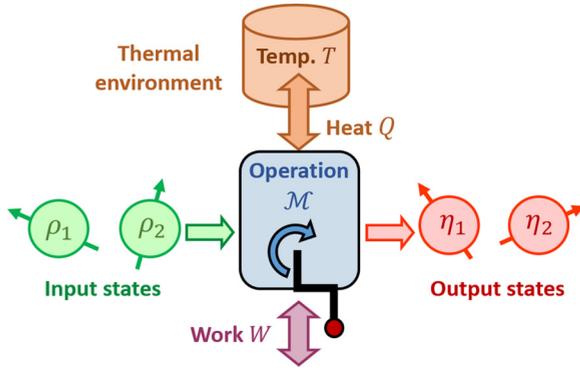
FIG. 1. A dual-purpose qubit operation transforms an input state, either ρ_1 or ρ_2 (occurring with probabilities p_1, p_2), into the respective output η_1 or η_2. This is implemented via some cyclic variation of the Hamiltonian, in a procedure which does not require knowledge of which of the states is present. During the process, the system exchanges heat with a fixed-temperature environment, and work with the driving field. Such operations are essential for information processing, and we would like to determine the energetic requirements.

The present paper undertakes a detailed thermodynamic analysis of single-qubit operations with two possible input states and corresponding outputs: *dual-purpose qubit operations*. We introduce a family of explicit protocols for carrying out such operations, as well as means to optimize for thermodynamic performance. These results are applied in a numerical case study of work extraction from coherence, highlighting the scale of the work penalty, as well as qualitative differences from the single-input case. Finally, we derive an analytic bound on work extraction, which is complementary to Bedingham and Maroney's result [32], despite arising from an entirely different theoretical treatment.

Our bound on work extraction is related to an effective speed limit: in order to preserve the output states' mutual distinguishability, any thermal contact with the environment must be limited in duration, avoiding complete thermalization. This is fundamentally at odds with the requirements for thermodynamic reversibility, where state transformations involving contact with the environment must happen extremely slowly, with the state remaining in thermal equilibrium. Our results call attention to the scale of the resulting energetic toll, and clearly illustrate the general principles preventing reversibility for multipurpose operations acting on systems of any size, with any number of inputs.

II. SETUP

We will primarily consider the following scenario. We are given a qubit which has been prepared in either state ρ_1 or ρ_2, and tasked with transforming it to a corresponding output state η_1 or η_2, depending on the input. We know what the possible states are, and the respective probabilities of their occurrence, p_1 and p_2, but not which of the states is present in a given instance. The objective is to carry out the mapping exactly for both inputs via some shared protocol, while maximizing the work extracted (or minimizing work expended) in the process. If our method extracts W_1 of work in transforming ρ_1 to η_1,

and W_2 in transforming ρ_2 to η_2, then we aim to maximize the mean work $\overline{W} = p_1 W_1 + p_2 W_2$.

Crucially, from an information perspective it is not sufficient to map the average input to the average output. For example, consider a scenario where the input states are $\rho_1 = |0\rangle\langle 0|$, $\rho_2 = |1\rangle\langle 1|$ and each occurs with probability $\frac{1}{2}$, so that the average input is $\overline{\rho} = \frac{\mathbb{I}}{2}$. If we consider only their action on the average state, then the qubit gates $\mathbb{I}, X, Y$, and Z have an identical effect, all producing the output $\overline{\eta} = \frac{\mathbb{I}}{2}$. But clearly it would not do to replace X with $\mathbb{I}$ in a circuit: the action on individual input states is important.

It will be assumed that we can vary the qubit Hamiltonian freely, and thereby control the system's unitary dynamics—subject to the condition that the Hamiltonian is reset to its initial value H_0 at the end of the protocol, so that the operation might be performed in a cycle (for example, to process a string of inputs). It is also assumed that if the Hamiltonian is held fixed at H, the qubit will thermalize over some finite timescale to the Gibbs state $\tau = \frac{1}{Z}\exp(-\frac{H}{k_B T})$, where the temperature T of the thermal environment is taken to be fixed throughout, and where $Z = \text{tr}[\exp(-\frac{H}{k_B T})]$ denotes the partition function.

As is standard practice in quantum thermodynamics [8,31,34], we define the work done by the system as equal to the reduction of its internal energy as a direct result of changing the Hamiltonian, and heat absorbed by the system as an increase in its internal energy due to a changing state:

$$\dot{W} = -\text{tr}[\rho\dot{H}],$$
$$\dot{Q} = \text{tr}[H\dot{\rho}]. \qquad (1)$$

We adopt the convention that positive work is done *by* the system against the driving field, and positive heat is absorbed by the system from the thermal environment. Storing or otherwise utilizing the extracted work is a separate issue [24,35] not treated here.

The Clausius inequality dictates that the work yield of a thermodynamic process is no greater than the reduction of free energy $F(\rho) = \text{tr}[H\rho] - k_B T\,S(\rho)$ between the initial and final state—here $S(\rho) = -\text{tr}[\rho\ln\rho]$ denotes the von Neumann entropy [36]. Applying this to our dual-purpose qubit operation, the mean work extraction is bounded by the free energy difference between the average input state $\overline{\rho} = p_1\rho_1 + p_2\rho_2$ and the average output $\overline{\eta} = p_1\eta_1 + p_2\eta_2$:

$$\overline{W} \leqslant F(\overline{\rho}) - F(\overline{\eta})$$
$$= \text{tr}[H_0(\overline{\rho} - \overline{\eta})] - k_B T[S(\overline{\rho}) - S(\overline{\eta})]. \qquad (2)$$

We will refer to protocols which saturate (2) as thermodynamically reversible—in that the average output could be transformed back to the average input at zero net work cost. Note that the process might still be *logically* irreversible: it might be impossible to deterministically recover ρ_1 from η_1 or ρ_2 from η_2 [37]. For example, Landauer erasure is logically irreversible but may saturate (2).

The above already implies a thermodynamic compromise. If we were able to choose our protocol based on prior knowledge of which of the states $\rho_{1,2}$ was present, we could extract work equal to $F(\rho_{1,2}) - F(\eta_{1,2})$. However, due to the concavity of the von Neumann entropy, the ideal dual-purpose work extraction given by (2) is generally

less than the weighted average: $F(\overline{\rho}) - F(\overline{\eta}) \leqslant p_1[F(\rho_1) - F(\eta_1)] + p_2[F(\rho_2) - F(\eta_2)]$.

This is *not* the compromise we will be exploring in detail here. Instead we will be investigating ways that a dual-purpose protocol might approach the bound given by (2), and analyzing the reasons why even that bound cannot in general be saturated. We base our approach on a well-studied type of reversible protocol, which maps a single, known initial state to some other final state [11]. Taking this as a starting point, we identify necessary extensions for a dual-purpose operation, and see where those extensions violate the conditions for thermodynamic reversibility.

A. How single-input operations are carried out reversibly

It is instructive to consider how an operation with a single input ρ and output η can be implemented reversibly. A protocol which appears in various forms in the literature [11,13,30,32] proceeds along the following lines.

(1) Begin with a system in the state ρ, with initial Hamiltonian H_0. Abruptly quench the Hamiltonian to $H_\rho = -k_B T(\ln \rho + \ln Z \, \mathbb{I})$. The system, still in state ρ, is in thermal equilibrium[1] at temperature T with respect to H_ρ, since $\rho = \frac{1}{Z} \exp(-\frac{H_\rho}{k_B T})$.

(2) Slowly and smoothly adjust the Hamiltonian from H_ρ to $H_\eta = -k_B T(\ln \eta + \ln Z \, \mathbb{I})$, over a period much longer than the thermalization timescale. The system undergoes quasistatic, isothermal evolution, tracking the instantaneous thermal state from ρ to η.

(3) Quench the Hamiltonian back to H_0, quickly enough that the system undergoes negligible evolution and remains in the intended output state η.

When the thermodynamic analysis is performed using Eq. (1), this method is found to saturate the Clausius inequality (2) for both classical and quantum states [11,30]. Key to its reversibility is the fact that the processes are either adiabatic, involving no exchange of heat with the environment (steps 1 and 3), or quasistatic, with the state at all times in thermal equilibrium with the environment (step 2). The system never undergoes irreversible thermalization from an initial state which is appreciably out of equilibrium [28].

Why can the same approach not be used for a multipurpose operation? Notice that the Hamiltonians H_ρ and H_η are fine tuned to the input and output state. This cannot be done simultaneously for more than one possible input, since if $\rho_1 \neq \rho_2$, then $H_{\rho_1} \neq H_{\rho_2}$.

Suppose that we optimize this method for the *average* input and output, replacing H_ρ and H_η with $H_{\overline{\rho}}$ and $H_{\overline{\eta}}$ in the above. The system, regardless of whether it began in state ρ_1 or ρ_2, thermalizes to $\overline{\rho}$ at the start of step 2, before slowly evolving from there to $\overline{\eta}$, finally remaining in that state as the Hamiltonian is reset in step 3.

This is unacceptable: we obtain the same output state $\overline{\eta}$ irrespective of the input. It does not suffice to produce the right

output state only on average.[2] A similar problem arises no matter what state we optimize for. In fact, any initial-state dependence is lost as soon as the system is allowed to completely thermalize, since the Gibbs state is uniquely determined by the Hamiltonian for a given temperature. This rules out any protocol involving quasistatic isothermal evolution, wherever we require two or more distinct outputs. For multipurpose operations, we need a framework to treat nonequilibrium processes.

B. Discrete quantum processes

Our approach is to decompose the operation as a sequence of discrete steps which are either unitary (and involve only work) or thermalizing (involving only heat transfer). This framework, introduced in [31], provides a means to quantify heat and work for processes which start or end in nonequilibrium configurations, and recovers continuous trajectories in the limit of many steps, including reversible quasistatic ones. However, it is necessary to extend that framework to account for *incomplete* thermalization, for the reason outlined above. We take the primitive operations to be the following.

1. Unitaries combined with quenches

These encompass operations which involve zero heat transfer, which can be further decomposed into unitary evolution generated by a fixed Hamiltonian, $U = e^{-\frac{i}{\hbar}Ht}$; and instantaneous quenches, which alter the Hamiltonian $H \mapsto H'$ without any immediate change in the qubit state. The action of the combined unitary and quench $\mathcal{U}$ is to map the state $\rho \mapsto U\rho U^\dagger$, while extracting work equal to $W = \mathrm{tr}[H\rho] - \mathrm{tr}[H'U\rho U^\dagger]$.

2. Partial thermalizations

Conversely, these represent processes where zero work is done. A partial thermalization $\mathcal{T} : \rho \mapsto \lambda\rho + (1-\lambda)\tau$ mixes the qubit state with the Gibbs thermal state, $\tau = \frac{1}{Z}\exp(-\frac{H}{k_B T})$ for a given Hamiltonian H, with mixing parameter $0 \leqslant \lambda \leqslant 1$. This represents a linear interpolation of the dynamics taking ρ to τ, and we need not consider any details of the environment other than its temperature.[3] The heat absorbed by the system is given by $Q = \mathrm{tr}\{H[\lambda\rho + (1 - \lambda)\tau]\} - \mathrm{tr}[H\rho]$.

We call $\mathcal{S}$ the set of all transformations that can be achieved as an arbitrary sequence of unitaries and partial thermalizations, where the Hamiltonian may be changed between one partial thermalization step and another. The objective, then, is to compose the dual-purpose qubit operation $\mathcal{M} \in \mathcal{S}$ which maps $\rho_1 \mapsto \eta_1$ and $\rho_2 \mapsto \eta_2$ and, assuming this can be done, to optimize its thermodynamic performance.

---

[1]This approach fails to work exactly if ρ or η are pure states, since $H_{\rho/\eta}$ would require an infinite energy gap. However, the operation can still be carried out *approximately*, mapping between mixed states ε close to ρ and η, for a finite energy cost.

[2]Unless the objective is to reset the system to a standard state, for example $\eta_1 = \eta_2 = |0\rangle\langle 0|$, as in Landauer erasure. In that case the above indeed represents an optimal, thermodynamically reversible strategy.

[3]This is clearly a simplification of open-system dynamics, but it is a useful model and commonly used in quantum thermodynamics [38–42]. Similar dynamics can be recovered from collisional models of thermalization [43].
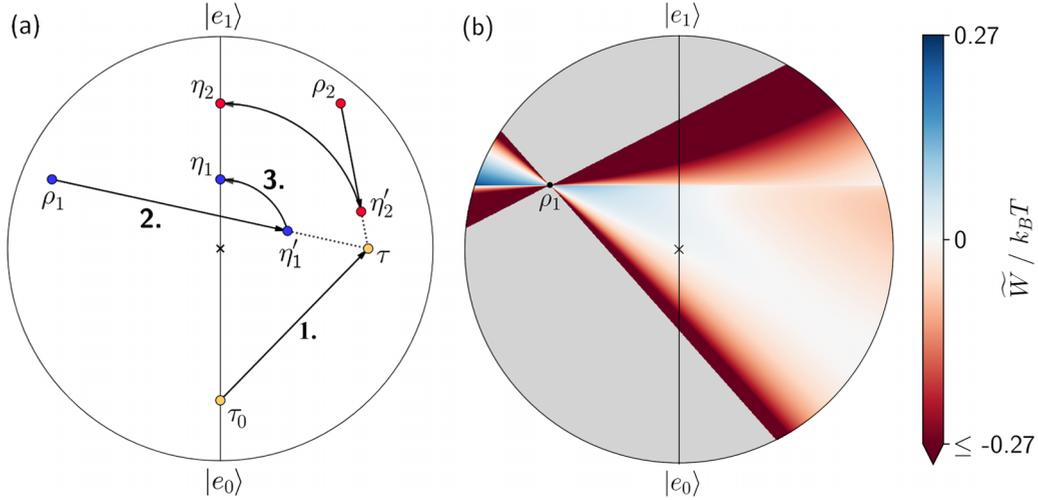
FIG. 2. (a) A work-from-coherence protocol represented in the Bloch sphere. The operation removes coherences in the energy eigenbasis for two possible input states ρ_1 and ρ_2 while preserving energy level populations, leaving the system in the final state η_1 or η_2. (1) The qubit Hamiltonian is quickly adjusted, changing the Gibbs state from τ_0 to τ. (2) The qubit is allowed to partially thermalize: the initial state—say, ρ_1—is transformed to a mixture with the Gibbs state: $\eta_1' = \lambda \rho_1 + (1 - \lambda)\tau$. (3) The Hamiltonian is quickly reset to its original value, and a unitary rotation U is applied, taking the qubit from η_1' to the final state η_1. The values of λ, τ, and U are uniquely chosen such that the same steps would also transform ρ_2 to η_2. (b) Extractable work plotted as a function of ρ_2, in the plane containing ρ_1, $|e_0\rangle$, and $|e_1\rangle$. In the gray region, it is not possible to construct a protocol out of unitaries and partial thermalizations that works for both inputs: in the geometric construction of panel (a), the thermal state τ would have to lie *outside* the Bloch sphere. This is in contrast with work-from-coherence protocols optimized for a single input: in that case, positive work extraction is possible for any state with coherences [11]. $\widetilde{W}$ is obtained by numerically optimizing Eq. (15) with $p_1 = p_2 = \frac{1}{2}$ and $N = 20$.

III. FEASIBILITY

Before turning to thermodynamic considerations, we examine whether it is actually possible to carry out a two-input qubit operation within the framework set out above. As we will find, the answer depends nontrivially on the input and output states. To see this, we first establish that the set $\mathcal{S}$ of possible transformations is identical to the set $\mathcal{S}_1 \subset \mathcal{S}$ of transformations that can be carried out with a *single* partial thermalization followed by a unitary. Given inputs and outputs for which the transformation is feasible, we are able to explicitly construct a protocol belonging to $\mathcal{S}_1$ [see Fig. 2(a)].

As a starting point, consider a generic qubit map $\mathcal{M} \in \mathcal{S}$ composed as a finite sequence of unitaries $\mathcal{U}$ and partial thermalizations $\mathcal{T}$, something of the form $\mathcal{M} = \mathcal{T}\mathcal{T}\mathcal{U}\mathcal{U}\mathcal{U}\mathcal{T}\mathcal{U}\ldots\mathcal{T}\mathcal{U}$. By combining consecutive unitaries and separating consecutive thermalizations with trivial unitaries $\mathcal{U} = \mathbb{I}$, such an operation can always be rewritten in the form $\mathcal{T}_N \mathcal{U}_N \ldots \mathcal{T}_2 \mathcal{U}_2 \mathcal{T}_1 \mathcal{U}_1$. Next we observe that any composite operation of the form $\mathcal{U}^{-1}\mathcal{T}\mathcal{U}$ carries out a mapping identical to a partial thermalization $\mathcal{T}'$ towards $\tau' = U^\dagger \tau U$. To see this, consider any state ρ, where we obtain

$$
\begin{aligned}
\mathcal{U}^{-1}\mathcal{T}\mathcal{U}(\rho) &= U^\dagger[\lambda[U\rho U^\dagger] + (1-\lambda)\tau]U \\
&= \lambda\rho + (1-\lambda)U^\dagger\tau U \\
&= \mathcal{T}'(\rho).
\end{aligned}
\tag{3}
$$

If τ is the Gibbs state for a Hamiltonian H, then τ' is the Gibbs state for $H' = U^\dagger H U$. In addition, the heat absorbed from the environment is the same for both processes, and equal to the change in internal energy during the partial thermalization

step:

$$
\begin{aligned}
Q &= \text{tr}\{H\left(\lambda[U\rho U^\dagger] + (1-\lambda)\tau\right)\} - \text{tr}[H[U\rho U^\dagger]] \\
&= \text{tr}\{H'\left[\lambda\rho + (1-\lambda)\tau'\right]\} - \text{tr}[H'\rho] \\
&= Q'.
\end{aligned}
\tag{4}
$$

For any state, $\mathcal{U}^{-1}\mathcal{T}\mathcal{U}$ and $\mathcal{T}'$ carry out the same transformation at the same energy cost, and so they are for our purposes equivalent. This can be employed to further simplify our ansatz operation. Writing $\mathcal{U}_i' = \mathcal{U}_i \ldots \mathcal{U}_2 \mathcal{U}_1$, and $\mathcal{T}_i' = \mathcal{U}_i'^{-1}\mathcal{T}_i\mathcal{U}_i'$, and using the fact that $\mathcal{U}_i'(\mathcal{U}_{i-1}')^{-1} = \mathcal{U}_i$,

$$
\begin{aligned}
\mathcal{T}_N &\mathcal{U}_N \mathcal{T}_{N-1}\mathcal{U}_{N-1} \ldots \mathcal{T}_2\mathcal{U}_2\mathcal{T}_1\mathcal{U}_1 \\
&= \mathcal{U}_N'\left(\mathcal{U}_N'^{-1}\mathcal{T}_N\mathcal{U}_N'\right)\left(\mathcal{U}_{N-1}'^{-1}\mathcal{T}_{N-1}\mathcal{U}_{N-1}'\right)\cdots \\
&\quad \cdots\left(\mathcal{U}_2'^{-1}\mathcal{T}_2\mathcal{U}_2'\right)\left(\mathcal{U}_1'^{-1}\mathcal{T}_1\mathcal{U}_1'\right) \\
&= \mathcal{U}_N'\,\mathcal{T}_N'\mathcal{T}_{N-1}' \ldots \mathcal{T}_2'\mathcal{T}_1'.
\end{aligned}
\tag{5}
$$

So, our original operation $\mathcal{M}$ could be implemented at equal energy cost by a sequence of consecutive partial thermalizations $\mathcal{T}_i'$ followed by a single unitary $\mathcal{U}'$. While the latter might pose greater practical challenges, the fundamental limitations are unchanged. Hence we can drop the primes and consider protocols of the form $\mathcal{U}\mathcal{T}_N \ldots \mathcal{T}_2\mathcal{T}_1$ for the rest of the paper, without loss of generality.[4] The next step is to recognize that $\mathcal{T}_N \ldots \mathcal{T}_2\mathcal{T}_1$ can be replaced by a single partial thermalization.

---

[4]Strictly speaking, a quench still occurs between successive partial thermalizations, but its immediate effect on the qubit state is the identity map, which is omitted here for conciseness.

Starting with $N = 2$, let $\mathcal{T}_1$ be characterized by (λ_1, τ_1) and $\mathcal{T}_2$ by (λ_2, τ_2). Then for any state ρ,

$$
\begin{aligned}
\mathcal{T}_2\mathcal{T}_1(\rho) &= \lambda_2[\lambda_1\rho + (1 - \lambda_1)\tau_1] + (1 - \lambda_2)\tau_2 \\
&= \lambda_2\lambda_1\rho + (1 - \lambda_2\lambda_1)\left[\frac{\lambda_2(1 - \lambda_1)}{1 - \lambda_2\lambda_1}\tau_1 \right. \\
&\quad \left. + \frac{1 - \lambda_2}{1 - \lambda_2\lambda_1}\tau_2\right] \\
&\equiv \lambda_2^{\text{eff}}\rho + \left(1 - \lambda_2^{\text{eff}}\right)\tau_2^{\text{eff}},
\end{aligned}
\tag{6}
$$

where we have defined $\lambda_2^{\text{eff}} = \lambda_2\lambda_1$ and $\tau_2^{\text{eff}} = \frac{\lambda_2(1-\lambda_1)}{1-\lambda_2\lambda_1}\tau_1 + \frac{1-\lambda_2}{1-\lambda_2\lambda_1}\tau_2$. It is straightforward to check that $\lambda_2^{\text{eff}} \in [0, 1]$, and that τ_2^{eff} is a convex combination of density operators and is therefore itself a density operator. Hence τ_2^{eff} can be understood to be the thermal state corresponding to a Hamiltonian $-k_{\mathrm{B}}T(\ln\tau_2^{\text{eff}} + \ln Z\,\mathbb{I})$. So, a single partial thermalization $\mathcal{T}_2^{\text{eff}}$ characterized by $(\lambda_2^{\text{eff}}, \tau_2^{\text{eff}})$ carries out the same mapping as $\mathcal{T}_2\mathcal{T}_1$.

The result extends inductively: if $\mathcal{T}_N \ldots \mathcal{T}_1$ reduces to $\mathcal{T}_N^{\text{eff}}$, then $\mathcal{T}_{N+1}\mathcal{T}_N \ldots \mathcal{T}_1$ can be written as $\mathcal{T}_{N+1}\mathcal{T}_N^{\text{eff}}$, which can be reduced to $\mathcal{T}_{N+1}^{\text{eff}}$, by the above. The mixing parameter and thermal state are given by

$$
\begin{aligned}
\lambda_N^{\text{eff}} &= \prod_{i=1}^{N}\lambda_i \\
\tau_N^{\text{eff}} &= \frac{\lambda_N\left(1 - \lambda_{N-1}^{\text{eff}}\right)}{1 - \lambda_N^{\text{eff}}}\tau_{N-1}^{\text{eff}} + \frac{1 - \lambda_N}{1 - \lambda_N^{\text{eff}}}\tau_N
\end{aligned}
\tag{7}
$$

where $\lambda_0^{\text{eff}} = 1$ and $\tau_0^{\text{eff}} = \mathbb{I}$. The upshot of this is that a two-step protocol $\mathcal{U}\mathcal{T}_N^{\text{eff}}$ carries out the same mapping as the $N{+}1$ step protocol $\mathcal{U}\mathcal{T}_N \ldots \mathcal{T}_1$.

What we have shown is that for every operation $\mathcal{M} \in \mathcal{S}$, we also have $\mathcal{M} \in \mathcal{S}_1$. There is an important caveat: while $\mathcal{U}\mathcal{T}_N^{\text{eff}}$ executes the same mapping as $\mathcal{U}\mathcal{T}_N \ldots \mathcal{T}_1$, the total heat absorbed during the process will typically differ. As we will see in Sec. IV, it is generally thermodynamically favorable to use $N \gg 1$ thermalizing steps.

We are now ready to decide on the existence of a suitable qubit transformation acting on two generic input states, and give details on how to construct it. Consider an operation $\mathcal{M} \in \mathcal{S}$, with input states ρ_1, ρ_2 and outputs η_1, η_2.[5] By the above, $\mathcal{M}$ can be written as $\mathcal{U}\mathcal{T}$ where $\mathcal{U}$ is characterized by some unitary matrix U, and $\mathcal{T}$ by a mixing parameter $\lambda \in [0, 1]$ and density matrix τ. Then λ, τ, and U must satisfy the following:

$$
\begin{aligned}
U[\lambda\rho_1 + (1 - \lambda)\tau]U^{\dagger} &= \eta_1 \\
U[\lambda\rho_2 + (1 - \lambda)\tau]U^{\dagger} &= \eta_2.
\end{aligned}
\tag{8}
$$

Subtracting the second line from the first we obtain

$$
\lambda U(\rho_1 - \rho_2)U^{\dagger} = \eta_1 - \eta_2,
\tag{9}
$$

and taking the trace norm[6] of both sides leads to an expression for the mixing parameter:

$$
\lambda = \frac{\|\eta_1 - \eta_2\|_1}{\|\rho_1 - \rho_2\|_1}.
\tag{10}
$$

The requirement that $\lambda \leqslant 1$ necessitates that $\|\eta_1 - \eta_2\|_1 \leqslant \|\rho_2 - \rho_1\|_1$. Since $\rho_1 - \rho_2$ is a traceless Hermitian operator on a qubit, it can be diagonalized as $p|\psi_+\rangle\langle\psi_+| - p|\psi_-\rangle\langle\psi_-|$, for some $p > 0$, and some orthonormal states $|\psi_+\rangle$ and $|\psi_-\rangle$. Likewise, we can write $\eta_1 - \eta_2 = q|\phi_+\rangle\langle\phi_+| - q|\phi_-\rangle\langle\phi_-|$: note that $\|\rho_1 - \rho_2\|_1 = 2p$ and $\|\eta_1 - \eta_2\|_1 = 2q$. Substituting these expressions into Eq. (9) leads to a formula for U:

$$
U = |\phi_+\rangle\langle\psi_+| + |\phi_-\rangle\langle\psi_-|.
\tag{11}
$$

Now that λ and U have been determined in terms of ρ_1, ρ_2, η_1, and η_2, a rearrangement of the first line[7] of Eq. (8) yields an expression for τ:

$$
\tau = \frac{1}{1 - \lambda}[U^{\dagger}\eta_1 U - \lambda\rho_1].
\tag{12}
$$

Equation (12) by itself guarantees that τ is a Hermitian operator with unit trace. However, we also require $\tau \geqslant 0$ in order that τ is a well-defined density matrix.

So, given an operation specified by inputs ρ_1, ρ_2 and outputs η_1, η_2, it is straightforward to determine whether the operation belongs in $\mathcal{S}_1$: we can evaluate λ, U, and τ using Eqs. (10)–(12), and check that $\lambda \leqslant 1$ and $\tau \geqslant 0$. If both these conditions hold, then the above construction provides an explicit protocol $\mathcal{U}\mathcal{T}$ to carry out the operation [see Fig. 2(a)]. Moreover, the operation can also be implemented using an extended protocol $\mathcal{U}\mathcal{T}_N \ldots \mathcal{T}_1 \in \mathcal{S}_N$ for which λ_N^{eff} and τ_N^{eff} as given in (7) are equal to λ and τ in Eqs. (10) and (12). Conversely, if $\lambda > 1$ or if τ has a negative eigenvalue, then the dual-purpose operation cannot be composed as a sequence of unitaries and partial thermalizations, no matter how many steps are used.[8]

IV. WORK EXTRACTION

The decomposition of a qubit operation into unitaries and partial thermalizations is not unique—and the various ways it can be done generally involve differing net transfers of work and heat. This leads to the question of optimal implementation: for a given set of inputs and outputs, how can the operation be carried out so as to yield as much work as possible (or expend as little as possible)? And how does the optimum compare against the fundamental bound set by the second law?

---

[5]We assume a few things about the inputs and outputs. First, $(\rho_1, \rho_2) \neq (\eta_1, \eta_2)$, so that the operation is not simply the identity. Second, $\eta_1 \neq \eta_2$: otherwise, we could simply adjust the Hamiltonian such that the Gibbs state coincides with η_1, and allow the qubit to completely thermalize. This second assumption also implies that $\rho_1 \neq \rho_2$, because CPTP quantum operations are never one-to-many.

[6]$\|\rho\|_1 = \text{tr}[\sqrt{\rho\rho^{\dagger}}]$.

[7]Doing the same for the second line (or any affine combination of the two lines) would give an equivalent expression.

[8]Not all CPTP channels can be composed in this way: the set of feasible operations can be expanded through the use of an ancilla system. To operate in a cycle, the ancilla's state must be reset, erasing thermodynamically relevant system-ancilla correlations [18,32]. For the simplicity of analysis, we here consider only those qubit operations which can be implemented without the use of ancillas.

We approach this problem by first deriving an expression for the average work extraction for a protocol involving N partial thermalization steps. This is employed in Sec. IV A to numerically solve for optimum *work from coherence* protocols for two inputs, allowing for comparison against the reversible single-input protocol in [11]. In Sec. IV B we derive a general upper bound on work extraction for two-input operations, opening up an interpretation of why thermodynamic reversibility cannot be achieved.

We choose as a figure of merit the average work yield, $\overline{W} = p_1 W_1 + p_2 W_2$, where p_1 is the probability that the input is ρ_1, and W_1 is the work extracted in transforming ρ_1 to η_1 (and likewise for the second input). Since the overall change in the qubit's expected energy is path independent, then by the first law ($\Delta \overline{U} = \overline{Q} - \overline{W}$), maximizing work extraction amounts to the same thing as maximizing the heat drawn from the environment.

First, let us evaluate the heat for a single partial thermalization, with mixing parameter λ and thermal state τ, acting on a state ρ. The heat absorbed is equal to the resulting difference in the qubit's expected energy:

$$Q = \mathrm{tr}[H(\lambda\rho + (1-\lambda)\tau)] - \mathrm{tr}[H\rho]$$
$$= -k_{\mathrm{B}}T(1-\lambda)\mathrm{tr}[(\tau - \rho)\ln\tau]. \quad (13)$$

In the second line, we used that $H = -k_{\mathrm{B}}T[\ln(\tau) + \ln Z\,\mathbb{I}]$. Now, consider a protocol $\mathcal{U}\mathcal{T}_N \ldots \mathcal{T}_1$: The average heat drawn from the environment is given by a sum over the N partial thermalization steps, averaged over the input states (the unitary step involves no heat transfer). Letting $\rho_n^i = \mathcal{T}_{i-1}\ldots\mathcal{T}_1(\rho_n)$ denote the transformed state of ρ_n prior to the ith partial thermalization, for $n = 1, 2$, we have

$$\overline{Q} = -k_{\mathrm{B}}T\sum_{n=1}^{2}p_n\sum_{i=1}^{N}(1-\lambda_i)\mathrm{tr}\left[(\tau_i - \rho_n^i)\ln\tau_i\right]$$
$$= -k_{\mathrm{B}}T\sum_{i=1}^{N}(1-\lambda_i)\mathrm{tr}[(\tau_i - \overline{\rho}_i)\ln\tau_i]$$
$$= k_{\mathrm{B}}T\sum_{i=1}^{N}(1-\lambda_i)[S(\tau_i) - S(\overline{\rho}_i) - S(\overline{\rho}_i||\tau_i)] \quad (14)$$

where $\overline{\rho}_i = \mathcal{T}_{i-1}\ldots\mathcal{T}_1(\overline{\rho}) \equiv p_1\rho_1^i + p_2\rho_2^i$ is the average state before the ith partial thermalization, and $S(\rho||\sigma) = \mathrm{tr}[\rho\ln\rho - \rho\ln\sigma]$ is the quantum relative entropy. Invoking the first law, the work yield is given by subtracting the overall change of expected energy from (14), to give

$$\overline{W} = k_{\mathrm{B}}T\sum_{i=1}^{N}(1-\lambda_i)[S(\tau_i) - S(\overline{\rho}_i) - S(\overline{\rho}_i||\tau_i)]$$
$$- \mathrm{tr}[H_0(\overline{\eta} - \overline{\rho})]. \quad (15)$$

The above is straightforward to evaluate given the parameters λ_i, τ_i. However, we are ultimately interested in the maximum possible work extraction, $\widetilde{W}$, over all suitable choices of protocol:

$$\widetilde{W} = \lim_{N\to\infty}\max_{\{\lambda_i\},\{\tau_i\}}\overline{W}, \quad (16)$$

where the maximization[9] is subject to the constraints that $\lambda_N^{\mathrm{eff}} = \lambda$ and $\tau_N^{\mathrm{eff}} = \tau$ as given in Eqs. (7), (10), and (12), ensuring that the protocol maps $\rho_1 \mapsto \eta_1$ and $\rho_2 \mapsto \eta_2$. Note that the expression for $\overline{W}$ (15) only depends on the trajectory of the average state $\overline{\rho}$: the individual inputs and outputs ρ_1, ρ_2, η_1, and η_2 enter Eq. (16) implicitly through the optimization constraints.

Recall that some operations are unfeasible because any suitable protocol would involve thermalization towards a state τ_i with a negative eigenvalue. For operations at the *boundary* of feasibility, τ_i must approach a pure state, with a zero eigenvalue. But in this limit, the relative entropy term $S(\overline{\rho}_i||\tau_i)$ appearing in (15)—and consequently the work cost of the operation—diverge towards $+\infty$.

While (16) does not admit a simple closed form, it does provide a starting point for numerical investigations, and will later provide the basis for an analytic upper bound on work extraction, (23). We effectively already have a lower bound on $\widetilde{W}$ stemming from the protocol derived at the end of Sec. III: the expressions for λ and τ in Eqs. (10) and (12) can be directly substituted into (15).

A. Work from coherence

In the previous subsection we derived a formula for the average work yield of a two-input qubit operation (15). In order to better understand its properties, let us apply it to a specific operation: work extraction from coherence.

By this we mean a process of work extraction which leaves unchanged the state's diagonal matrix elements $\langle e_n|\rho|e_n\rangle$ in the eigenbasis of the initial Hamiltonian H_0, and which involves a single heat bath at fixed temperature T and a cyclic variation of the Hamiltonian, returning it to H_0 at the end of the process [11,13,28]. In this context, coherences—that is, off-diagonal elements $\langle e_m|\rho|e_n\rangle$—represent a thermodynamic resource. A state ρ with nonzero coherences has lower von Neumann entropy than the corresponding *decohered* state $\eta = \mathrm{diag}(\rho) := |e_0\rangle\langle e_0|\rho|e_0\rangle\langle e_0| + |e_1\rangle\langle e_1|\rho|e_1\rangle\langle e_1|$. Since the expected energy is unchanged, i.e., $\mathrm{tr}[H_0(\eta - \rho)] = 0$, transforming ρ to η results in a reduction in free energy $\Delta F = k_{\mathrm{B}}T[S(\eta) - S(\rho)] \leqslant 0$. Consequently, if carried out reversibly, the operation can extract positive work, such as in the protocol developed in [11] for a single input state.

Still, the question remains how closely a *dual-purpose* work from coherence protocol can approach the reversible work yield. We consider an operation with two inputs $\rho_{1,2}$, and completely decohered outputs:[10]

$$\eta_{1,2} = \mathrm{diag}(\rho_{1,2}). \quad (17)$$

Figure 2(a) represents a protocol which carries out this mapping using a single partial thermalization step. If the process

---

[9]The optimum work extraction must increase monotonically with the number of partial thermalization steps N: for any N-step protocol we can construct an $N+1$ step protocol for the same operation with identical work yield by decomposing one of its partial thermalization steps $\mathcal{T}_i = (\lambda_i, \tau_i)$ into two consecutive steps $\mathcal{T}_i^{(1)} = \mathcal{T}_i^{(2)} = (\sqrt{\lambda_i}, \tau_i)$.

[10]By linearity, $\overline{\eta} = \mathrm{diag}(\overline{\rho})$.
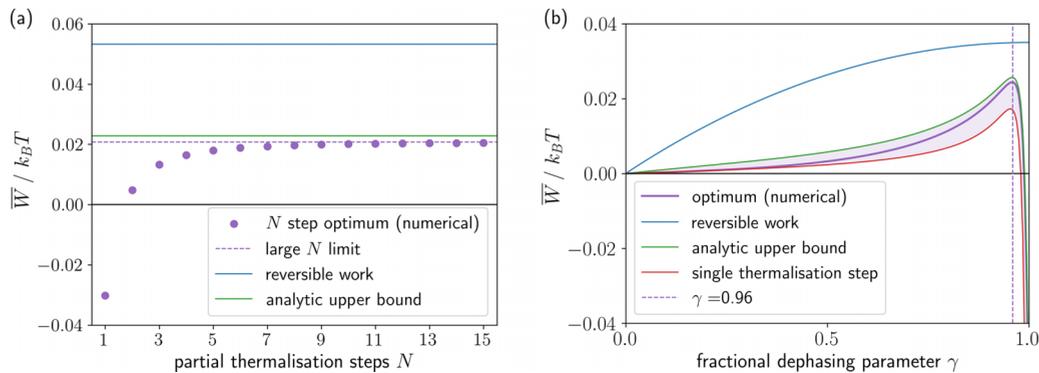
FIG. 3. (a) Coherence work extraction for protocols involving increasing numbers of partial thermalization steps N, obtained via numerical optimization (purple scatter). Thermodynamic performance rapidly improves with N, converging towards an upper limit (dotted line) which is still significantly lower than the work extraction for a thermodynamically reversible operation, given by the free energy reduction $k_{\mathrm{B}}T[S(\overline{\eta}) - S(\overline{\rho})]$ (blue line). The improved analytic bound Eq. (23) (green line) provides a closer estimate of the dual-purpose work extraction, accounting for the constraint of preserving distinguishability between the two input states. The scatter was obtained by numerically maximizing Eq. (15) over $\tau_1, \ldots, \tau_N$, for input states with Bloch vectors $\rho_1 = (0.735, 0.273, -0.286)$ and $\rho_2 = (-0.496, -0.470, -0.294)$, and outputs $\eta_1 = (0, 0, -0.286)$ and $\eta_2 = (0, 0, -0.294)$, with $p_1 = p_2 = \frac{1}{2}$. The initial Hamiltonian is taken as $H_0 = E_0\sigma_z$. (b) Work extraction for an operation which *partially* removes coherences in the energy eigenbasis (purple curve). Off-diagonal elements $\langle e_m|\rho|e_n\rangle$ are scaled by a factor $1 - \gamma$, such that $\gamma = 0$ leaves the state unchanged and $\gamma = 1$ corresponds to complete dephasing. Surprisingly, for this pair of input states, maximum work extraction is *not* achieved for complete dephasing, where the reduction of free energy (blue curve) is greatest: in fact the maximum is at $\gamma \approx 0.96$ (dotted line). The analytic bound (23) (green curve) better captures the dependence of work extraction on γ. On the other hand, the single step work extraction (red curve) provides a lower bound on the optimum, and since λ (10) and τ (12) are uniquely determined in this case we can constrain $\widetilde{W}$ to the shaded window in the figure without performing any numerical optimization. For this plot, the input Bloch vectors were $\rho_1 = (0.249, 0.183, 0.494)$ and $\rho_2 = (-0.044, -0.640, 0.508)$, and the outputs $\eta_1 = (0.249(1-\gamma), 0.183(1-\gamma), 0.494)$ and $\eta_2 = (-0.044(1-\gamma), -0.640(1-\gamma), 0.508)$, with $p_1 = p_2 = \frac{1}{2}$. The optimum two-input work output was computed by maximizing (15) using $N=20$ steps: as can be seen in panel (a), this is sufficient for the work extraction to converge with reasonable precision. For both panels, the input states were generated randomly and postselected to illustrate the sign change for work extraction with 2+ thermalization steps in panel (a), and to show the thermodynamic advantage of incomplete dephasing in panel (b).

could be carried out in a thermodynamically reversible way, it would extract work equal to the free energy reduction $-\Delta\overline{F} = k_{\mathrm{B}}T[S(\overline{\eta}) - S(\overline{\rho})] \geqslant 0$, on average over the inputs.

On the other hand, the actual maximum work extraction can be estimated numerically using Eq. (15). Fixing the states ρ_1, ρ_2 and the number of partial thermalization steps N, the work output $\overline{W}$ can be maximized as a function of the thermal states $\tau_1, \ldots, \tau_{N-1}$ and mixing parameters $\lambda_1, \ldots, \lambda_{N-1}$.[11] If N is chosen sufficiently large, the resulting value of $\overline{W}$ will be a close approximation to the overall optimum $\widetilde{W}$. Figure 3(a) shows convergent behavior of $\overline{W}$ with increasing N. Typically, 20 thermalization steps are sufficient for the work yield to converge within a 1% margin of its limiting value. As an initial guess for optimization, we took that for all i, $\tau_i = \tau$ as given in Eq. (12). It was assumed[12] throughout that $\lambda_i = (\frac{\|\eta_1 - \eta_2\|_1}{\|\rho_1 - \rho_2\|_1})^{\frac{1}{N}}$.

The resulting numerical approximation to $\widetilde{W}$ is conspicuously lower than the change in free energy. In fact, for many pairs of input states, work extraction is not possible at all. Uniform sampling of the Bloch sphere reveals that the

operation (17) is only *feasible* for 62% of pairs of input states: the remainder violate the conditions established at the end of Sec. III, and there is no way to carry out the intended mapping for both inputs using unitaries and partial thermalizations. Positive work extraction is only possible for 10% of the feasible input pairs with $p_1 = p_2 = \frac{1}{2}$. The heatmap in Fig. 2(b) gives an indication of the relative scale of these regimes: unfeasible pairs are marked in gray, and ones where work can be extracted are marked in blue. For the majority of states, it is possible to remove coherences while preserving the energy distribution, but it costs work to do so—these pairs lie in the red regions.

This is a stark departure from single input work from coherence, which can be performed reversibly for any state, and can *always* be used to extract work. Evidently, the compromises necessary for a dual-purpose operation impose a severe energetic penalty. In Sec. IV B we will investigate the thermodynamic explanation for this penalty, and derive an improved bound on work extraction, Eq. (23) (plotted as the green lines in Fig. 3).

In the meantime there is an interesting caveat to explore. Up to this point we have considered an operation which completely removes energy-basis coherences. However, in a broader sense, work from coherence can include any operation which leaves the energy level populations unchanged—that is, one for which $\mathrm{diag}(\eta_{1,2}) = \mathrm{diag}(\rho_{1,2})$. This could mean simply scaling the coherences by some factor $1 - \gamma \in [0, 1]$, such that $\gamma = 0$ leaves the state unchanged and $\gamma = 1$ corresponds

---

[11]For the final partial thermalization, λ_N and τ_N are fixed by the constraint that $\lambda_N^{\mathrm{eff}} = \lambda$ and $\tau_N^{\mathrm{eff}} = \tau$.

[12]While this is a nontrivial assumption for finite N, it should make no difference in the limit as $N \to \infty$. Any partial thermalization step (λ, τ) can be approximated by n consecutive steps $((\frac{\|\eta_1 - \eta_2\|_1}{\|\rho_1 - \rho_2\|_1})^{\frac{1}{N}}, \tau)$, where n is chosen such that $(\frac{\|\eta_1 - \eta_2\|_1}{\|\rho_1 - \rho_2\|_1})^{\frac{n}{N}} \approx \lambda$.

to complete removal of coherences:

$$\eta_{1,2} = (1 - \gamma)\rho_{1,2} + \gamma \, \mathrm{diag}(\rho_{1,2}). \tag{18}$$

The work yield for an operation of this type is plotted as a function of γ in Fig. 3(b). Surprisingly, maximum work extraction is *not* necessarily achieved by completely removing coherences—even though that would result in a greater reduction in free energy. For some pairs of inputs, such as those in Fig. 3(b), a partial decoherence operation may extract positive work where the full removal of coherences would cost work to perform, or even where full removal of coherences is not feasible.

B. Bound on extractable work

For operations with a single input and output state, thermodynamic reversibility can be recovered in the limit of large N: crucially, the system's state must remain infinitesimally close to thermal equilibrium whenever it is in contact with the environment [31]. This is the familiar quasistatic limit. Why can we not do the same for a multipurpose operation by ensuring that the average state undergoes a quasistatic evolution?

Consider the term $\sum_{i=1}^{N}(1 - \lambda_i)[S(\tau_i) - S(\overline{\rho_i})]$ appearing in the equation for $\overline{W}$ (15). Due to the concavity of the von Neumann entropy, for all i,

$$\begin{aligned}
(1-\lambda_i)[S(\tau_i) - S(\overline{\rho_i})] &= [\lambda_i S(\overline{\rho_i}) + (1-\lambda_i)S(\tau_i)] - S(\overline{\rho_i}) \\
&\leqslant S(\lambda_i \overline{\rho_i} + (1-\lambda_i)\tau_i) - S(\overline{\rho_i}) \\
&= S(\overline{\rho}_{i+1}) - S(\overline{\rho_i}).
\end{aligned} \tag{19}$$

So, the sum over i is bounded by the following:

$$\begin{aligned}
\sum_{i=1}^{N}(1 - \lambda_i)[S(\tau_i) - S(\overline{\rho_i})] &\leqslant \sum_{i=1}^{N}[S(\overline{\rho}_{i+1}) - S(\overline{\rho_i})] \\
&= S(\overline{\eta}) - S(\overline{\rho}),
\end{aligned} \tag{20}$$

where we have used that the average state prior to the first partial thermalization step is equal to the average input state, $\overline{\rho}_1 = \overline{\rho}$, and that following the final partial thermalization the state is related by a unitary to the average output, $\overline{\rho}_{N+1} = U^{\dagger}\overline{\eta}U$, where U is given in Eq. (11). So, the overall work extraction is bounded by

$$\begin{aligned}
\widetilde{W} &\leqslant k_{\mathrm{B}}T[S(\overline{\eta}) - S(\overline{\rho})] - \mathrm{tr}[H_0(\overline{\eta} - \overline{\rho})] \\
&\quad - k_{\mathrm{B}}T \sum_{i=1}^{N}(1 - \lambda_i)S(\overline{\rho_i}||\tau_i).
\end{aligned} \tag{21}$$

The first two terms in (21) represent the free energy reduction $-\Delta\overline{F}$. The final term is a contribution to irreversible entropy production (wasted free energy) $k_{\mathrm{B}}T\,\Sigma = -\Delta\overline{F} - \widetilde{W}$, resulting from *lag* of the average state behind the thermal state during partial thermalizations. Similar lag terms appear in other works but are usually associated with suboptimal protocols, for example due to inaccurate knowledge of the initial state, limited control, or finite speed precluding ideal quasistatic processes [28,44–46]. Here we will show that the lag is necessary to preserve dependence on the initial state (see Fig. 4), and the resulting energy penalty is unavoidable for dual-purpose operations. To do so we employ Pinsker's
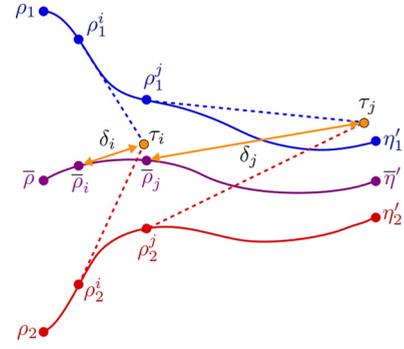
FIG. 4. State trajectories during a sequence of many partial thermalization steps, approximating a smooth curve. In the ith step, the qubit state (ρ_1^i or ρ_2^i) evolves a small fraction of the distance towards the thermal state τ_i, which varies throughout the protocol: the tangents to the states' trajectories point towards τ_i. The lag $\delta_i = \|\overline{\rho_i} - \tau_i\|$ of the average state $\overline{\rho_i}$ behind the thermal state τ_i leads to irreversible entropy production: the process can only approach thermodynamic reversibility if δ_i can be made arbitrarily small for *every* partial thermalization step (21). However, if δ_i is made too small, the distance between the two input states $\|\rho_1^i - \rho_2^i\|$ will decay much faster than the remaining distance $\|\overline{\rho_i} - \overline{\eta}'\|$ between the present average state $\overline{\rho_i}$ and the target average state $\overline{\eta}'$ at the end of the process. If δ_i is too small during one thermalizing step, it must be compensated for with a larger δ_j in another step, in order to reach the target output states while preserving the distinction between ρ_1 and ρ_2. The average lag δ throughout the process is lower bounded, meaning that entropy production and the resulting irreversible dissipation of free energy cannot be eliminated (23).

inequality $S(\rho||\sigma) \geqslant \frac{1}{2}\|\sigma - \rho\|_1^2$ [47] to relate the quantum relative entropy to the trace distance, and thereby bound the entropy production in (21) by

$$\begin{aligned}
\Sigma &\geqslant \frac{1}{2}\sum_{i=1}^{N}(1 - \lambda_i)\|\tau_i - \overline{\rho_i}\|_1^2 \\
&= \frac{1}{2}\sum_{i=1}^{N}\frac{1}{1 - \lambda_i}\|\overline{\rho}_{i+1} - \overline{\rho_i}\|_1^2,
\end{aligned} \tag{22}$$

where the second line was obtained by substituting $\overline{\rho}_{i+1} = \lambda_i \overline{\rho_i} + (1 - \lambda_i)\tau_i$. At a glance it might appear that the above can be made arbitrarily small. For the average step, $\|\overline{\rho}_{i+1} - \overline{\rho_i}\|_1$ need not exceed $\frac{1}{N}\|U^{\dagger}\overline{\eta}U - \overline{\rho}\|_1$: so the sum of squares in (22) might vanish in the limit of large N, provided that $\frac{1}{1-\lambda_i}$ remains bounded. However, each partial thermalization step reduces the distance $\|\rho_1^i - \rho_2^i\|$ between the input states by a factor λ_i—so the geometric mean of λ_i must equal $\left(\frac{\|\eta_1 - \eta_2\|_1}{\|\rho_1 - \rho_2\|_1}\right)^{\frac{1}{N}}$ [see Eq. (10)]. As a result, $\frac{1}{1-\lambda_i}$ does indeed diverge for large N. Put differently, while $\|\overline{\rho}_{i+1} - \overline{\rho_i}\|_1$ can vanish for all steps, the distance between the average state and the thermal state $\|\tau_i - \overline{\rho_i}\|_1 = \frac{1}{1-\lambda_i}\|\overline{\rho}_{i+1} - \overline{\rho_i}\|_1$ cannot, and this is ultimately the thermodynamically relevant one. Figure 4 illustrates the dilemma of minimizing irreversible dissipation while preserving distinguishability between the individual output states.

Applying standard geometric inequalities to Eq. (22) (see the Appendix), the extractable work (21) can be upper

bounded in terms of the input and output states:

$$\widetilde{W} \leqslant -\Delta \overline{F} - \frac{k_{\mathrm{B}}T}{2} \frac{\| U^{\dagger}\overline{\eta} U - \overline{\rho} \|_1^2}{\ln \left(\frac{\|\rho_1 - \rho_2\|_1}{\|\eta_1 - \eta_2\|_1} \right)}, \qquad (23)$$

where the unitary U is given by Eq. (11). The final term, representing unavoidable dissipation, fits well with the intuitive picture presented in Fig. 4. Looking at the numerator, the dissipation is greater for operations which require the average state $\overline{\rho}$ to be moved a greater distance through the Hilbert space to $U^{\dagger}\overline{\eta}U$ during the thermalizing portion of the protocol. From the denominator, we see that it is costly to preserve a relatively large trace distance between the individual outputs η_1 and η_2.

On the other hand, an operation can be made thermodynamically reversible only if the dissipation term (23) vanishes. Setting the numerator to zero implies that the average output is related to the average input by a unitary.[13] Looking at the denominator, dissipation asymptotically vanishes as $\|\eta_1 - \eta_2\|_1 \to 0$, with reversibility recovered in the limiting case where both inputs are mapped to the same output. Most prominently, Landauer erasure falls into this latter category, resetting the system from any initial state to a standard $|0\rangle\langle 0|$ state. The exceptional cases where the wasted free energy may vanish are in line with those identified by Bedingham and Maroney [32]. However, we are able to go further by specifying a procedure to carry out the operation, as well as methods to optimize for maximum work extraction.

The irreversible work cost implied by (23) does *not* disappear as the probabilities p_1 or p_2 approach zero. In fact we see a discontinuous jump in $\widetilde{W}$ between the single-input operation where input state ρ_1 occurs with certainty, and the dual-purpose operation where ρ_1 occurs with probability $1 - \varepsilon$ and ρ_2 with probability ε.

While (23) is not a tight bound, it gives a significantly closer estimate to the optimum work extraction $\widetilde{W}$ than the free energy reduction alone. This is evident in both panels of Fig. 3, where the improved bound is plotted alongside the reversible work $-\Delta \overline{F}$, as well as numerically optimized protocols.

In practice, the need to avoid complete thermalization means that thermal contact with the environment must be limited in duration. In a simple model, we might treat the qubit as approaching the Gibbs state at an exponential rate with a fixed timescale t_{th}, such that for each partial thermalization step, the mixing parameter λ_i can be expressed in terms of the duration of the step t_i, as $\lambda_i = \exp(-\frac{t_i}{t_{\mathrm{th}}})$.

In that case, the total duration $t_{\mathrm{tot}} = \sum_{i=1}^{N} t_i$ of thermal contact is determined by the reduction in trace distance between

the inputs:

$$\frac{t_{\mathrm{tot}}}{t_{\mathrm{th}}} = -\ln \prod_{i=1}^{N} \lambda_i = \ln \left(\frac{\|\rho_1 - \rho_2\|_1}{\|\eta_1 - \eta_2\|_1} \right). \qquad (24)$$

This in turn means that the irreversible dissipation of free energy is inversely proportional to t_{tot}. From (23), we have

$$-\Delta \overline{F} - \widetilde{W} \geqslant \frac{k_{\mathrm{B}}T}{2} \frac{t_{\mathrm{th}}}{t_{\mathrm{tot}}} \| U^{\dagger}\overline{\eta}U - \overline{\rho} \|_1^2. \qquad (25)$$

Dissipation could vanish in the limit of large t_{tot}, but the trace distance between the outputs $\|\eta_1 - \eta_2\|_1$ would vanish too, as can be seen from Eq. (24). This means that operations with more than one output *cannot* be carried out arbitrarily slowly, even though it would be thermodynamically favorable to do so.

V. CONCLUSION AND OUTLOOK

In the context of information processing, it is necessary to consider physical processes which produce different final states depending on how the system is initialized. We have examined the thermodynamics of a simple case: qubit transformations with two possible input states. For what might appear a straightforward extension beyond single-input processes, the additional constraints have drastic consequences.

In Sec. III we characterized the class of two-input qubit operations which can be implemented through a combination of unitary and partial thermalization steps, and determined a three step procedure for carrying them out. In Sec. IV, we found that even for thermodynamically optimal protocols, a considerable fraction of the free energy change must be irreversibly dissipated. This is attributed to the requirement that the system is significantly out of equilibrium during thermal contact with the environment. By quantifying the resultant entropy production, we derived an improved upper bound on the extractable work [Eq. (23)], and related it to an effective minimum speed limit.

Naturally, there arises the question of extending the present quantitative results to general n-input multipurpose operations. We expect this extension to be nontrivial, not least because for any three states ρ_1, ρ_2, ρ_3 of a d-dimensional system, the generalized angle $\frac{\mathrm{tr}[(\rho_1 - \rho_2)(\rho_1 - \rho_3)]}{\mathrm{tr}[\rho_1 - \rho_2]\mathrm{tr}[\rho_1 - \rho_3]}$ is preserved under both unitaries and partial thermalizations. If this angle were to differ between a triplet of input states and the corresponding outputs, the transformation would be impossible without expanding the set of primitive operations. For 3+ dimensional systems, this could be approached by incorporating *partial level thermalizations*, whereby partial thermalization occurs within energy subspaces [41]. There is also the prospect of addressing the same problem using a more sophisticated treatment of thermalization dynamics, for example by accounting for non-Markovian effects.

On the other hand, the arguments in Sec. II A which rule out reversible quasistatic processes extend readily, and we expect that qualitatively similar results may hold for quantum operations of higher complexity. At a basic level, more input states will always mean a more constrained thermodynamic optimization, and can only mean greater energy penalties: in

---

[13]Moreover, it must be the *same* unitary U (11) that maps the eigenbasis of $\rho_1 - \rho_2$ to that of $\eta_1 - \eta_2$. However, this does not imply that the overall operation is purely unitary, in that $\eta_{1,2} = U \rho_{1,2} U^{\dagger}$. For example, the operation could remain thermodynamically reversible if the qubit underwent partial thermalization towards a thermal state τ identical to the average state $\overline{\rho}$.

this sense, much of the insight can already be gained from the two input case analyzed here.

Sometimes, quantum investigations uncover analogous classical effects [13]. Arguably the only truly quantum principle in play here is the inability to freely copy or measure the state[14] and choose the protocol accordingly. We might imagine an analogous classical scenario where practical constraints prevent us from incorporating measurement feedback in our process: in that case we anticipate that a similar thermodynamic handicap might apply.

As we have seen, optimizing a quantum thermodynamic process for more than one input is more costly than the free energy difference would suggest. Conversely, holding prior knowledge of the system's state confers a greater advantage, emphasizing the fuel value of information [18,49]. This adds to a growing recognition that the Clausius inequality and Landauer bound do not tell the whole story with respect to irreversibility [29,32,46,50,51].

The code used to produce the data in Figs. 2 and 3 is available upon reasonable request to JD, j.dunlop@exeter.ac.uk.

APPENDIX: DERIVING THE BOUND ON ENTROPY PRODUCTION

Applying the Cauchy-Schwarz inequality to (22), the entropy production Σ can be bounded by

$$\Sigma \geqslant \frac{1}{2} \sqrt{\sum_{i=1}^{N} \frac{1}{(1-\lambda_i)^2}} \sqrt{\sum_{i=1}^{N} \|\overline{\rho}_{i+1} - \overline{\rho}_i\|_1^4}. \qquad (A1)$$

The first square-root term in the above can be bounded using the power mean inequality and the fact that $\ln x \leqslant x - 1$ for all $x > 0$:

$$\left(\frac{1}{N} \sum_{i=1}^{N} (1-\lambda_i)^{-2} \right)^{-\frac{1}{2}} \leqslant \frac{1}{N} \sum_{i=1}^{N} (1-\lambda_i)$$

$$= 1 - \frac{1}{N} \sum_{i=1}^{N} \lambda_i$$

---

[14]These are necessary ingredients of Bennett's reversible computation [48] and of unitary quantum computation, if the result is to be read out.

$$\leqslant 1 - \prod_{i=1}^{N} \lambda_i^{\frac{1}{N}}$$

$$\leqslant -\frac{1}{N} \ln \left(\prod_{i=1}^{N} \lambda_i \right)$$

$$\Rightarrow \sqrt{\sum_{i=1}^{N} \frac{1}{(1-\lambda_i)^2}} \geqslant \frac{N\sqrt{N}}{-\ln \left(\prod_{i=1}^{N} \lambda_i \right)}. \qquad (A2)$$

Turning to the second square-root term in (A1), we employ the fact that for any collection of positive real a_i, $\sum_{i=1}^{N} a_i^2 \geqslant \frac{1}{N}(\sum_{i=1}^{N} a_i)^2$:

$$\sqrt{\sum_{i=1}^{N} \|\overline{\rho}_{i+1} - \overline{\rho}_i\|_1^4} \geqslant \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \|\overline{\rho}_{i+1} - \overline{\rho}_i\|_1^2$$

$$\geqslant \frac{1}{N\sqrt{N}} \left(\sum_{i=1}^{N} \|\overline{\rho}_{i+1} - \overline{\rho}_i\|_1 \right)^2$$

$$\geqslant \frac{1}{N\sqrt{N}} \|\overline{\rho}_{N+1} - \overline{\rho}_1\|_1^2, \qquad (A3)$$

where the last line expresses the triangle inequality. Substituting (A2) and (A3) into (A1),

$$\Sigma \geqslant \frac{\|\overline{\rho}_{N+1} - \overline{\rho}_1\|_1^2}{-2\ln \left(\prod_{i=1}^{N} \lambda_i \right)}. \qquad (A4)$$

Finally, we note that $\overline{\rho}_1 = \overline{\rho}$ and that $U\overline{\rho}_{N+1}U^\dagger = \overline{\eta}$, where U is given by Eq. (11), and that $\prod_{i=1}^{N} \lambda_i = \frac{\|\eta_1 - \eta_2\|_1}{\|\rho_1 - \rho_2\|_1}$ (10). This allows us to express (A4) in a way that depends only on the input and output states,

$$\Sigma \geqslant \frac{1}{2} \frac{\|U^\dagger \overline{\eta} U - \overline{\rho}\|_1^2}{\ln \left(\frac{\|\rho_1 - \rho_2\|_1}{\|\eta_1 - \eta_2\|_1} \right)}, \qquad (A5)$$

finally leading to an upper bound on extractable work:

$$\widetilde{W} \leqslant -\Delta \overline{F} - \frac{k_B T}{2} \frac{\|U^\dagger \overline{\eta} U - \overline{\rho}\|_1^2}{\ln \left(\frac{\|\rho_1 - \rho_2\|_1}{\|\eta_1 - \eta_2\|_1} \right)}. \qquad (A6)$$

[1] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[2] A. S. Holevo, Quantum coding theorems, Russ. Math. Surv. **53**, 1295 (1998).

[3] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).

[4] C. H. Bennett, Notes on Landauer's principle, reversible computation, and Maxwell's demon, Stud. Hist. Philos. Sci. Part B **34**, 501 (2003).

[5] W. H. Zurek, Maxwell's demon, Szilard's engine and quantum measurements, arXiv:quant-ph/0301076.

[6] K. Maruyama, F. Nori, and V. Vedral, *Colloquium:* The physics of Maxwell's demon and information, Rev. Mod. Phys. **81**, 1 (2009).

[7] I. H. Deutsch, Harnessing the power of the second quantum revolution, PRX Quantum **1**, 020101 (2020).

[8] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—A topical review, J. Phys. A **49**, 143001 (2016).

[9] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

[10] N. Linden, S. Popescu, and P. Skrzypczyk, How small can thermal machines be? The smallest possible refrigerator, Phys. Rev. Lett. **105**, 130401 (2010).

[11] P. Kammerlander and J. Anders, Coherence and measurement in quantum thermodynamics, Sci. Rep. **6**, 22174 (2016).

[12] G. Francica, F. C. Binder, G. Guarnieri, M. T. Mitchison, J. Goold, and F. Plastina, Quantum coherence and ergotropy, Phys. Rev. Lett. **125**, 180603 (2020).

[13] A. Smith, K. Sinha, and C. Jarzynski, Quantum coherences and classical inhomogeneities as equivalent thermodynamics resources, Entropy **24**, 474 (2022).

[14] P. A. Camati, J. F. G. Santos, and R. M. Serra, Coherence effects in the performance of the quantum Otto heat engine, Phys. Rev. A **99**, 062103 (2019).

[15] M. Lostaglio, D. Jennings, and T. Rudolph, Description of quantum coherence in thermodynamic processes requires constraints beyond free energy, Nat. Commun. **6**, 6383 (2015).

[16] K. Korzekwa, M. Lostaglio, J. Oppenheim, and D. Jennings, The extraction of work from quantum coherence, New J. Phys. **18**, 023045 (2016).

[17] M. O. Scully, M. S. Zubairy, G. S. Agarwal, and H. Walther, Extracting work from a single heat bath via vanishing quantum coherence, Science **299**, 862 (2003).

[18] L. d. Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, The thermodynamic meaning of negative entropy, Nature (London) **474**, 61 (2011).

[19] L. Bresque, P. A. Camati, S. Rogers, K. Murch, A. N. Jordan, and A. Auffèves, Two-qubit engine fueled by entanglement and local measurements, Phys. Rev. Lett. **126**, 120605 (2021).

[20] K. V. Hovhannisyan, M. Perarnau-Llobet, M. Huber, and A. Acín, Entanglement generation is not necessary for optimal work extraction, Phys. Rev. Lett. **111**, 240401 (2013).

[21] X. L. Huang, H. Xu, X. Y. Niu, and Y. D. Fu, A special entangled quantum heat engine based on the two-qubit Heisenberg XX model, Phys. Scr. **88**, 065008 (2013).

[22] A. Hewgill, A. Ferraro, and G. De Chiara, Quantum correlations and thermodynamic performances of two-qubit engines with local and common baths, Phys. Rev. A **98**, 042102 (2018).

[23] L. A. Williamson and M. J. Davis, Many-body enhancement in a spin-chain quantum heat engine, Phys. Rev. B **109**, 024310 (2024).

[24] M. Perarnau-Llobet, K. V. Hovhannisyan, M. Huber, P. Skrzypczyk, N. Brunner, and A. Acín, Extractable work from correlations, Phys. Rev. X **5**, 041011 (2015).

[25] A. Auffèves, Quantum technologies need a quantum energy initiative, PRX Quantum **3**, 020101 (2022).

[26] L. Gammaitoni, D. Chiuchiú, M. Madami, and G. Carlotti, Towards zero-power ICT, Nanotechnology **26**, 222001 (2015).

[27] Y.-Z. Zhen, D. Egloff, K. Modi, and O. Dahlsten, Universal bound on energy cost of bit reset in finite time, Phys. Rev. Lett. **127**, 190602 (2021).

[28] M. H. Mohammady, A. Auffèves, and J. Anders, Energetic footprints of irreversibility in the quantum regime, Commun. Phys. **3**, 89 (2020).

[29] P. Taranto, F. Bakhshinezhad, A. Bluhm, R. Silva, N. Friis, M. P. E. Lock, G. Vitagliano, F. C. Binder, T. Debarba, E. Schwarzhans, F. Clivaz, and M. Huber, Landauer versus Nernst: What is the true cost of cooling a quantum system? PRX Quantum **4**, 010332 (2023).

[30] K. Takara, H. Hasegawa, and D. Driebe, Generalization of the second law for a transition between nonequilibrium states, Phys. Lett. A **375**, 88 (2010).

[31] J. Anders and V. Giovannetti, Thermodynamics of discrete quantum processes, New J. Phys. **15**, 033022 (2013).

[32] D. J. Bedingham and O. J. E. Maroney, The thermodynamic cost of quantum operations, New J. Phys. **18**, 113050 (2016).

[33] G. Chiribella, F. Meng, R. Renner, and M.-H. Yung, The nonequilibrium cost of accurate information processing, Nat. Commun. **13**, 7155 (2022).

[34] R. Alicki, The quantum open system as a model of the heat engine, J. Phys. A **12**, L103 (1979).

[35] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, Nat. Commun. **5**, 4185 (2014).

[36] M. Esposito and C. V. den Broeck, Second law and Landauer principle far from equilibrium, Europhys. Lett. **95**, 40004 (2011).

[37] T. Sagawa, Thermodynamic and logical reversibilities revisited, J. Stat. Mech. (2014) P03025.

[38] E. A. Aguilar, H. Wojewódka-Ściążko, M. Stankiewicz, C. Perry, P. Ćwikliński, A. Grudka, K. Horodecki, and M. Horodecki, Thermal operations in general are not memoryless, arXiv:2009.03110.

[39] Á. M. Alhambra, M. Lostaglio, and C. Perry, Heat-bath algorithmic cooling with optimal thermalization strategies, Quantum **3**, 188 (2019).

[40] M. Lostaglio, Á. M. Alhambra, and C. Perry, Elementary thermal operations, Quantum **2**, 52 (2018).

[41] C. Perry, P. Ćwikliński, J. Anders, M. Horodecki, and J. Oppenheim, A sufficient set of experimentally implementable thermal operations for small systems, Phys. Rev. X **8**, 041049 (2018).

[42] M. Quadeer, K. Korzekwa, and M. Tomamichel, Work fluctuations due to partial thermalizations in two-level systems, Phys. Rev. E **103**, 042141 (2021).

[43] V. Scarani, M. Ziman, P. Štelmachovič, N. Gisin, and V. Bužek, Thermalizing quantum machines: Dissipation and entanglement, Phys. Rev. Lett. **88**, 097905 (2002).

[44] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Dissipation: The phase-space perspective, Phys. Rev. Lett. **98**, 080602 (2007).

[45] M. Campisi and R. Fazio, Dissipation, correlation and lags in heat engines, J. Phys. A **49**, 345002 (2016).

[46] A. Kolchinsky and D. H. Wolpert, Work, entropy production, and thermodynamics of information under protocol constraints, Phys. Rev. X **11**, 041024 (2021).

[47] M. Ohya and N. Watanabe, Quantum entropy and its applications to quantum communication and statistical physics, Entropy **12**, 1194(2010).

[48] C. H. Bennett, Logical reversibility of computation, IBM J. Res. Dev. **17**, 525 (1973).

[49] J. J. Park, K.-H. Kim, T. Sagawa, and S. W. Kim, Heat engine driven by purely quantum information, Phys. Rev. Lett. **111**, 230402 (2013).

[50] M. Violaris and C. Marletto, Irreversibility beyond Landauer's principle: Is it possible to perform erasure using the quantum homogenizer? New J. Phys. **24**, 113030 (2022).

[51] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, Inadequacy of von Neumann entropy for characterizing extractable work, New J. Phys. **13**, 053015 (2011).