Thermodynamically Ideal Quantum State Inputs to Any Device

Paul M. Riechers,[1,*] Chaitanya Gupta,[2] Artemy Kolchinsky,[3,4] and Mile Gu[5,6,7,†]
[1] *Beyond Institute for Theoretical Science, San Francisco, California 94110, USA*
[2] *H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*
[3] *ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain*
[4] *Universal Biology Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[5] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639798, Singapore*
[6] *CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore 117543, Singapore*
[7] *Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore*

(Received 5 August 2023; accepted 27 June 2024; published 29 July 2024)

We investigate and ascertain the ideal inputs to any finite-time physical process. We demonstrate that the expectation values of entropy flow, heat, and work can all be determined via Hermitian observables of the initial state. These Hermitian operators encapsulate the breadth of behavior and the ideal inputs for common thermodynamic objectives. We show how to construct these Hermitian operators from measurements of thermodynamic output from a finite number of effectively arbitrary inputs. The behavior of a small number of test inputs thus determines the full range of thermodynamic behavior from all inputs. For any process, entropy flow, heat, and work can all be extremized by pure input states—eigenstates of the respective operators. In contrast, the input states that minimize entropy production or maximize the change in free energy are nonpure mixed states obtained from the operators as the solution of a convex-optimization problem. To attain these, we provide an easily implementable gradient-descent method on the manifold of density matrices, where an analytic solution yields a valid direction of descent at each iterative step. Ideal inputs within a limited domain, and their associated thermodynamic operators, are obtained with less effort. This allows analysis of ideal thermodynamic inputs within quantum subspaces of infinite-dimensional quantum systems; it also allows analysis of ideal inputs in the classical limit. Our examples illustrate the diversity of "ideal" inputs: distinct initial states minimize entropy production, extremize the change in free energy, and maximize work extraction.

DOI: 10.1103/PRXQuantum.5.030318

I. INTRODUCTION

Throughout its history, thermodynamics has primarily investigated the efficiency of various control processes for implementing a desired functionality. However, the complementary question of *which initial physical states produce the best thermodynamic behavior* remains relatively unexplored. Indeed, there is a historical reason for this: in equilibrium transformations, the system always stays infinitesimally close to equilibrium, so there is no sense in asking about alternative inputs to the process. Yet modern devices transform quantum and classical systems rapidly. These finite-time nonequilibrium transformations have a highly nontrivial initial-state dependence. Here, we explore the ideal thermodynamic inputs to such devices, where the system can be arbitrarily far from equilibrium throughout the transformation.

The initial-state dependence of entropy production and associated thermodynamic quantities has been explored only recently in relation to the ideal inputs, via mismatch costs [1–5]. However, the minimally dissipative input has only been characterized in the case of reset processes [4,5] and, even then, a construction has only been given for qubits [4]. In the following, we constructively identify the thermodynamically ideal inputs for a much broader class of objectives, including heat minimization, maximizing work extraction, and maximizing gain in free energy. Moreover, the ideal inputs are characterized and constructively identified for systems of arbitrary finite dimensions, for any finite-time process.

---

[*] Contact author: pmriechers@gmail.com
[†] Contact author: mgu@quantumcomplexity.org

We find that common expectation values in thermodynamics—heat, work, entropy flow, entropy production, free-energy gain, etc.—all fall into two camps: (i) a linear functional of the initial state or (ii) a linear functional of the initial state plus a change in its von Neumann entropy. The former is extremized by an eigenstate of a Hermitian operator that can be inferred from observed behavior, while the latter is extremized by a nonpure mixed state obtained as the solution to a convex constrained-optimization problem (for which we provide a new algorithm).

With our suite of results in hand, it is finally possible to compare against the many "optimization principles" that have been proposed in thermodynamics over a century [6,7]. For example, we show how to find the initial state that extremizes entropy production. One can then check if this extremal state is the same as the steady state. In a trivial case, yes: the equilibrium state minimizes entropy production for a static protocol in the presence of a single bath. In other cases, no: in the case of multiple baths with different temperatures or chemical potentials, the nonequilibrium steady state (NESS) generally neither maximizes nor minimizes entropy production.

In addition to identifying ideal inputs, we also identify the full breadth of thermodynamic behavior that can be expected from different inputs to any quantum process. These results can usefully be applied to quantum subspaces or classical systems as well. The results thus apply broadly, from biology to electronics to cosmology, wherever finite-time thermodynamics is relevant.

II. OVERVIEW OF FRAMEWORK

It is often desirable for a physical device to implement a fixed transformation on arbitrary input. For example, in quantum technology, we often want to construct a device that implements a prescribed completely positive and trace preserving (CPTP) map. The device can achieve this via a time-dependent protocol that partially controls the Hamiltonian of the system and its interactions with its environment.

It is important to note that thermodynamics is not determined by the CPTP map alone. However, a device implies a fixed *implementation* of the CPTP map, which is determined by both the initial state of the environment and the joint unitary acting on the system-environment supersystem. The choice of device thus has thermodynamic implications. Here, we explore the thermodynamics, optimal inputs, and breadth of behavior from any fixed device for implementing a CPTP map.

More explicitly, we assume that the control protocol and the initial state of the environment are fixed, while we are free to select the initial (mixed or pure) state of the system to input to our device. This implies that the initial joint state of the system-environment supersystem is uncorrelated: $\rho_0^{\mathrm{tot}} = \rho_0 \otimes \rho_0^{\mathrm{env}}$. Nevertheless, correlations can be established during the protocol due to system-environment coupling and the system may have non-Markovian evolution. In particular, the joint system-baths supersystem evolves unitarily via $U_{0:t}$, such that the joint state at any later time t is given as $\rho_t^{\mathrm{tot}} = U_{0:t} \rho_0^{\mathrm{tot}} U_{0:t}^\dagger$. The reduced states of the system and environment are given at any time by the appropriate partial trace of the joint state: $\rho_t = \mathrm{tr}_{\mathrm{env}}(\rho_t^{\mathrm{tot}})$ and $\rho_t^{\mathrm{env}} = \mathrm{tr}_{\mathrm{sys}}(\rho_t^{\mathrm{tot}})$. Any CPTP transformation of the system can be achieved this way, since such an interaction with the environment instantiates Stinespring dilation [8].

In the following sections, we will discover the thermodynamically ideal inputs to any such implementation of any CPTP transformation. To achieve this, we first introduce *thermodynamic operators* in Sec. III, *generalized Bloch vectors* in Sec. IV, and *thermodynamic vectors* in Sec. V. We then show how to construct these thermodynamic vectors and operators from experimental observations in Sec. VI. This finally allows us to discuss and construct the thermodynamically ideal inputs in Secs. VII and VIII. Section IX shows how these results apply to restricted subspaces—which is important for understanding the classical limit and for applying the results to low-energy subspaces of infinite-dimensional systems. Finally, the examples in Secs. X–XII illustrate physical implications of our results.

III. THERMODYNAMIC OPERATORS

Thermodynamic quantities such as work, heat, and entropy flow are notoriously path-dependent quantities. Even their average values depend on the time-dependent control protocol and the time-dependent density matrices of system and environment. In fact, there are many competing definitions for work, heat, and entropy flow that may be more or less relevant in various scenarios [9]—our general results will apply to them all. The point for now is that these are all path-dependent quantities. For example, the expectation value of entropy flow can very generally be calculated as [4]

$$\langle \Phi \rangle_{\rho_0} = -k_B \int_0^\tau \mathrm{tr}\left(\dot{\rho}_t^{\mathrm{env}} \ln \boldsymbol{\pi}_t^{\mathrm{env}}\right) dt, \tag{1}$$

where $\boldsymbol{\pi}_t^{\mathrm{env}}$ is a tensor product of local-equilibrium reference states for the environment at time t and k_B is Boltzmann's constant. Familiar expressions such as $\langle \Phi \rangle_{\rho_0} = \int (\langle \delta Q^{(b)} \rangle / T_t^{(b)})\, dt$ and $\langle \Phi \rangle_{\rho_0} = \langle Q^{(b)} \rangle / T^{(b)} + \langle Q^{(b')} \rangle / T^{(b')}$ are special cases of Eq. (1), where $Q^{(b)}$ is the change in energy of bath b and $T_t^{(b)}$ is the temperature of bath b at time t [10,11]. Notably, entropy flow to the environment, plus the change in entropy of the system, yields entropy production—the notorious quantity of the famous second law of thermodynamics—from which we can obtain Landauer's bound, Carnot's bound, and much more [3,12–14]. With H_t as the time-dependent Hamiltonian of the system, similar integral expressions for work—typically defined in the weak-coupling limit as

$$\langle W\rangle_{\rho_0} = \int_0^\tau \mathrm{tr}\left(\rho_t \dot{H}_t\right)\, dt \tag{2}$$

—and heat absorbed by the system

$$\langle Q\rangle_{\rho_0} = \int_0^\tau \mathrm{tr}\,(\dot{\rho}_t H_t)\, dt \tag{3}$$

are familiar centerpieces of the thermodynamic arsenal. In this path-dependent spirit, it has been famously emphasized that "work is not an observable" [15].

It is therefore perhaps surprising that the expectation values of work, heat, and entropy flow can all be determined via Hermitian observables of the initial state. As we show in Appendix A, since the above expectation values are all linear functionals of the initial density matrix of the system, they can be expressed as

$$\langle \Phi\rangle_{\rho_0} = \mathrm{tr}(\rho_0 \boldsymbol{\Phi}), \quad \langle W\rangle_{\rho_0} = \mathrm{tr}(\rho_0 \mathcal{W}),$$
$$\text{and} \quad \langle Q\rangle_{\rho_0} = \mathrm{tr}(\rho_0 \mathcal{Q}), \tag{4}$$

where $\boldsymbol{\Phi}$, $\mathcal{W}$, and $\mathcal{Q}$ are Hermitian operators, which we refer to as the *expected-entropy-flow operator*, the *expected-work operator*, and the *expected-heat operator*, respectively.

Importantly, we do not assume that these *thermodynamic operators* are known *a priori*. Instead, we will show how they can be constructed from experimental observations using a finite number of arbitrary inputs, given a general finite-time physical process. These thermodynamic operators in turn reveal the special collection of initial states that minimize entropy flow, minimize heat, and maximize work extraction. The operators also allow direct calculation of thermodynamic expectation values from any initial state.

The expressions in Eq. (4) all clearly take the form

$$\langle X\rangle_{\rho_0} = \mathrm{tr}(\rho_0 \mathcal{X}), \tag{5}$$

which we will study in general. We will say that a thermodynamic quantity is "type I" when its expectation value can be expressed as a linear functional of the initial state $\mathrm{tr}(\rho_0 \mathcal{X})$, as in Eq. (5). As an added benefit of studying this general formulation, our methods can also be applied to reconstruct the Hamiltonian, the expected-change-of-energy operator, and infinitely many other linear operators that conform to this general linear structure for expectation values. If the expectation value is always real valued, then $\mathcal{X}$ is guaranteed to be Hermitian (see Ref. [16, Theorem 2.4.3]).

Notably, Eq. (5) accommodates any of the prominent definitions for work and heat, including those discussed in Refs. [17–20], some of which are relevant for arbitrarily strong coupling between system and baths, and with nonequilibrium environments. For any candidate definition of work and heat, one only needs to verify that its expectation value is a linear functional of the initial state to confirm that it is indeed a type-I quantity amenable to our analysis.

To give a concrete example, an alternative definition of work in the strong-coupling regime is $W^{(*)}$: the net change in energy of the system-bath supersystem during a protocol of duration τ [12]. Let H_t^{tot} be the total Hamiltonian including the system, environment, and coupling contributions; over time, it induces the net unitary evolution operator $U_{0:\tau}$. The expectation value for this version of work is thus

$$\langle W^{(*)}\rangle_{\rho_0} = \mathrm{tr}(\rho_\tau^{\mathrm{tot}} H_\tau^{\mathrm{tot}}) - \mathrm{tr}(\rho_0^{\mathrm{tot}} H_0^{\mathrm{tot}}), \tag{6}$$

which, as we show in Appendix A, can be written as a linear functional of the initial state of the system alone when the initial state of the environment is fixed. By our Theorem 6 of Appendix A, this thus implies that there exists an input-independent thermodynamic operator $\mathcal{W}^{(*)}$ such that

$$\langle W^{(*)}\rangle_{\rho_0} = \mathrm{tr}(\rho_0 \mathcal{W}^{(*)}). \tag{7}$$

It is known that invasive measurements can change the expected value of work and other thermodynamic variables [21]. We emphasize that different measurement schemes must be treated as distinct quantum processes, since each implies a unique sequence of dynamic interventions. Accordingly, each measurement scheme induces its own set of thermodynamic operators. For example, in Appendix B, we address the two-point-measurements (TPM) scheme and give an explicit construction of the expected-TPM-work operator $\mathcal{W}_{\mathrm{TPM}}$, for which

$$\langle W_{\mathrm{TPM}}\rangle_{\rho_0} = \mathrm{tr}(\rho_0 \mathcal{W}_{\mathrm{TPM}}). \tag{8}$$

Indeed, for any measurement scheme—TPM [22], one-point measurement [23–25], or any other scheme—thermodynamic operators can be constructed and our framework can be applied to identify both the breadth of behavior and the ideal inputs within the scheme.

It is important to note that each of these thermodynamic operators $\mathcal{X}$ contains all the information needed for expectation values of the relevant thermodynamic quantity X but does not contain the information that would be needed for expectation values of functions of the thermodynamic quantity, such as X^2 or e^X. As a point of nomenclature, we note that this is not about whether the operators are "observable," as has been suggested [15], but is, rather, an elementary fact about expectation values—$\langle X \rangle$ does not determine $\langle f(X) \rangle$. The source of insufficiency is simple to see if we take work as an example: the full work distribution of a process, whether classical or quantum, has support on a space much larger than the dimension d of the system. Accordingly, the d eigenvalues of the thermodynamic operator cannot represent the full probability distribution of the thermodynamic quantity.

Instantaneous thermodynamic quantities—such as energy, position, momentum, angular momentum, or spin—offer a familiar exception, since they describe variables with the same (or smaller) dimension as the system. Accordingly, their $d \times d$ operators can be used to calculate all moments of their representative quantity. Similarly, sufficiently high-dimensional representations of other thermodynamic operators would allow their encapsulation of higher moments [26]. In future studies, it would be worth exploring how quasiprobabilities relate to these higher-dimensional work operators, since each enables an investigation of fluctuations [27,28]. Indeed, quasiprobabilities are linear functionals of the initial state [28] and so the real and imaginary parts of each quasiprobability are thus type-I expectation values amenable to our framework.

To address a final nuance, we note that a random variable X in the quantum domain may depend on further specification of subensembles, since a density matrix can be decomposed in many ways [29]. Nevertheless, as discussed in Appendix C, all decompositions of the initial density matrix lead to the same expectation value for a fixed quantum process. Accordingly, our notation "$\langle X \rangle_{\rho_0}$" unambiguously refers to the expectation value of the relevant thermodynamic quantity X for all possible decompositions of the quantum state.

Despite these caveats, the thermodynamic operators serve immense utility. Their d eigenstates represent the only d features that influence the expected value of the thermodynamic variable. Once inferred for a process, the thermodynamic operators tell the full breadth of expected behavior from any input and identify the unique pure states leading to extremal behavior. Moreover, the performance of any of the infinitely many possible inputs can be calculated simply and directly from the thermodynamic operator, without needing to run a new experiment each time.

IV. GENERALIZED BLOCH VECTOR

We have promised that thermodynamic operators are the key to identifying thermodynamically ideal inputs to any process. It will be important, then, to be able to construct these operators. Our construction will lean on generalized Bloch vectors and related thermodynamic vectors, introduced in this section and in Sec. V, respectively.

It is well known that the state of a qubit ρ_t can be expressed via its Bloch vector $\vec{a}_t$:

$$\rho_t = I/2 + \vec{a}_t \cdot \vec{\sigma}/2, \tag{9}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices. For a quantum system of arbitrary finite dimension—i.e., a qudit ρ_t acting on a d-dimensional vector space $\mathcal{V}_d$—we achieve something similar via a slight adaptation of Ref. [30]. We choose any complete basis $(I/d, \Gamma_1, \Gamma_2, \ldots, \Gamma_{d^2-1})$ for linear operators acting on $\mathcal{V}_d$, such that the Hermitian operators Γ_n are all traceless and mutually orthogonal, satisfying

$$\mathrm{tr}(\Gamma_n) = 0 \quad \text{and} \tag{10a}$$

$$\mathrm{tr}(\Gamma_m \Gamma_n) = \eta\, \delta_{m,n}, \tag{10b}$$

where we will choose the normalizing constant to be $\eta = (d-1)/d$. Any density matrix then has a unique decomposition in the operator basis $\vec{\Gamma} = (\Gamma_1, \Gamma_2, \ldots, \Gamma_{d^2-1})$, described by the *generalized Bloch vector* $\vec{b}_t \in \mathbb{R}^{d^2-1}$ via

$$\rho_t = I/d + \vec{b}_t \cdot \vec{\Gamma}. \tag{11}$$

Since the magnitude of the Bloch vector is $b_t = \sqrt{[\mathrm{tr}(\rho_t^2)d - 1]/(d-1)}$, the density matrix represents a pure state if and only if the magnitude of its corresponding Bloch vector is one. For $d > 2$, not all points in the Bloch ball correspond to physical states, but the set of all physical states is nevertheless a convex set—the convex hull of the pure states, which all lie on a $2(d-1)$-dimensional submanifold of the (d^2-2)-dimensional surface of the Bloch sphere [30].

For concreteness, we can choose the ordered operator basis to be a scaled ordering of generalized Gell-Mann matrices—the generators of SU(d) (for details, see Appendix E). The standard Bloch vector is then recovered in the familiar two-dimensional case of a qubit, where then $\vec{\Gamma} = \vec{\sigma}/2 = (\sigma_x/2, \sigma_y/2, \sigma_z/2)$ and $\eta = 1/2$. Alternatively, Appendix F shows how to construct valid composite operator bases.

V. THERMODYNAMIC VECTORS

Leveraging the general Bloch decomposition in Eq. (11) of the initial state, we find that we can express each expectation value in Eq. (4) as

$$\langle X \rangle_{\rho_0} = \langle X \rangle_{I/d} + \vec{b}_0 \cdot \vec{x}, \tag{12}$$

where $\vec{x} \in \mathbb{R}^{d^2-1}$ is the relevant *thermodynamic vector*

$$\vec{x} = \mathrm{tr}(\vec{\Gamma}\mathcal{X}). \tag{13}$$

In particular, the thermodynamic vector could be the *entropy-flow vector* $\vec{\varphi} = \mathrm{tr}(\vec{\Gamma}\boldsymbol{\Phi})$, the *work vector* $\vec{w} = \mathrm{tr}(\vec{\Gamma}\mathcal{W})$, or the *heat vector* $\vec{q} = \mathrm{tr}(\vec{\Gamma}\mathcal{Q})$.

Conversely, the thermodynamic operators can be constructed from the thermodynamic vectors:

$$\mathcal{X} = \langle X \rangle_{I/d} I + \vec{x} \cdot \vec{\Gamma}/\eta. \tag{14}$$

Using Eqs. (10a) and (10b), it is easy to verify that Eq. (14) satisfies $\mathrm{tr}(\rho_0 \mathcal{X}) = \langle X \rangle_{I/d} + \vec{b}_0 \cdot \vec{x}$.

In the next sections, we show how both $\langle X \rangle_{I/d}$ and the thermodynamic vector $\vec{x}$ can be obtained via linear algebra from experimental measurements of thermodynamic output from a finite number of almost arbitrary inputs [31]. Via Eq. (14), this allows us to experimentally reconstruct the thermodynamic operators from observations of any process.

VI. CONSTRUCTING THERMODYNAMIC VECTORS AND OPERATORS FROM OBSERVATIONS

Suppose that an experimentalist has an apparatus to transform the state of a finite-dimensional quantum system. (We will address infinite-dimensional systems in Sec. IX.) This experimentalist measures the expectation value of the thermodynamic random variable X that results from each of d^2 linearly independent inputs to their device; i.e., they record the average quantity from each of the initial states $(\rho_0^{(n)})_{n=1}^{d^2}$ with corresponding generalized Bloch vectors $(\vec{b}_0^{(n)})_{n=1}^{d^2}$. Note that the generalized Bloch vectors can be obtained as $\vec{b}_0^{(n)} = \mathrm{tr}(\rho_0^{(n)}\vec{\Gamma})/\eta$. From Eq. (12), we see that

$$\underbrace{\begin{bmatrix} 1 & \vec{b}_0^{(1)} \\ 1 & \vec{b}_0^{(2)} \\ \vdots & \vdots \\ 1 & \vec{b}_0^{(d^2)} \end{bmatrix}}_{=:B} \begin{bmatrix} \langle X \rangle_{I/d} \\ \vec{x} \end{bmatrix} = \begin{bmatrix} \langle X \rangle_{\rho_0^{(1)}} \\ \langle X \rangle_{\rho_0^{(2)}} \\ \vdots \\ \langle X \rangle_{\rho_0^{(d^2)}} \end{bmatrix}. \tag{15}$$

Note that the B matrix defined in Eq. (15) is invertible, since the d^2 initial states are all linearly independent. Hence, with the d^2 measurements in hand and some simple linear algebra, we can find both (i) the expectation value from the reference input $\langle X \rangle_{I/d}$ and (ii) the input-independent thermodynamic vector $\vec{x}$:

$$\begin{bmatrix} \langle X \rangle_{I/d} \\ \vec{x} \end{bmatrix} = B^{-1} \begin{bmatrix} \langle X \rangle_{\rho_0^{(1)}} \\ \langle X \rangle_{\rho_0^{(2)}} \\ \vdots \\ \langle X \rangle_{\rho_0^{(d^2)}} \end{bmatrix}. \tag{16}$$

$\langle X \rangle_{I/d}$ and $\vec{x}$ can now be used in Eq. (14) to construct the thermodynamic operator of interest: $\mathcal{X} = \langle X \rangle_{I/d} I + \vec{x} \cdot \vec{\Gamma}/\eta$. This constitutes a type of "operator tomography," which is distinct but reminiscent of both quantum state tomography and process tomography.

It is worth briefly noting that our framework also offers an alternative approach to standard quantum process tomography [32], since we can directly reconstruct the generalized Bloch representation of the process using the above Bloch matrix B. As always, we can identify any quantum channel through the preparation and evolution of any d^2 linearly independent inputs to the channel. Let $\mathcal{T}$ denote the unknown d^2-by-d^2 matrix that maps (input) generalized Bloch vectors to (output) generalized Bloch vectors via

$$\begin{bmatrix} 1 & \vec{b}_\tau \end{bmatrix} = \begin{bmatrix} 1 & \vec{b}_0 \end{bmatrix} \mathcal{T}. \tag{17}$$

Stacking the solutions for d^2 linearly independent inputs, with output vectors $\vec{b}_\tau^{(n)} = \mathrm{tr}(\rho_\tau^{(n)}\vec{\Gamma})/\eta$, we can invert B to construct this linear Bloch representation of the channel:

$$\mathcal{T} = B^{-1}B', \tag{18}$$

where the nth row of the matrix B' is $\begin{bmatrix} 1 & \vec{b}_\tau^{(n)} \end{bmatrix}$. Spectral analysis of $\mathcal{T}$ yields modes of decay, unitary subspaces, and stationary states of the map, if they exist.

We are now equipped to identify and construct the thermodynamically ideal inputs to any device.

VII. IDEAL INPUTS FOR TYPE-I OBJECTIVES

Recall that a thermodynamic quantity is "type I" when its expectation value can be expressed as a linear functional of the initial state $\mathrm{tr}(\rho_0 \mathcal{X})$. It is often desirable to minimize or maximize type-I quantities, e.g., to minimize heat or maximize work extraction.

An immediate observation can be drawn from Eq. (5) for any finite-dimensional quantum system, by considering the spectral decomposition of the bounded operator $\mathcal{X} = \sum_{n=1}^{d} \lambda_n |v_n\rangle\langle v_n|$, where $\Lambda_{\mathcal{X}} = (\lambda_n)_n$ is the tuple of $\mathcal{X}$'s eigenvalues with corresponding eigenstates $V_{\mathcal{X}} = (|v_n\rangle)_n$.

Theorem 1. There is always a pure-state input that extremizes the expectation value of type-I quantities $\langle X \rangle_{\rho_0} = \mathrm{tr}(\rho_0 \mathcal{X})$. This pure state corresponds to an eigenstate of the thermodynamic operator $\mathcal{X}$ with extremal eigenvalue.

For example: there is always a pure-state input that minimizes heat; there is always a pure-state input that maximizes heat; there is always a pure-state input that minimizes entropy flow; and so on.

The implications are important and somewhat surprising. For instance, consider the work extracted $W_{\text{extracted}} = -W$ during a finite-time cyclic work-extraction protocol from a finite-dimensional quantum system. Theorem 1 asserts that maximal work can be extracted from the pure-state input $|w_{\min}\rangle \in \operatorname{argmin}_{|v\rangle \in V_{\mathcal{W}}} \{\langle v|\mathcal{W}|v\rangle\}$ [33]. When the minimal-eigenvalue eigenspace of $\mathcal{W}$ is degenerate, maximal work extraction can be achieved by both pure and mixed inputs. In contrast, as we will see later, the state that minimizes entropy production is generically a mixed state. Therefore, *pure-state inputs that maximize work extraction imply nonminimal entropy production.* Similarly, *pure-state inputs that minimize heat imply nonminimal entropy production.*

One broad lesson from this analysis is that maximizing work extraction, minimizing heat, minimizing entropy production, etc., are all distinct concepts that should not be conflated. There is a great diversity in optimality among different thermodynamic goals. This will be emphasized again in our first example in Sec. X.

It is worth noting that the smallest and largest eigenvalues of each thermodynamic operator demarcate the range of corresponding expectation values that can be achieved via alternative inputs:

$$\langle X\rangle_{\rho_0} \in [\min(\Lambda_{\mathcal{X}}), \max(\Lambda_{\mathcal{X}})]. \tag{19}$$

All values in this continuous range are achievable by some input.

VIII. IDEAL INPUTS FOR TYPE-II OBJECTIVES

Not all thermodynamic expectation values can be expressed like Eq. (5), as a linear functional of the initial state. In particular, entropy production, reduction in nonequilibrium free energy, and change in entropy all have expectation values that are nonlinear functions of the initial state. However, in each of these three cases, the nonlinearity is of the same form, since it derives from the change of von Neumann entropy $S(\rho) = -\mathrm{tr}(\rho \ln \rho)$. Accordingly, the initial states that extremize this second class of thermodynamic quantities all share similar features.

The expectation value of entropy production, $\langle \Sigma \rangle_{\rho_0} = \langle \Phi \rangle_{\rho_0} + k_B \Delta S(\rho_t)$, plays a central role in nonequilibrium thermodynamics. When the environment begins in local equilibrium and is uncorrelated with both itself and the system, then the famous second law of thermodynamics is valid: $\langle \Sigma \rangle_{\rho_0} \geq 0$. In the appropriate circumstances, entropy production can alternatively be expressed as the work performed beyond the change in nonequilibrium free energy $T\langle \Sigma \rangle_{\rho_0} = \langle W \rangle_{\rho_0} - \Delta \mathcal{F}_t$, where T is the initial temperature of the environment.

The expectation values of (i) entropy production, (ii) change in free energy, and (iii) change in entropy are each proportional to

$$f_{\rho_0}^{(\mathcal{X})} := \mathrm{tr}(\rho_0 \mathcal{X}) + S(\rho_\tau) - S(\rho_0) \tag{20}$$

for the appropriate linear operators $\mathcal{X}$. We will say that a thermodynamic quantity is "type II" when its expectation value is proportional to $f_{\rho_0}^{(\mathcal{X})}$ for some linear operator $\mathcal{X}$ and some positive constant of proportionality. In particular,

$$\langle \Sigma \rangle_{\rho_0} = k_B f_{\rho_0}^{(\Phi/k_B)} \quad \text{and} \quad \Delta S_t = f_{\rho_0}^{(0)} \quad \text{and}$$
$$-\Delta \mathcal{F}_t = k_B T f_{\rho_0}^{(-(\mathcal{Q}+\mathcal{W})/k_B T)}. \tag{21}$$

A. General processes

1. Inputs that minimize type-II functions

Theorem 2. An initial state that locally minimizes a type-II quantity also globally minimizes it.

This is because type-II quantities are convex in the initial state.

For general processes, the nonlinearity of type-II objectives makes it difficult to find a closed-form expression for ideal inputs $\operatorname{argmin}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$. Nevertheless, since type-II quantities are convex in the initial state, any number of simple algorithms, including gradient descent and related variations, are guaranteed to converge to the ideal input upon iteration. By convexity, a local minimum in $f_{\rho_0}^{(\mathcal{X})}$, when minimizing over the set of density matrices, is also the global minimum.

However, constrained optimization—restricting to the set of density matrices in this case—is nontrivial. Although some techniques for gradient descent on the manifold of density matrices have been developed (see, e.g., Refs. [34,35] and references therein), we have found a more direct solution to our problem, which we provide in this section. Our resulting algorithm for gradient descent provides a quantum generalization of the Frank-Wolfe algorithm [36], with a simple analytically solvable direction for descent at each step.

Using techniques introduced in Ref. [3], we can analytically calculate the gradient of type-II expectation values around any initial state, given any parametrization of the state space. If we consider arbitrary infinitesimal changes in the initial generalized Bloch vector, then the partial derivative of $f_{\rho_0}^{(\mathcal{X})}$, with respect to each Bloch-vector component, can be expressed as

$$\frac{\partial}{\partial b_{0_m}} f_{\rho_0}^{(\mathcal{X})} = \mathrm{tr}(\Gamma_m \mathcal{X}) + \mathrm{tr}(\Gamma_m \ln \rho_0)$$
$$- \mathrm{tr}\left\{\mathrm{tr}_{\text{env}}\left[U(\Gamma_m \otimes \rho_0^{\text{env}})U^\dagger\right] \ln \rho_\tau\right\}. \tag{22}$$

This follows via an adaptation of the derivation that led to Eqs. (C16) and (O4) in Ref. [3].

It is useful to note how the elements of an arbitrary matrix basis can be expressed as a linear combination of the d^2 test inputs:

$$\Gamma_m = \sum_{n=1}^{d^2} (B^{-1})_{m+1,n}\, \rho_0^{(n)}, \tag{23}$$

where we have included the reference state as $\Gamma_0 = I/d$. We can likewise express the time evolution of these matrices, as the same linear combination of the time-evolved test inputs:

$$\Gamma'_m := \mathrm{tr}_{\mathrm{env}}\left[U(\Gamma_m \otimes \rho_0^{\mathrm{env}})U^\dagger\right] = \sum_{n=1}^{d^2} (B^{-1})_{m+1,n}\, \rho_\tau^{(n)}. \tag{24}$$

Partial derivatives of $f_{\rho_0}^{(\mathcal{X})}$ can thus be calculated via

$$\frac{\partial}{\partial b_{0m}} f_{\rho_0}^{(\mathcal{X})} = \mathrm{tr}(\Gamma_m \mathcal{X}) + \mathrm{tr}(\Gamma_m \ln \rho_0) - \mathrm{tr}(\Gamma'_m \ln \rho_\tau), \tag{25}$$

where $\rho_\tau = \Gamma'_0 + 1/\eta \sum_{n=1}^{d^2-1} \mathrm{tr}(\rho_0 \Gamma_n)\, \Gamma'_n$.

A gradient can be constructed as the Hermitian operator,

$$\vec{\nabla} f_{\rho_0}^{(\mathcal{X})} = \sum_{m=1}^{d^2-1} \Gamma_m \frac{\partial}{\partial b_{0m}} f_{\rho_0}^{(\mathcal{X})}. \tag{26}$$

However, extra care must be taken to stay along the manifold of valid density matrices. To achieve this, one can use the directional derivative

$$\frac{\rho'_0 - \rho_0}{\|\rho'_0 - \rho_0\|} \cdot \vec{\nabla} f_{\rho_0}^{(\mathcal{X})} = \mathrm{tr}\left[(\rho'_0 - \rho_0)\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\right] / \|\rho'_0 - \rho_0\|, \tag{27}$$

which is the linear change in $f_{\rho_0}^{(\mathcal{X})}$ at ρ_0 when moving in the direction of ρ'_0. Since density matrices are a convex set, a change in this direction is guaranteed to move along the manifold of density matrices. For simplicity, $\|\cdot\|$ can be chosen to be the trace norm.

Recall that since $f_{\rho_0}^{(\mathcal{X})}$ is convex over initial states, descent always benefits its global minimization; there are no nonglobal local minima in which to get stuck. If ρ_0 does not minimize $f_{\rho_0}^{(\mathcal{X})}$, then we can find directions (along the manifold of density matrices) with negative slope and these directions will lead toward the minimizer. Note that moving infinitesimally from ρ_0 toward ρ'_0 will reduce $f_{\rho_0}^{(\mathcal{X})}$ whenever $\mathrm{tr}\left[(\rho'_0 - \rho_0)\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\right] < 0$. Accordingly, we can always find a valid direction of descent by identifying the ρ'_0 that minimizes $\mathrm{tr}\left[(\rho'_0 - \rho_0)\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\right]$. Fortunately, this desired ρ'_0—call it $\sigma_{\min}$—can be found explicitly and analytically, since

$$\sigma_{\min} := \mathrm{argmin}_{\rho'_0}\, \mathrm{tr}\left[(\rho'_0 - \rho_0)\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\right] \tag{28}$$

$$= \mathrm{argmin}_{\rho'_0}\, \mathrm{tr}\left(\rho'_0 \vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\right) \tag{29}$$

$$= \frac{|\xi\rangle\langle\xi|}{\langle\xi|\xi\rangle}, \tag{30}$$

where

$$|\xi\rangle = \mathrm{argmin}_{|\lambda\rangle}\, \langle\lambda|\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}|\lambda\rangle \tag{31}$$

is the minimal-eigenvalue eigenstate of the Hermitian operator $\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}$.

Putting this all together, we can now propose a simple descent method for finding the optimal quantum state:

Algorithm to obtain $\mathrm{argmin}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$

From d^2 linearly independent test inputs, record the initial Bloch matrix B, construct the thermodynamic operator $\mathcal{X}$, and record the time-evolved test states $(\rho_\tau^{(n)})_{n=1}^{d^2}$ to obtain $(\Gamma'_n)_{n=0}^{d^2-1}$.
Choose an arbitrary initial density matrix ρ_0, which will be updated iteratively.
At each iterative step k:

1. Calculate the gradient $\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}$ via Eq. (25).
2. Determine the descent direction $\hat{n} = (\sigma_{\min} - \rho_0)/\|\sigma_{\min} - \rho_0\|$, with $\sigma_{\min} = |\xi\rangle\langle\xi|/\langle\xi|\xi\rangle$, from the minimal-eigenvalue eigenstate $|\xi\rangle$ of the Hermitian operator $\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}$.
3. Update the initial state ρ_0 to approach $\mathrm{argmin}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$, according to:

$$\rho_0 \mapsto \rho_0 - a_k \hat{n}\, \mathrm{tr}\big(\hat{n}\, \vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\big), \tag{32}$$

where a_k is a small positive value that diminishes with large k.

Note that $\hat{n}$ is a traceless operator. We take $a_k = 2/(k+2)$, as suggested by the Frank-Wolfe algorithm [36]. In the limit of many iterations, the algorithm then converges toward its unique fixed point $\alpha_0 = \mathrm{argmin}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$ with error on the order of $\mathcal{O}(1/k)$.

The minimizing input is generically a mixed state with full support. In these cases, success in the type-II minimization can be verified through the mismatch theorem:

$$f_{\rho_0}^{(\mathcal{X})} - f_{\alpha_0}^{(\mathcal{X})} = D[\rho_0 \| \alpha_0] - D[\rho_\tau \| \alpha_\tau], \tag{33}$$

where $D[\rho\|\alpha] = \mathrm{tr}(\rho \ln \rho) - \mathrm{tr}(\rho \ln \alpha)$ is the quantum relative entropy, and $\alpha_0 = \mathrm{argmin}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$ is the ideal input [3,5].

2. Inputs that maximize type-II functions

Theorem 3. A pure state input maximizes a type-II quantity.

This is because type-II quantities are convex functions of the initial state and the maximum of a convex function over a convex set (set of mixed states) is achieved on the boundary of the convex set (the pure states).

Local maximization can be achieved by a slight adaptation of the above algorithm, by ascending rather than descending the gradient. The only difference is that the ascent direction is obtained via the *maximal*-eigenvalue eigenstate of the Hermitian operator $\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}$, whereas gradient descent uses its minimal-eigenvalue eigenstate.

However, local maxima of type-II quantities are not necessarily global maxima. In our simulations, we have found the global maxima by seeding our algorithm with $\mathrm{argmax}_{\rho_0} \mathrm{tr}(\rho_0 \mathcal{X})$, which has been found via spectral decomposition of the thermodynamic operator $\mathcal{X}$.

For completeness and clarity, we provide this ascent algorithm explicitly:

Algorithm to obtain $\mathrm{argmax}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$

From d^2 linearly independent test inputs, record the initial Bloch matrix B, construct the thermodynamic operator $\mathcal{X}$, and record the time-evolved test states $(\rho_\tau^{(n)})_{n=1}^{d^2}$ to obtain $(\Gamma'_n)_{n=0}^{d^2-1}$.

Choose the initial density matrix to be $\rho_0 = \mathrm{argmax}_{\rho_0} \mathrm{tr}(\rho_0 \mathcal{X})$, which will be updated iteratively.

At each iterative step k:

(1) Calculate the gradient $\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}$ via Eq. (25).

(2) Determine the ascent direction $\hat{n} = (\sigma_{\max} - \rho_0)/\|\sigma_{\max} - \rho_0\|$, with $\sigma_{\max} = (|\psi\rangle\langle\psi|/\langle\psi|\psi\rangle)$, from the maximal-eigenvalue eigenstate $|\psi\rangle$ of the Hermitian operator $\vec{\nabla} f_{\rho_0}^{(\mathcal{X})}$.

(3) Update the initial state ρ_0 to approach $\mathrm{argmax}_{\rho_0} f_{\rho_0}^{(\mathcal{X})}$, according to:

$$\rho_0 \mapsto \rho_0 + a_k \hat{n}\, \mathrm{tr}\big(\hat{n}\, \vec{\nabla} f_{\rho_0}^{(\mathcal{X})}\big), \tag{34}$$

where a_k is a small positive value that diminishes with large k.

In the limit of many iterations, the algorithm converges toward a local maximum of $f_{\rho_0}^{(\mathcal{X})}$.

B. Overwriting processes

Processes that overwrite the physical state of the system are an important class of processes, including memory reset, state preparation, work extraction, and processes that lead to either equilibrium or nonequilibrium steady states. We find that their thermodynamically ideal inputs can all be found directly and analytically.

For reliable overwriting processes, for which the final state r_τ is very nearly independent of the input, $S(\rho_\tau) = S(r_\tau)$ will effectively be a constant. In this case, as shown in Appendix D, $f_{\rho_0}^{(\mathcal{X})}$ can be expressed as $f_{\rho_0}^{(\mathcal{X})} = D[\rho_0 \| \omega^{(\mathcal{X})}] - \ln[\mathrm{tr}(e^{-\mathcal{X}})] + S(r_\tau)$ where $\omega^{(\mathcal{X})} := e^{-\mathcal{X}}/\mathrm{tr}(e^{-\mathcal{X}})$. Thus, for reliable overwriting processes, it is clear that $\omega^{(\mathcal{X})}$ uniquely minimizes $f_{\rho_0}^{(\mathcal{X})}$. It is worth noting that if $\mathcal{X}$ is a bounded finite-dimensional operator, $\omega^{(\mathcal{X})}$ has full rank. Physically, this tells us that $\omega^{(\mathcal{X})}$ is a nonpure mixed state.

Theorem 4. For any overwriting process, the expectation value of any type-II quantity $\langle X\rangle_{\rho_0} = p f^{(\mathcal{X})}_{\rho_0} = p[\mathrm{tr}(\rho_0\mathcal{X}) + S(\rho_\tau) - S(\rho_0)]$ is uniquely minimized by the mixed-state input

$$\omega^{(\mathcal{X})} = e^{-\mathcal{X}}/\mathrm{tr}(e^{-\mathcal{X}}). \tag{35}$$

The corresponding minimal value is $\langle X\rangle_{\omega^{(\mathcal{X})}} = pS(r_\tau) - p\ln[\mathrm{tr}(e^{-\mathcal{X}})]$, where r_τ is the input-independent final state of the overwriting process.

Recall that the thermodynamic operator can be expressed as $\mathcal{X} = \langle X\rangle_{I/d} I + \vec{x}\cdot\vec{\Gamma}/\eta$. In terms of the thermodynamic vector $\vec{x}$, we find that

$$\omega^{(\mathcal{X})} = e^{-\vec{x}\cdot\vec{\Gamma}/\eta}/\mathrm{tr}(e^{-\vec{x}\cdot\vec{\Gamma}/\eta}). \tag{36}$$

For a qubit in the standard Pauli-matrix basis $\vec{\Gamma} = \vec{\sigma}/2$, this further reduces to

$$\omega^{(\mathcal{X})} = e^{-\vec{x}\cdot\vec{\sigma}}/\mathrm{tr}(e^{-\vec{x}\cdot\vec{\sigma}}) \tag{37}$$

$$= I/2 - \frac{\hat{x}\cdot\vec{\sigma}}{2}\tanh x. \tag{38}$$

Intuitively, Eq. (38) tells us that the minimizing Bloch vector $\vec{a}^* = -\tanh(x)\hat{x}$ points in the opposite direction to the thermodynamic vector $\vec{x}$ to reduce entropy flow or promote energy gain, as the case may be. However, this tendency to reduce heat or increase energy is balanced against the entropy gain incurred when tarnishing a pure state. As the magnitude of the thermodynamic vector x grows beyond unity, the minimizing Bloch vector converges exponentially to the edge of the Bloch sphere.

For example, for any reliable overwriting process operating on a qubit, this tells us that the unique initial state leading to minimal entropy production is $I/2 - (\hat{\varphi}\cdot\vec{\sigma}/2)\tanh\varphi$, in agreement [37] with Ref. [Eq. (25)] [4] obtained by different means. Whereas Ref. [4] has identified the initial state of a qubit that would lead to minimal entropy production during any qubit-reset process, the current result provides a significant generalization. We now identify the ideal input to any overwriting process in any finite dimension for any type-II objective, including maximizing free-energy gain.

While the initial states leading to minimal entropy production or maximal gain in free energy are nontrivial mixed states, the input leading to the largest reduction in entropy is always the same for any overwriting process. Note that the largest reduction in entropy is achieved by the fully mixed input state $\omega^{(0)} = I/d$ for any overwriting process in any dimension d.

We have so far found the ideal inputs to *minimize* the expectation values of type-II quantities during an overwriting process—e.g., the inputs leading to minimal entropy production or maximal increase in nonequilibrium free energy. It is just as natural and important to ask: Which inputs *maximize* the expectation values of type-II quantities, leading, e.g., to maximal entropy production or the biggest reduction in nonequilibrium free energy?

Since $S(\rho_\tau) = S(r_\tau)$ is independent of the input to an overwriting process, maximizing $f^{(\mathcal{X})}_{\rho_0} = \mathrm{tr}(\rho_0\mathcal{X}) + S(\rho_\tau) - S(\rho_0)$ asks us to simultaneously minimize $S(\rho_0)$ and maximize $\mathrm{tr}(\rho_0\mathcal{X})$. Note that $S(\rho_0)$ is minimized for any pure state, whereas $\mathrm{tr}(\rho_0\mathcal{X})$ is maximized by some pure state, according to Theorem 1. Accordingly, the two objectives can be simultaneously satisfied by identifying the pure state that maximizes $\mathrm{tr}(\rho_0\mathcal{X})$. We thus inherit an answer from Theorem 1, when seeking to maximize the expectation value of a type-II quantity.

Theorem 5. For any overwriting process, the expectation value of any type-II quantity $\langle X\rangle_{\rho_0} = p f^{(\mathcal{X})}_{\rho_0} = p[\mathrm{tr}(\rho_0\mathcal{X}) + S(\rho_\tau) - S(\rho_0)]$ is maximized by a pure-state input that maximizes $\mathrm{tr}(\rho_0\mathcal{X})$. This is satisfied by any eigenstate of $\mathcal{X}$ with maximal eigenvalue.

For example, for any overwriting process, the input that maximizes entropy production is the pure state that maximizes entropy flow, which is an eigenstate of the expected-entropy-flow operator $\mathbf{\Phi}$ with maximal eigenvalue. Similarly, the biggest reduction in nonequilibrium free energy will be achieved for any overwriting process from the pure-state input that maximizes the reduction in energy, which is an eigenstate of the change-of-energy operator $\mathcal{Q} + \mathcal{W}$ with minimal eigenvalue.

C. Perturbative correction to type-II minimizers

Here, we find a closed-form expression for the approximate minimizer of any type-II functional—e.g., the input state that minimizes entropy production or maximizes the change in nonequilibrium free energy—for any process, via a second-order expansion of the type-II functional around a reference state. This method becomes exact in the limit that the true minimizer is a small perturbation from the reference state—e.g., in the case of small changes to a protocol with known minimizer.

To obtain our result, we minimize $f^{(\mathcal{X})}_{\rho_0}$ over the space of valid generalized Bloch vectors. It is convenient to introduce the function $\tilde{\rho}$ with $\tilde{\rho}(\vec{b}) = I/d + \vec{b}\cdot\vec{\Gamma}$, which maps Bloch vectors to their corresponding density matrices, and the function $\tilde{\rho}'$ with $\tilde{\rho}'(\vec{b}) = \Gamma'_0 + \vec{b}\cdot\vec{\Gamma}'$, which maps initial Bloch vectors (at time 0) to their corresponding time-evolved density matrices at time τ. To simplify the notation, we introduce the function $\tilde{f}$ such that $\tilde{f}(\vec{b}) = f^{(\mathcal{X})}_{\tilde{\rho}(\vec{b})}$.

We seek the optimal $\vec{b}^* = \mathrm{argmin}_{\vec{b}}\tilde{f}(\vec{b})$ in the *perturbative regime*, where the optimal $\tilde{\rho}(\vec{b}^*)$ is a small perturbation away from some reference initial state $\tilde{\rho}(\vec{\pi})$. For instance,

$\tilde{f}$ may represent entropy production and $\tilde{\rho}(\vec{\pi})$ may be an initial equilibrium state such that $\tilde{\rho}(\vec{b}^*) \approx \tilde{\rho}(\vec{\pi})$ in the limit of slow time-dependent driving.

Assume that the density matrix $\rho(\vec{\pi})$ is positive definite. By continuity, for any $\vec{b}$ sufficiently close to $\vec{\pi}$, $\tilde{\rho}(\vec{b})$ is also positive definite and therefore a valid density matrix [38]. Now expand $\tilde{f}$ to second order in $\vec{\epsilon} := \vec{b} - \vec{\pi}$,

$$\tilde{f}(\vec{b}) \approx \tilde{f}(\vec{\pi}) + \vec{\epsilon}^\top \vec{j} + \frac{1}{2}\vec{\epsilon}^\top \mathbf{H}\vec{\epsilon}, \tag{39}$$

where $\vec{j}$ is the gradient vector and $\mathbf{H}$ is the Hessian matrix of $\tilde{f}$ evaluated at $\vec{\pi}$. The elements of the gradient vector are given by

$$\begin{aligned} j_n &= \partial_{b_n}\tilde{f}\,\Big|_{\vec{\pi}} \\ &= \mathrm{tr}(\Gamma_n \mathcal{X}) + \mathrm{tr}\left[\Gamma_n \ln\tilde{\rho}(\vec{\pi})\right] - \mathrm{tr}\left[\Gamma'_n \ln\tilde{\rho}'(\vec{\pi})\right]. \end{aligned} \tag{40}$$

The elements of the Hessian matrix are $\mathbf{H}_{m,n} = (\partial_{b_m}\partial_{b_n}\tilde{f})|_{\vec{\pi}}$. The Hessian follows by considering the second derivatives of the von Neumann entropy, which we calculate in Appendix H, to find

$$\partial_{b_m}\partial_{b_n} S(\tilde{\rho})\big|_{\vec{\pi}} = -\sum_{k,\ell} \phi(\nu_k, \nu_\ell)\langle k|\Gamma_n|\ell\rangle\langle\ell|\Gamma_m|k\rangle, \tag{41}$$

where we have used the eigendecomposition $\tilde{\rho}(\vec{\pi}) = \sum_k \nu_k |k\rangle\langle k|$ and defined $\phi(a,b) := (\ln a - \ln b)/(a-b)$ (with $\phi(a,a) = 1/a$ by continuity), which is the reciprocal of the logarithmic mean. We note that this second-order expansion has some resemblance to the Kubo-Mori-Bogoliubov metric of quantum information geometry [39]. We derive a similar expression for $\partial_{b_m}\partial_{b_n} S(\tilde{\rho}')$, which gives the following form for the matrix elements of the Hessian:

$$\begin{aligned} \mathbf{H}_{m,n} = &\left[\sum_{k,\ell} \phi(\nu_k, \nu_\ell)\langle k|\Gamma_n|\ell\rangle\langle\ell|\Gamma_m|k\rangle\right] \\ &- \left[\sum_{k,\ell} \phi(\nu'_k, \nu'_\ell)\langle k'|\Gamma'_n|\ell'\rangle\langle\ell'|\Gamma'_m|k'\rangle\right], \end{aligned} \tag{42}$$

where we have used the eigendecomposition $\tilde{\rho}'(\vec{\pi}) = \sum_k \nu'_k |k'\rangle\langle k'|$. Note that $\mathbf{H}$ is positive semidefinite since $\tilde{f}$ is a convex function. For simplicity, assume for now that $\tilde{f}$ is strictly convex, in which case $\mathbf{H}$ is positive definite and has an inverse $\mathbf{H}^{-1}$. (We give the generalization later.)

Finally, we minimize Eq. (39) in closed form. First, complete the square to write

$$\begin{aligned} \vec{\epsilon}^\top \vec{j} + \frac{1}{2}\vec{\epsilon}^\top \mathbf{H}\vec{\epsilon} &= \frac{1}{2}(\vec{\epsilon} + \mathbf{H}^{-1}\vec{j})^\top \mathbf{H}(\vec{\epsilon} + \mathbf{H}^{-1}\vec{j}) \\ &\quad - \frac{1}{2}\vec{j}^\top \mathbf{H}^{-1}\vec{j} \qquad (43) \\ &\geq -\frac{1}{2}\vec{j}^\top \mathbf{H}^{-1}\vec{j}, \qquad (44) \end{aligned}$$

where the last inequality is achieved by setting $\vec{\epsilon}^* = -\mathbf{H}^{-1}\vec{j}$. This implies that in the perturbative regime, the optimal state is given by

$$\tilde{\rho}(\vec{b}^*) = \tilde{\rho}(\vec{\pi}) - (\mathbf{H}^{-1}\vec{j}) \cdot \vec{\Gamma}, \tag{45}$$

which achieves the optimal value

$$\min \tilde{f}(\vec{b}) = \tilde{f}(\vec{\pi}) - \frac{1}{2}\vec{j}^\top \mathbf{H}^{-1}\vec{j}. \tag{46}$$

The above expressions for the optimal state generalize to allow for a singular Hessian if we replace $\mathbf{H}^{-1}$ with the Drazin inverse $\mathbf{H}^{\mathcal{D}}$ (or the group inverse, since we can diagonalize the Hessian) [40].

IX. GENERALIZED BLOCH VECTORS, THERMODYNAMIC OPERATORS, AND IDEAL INPUTS WITHIN A RESTRICTED SUBSPACE

It will often be useful to know the ideal input to a device, within a restricted subspace of possible inputs. For example, when operating on a system with a countably infinite number of energy eigenstates, we may care about inputs with nonzero probability amplitude only in the lowest N energy eigenstates, for some finite N. Or, we may be interested in the best classical inputs to a device, when coherent states cannot be readily prepared. In such cases, we can find the thermodynamically ideal input to the transformation within this finite-dimensional subspace via a very straightforward adaptation of the above techniques.

Let $\mathcal{P}_P$ be the convex subspace of the possible density matrices $\mathcal{P}$ of the system, induced by the set of orthogonal projectors $P = \{\Pi_j\}_j$ with $\Pi_j \Pi_k = \delta_{j,k}\Pi_j$, such that

$$\mathcal{P}_P := \left\{\rho \in \mathcal{P} : \rho = \sum_{\Pi \in P} \Pi\rho\Pi\right\}. \tag{47}$$

Density matrices in this subspace act on a d_P-dimensional vector space $\mathcal{V}_P$, where $d_P = \sum_{\Pi\in P}\mathrm{tr}(\Pi)$. The identity operator on $\mathcal{V}_P$ is given by $I_P = \sum_{\Pi\in P}\Pi$.

Each projector $\Pi_j \in P$ has an associated dimension $d_j = \mathrm{tr}(\Pi_j)$. The restricted subspace $\mathcal{P}_P$ is spanned by a basis of L linearly independent density matrices, where $L = \sum_{j=1}^{|P|} d_j^2$. In general, $d_P \leq L \leq d_P^2$.

By creating an appropriate operator basis and generalized Bloch vector for initial density matrices restricted to $\mathcal{P}_P$, we can easily adapt and leverage all results of previous sections in this restricted setting. To achieve this, the following development closely parallels the previous introduction of generalized Bloch vectors.

When the initial density matrix ρ_0 is restricted to $\mathcal{P}_P$, it is useful to choose a complete basis $(I_P/d_P, \Gamma_1, \Gamma_2, \ldots, \Gamma_{L-1})$ for linear operators acting on $\mathcal{V}_P$, such that the Hermitian operators Γ_n are all traceless and mutually orthogonal, satisfying $\text{tr}(\Gamma_n) = 0$ and $\text{tr}(\Gamma_m \Gamma_n) = \eta_P\, \delta_{m,n}$, where we will choose the normalizing constant to be $\eta_P = (d_P - 1)/d_P$. Any initial density matrix then has a unique decomposition in the operator basis $\vec{\Gamma}_P = (\Gamma_1, \Gamma_2, \ldots, \Gamma_{L-1})$, described by the generalized Bloch vector $\vec{b}_0 \in \mathbb{R}^{L-1}$ via

$$\rho_0 = I_P/d_P + \vec{b}_0 \cdot \vec{\Gamma}_P. \tag{48}$$

Since the magnitude of the Bloch vector is $b_0 = \sqrt{[\text{tr}(\rho_0^2) d_P - 1]/(d_P - 1)}$, the density matrix represents a pure state if and only if the magnitude of its corresponding Bloch vector is one.

For concreteness, we can choose the ordered operator basis to begin with a scaled ordering of the d_P diagonal generalized Gell-Mann matrices, followed by $d_j(d_j - 1)/2$ nondiagonal symmetric Gell-Mann matrices and $d_j(d_j - 1)/2$ antisymmetric Gell-Mann matrices for each projector $\Pi_j \in P$ with $d_j > 1$. Recall that these matrices are given explicitly in Appendix E.

It is now productive to consider the restriction of a thermodynamic operator to this subspace $\mathcal{X}_P := \sum_{\Pi \in P} \Pi \mathcal{X} \Pi$. It is easy to check that the expectation value for any type-I thermodynamic quantity satisfies

$$\text{tr}(\rho_0 \mathcal{X}_P) = \text{tr}(\rho_0 \mathcal{X}) = \langle X \rangle_{\rho_0} \quad \text{for all } \rho_0 \in \mathcal{P}_P. \tag{49}$$

Notably, $\mathcal{X}_P$ can be constructed directly, just as $\mathcal{X}$ was in the previous sections—but now in the finite-dimensional restricted subspace $\mathcal{P}_P$, using the operator basis $\vec{\Gamma}_P$, the generalized Bloch vector, and the thermodynamic operator $\vec{x}_P$ associated with this restricted subspace. For general processes, the thermodynamic operator restricted to this subspace is constructed as

$$\mathcal{X}_P = \langle X \rangle_{I_P/d_P} I_P + \vec{x}_P \cdot \vec{\Gamma}_P/\eta_P. \tag{50}$$

The thermodynamically ideal inputs within the subspace of interest are constructed just as before, with the extremal eigenvalues and associated eigenstates of $\mathcal{X}_P$ playing the special role indicated in Theorems 1 and 5. For overwriting processes, the minimal value of type-II quantities is achieved within subspace $\mathcal{P}_P$ by $\omega^{(\mathcal{X}_P)} = e^{-\mathcal{X}_P}/\text{tr}(e^{-\mathcal{X}_P}) = e^{-\vec{x}_P \cdot \vec{\Gamma}_P/\eta_P}/\text{tr}(e^{-\vec{x}_P \cdot \vec{\Gamma}_P/\eta_P})$, which extends an analogous result of Ref. [5] that was formulated for the case of entropy production.

It is interesting to note that these methods—restricting to the d_P-dimensional subspace—work even though the states are not restricted to this subspace during their evolution. Indeed, the protocol can spread these initially restricted inputs across infinite dimensions but our finite-dimensional inference of the ideal input within this initial subspace remains valid.

To pursue the example of inputs with support restricted to the N lowest-energy eigenstates $\{|E_n\rangle\}_{n=1}^N$ of some initial system Hamiltonian H_0, we would consider the density matrices with support on the Hilbert subspace of the system, $\mathcal{H}_\text{sub} = \text{span}\left(\{|E_n\rangle\}_{n=1}^N\right)$; i.e., we would consider the restricted set of initial density matrices $\mathcal{P}_\text{sub} = \left\{\rho = \sum_\ell p_\ell (|\psi_\ell\rangle\langle\psi_\ell| / \langle\psi_\ell|\psi_\ell\rangle) : p_\ell \in (0,1], \sum_\ell p_\ell = 1, |\psi_\ell\rangle \in \mathcal{H}_\text{sub}\right\}$, which is induced by the single rank-N projector $\Pi = \sum_{n=1}^N |E_n\rangle\langle E_n|$. Within this subspace, there are $L = d_{\{\Pi\}}^2 = N^2$ linearly independent initial states (compared to the $d^2 = \infty$ linearly independent initial states within the full Hilbert space).

A. Classical systems

If we define classical inputs as those states restricted to be initially incoherent in a particular "classical" orthonormal basis $\mathcal{C}$, then we can see how the general quantum problem simplifies significantly, if we seek the classical thermodynamic operators and ideal classical inputs. Classical density matrices are those induced by the set of orthonormal rank-1 projectors $P = \{|b\rangle\langle b|\}_{b \in \mathcal{C}}$. Classical density matrices are diagonal in the classical basis and can be regarded as a representation of a probability distribution over these classical states. Note that for d-dimensional systems, there are only $L = d$ linearly independent classical density matrices, in contrast to the d^2 linearly independent quantum density matrices over the same space. Accordingly, each classical thermodynamic vector has $d - 1$ components, rather than the $d^2 - 1$ components of its quantum counterpart. Similarly, each classical thermodynamic operator will be fully determined by the behavior of d, rather than d^2, initial states.

X. EXAMPLE 1: NONEQUILIBRIUM THERMODYNAMICS OF A QUBIT-RESET DEVICE

Quantum computing requires a mechanism for resetting each qubit to the computational-basis state $|0\rangle = \sigma_z |0\rangle$. Different implementations of the same task will, however, have distinct sets of thermodynamically ideal inputs. Nonequilibrium thermodynamic quantities are determined less by *what* you do than *how* you do it.

The following provides a paradigmatic illustration of our results, for a straightforward qubit-reset mechanism, highlighting (i) the diversity of ideal inputs depending on the thermodynamic objective and (ii) how the behavior of a small number of test inputs bounds the thermodynamic behavior of all possible inputs to the device.

Section X A gives the particular dynamics for our example. Section X B covers how to build the thermodynamic vectors and operators, for any operation on a qubit. Subsequent subsections then demonstrate the diversity of ideal inputs and that the breadth of possible thermodynamic behaviors is determined by the behavior of a small number of inputs.

A. Dynamical equations for a qubit-reset device

As in Refs. [4,41], our example qubit-reset device works by changing both the energy gap and spatial orientation of the energy eigenstates of the qubit while it is in weak contact with a thermal reservoir at inverse temperature $\beta = 1/(k_B T)$. Over the finite-time protocol, from time 0 through $\tau = 50\beta\hbar$, the time-varying Hamiltonian is

$$H_t = \frac{E_t}{2}\left[\cos(\theta_t)\sigma_z + \sin(\theta_t)\sigma_x\right], \tag{51}$$

where $E_t = k_B T[1 + 49\sin^2(\pi t/100\beta\hbar)]/5$ and $\theta_t = \pi t/(50\beta\hbar)$. While E_t quantifies the energy gap between the instantaneous energy eigenstates of the system, θ_t parametrizes the instantaneous orientation of the energy eigenbasis relative to the “computational” z-basis. We assume idealized conditions (large baths, weak coupling, etc.) such that the dynamics are well described by a time-dependent quantum master equation,

$$\dot{\rho}_t = \mathcal{L}_t(\rho_t) = \frac{i}{\hbar}[\rho_t, H_t] + \frac{cE_t}{\hbar}(N_t + 1)\mathcal{D}[L_t](\rho_t) + \frac{cE_t}{\hbar}N_t\mathcal{D}[L_t^\dagger](\rho_t), \tag{52}$$

where $\mathcal{D}[L](\rho) = L\rho L^\dagger - \frac{1}{2}\{L^\dagger L, \rho\}, N_t = (e^{\beta E_t} - 1)^{-1}$, and $c = 1/5$ is the coupling strength to the bath. The time-dependent lowering operator can be represented as

$$L_t = \frac{1}{2}\left[\cos(\theta_t)\sigma_x - i\sigma_y - \sin(\theta_t)\sigma_z\right] \tag{53}$$

and satisfies the detailed-balance condition $[L_t, H_t] = E_t L_t$ [41,42]. Transitions thus occur between instantaneous energy eigenstates of the system. In particular, $\mathcal{D}[L_t](\rho_t)$ takes the excited population and shifts it to the ground state, while $\mathcal{D}[L_t^\dagger](\rho_t)$ takes the ground-state population and shifts it to the excited state. Moreover, the ratio of transition rates between the excited- and ground-state populations satisfies detailed balance since $(N_t + 1)/N_t = e^{\beta E_t}$.

B. Thermodynamic operators for a qubit

When the system of interest is a single qubit then, regardless of the size of the environment or the dynamics under consideration, a standard Bloch matrix can be employed. We can choose $\rho_0^{(1)} = I/2$, $\rho_0^{(2)} = I/2 + \sigma_x/2 = |+\rangle\langle+|$, $\rho_0^{(3)} = I/2 + \sigma_y/2$, and $\rho_0^{(4)} = I/2 + \sigma_z/2 = |0\rangle\langle 0|$. Recall that the standard Bloch vectors can be obtained as $\vec{b}_0^{(n)} = \mathrm{tr}(\rho_0^{(n)}\vec{\sigma})$. From our proposed choice of initial density matrices above, we then obtain a standard Bloch matrix for qubits:

$$B = \begin{bmatrix} 1 & \vec{b}_0^{(1)} \\ 1 & \vec{b}_0^{(2)} \\ 1 & \vec{b}_0^{(3)} \\ 1 & \vec{b}_0^{(4)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \tag{54}$$

which is easily inverted to yield

$$B^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}. \tag{55}$$

To find the thermodynamic vectors and thermodynamic operators, we need to now either experimentally or analytically obtain the expectation values for the thermodynamic variables under these four initial system states, $\langle X\rangle_{\rho_0^{(n)}}$ for $n \in \{1, 2, 3, 4\}$. We then find the thermodynamic vectors via Eq. (16),

$$\begin{bmatrix} \langle X\rangle_{I/2} \\ \vec{x} \end{bmatrix} = B^{-1}\begin{bmatrix} \langle X\rangle_{\rho_0^{(1)}} \\ \langle X\rangle_{\rho_0^{(2)}} \\ \langle X\rangle_{\rho_0^{(3)}} \\ \langle X\rangle_{\rho_0^{(4)}} \end{bmatrix}, \tag{56}$$

and finally obtain the thermodynamic operators via Eq. (14):

$$\mathcal{X} = \langle X \rangle_{I/2} I + \vec{x} \cdot \vec{\sigma}. \tag{57}$$

C. Diversity among ideal inputs for thermodynamic objectives

To determine how the thermodynamic behavior of the device depends on the input state, we track the evolution of four randomly sampled initial density matrices, together with the thermodynamic output from each of these four inputs. From the Bloch matrix B and the measured thermodynamic output, we construct the Hermitian thermodynamic operators. For example, the expected-heat and expected-work operators, $\mathcal{Q}$ and $\mathcal{W}$, allow us to determine (i) the ideal inputs leading to minimal and maximal heat and work and (ii) the full range of heat and work that can be attained by *any* input to the device. These are obtained from the extremal eigenvalues and associated eigenstates of the thermodynamic operators.

Simple combination and manipulation of the heat and work operators reveals the diversity of ideal inputs for a multitude of different thermodynamic objectives, as shown in Fig. 1.

In the figures, we emphasize the heat exhausted to the environment $-Q$, rather than the heat Q directly. In this example, with a single environmental bath at constant temperature T, note that entropy flow to the environment is simply related to heat out of the system via $\Phi = -Q/T$. The expected-entropy-flow operator is thus simply related to the expected-heat operator in this case, via $\boldsymbol{\Phi} = -\mathcal{Q}/T$. Meanwhile, the expected-energy-change operator is simply $\mathcal{Q} + \mathcal{W}$.

D. Bounding the behavior of all inputs

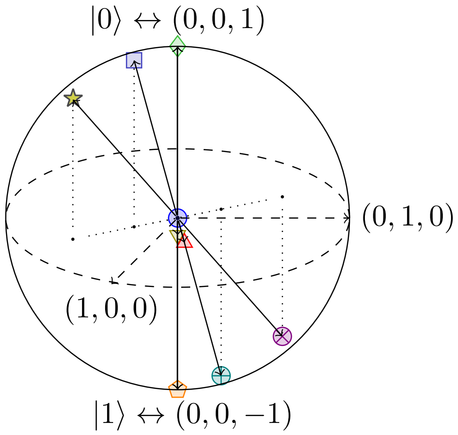
Continuing our example of the qubit-reset dynamics, we now leverage our results to identify the extremal thermodynamic behavior that can be attained by any input throughout the process.

In Fig. 2, we demonstrate that thermodynamic observations from just four inputs yield the full range of thermodynamic behavior from any input. For example, the minimum and maximum expected work at any time $t \in [0, \tau]$, obtainable from alternative inputs, is determined by the expected-work operator at that time. The expected-work operator at any time is constructed from the expected work performed on each of the four test inputs up to that time. Determining the ranges of work, energy change, and heat thus reduces to determining eigenvalue ranges of the respective Hermitian operators.

Determining the range of entropy production throughout the process is somewhat more complicated, although it still only requires the data from four test inputs. Notably, in Fig. 2(b), we find the states of minimal and maximal entropy production at times before the state is fully reset. This employs the gradient-descent algorithm developed in Sec. VIII A.

XI. EXAMPLE 2: STRONG INTERACTIONS WITH SMALL ENVIRONMENTS

We emphasize that we can find the ideal inputs and breadth of behavior in terms of thermodynamic operators that act on the system alone, even when the system is strongly coupled to an arbitrarily small (or arbitrarily



	Objective	Ideal input		
△	min. entropy production	$\mathrm{argmin}_{\rho_0} \langle \Sigma \rangle_{\rho_0}$	mixed	$\vec{a}_0 \approx (-0.02, 0.03, -0.15)$
▽	max. free-energy gain	$\mathrm{argmax}_{\rho_0} \Delta \mathcal{F}_t$	mixed	$\vec{a}_0 \approx (0, 0, -0.10)$
○	min. entropy change	$\mathrm{argmax}_{\rho_0} \Delta S_t$	mixed	$\vec{a}_0 = (0, 0, 0)$
	max. heat exhausted	$\mathrm{argmax}_{\rho_0} \langle -Q \rangle_{\rho_0} =$	pure	$\vec{a}_0 \approx (0.13, -0.20, 0.97)$
□	max. entropy production	$\mathrm{argmax}_{\rho_0} \langle \Sigma \rangle_{\rho_0} =$		
	max. entropy flow	$\mathrm{argmax}_{\rho_0} \langle \Phi \rangle_{\rho_0}$		
⊕	min. heat exhausted	$\mathrm{argmin}_{\rho_0} \langle -Q \rangle_{\rho_0} =$	pure	$\vec{a}_0 \approx (-0.13, 0.20, -0.97)$
	min. entropy flow	$\mathrm{argmin}_{\rho_0} \langle \Phi \rangle_{\rho_0}$		
◇	min. free-energy change	$\mathrm{argmin}_{\rho_0} \Delta \mathcal{F}_t =$	pure	$\vec{a}_0 \approx (0, 0, 1)$
	min. energy change	$\mathrm{argmin}_{\rho_0} \langle Q + W \rangle_{\rho_0}$		
⬠	max. energy change	$\mathrm{argmax}_{\rho_0} \langle Q + W \rangle_{\rho_0}$	pure	$\vec{a}_0 \approx (0, 0, -1)$
⊗	min. work	$\mathrm{argmin}_{\rho_0} \langle W \rangle_{\rho_0}$	pure	$\vec{a}_0 \approx (-0.32, 0.49, -0.81)$
★	max. work	$\mathrm{argmax}_{\rho_0} \langle W \rangle_{\rho_0}$	pure	$\vec{a}_0 \approx (0.32, -0.49, 0.81)$
	max. entropy change	$\mathrm{argmax}_{\rho_0} \Delta S_t$	pure	anywhere on Bloch shell

FIG. 1. The diversity of ideal inputs for a finite-time qubit-reset process, displayed on and in the Bloch sphere. The states extremizing heat, work, and energy change all lie on the surface of the Bloch sphere, in the direction of a maximal eigenstate of the corresponding thermodynamic operators. The entire surface of the Bloch sphere maximizes entropy gain. Minimal entropy production and maximal free energy gain are achieved by nontrivial mixed-state inputs. The change in entropy is minimized by the fully mixed input. Entropy production is maximized by the same pure-state input that maximizes heat exhaustion. The greatest loss of free energy occurs for the same pure-state input that loses the most energy.

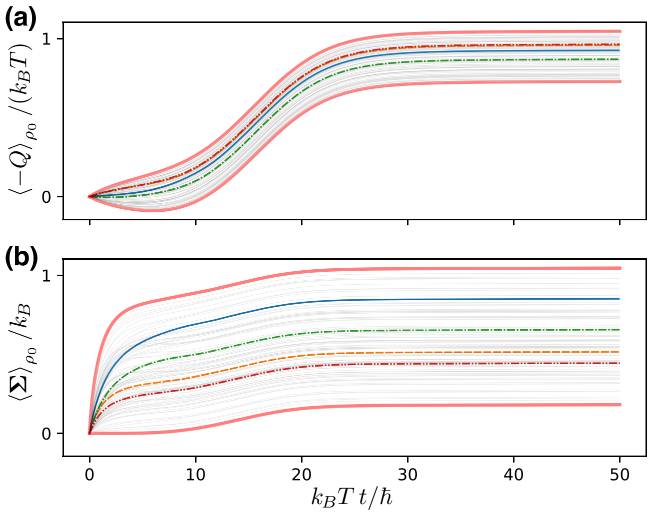


FIG. 2. Tracking the behavior of four inputs is enough to bound the behavior of all other inputs to a qubit process. Here, we show the range of expectation values for exhausted heat and entropy production throughout a finite-time qubit-reset process. The expectation values from four random inputs are shown as dashed lines. This allows construction of the thermodynamic operator $\mathcal{Q}$ throughout time. (a) The maximal and minimal heat, corresponding to extremal eigenvalues of $\mathcal{Q}$, shown as thick red solid lines. (b) The maximal and minimal entropy production, obtained from gradient descent or ascent, shown as thick red solid lines. These extrema bound the behavior of all other inputs, including the behavior of 100 other random initial conditions shown as thin gray solid lines.

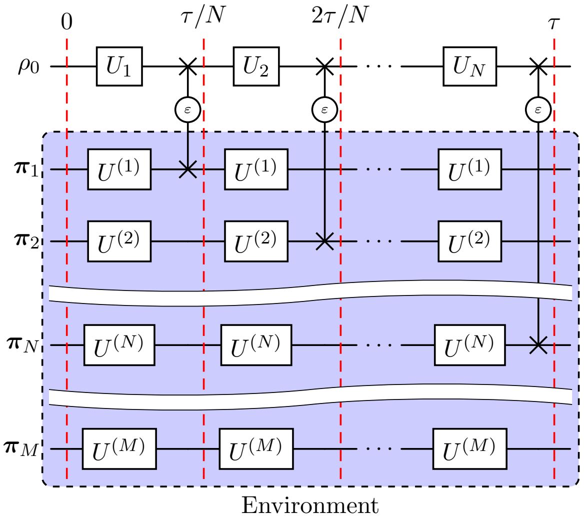


FIG. 3. A flexible toy model for interactions with an arbitrarily small environment, with time-dependent system dynamics regularly interrupted by partial SWAPs with a sequence of (potentially distinct) Gibbs states. Vertical slices through the quantum circuit indicate equal-time segments of duration τ/N. Each bath has a time-homogenous Hamiltonian inducing the unitary time evolution $U^{(m)} = e^{-iH^{(m)}\tau/\hbar N}$, while the system can be controlled to evolve more arbitrarily according to some $(U_n)_{n=1}^N$.

large) environment. The following example aims to make this clear, by analyzing a sequence of partial SWAPs of system and environmental subsystems, where each is represented by a single qubit.

A. Partial SWAP with the environment

As its name implies, the SWAP gate is defined by the unitary $U_{\text{SWAP}} |\alpha\rangle |\beta\rangle = |\beta\rangle |\alpha\rangle$ for any two qudit states $|\alpha\rangle$ and $|\beta\rangle$ of the same dimension [43]. More generally, the partial SWAP gate is defined by the unitary [44,45]

$$U^{\varepsilon}_{\text{SWAP}} = e^{-i\varepsilon\pi/2} \left(cI + isU_{\text{SWAP}} \right), \tag{58}$$

where $c := \cos(\varepsilon\pi/2)$ and $s := \sin(\varepsilon\pi/2)$, although the overall phase $e^{-i\varepsilon\pi/2}$ is physically irrelevant. For qubits, the Heisenberg exchange Hamiltonian naturally implements this partial SWAP operation [46]. When $\varepsilon = 1/2$, the $\sqrt{\text{SWAP}}$ gate is implemented which, together with single-qubit gates, is capable of universal quantum computation [46].

We consider an environment made up of M qudits that are of the same finite dimension d as the system. The system has some time-dependent unitary evolution interrupted at regular intervals by a sequence of $N \leq M$ nearly instantaneous partial SWAPs with different parts of the environment. Each part of the environment is in local equilibrium but each can be a Gibbs state $\boldsymbol{\pi}_m = e^{-H^{(m)}/k_B T_m} / \text{tr}\left(e^{-H^{(m)}/k_B T_m} \right)$ relative to its own Hamiltonian $H^{(m)}$ at its own temperature T_m. The initial local-equilibrium state of the environment is thus $\boldsymbol{\pi}^{\text{env}} = \bigotimes_{m=1}^M \sigma_0^{(m)} = \bigotimes_{m=1}^M \boldsymbol{\pi}_m$, while $\sigma_t^{(m)}$ will denote the reduced state of each bath throughout the protocol. The general setup for this physical interaction between system and environment is depicted via the annotated circuit diagram of Fig. 3. In various special cases, this setup can be related to the frameworks of thermal operations [47], "repeated interactions" [48], and collision models [43].

Assuming an implicit work reservoir, we can quantify the work supplied or extracted as the change in total energy of the system-environment supersystem, as in Eq. (6):

$$\begin{aligned}\langle W^{(*)} \rangle_{\rho_0} &= \text{tr}(\rho_\tau^{\text{tot}} H_\tau^{\text{tot}}) - \text{tr}(\rho_0^{\text{tot}} H_0^{\text{tot}}) \\ &= \text{tr}(\rho_\tau H_\tau) - \text{tr}(\rho_0 H_0) + \sum_{n=1}^{N} \text{tr}\left[(\sigma_\tau^{(n)} - \boldsymbol{\pi}_n) H^{(n)} \right]\end{aligned} \tag{59}$$

$$\begin{aligned}&= \text{tr}(\rho_\tau H_\tau) - \text{tr}(\rho_0 H_0) \\ &\quad + \sum_{n=1}^{N} \text{tr}\left[(\sigma_{n\tau/N}^{(n)} - \boldsymbol{\pi}_n) H^{(n)} \right],\end{aligned} \tag{60}$$

where the last line follows from $U^{(n)\dagger}H^{(n)}U^{(n)} = H^{(n)}$. To make further progress, we find that

$$\begin{aligned}\sigma^{(n)}_{n\tau/N} &= \mathrm{tr}_1\left[U^{\varepsilon}_{\mathrm{SWAP}}\left(\breve{\rho}_{(n-1)\tau/N}\otimes\boldsymbol{\pi}_n\right)U^{\varepsilon\,\dagger}_{\mathrm{SWAP}}\right] && (61)\\ &= \mathrm{tr}_1\left[(cI + isU_{\mathrm{SWAP}})\left(\breve{\rho}_{(n-1)\tau/N}\otimes\boldsymbol{\pi}_n\right)\right.\\ &\qquad \left.\times\,(cI - isU_{\mathrm{SWAP}})\right] && (62)\\ &= c^2\boldsymbol{\pi}_n + s^2\breve{\rho}_{(n-1)\tau/N} + ics\left[\breve{\rho}_{(n-1)\tau/N},\boldsymbol{\pi}_n\right] && (63)\end{aligned}$$

and

$$\begin{aligned}\rho_{n\tau/N} &= \mathrm{tr}_2\left[U^{\varepsilon}_{\mathrm{SWAP}}\left(\breve{\rho}_{(n-1)\tau/N}\otimes\boldsymbol{\pi}_n\right)U^{\varepsilon\,\dagger}_{\mathrm{SWAP}}\right] && (64)\\ &= \mathrm{tr}_2\left[(cI + isU_{\mathrm{SWAP}})\left(\breve{\rho}_{(n-1)\tau/N}\otimes\boldsymbol{\pi}_n\right)\right.\\ &\qquad \left.\times\,(cI - isU_{\mathrm{SWAP}})\right] && (65)\\ &= c^2\breve{\rho}_{(n-1)\tau/N} + s^2\boldsymbol{\pi}_n + ics\left[\boldsymbol{\pi}_n,\breve{\rho}_{(n-1)\tau/N}\right], && (66)\end{aligned}$$

where we have used the shorthand notation $\breve{\rho}_{n\tau/N} := U_n\rho_{n\tau/N}U_n^\dagger$, and we use the fact that

$$\mathrm{tr}_1\left([U_{\mathrm{SWAP}},\rho\otimes\sigma]\right) = [\rho,\sigma] \tag{67}$$

and

$$\mathrm{tr}_2\left([U_{\mathrm{SWAP}},\rho\otimes\sigma]\right) = [\sigma,\rho], \tag{68}$$

where $\mathrm{tr}_1(\cdot)$ traces over the first subsystem (here, the "system" of interest) and $\mathrm{tr}_2(\cdot)$ traces over the second subsystem (here, the relevant bath subsystem).

For any choice of baths $(\boldsymbol{\pi}_n)_{n=1}^N$ and time-inhomogeneous unitaries $(U_n)_{n=1}^N$ acting on the system, the system state can be simply iterated forward via Eq. (66) as a time-inhomogeneous transition operator (which is most obvious to see in the matrix representation acting on the generalized Block vector). Moreover, the relevant bath states, after interacting with the system, can be obtained from Eq. (63), which allows us to calculate the work in Eq. (60).

1. A simple $\sqrt{\text{SWAP}}$ engine

As an explicit example, we analyze the simple thermodynamic process shown in Fig. 4, which can function as an engine. We take $U_1 = U_2 = I$, while $\varepsilon = 1/2$ implements the $\sqrt{\text{SWAP}}$ gate with $c = s = 1/\sqrt{2}$. The system first interacts with a hot bath, with Hamiltonian $H^{(1)} = H_{\tau/2} = 2E\,|1\rangle\langle 1|$ and temperature $T_1 = 10T_2$, where $E/k_BT_1 > 0$. At the end of the cycle, the system interacts with a cold bath, with Hamiltonian $H^{(2)} = H_0 = H_\tau = E\,|1\rangle\langle 1|$ and temperature T_2. Under these conditions, we find that the work operator is diagonal in the computational basis,

$$\mathcal{W}^{(*)} = w_{\min}\,|0\rangle\langle 0| + w_{\max}\,|1\rangle\langle 1|, \tag{69}$$

where $w_{\min} \approx -0.225E$ and $w_{\max} \approx 0.275E$. Maximum work can be extracted from the pure input state $|0\rangle$ but it is natural to wonder whether work can also be extracted asymptotically in the steady state if the channel is applied repeatedly.

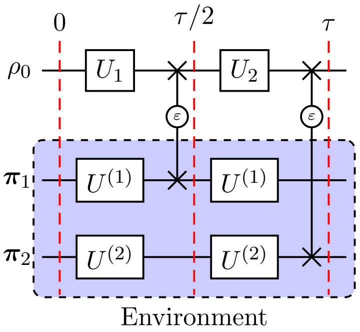


FIG. 4. One cycle through a simple thermodynamic process that can function as a quantum engine. Fresh environmental subsystems would be used for subsequent cycles. The system Hamiltonian is $H_t = w_t\,|1\rangle\langle 1|$, where w_t increases from E to $2E$ during the first stroke and decreases from $2E$ to E after the first partial SWAP, with the net effect that $U_1 = U_2 = I$. The system interacts first with a hot bath, with Hamiltonian $H^{(1)} = H_{\tau/2} = 2E\,|1\rangle\langle 1|$ and temperature $T_1 = 10T_2$. At the end of the cycle, the system interacts with a cold bath, with Hamiltonian $H^{(2)} = H_0 = H_\tau = E\,|1\rangle\langle 1|$ and temperature T_2.

Combining Eqs. (66) and (18) yields the matrices that update the Bloch vector of the system: $\begin{bmatrix}1 & \vec{b}_{\tau/2}\end{bmatrix} = \begin{bmatrix}1 & \vec{b}_0\end{bmatrix}\mathcal{T}_1$ and $\begin{bmatrix}1 & \vec{b}_\tau\end{bmatrix} = \begin{bmatrix}1 & \vec{b}_{\tau/2}\end{bmatrix}\mathcal{T}_2$. Adapting Eq. (18) of Ref. [44], we find that

$$\mathcal{T}_n = \begin{bmatrix}1 & 0 & 0 & q_n/2\\ 0 & 1/2 & -q_n & 0\\ 0 & q_n & 1/2 & 0\\ 0 & 0 & 0 & 1/2\end{bmatrix}, \quad \text{where } q_n = \mathrm{tr}(\boldsymbol{\pi}_n\sigma_z). \tag{70}$$

Each $\mathcal{T}_n$ has eigenvalues $\Lambda_n = \{1, 1/2, 1/2 \pm iq_n\}$, while the composite map

$$\begin{aligned}\mathcal{T} &= \mathcal{T}_1\mathcal{T}_2\\ &= \begin{bmatrix}1 & 0 & 0 & q_1/4 + q_2/2\\ 0 & 1/4 - q_1q_2 & -(q_1+q_2)/2 & 0\\ 0 & (q_1+q_2)/2 & 1/4 - q_1q_2 & 0\\ 0 & 0 & 0 & 1/4\end{bmatrix}\end{aligned} \tag{71}$$

has eigenvalues $\Lambda = \{1, 1/4, 1/4 - q_1q_2 \pm i(q_1+q_2)/2\}$. Moreover, the left eigenvector of $\mathcal{T}$ associated with the eigenvalue of unity yields the stationary Bloch vector of the system after full application of the protocol: $\begin{bmatrix}1 & \vec{b}^*\end{bmatrix}\mathcal{T} = \begin{bmatrix}1 & \vec{b}^*\end{bmatrix} = [1\ \ 0\ \ 0\ \ (q_1 + 2q_2)/3]$, revealing the stationary quantum state to this channel, $\rho^* = I/2 + (q_1 + 2q_2)\sigma_z/6$, which is an attractor upon repeated application of the channel. Work can indeed be extracted in the steady state, regardless of the initial input, since $\langle W^{(*)} \rangle_{\rho^*} = \mathrm{tr}(\rho^* \mathcal{W}^{(*)}) \approx -0.06E$ of work can be extracted asymptotically in each cycle.

For any initial input ρ_0 with Bloch vector $\vec{b}_0$, the state of the system after $k \in \{0, 1, 2, \ldots\}$ applications of the channel will be

$$\rho_{k\tau} = \left[\left(\begin{bmatrix} 1 & \vec{b}_0 \end{bmatrix} \mathcal{T}^k \right) \otimes I \right] \begin{bmatrix} I/2 \\ \sigma_x \\ \sigma_y \\ \sigma_z \end{bmatrix}. \tag{72}$$

This assumes that we use fresh bath subsystems for each application of the channel, but reuse the system in its new state as the "working medium" of the engine. We can then analytically determine the work expected upon the kth application of the channel:

$$\begin{aligned} \langle W^{(*)} \rangle_{\rho_{k\tau}} &= \mathrm{tr}(\rho_{k\tau} \mathcal{W}^{(*)}) \\ &= \begin{bmatrix} 1 & \vec{b}_0 \end{bmatrix} \mathcal{T}^k \begin{bmatrix} \langle W^{(*)} \rangle_{I/2} \\ \vec{w}^{(*)} \end{bmatrix} \qquad (73) \\ &= \sum_{\lambda \in \Lambda} \lambda^k \begin{bmatrix} 1 & \vec{b}_0 \end{bmatrix} \mathcal{T}_\lambda \begin{bmatrix} \langle W^{(*)} \rangle_{I/2} \\ \vec{w}^{(*)} \end{bmatrix}, \qquad (74) \end{aligned}$$

where $\mathcal{T}_\lambda = \vec{\lambda}_{\mathrm{R}} \vec{\lambda}_{\mathrm{L}} / \vec{\lambda}_{\mathrm{L}} \vec{\lambda}_{\mathrm{R}}$ is the spectral projection operator of $\mathcal{T}$ constructed from the outer product of right and left eigenvectors $\vec{\lambda}_{\mathrm{R}} \vec{\lambda}_{\mathrm{L}}$ normalized by their inner product $\vec{\lambda}_{\mathrm{L}} \vec{\lambda}_{\mathrm{R}}$, where $\lambda \vec{\lambda}_{\mathrm{R}} = \mathcal{T} \vec{\lambda}_{\mathrm{R}}$ is a column vector, and $\lambda \vec{\lambda}_{\mathrm{L}} = \vec{\lambda}_{\mathrm{L}} \mathcal{T}$ is a row vector [40]. Equation (74) shows how the extractable work per application of the channel converges to the steady-state value $\langle W^{(*)} \rangle_{\rho^*}$ via a sum of decaying exponentials for any initial input.

XII. EXAMPLE 3: EXTRACTING WORK FROM A SPATIALLY EXTENDED STATE

Let us now turn to the analysis of an infinite-dimensional quantum system: a charged particle in a time-dependent potential, in the presence of electromagnetic background radiation. This can be interpreted as an idealized model of a double quantum dot in the single-particle zero-current regime. In this example, we use time-dependent Lindbladian dynamics, with a time-modulated double well of potential energy across one spatial dimension. The protocol is capable of extracting work from some states that are initially localized in the right well. We consider the ideal input (for the task of work extraction) and how it compares to the performance of other inputs, within the Hilbert space spanned by the eight lowest-lying energy eigenstates of the initial Hamiltonian.

We will quantify work extraction simply as the negative work performed. In principle, this work can be extracted by a more explicit mechanism, as proposed in Ref. [49].

Given any protocol for work extraction, we can construct the expected-work operator to determine its breadth of behavior on all possible inputs. The device will be best at extracting work from the lowest-lying eigenstate of the expected-work operator; when the corresponding minimal eigenvalue is negative, its magnitude reports the maximal possible expected value of extractable work.

In particular, we simulate a work-extraction protocol via the time-dependent system Hamiltonian $H_t = (p_x^2/2m) + V_t$ where $p_x = -i\hbar\partial_x$ is the momentum operator and m is the mass of the single-particle system. The time-dependent potential-energy landscape is

$$V_t = 16 h_0 \left(\frac{x}{w_0} \right)^4 - 8 h_0 \left(\frac{x}{w_0} \right)^2 g_t - h_0 \frac{x}{w_0} f_t, \tag{75}$$

where g_t and f_t are non-negative scalar functions of time, with $g_0 = g_\tau = 1$ and $f_0 = f_\tau = 0$. To enter an interesting thermodynamic regime, we choose the initial energy barrier between the two wells to be $h_0 = 8 k_B T$, where $k_B T = 1/\beta$ is the thermal energy of the environment. To enter an interesting quantum regime, we choose the initial separation between the bottom of the two wells to be $w_0 = 3\lambda_{\mathrm{th}}$, where $\lambda_{\mathrm{th}} = \hbar\sqrt{2\pi\beta/m}$ is the thermal de Broglie wavelength.

As depicted in Fig. 5, g_t controls the height of the barrier while f_t controls the tilt of the energetic landscape throughout the protocol. For a semiconductor-based double quantum dot, the barrier height and tilt could plausibly be modulated by applied voltages, gating the barrier and the bias across the device, respectively [50]. Note that the protocol is cyclic, since the initial and final Hamiltonian are both the same symmetric double-well potential. During the protocol, the potential is tilted to the right, the barrier is lowered, the potential is untilted, and the barrier is then reintroduced, in that order [51]. The exact form of the control protocol is given in Appendix I.

The system is coupled to a thermal environment throughout the protocol. For our demonstration, we model the excitation and relaxation dynamics when the system interacts with the photon bath of an ideal black body at temperature T. The resulting dynamics of the system during the protocol can be described by a Lindblad master equation $\dot{\rho}_t = \mathcal{L}_t(\rho_t)$, given explicitly in Appendix I, which satisfies detailed balance in bath-mediated transitions between instantaneous energy eigenstates. Our finite-time protocol induces nonequilibrium quantum dynamics but we assume that the driving is sufficiently slow such that the assumptions that justify the Lindbladian description are justified.

We simulate a regime—$w_0 \approx 13$-nm interwell separation, and thermal relaxation rate of $\gamma = 366$ kHz at room temperature—that could plausibly be demonstrated in a

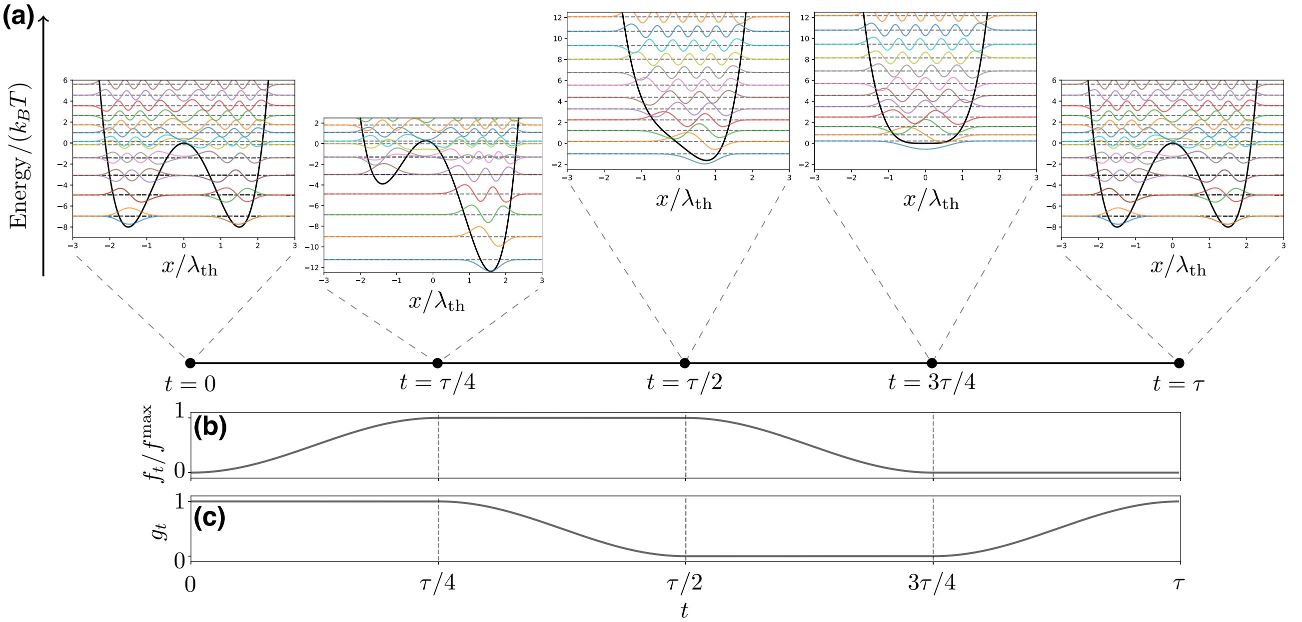


FIG. 5. The control protocol to extract work from the nonequilibrium state of a particle on the right-hand side of a double-well potential, via a time-varying one-dimensional potential-energy landscape in the quantum regime. (a) The potential-energy landscape at five representative times along the protocol, staggered so that they share the same energy axis. We show a spatial representation of the lowest-lying energy eigenstates in these potentials, at the height of their respective energy eigenvalues. (b),(c) The protocol for tilting the potential and lowering the energetic barrier, respectively.

room-temperature laboratory experiment using a nanofabricated device. The resulting dynamics exhibits a separation of time scales, where the relaxation dynamics are much slower than the time scale of coherent oscillations between energy eigenstates. (Dissipative transitions occur approximately only once per 100 million periods of phase oscillation between energy levels with a k_BT energy spacing.) We develop special methods to efficiently simulate the quantum dynamics (Appendix I 1) and thermodynamics (Appendix I 2) over the extended duration $\tau = 4 \times 10^9 \beta\hbar \approx 100$ μs of the nonequilibrium protocol.

Which initial state yields the most extracted work given this protocol? To determine this, we construct the expected-work operator $\mathcal{W}$ acting on the Hilbert space spanned by the eight lowest-lying initial energy eigenstates, via simulating the behavior of 64 random linearly independent initial states within this subspace. In our simulations, these initially restricted states evolve through the Hilbert space spanned by the 20 lowest-lying instantaneous energy eigenstates throughout the protocol.

These energy eigenstates, in turn, are determined numerically via an eigendecomposition of a discretized approximation of the instantaneous Hamiltonian represented in the spatial basis. In our numerical simulations, we partition the continuous spatial basis into 249 discrete spatial states of width $\delta x = 2\lambda_{\text{th}}/83$. In this discretization, the quantum state $|x, x+\delta x\rangle$ describes the system if it were localized in the spatial bin between location x and $x+\delta x$. In general, a pure quantum state (like one of the energy eigenstates or an eigenstate of the expected-work operator) is delocalized in a superposition of these spatial states.

In Fig. 6, we depict the eigenvalues and eigenstates of the expected-work operator, via their spatial probability density functions. Two eigenmodes of the expected-work operator allow for work extraction, while the other six

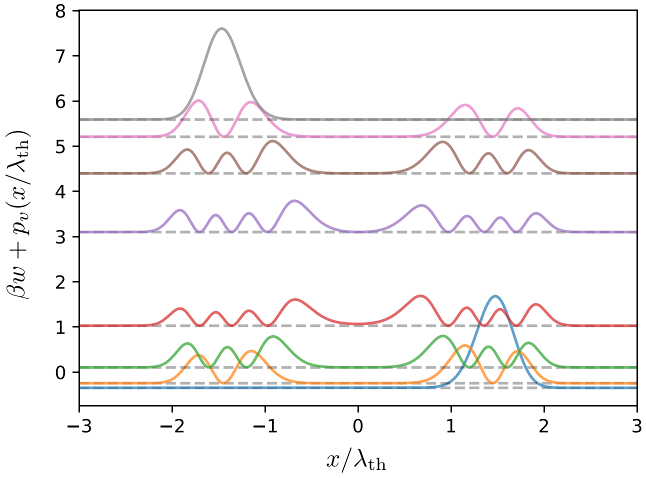


FIG. 6. The spatial probability density $p_v(x/\lambda_{\text{th}}) = \lim_{\delta x\to 0}(\lambda_{\text{th}}/\delta x)\,|\langle x, x+\delta x|v\rangle|^2$ for the eigenstates $|v\rangle$ of the expected-work operator $\mathcal{W}$, offset by the corresponding work eigenvalue w, such that $\mathcal{W}\,|v\rangle = w\,|v\rangle$.

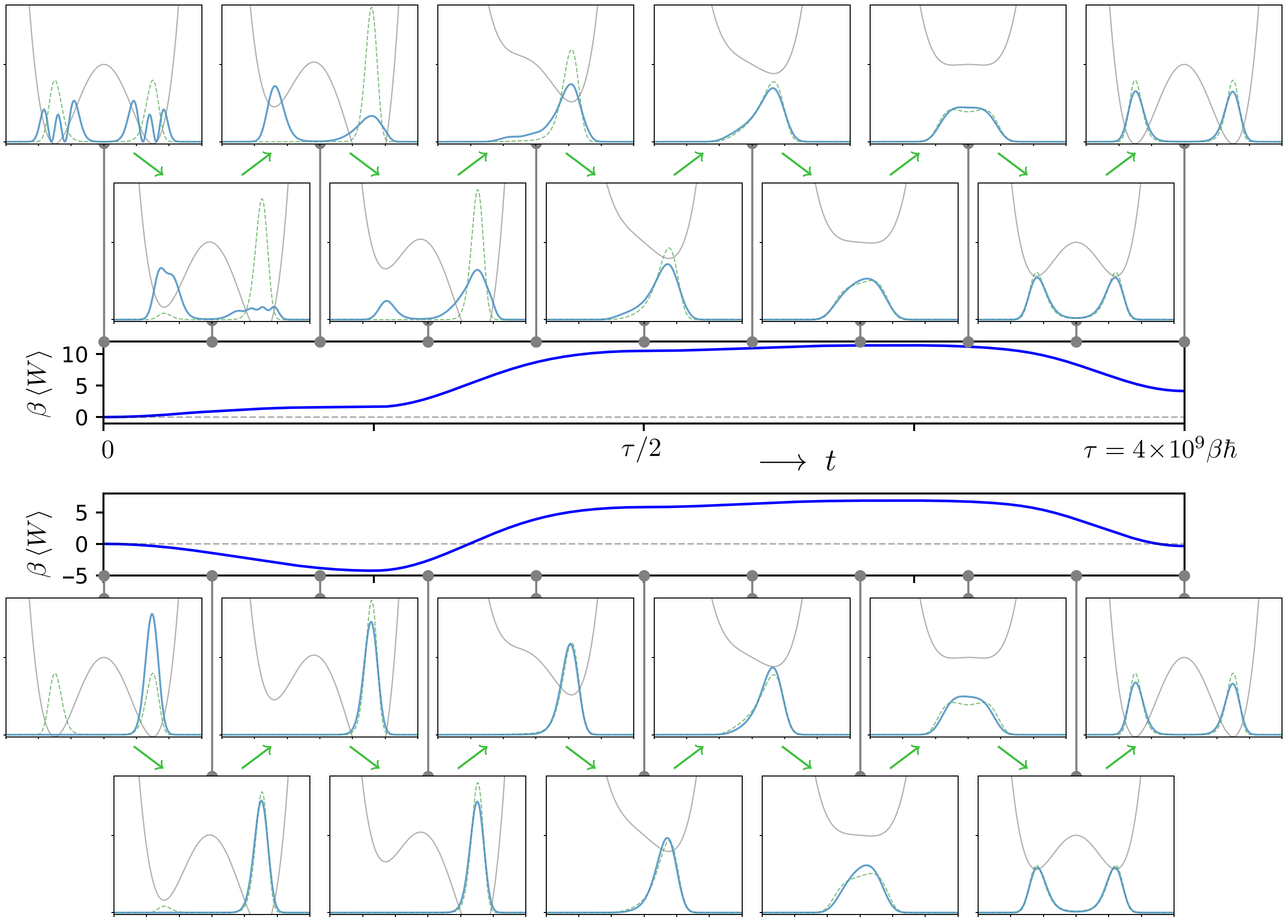


FIG. 7. Upper panels: snapshots of the spatial evolution of a random input to the protocol, along with the resulting trajectory of expected work. Bottom panels: snapshots of the spatial evolution of the ideal input to the work-extraction protocol, along with the resulting trajectory of expected work. The top and bottom panels show the spatial probability density of the state in thick solid blue, compared to the probability density for the instantaneous equilibrium state in thin dashed green; the instantaneous potential-energy landscape, which starts and ends as a symmetric double well, is shown in thin solid gray. The two long middle panels show the time evolution of the expected work from each of these two initial states.

modes require a work investment. As one might expect, the ideal initial state leading to maximal work extraction is initially localized in the right well. This state is approximately formed from equal parts of the two lowest-lying energy eigenstates. It is also intuitive that the worst input begins purely in the left well.

The remaining six nonextremal eigenstates of the expected-work operator are less immediately intuitive—yet they yield interesting insights once properly understood. As suggested by the spatial probability density functions in Fig. 6, each work eigenstate is approximately a linear combination of up to two energy eigenstates (although in detail, they contain contributions from all). Some investigation has revealed that each work eigenstate roughly corresponds to a relative phase combination of energy eigenstates that leaves the particle either mostly in the right well (for small eigenwork) or mostly in the left well (for large eigenwork) during the first half of the protocol. In Fig. 6, we display some kind of symmetry in the vertical spacing of work eigenvalues that could likely be related to fluctuation relations, although that investigation remains an open opportunity.

The ideal input to this particular work-extraction protocol yields approximately $0.3472\,k_BT$ of *extracted* work ($\langle W\rangle = -0.3472\,k_BT$), which is about half of the nonequilibrium addition to free energy (approximately $k_BT\ln 2$) available from that initial state. This work-extraction value has initially been found via the minimal eigenvalue of the expected-work operator and subsequently verified via direct simulation of this ideal input. The spatial evolution of this ideal state and the corresponding work trajectory are shown in the bottom half of Fig. 7. There, it is apparent that the ideal state remains on the right well for the first half of the work-extraction protocol.

For comparison, the top half of Fig. 7 shows the spatial evolution and corresponding work trajectory of a random input (in particular, we have chosen one of the initial energy eigenstates). No work is extracted when operating on this state; rather, about $4.138\,k_BT$ of work must be *expended* in the process. We have verified that this same work value is obtained (through ten significant digits) both from the full simulation and from the simple algebraic employment of the expected-work operator [52]: $\mathrm{tr}(\rho_0\mathcal{W}) \approx 4.138\,k_BT$. The series of panels in the top half of Fig. 7 reveal the work penalty for populating the left well during the first half of the protocol. The chronologically second snapshot reveals a slow local equilibration within each well. Subsequent snapshots reveal an even slower tendency toward global equilibration throughout the protocol.

As anticipated by the general theory, the initial state leading to maximal work extraction is a pure state. In this case, this state is close to, but distinct from, the input leading to minimal entropy production—the latter of which is a nonpure mixed state. As a final note, we emphasize that we have been able to identify the ideal input within the initial low-energy subspace, although the evolving states have not been limited to this subspace during their natural dynamics.

XIII. CONCLUSIONS

We have determined the ideal inputs that minimize or maximize various thermodynamic quantities for any fixed process that transforms a physical system in finite time. Many of these optimal inputs turn out to be pure states corresponding to eigenstates of Hermitian thermodynamic operators. We have shown how to reconstruct these operators via observed behavior from a finite number of experimentally accessible input states. Another class of thermodynamic quantities, based on entropies, have mixed-state minimizers but pure-state maximizers. The Hermitian thermodynamic operators determine these ideal states too. Our examples illustrate the incompatibility of common objectives: the “ideal” input depends on whether one intends to minimize heat, minimize entropy production, maximize free-energy gain, maximize work extraction, etc. As simple demonstrations, our examples have made simplifying assumptions of Markovianity and sometimes weak coupling; however, we emphasize that our results apply much more generally in non-Markovian scenarios with strong coupling, so long as the system and environment begin uncorrelated.

This investigation of ideal initial states complements the centuries-old tradition of rather seeking ideal protocols with an assumed initial state. Whether or not a protocol is ideal, our results highlight the initial-state dependence of the performance of a device across thermodynamic metrics and expose the breadth of its possible behavior. These results thus enable a full thermodynamic characterization of any classical or quantum device—including which inputs will extremize its thermodynamic behavior. Meanwhile, this set of results enables concrete answers to foundational questions in nonequilibrium thermodynamics, where sometimes conflicting claims of optimality principles have become folklore.

While we have emphasized thermodynamics, the results of this paper extend easily to other domains—in which the ideal inputs, as judged by some other criteria, such as maximizing the yield of a desired output state [53], will be obtained from the linear operators induced by those criteria.

ACKNOWLEDGMENTS

A.K. thanks Sosuke Ito for support. This work is supported by the Singapore Ministry of Education Tier 1 Grant No. RG146/20, by Grant No. FQXi-RFP-IPW-1903 (“Are quantum agents more energetically efficient at making predictions?”) from the Foundational Questions Institute, by the Fetzer Franklin Fund (a donor-advised fund of Silicon Valley Community Foundation), by the National Research Foundation, Singapore, by the Agency for Science, Technology and Research (A*STAR) under its CQT Bridging Grant and QEP2.0 program (NRF2021-QEP2-02-P06), and by the Singapore Ministry of Education Tier 2 Grant No. MOE-T2EP50221-0005. This project was partly made possible through the support of Grant No. 62417 from the John Templeton Foundation.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

APPENDIX A: EXISTENCE OF THERMODYNAMIC OPERATORS

Type-I expectation values—for expected work, heat, entropy flow, etc.—are linear functionals of the initial state, since they can all be written as

$$\langle X\rangle_{\rho_0} = \mathrm{tr}[\Xi_X(\rho_0)], \tag{A1}$$

where Ξ_X is a linear superoperator acting on the initial density matrix of the system. $\mathrm{tr}[\Xi_X(\cdot)]$ is a linear functional since Ξ_X and the trace operation are both linear and the trace returns a scalar value. The superoperator Ξ_X is, however, relatively unwieldy, since it acts on the d^2-dimensional vector space spanned by density matrices that in turn act on a d-dimensional vector space $\mathcal{V}$. The following theorem shows that the same linear functional can be represented via the lower-dimensional operator $\mathcal{X}$ acting on $\mathcal{V}$. This assures the existence of our thermodynamic operators.

Theorem 6. For any linear functional ℓ acting on a finite d^2-dimensional vector space $\mathcal{V}_{\mathrm{big}}$ of linear operators, which in turn operate on a d-dimensional vector space $\mathcal{V}_{\text{small}}$, there exists a linear operator $\mathcal{X}$ acting on $\mathcal{V}_{\text{small}}$ such that

$$\ell(\cdot) = \text{tr}(\mathcal{X}\,\cdot\,). \tag{A2}$$

Proof. There are only d^2 linearly independent linear functionals acting on the d^2-dimensional vector space $\mathcal{V}_{\text{big}}$. Let $K = \{|k\rangle\}_{k=1}^{d}$ be an arbitrary orthonormal basis for $\mathcal{V}_{\text{small}}$, with $\langle k| = |k\rangle^\dagger$. Note that the d^2 linearly independent linear functionals $\{\text{tr}(\,|j\rangle\,\langle k|\,\cdot\,)\}_{|j\rangle,|k\rangle\in K}$ form a complete basis for the dual space of $\mathcal{V}_{\text{big}}$. An arbitrary linear functional acting on $\mathcal{V}_{\text{big}}$ can thus be written as a linear combination of these basis functionals:

$$\ell(\cdot) = \sum_{j,k} x_{j,k}\text{tr}(\,|j\rangle\,\langle k|\,\cdot\,) = \text{tr}(\mathcal{X}\,\cdot\,) \tag{A3}$$

where $\mathcal{X} = \sum_{j,k} x_{j,k}\,|j\rangle\,\langle k|$ and $x_{j,k} = \ell(|k\rangle\,\langle j\,|)$. ■

Hence, for type-I expectation values:

$$\langle X\rangle_{\rho_0} = \text{tr}[\Xi_X(\rho_0)] = \text{tr}(\rho_0\mathcal{X}). \tag{A4}$$

Whenever the random variable X is real valued, its expectation value must also be real valued and so $\mathcal{X}$ is guaranteed to be Hermitian (see Ref. [16, Theorem 2.4.3]).

Sometimes, the linear superoperator Ξ_X can appear very complicated; but only its existence matters to guarantee the existence of the simpler operator $\mathcal{X}$. For example, entropy flow can generically be written as

$$\langle\Phi\rangle_{\rho_0} = -k_B\int_0^\tau \text{tr}\left(\dot{\rho}_t^{\text{env}}\ln\boldsymbol{\pi}_t^{\text{env}}\right)\,dt,$$

where $\boldsymbol{\pi}_t^{\text{env}}$ is assumed to be independent of ρ_0. If we write out $\dot{\rho}_t^{\text{env}}$ more explicitly, we find that

$$\begin{aligned}\langle\Phi\rangle_{\rho_0} &= -k_B\int_0^\tau \text{tr}\left(\dot{\rho}_t^{\text{env}}\ln\boldsymbol{\pi}_t^{\text{env}}\right)\,dt\\ &= -k_B\int_0^\tau \text{tr}\left[\left(\frac{d}{dt}\rho_t^{\text{env}}\right)\ln\boldsymbol{\pi}_t^{\text{env}}\right]\,dt \qquad\qquad (\text{A5})\\ &= -k_B\int_0^\tau \text{tr}\left\{\left[\frac{d}{dt}\text{tr}_{\text{sys}}(U_t\rho_0\otimes\rho_0^{\text{env}}U_t^\dagger)\right]\ln\boldsymbol{\pi}_t^{\text{env}}\right\}\,dt \qquad (\text{A6})\\ &= \text{tr}\left\{-k_B\int_0^\tau\left[\frac{d}{dt}\text{tr}_{\text{sys}}(U_t\rho_0\otimes\rho_0^{\text{env}}U_t^\dagger)\right]\ln\boldsymbol{\pi}_t^{\text{env}}dt\right\}. \qquad (\text{A7})\end{aligned}$$

Hence, we have found the relevant linear superoperator $\Xi_\Phi(\cdot) = -k_B\int_0^\tau\left[(d/dt)\text{tr}_{\text{sys}}(U_t\,\cdot\,\otimes\rho_0^{\text{env}}U_t^\dagger)\right]\ln\boldsymbol{\pi}_t^{\text{env}}\,dt$, which is somewhat complicated. No matter: its mere existence proves the existence of some simpler linear operator $\boldsymbol{\Phi}$ such that $\langle\Phi\rangle_{\rho_0} = \text{tr}(\rho_0\boldsymbol{\Phi})$ for any initial state ρ_0.

A similar procedure reveals the existence of the relevant linear superoperator $\Xi(\cdot)$ for work and heat, which in turn implies the existence of the simpler expected-work operator $\mathcal{W}$ and expected-heat operator $\mathcal{Q}$.

Expanding the expression for expected work, we find

$$\begin{aligned}\langle W\rangle_{\rho_0} &= \int_0^\tau \text{tr}\left(\rho_t\dot{H}_t\right)\,dt\\ &= \text{tr}\left[\int_0^\tau \text{tr}_{\text{env}}(U_t\rho_0\otimes\rho_0^{\text{env}}U_t^\dagger)\dot{H}_t\,dt\right].\end{aligned} \tag{A8}$$

Hence, the linear superoperator $\Xi_W(\cdot) = \int_0^\tau \text{tr}_{\text{env}}(U_t\,\cdot\,\otimes\rho_0^{\text{env}}U_t^\dagger)\dot{H}_t\,dt$ implies the existence of the simpler expected-work operator $\mathcal{W}$.

Expanding the expression for expected heat, we find

$$\begin{aligned}\langle Q\rangle_{\rho_0} &= \int_0^\tau \text{tr}\left(\dot{\rho}_tH_t\right)\,dt\\ &= \text{tr}\left\{\int_0^\tau\left[\frac{d}{dt}\text{tr}_{\text{env}}(U_t\rho_0\otimes\rho_0^{\text{env}}U_t^\dagger)\right]H_t\,dt\right\}.\end{aligned} \tag{A9}$$

Hence, the linear superoperator $\Xi_Q(\cdot) = \int_0^\tau\left[(d/dt)\text{tr}_{\text{env}}(U_t\,\cdot\otimes\rho_0^{\text{env}}U_t^\dagger)\right]H_t\,dt$ implies the existence of the simpler expected-heat operator $\mathcal{Q}$.

An alternative definition of work in the strong-coupling regime is $W^{(*)}$: the net change in energy of the system-bath supersystem during a protocol of duration τ [12]. Let H_t^{tot} be the total Hamiltonian including the system, the environment, and the coupling contributions; over time, it induces the net unitary evolution operator U_τ. The expectation value for this version of work is thus

$$\begin{aligned}\langle W^{(*)}\rangle_{\rho_0} &= \text{tr}(\rho_\tau^{\text{tot}}H_\tau^{\text{tot}}) - \text{tr}(\rho_0^{\text{tot}}H_0^{\text{tot}})\\ &= \text{tr}\left[U_\tau(\rho_0\otimes\rho_0^{\text{env}})U_\tau^\dagger H_\tau^{\text{tot}}\right] - \text{tr}\left[(\rho_0\otimes\rho_0^{\text{env}})H_0^{\text{tot}}\right] \qquad (\text{A10})\\ &= \text{tr}\left[(\rho_0\otimes\rho_0^{\text{env}})(U_\tau^\dagger H_\tau^{\text{tot}}U_\tau - H_0^{\text{tot}})\right], \qquad (\text{A11})\end{aligned}$$

which is a linear functional of the initial state of the system. By our Theorem 6, this thus implies that there exists a thermodynamic operator $\mathcal{W}^{(*)}$ such that

$$\langle W^{(*)}\rangle_{\rho_0} = \text{tr}(\rho_0\mathcal{W}^{(*)}).$$

We note that this is a type-I expectation value on the system subspace alone, even though the d^2 physical measurements (of work required to infer this thermodynamic operator) are not restricted to act on the d-dimensional system subspace.

APPENDIX B: THERMODYNAMIC OPERATORS CHARACTERIZE ANY MEASUREMENT SCHEME

Invasive measurements change not only the probability distribution for work but also the expected value for work under the prescribed measurement protocol [21]. Our framework applies just as well to these alternative protocols with measurement interventions. For example, let us briefly describe how our framework applies to the famous two-point-measurement (TPM) scheme for driven isolated quantum systems, which projects the state onto the instantaneous energy eigenbasis at both the beginning and end of the protocol [22].

If we are interested in the TPM scheme for a driven isolated quantum system, the relevant *expected-TPM-work operator* is

$$\mathcal{W}_{\rm TPM} = \sum_{|E\rangle \in V_H} |E\rangle\langle E| \sum_{|E'\rangle \in V_{H'}} (E' - E) \left|\langle E'|U_\tau|E\rangle\right|^2, \tag{B1}$$

where H and H' are the initial and final Hamiltonians for the system, respectively, and U_τ unitarily evolves the system between the two projective measurements. In the TPM scheme, the system transitions from initial energy eigenstate $|E\rangle$ to final energy eigenstate $|E'\rangle$ with probability $\langle E|\rho_0|E\rangle \left|\langle E'|U_\tau|E\rangle\right|^2$, resulting in a TPM work value of $E' - E$. It is easy to verify that

$$\langle W_{\rm TPM}\rangle_{\rho_0} = \sum_w w \Pr(W_{\rm TPM} = w) = \mathrm{tr}(\rho_0 \mathcal{W}_{\rm TPM}), \tag{B2}$$

where $W_{\rm TPM}$ is the random variable for the TPM work outcome and the sum runs over all $w \in \{E' - E : E \in \Lambda_H, E' \in \Lambda_{H'}\}$.

Reference [21] shows that $\mathcal{W} \neq \mathcal{W}_{\rm TPM}$. Invasive measurements can change the expectation value for work (and, indeed, for other thermodynamic variables too).

If desired, our framework can be employed to find the ideal inputs leading to minimal or maximal TPM work, from the minimal and maximal eigenstates of $\mathcal{W}_{\rm TPM}$. Indeed, for any measurement scheme—including also the one-time-measurement scheme [23–25] or any other—thermodynamic operators can be constructed and our framework can be applied to identify both the breadth of behavior and the ideal inputs within the scheme.

APPENDIX C: EXPECTATION-VALUE INVARIANCE TO SUBENSEMBLE DECOMPOSITION

Our contributions in the main text exclusively involve expectation values and so are invariant to density-matrix decompositions. Here, we show this invariance explicitly. As an added benefit, it can be useful to see how the quantum random variables induced by a subensemble decomposition relate to and generalize their classical counterparts.

In the classical case, fluctuations are primarily due to subjective uncertainty about which state the environment is in (although uncertainty of the system state also plays a role). Since quantum results should include classical results as a special case, we must also consider the role of subjective uncertainty in the quantum case. In general, this can be addressed via a probability distribution over (possibly nonorthogonal) subensembles of the joint system-environment supersystem [29].

There are many ways to decompose the initial joint density matrix of the system-environment supersystem. In fact, for a nonpure mixed state $\rho_0^{\rm tot}$, there are infinitely many pairs $((p_n)_n, (\sigma_n)_n)$ of probability distributions $(p_n)_n$ over constituent density matrices $(\sigma_n)_n$, for which $\rho_0^{\rm tot} = \sum_n \sigma_n p_n$. Each of these decompositions can represent a physically relevant preparation of the supersystem [29,54].

Each subensemble decomposition $((p_n)_n, (\sigma_n)_n)$ induces a random variable $X((p_n)_n, (\sigma_n)_n)$ [55]. With probability p_n, the subensemble σ_n will be realized, in which case the random variable X takes on the value $x(\sigma_n)$. When the thermodynamic quantity is a linear functional of the initial state—as is the case for work, heat, entropy flow, etc.—it is easy to see that

$$\begin{aligned}\langle X((p_n)_n, (\sigma_n)_n)\rangle_n &= \sum_n x(\sigma_n) p_n = x\left(\sum_n \sigma_n p_n\right) \\ &= x(\rho_0^{\rm tot}),\end{aligned} \tag{C1}$$

independent of the decomposition. In the main text, we always consider $\rho_0^{\rm tot} = \rho_0 \otimes \rho_0^{\rm env}$ for a fixed initial state of the environment $\rho_0^{\rm env}$. For these thermodynamic linear functionals, we thus use the shorthand notation

$$\langle X\rangle_{\rho_0} = x(\rho_0 \otimes \rho_0^{\rm env}). \tag{C2}$$

As a consequence, the expectation value $\langle X\rangle_{\rho_0}$ and thermodynamic operator $\mathcal{X}$ are both invariant to subensemble decompositions.

Let us take thermodynamic work (without measurement interventions) as an example. From Eq. (A8), we see that work can be quantified for any initial joint state σ of the supersystem via

$$w(\sigma) = \mathrm{tr}\left[\int_0^\tau \mathrm{tr}_{\rm env}(U_t \sigma U_t^\dagger) \dot{H}_t \, dt\right]. \tag{C3}$$

If the system and environment are unmeasured, then it is natural for eigenvalues of the uncorrelated density matrix to represent subjective uncertainty. We can thus take $((p_n)_n, |n\rangle\langle n|)$ as our subensemble decomposition, where

$|n\rangle$ is an eigenstate of ρ_0^{tot}. In this case, work will be a random variable W which takes on the value $w^{(n)} = w\left(|n\rangle\langle n|\right)$ with probability $\sum_m p_m \delta_{w^{(m)},w^{(n)}}$. We can then write

$$\langle W\rangle_{\rho_0} = \sum_n w^{(n)} p_n = w(\rho_0 \otimes \rho_0^{\mathrm{env}}) = \mathrm{tr}(\rho_0 \mathcal{W}). \quad \text{(C4)}$$

We see that the expectation value and the thermodynamic operator are both independent of the chosen subensemble decomposition.

APPENDIX D: OPERATOR EXPRESSIONS FOR TYPE-II EXPECTATION VALUES

The expectation values of entropy production, the reduction in nonequilibrium free energy, and the change in von Neumann entropy can all be written as

$$\begin{aligned} f_{\rho_0}^{(\mathcal{X})} &= \mathrm{tr}(\rho_0 \mathcal{X}) + S(\rho_\tau) - S(\rho_0) && \text{(D1)}\\ &= \mathrm{tr}(\rho_0 \ln \rho_0) - \mathrm{tr}(\rho_0 \ln e^{-\mathcal{X}}) + S(\rho_\tau) && \text{(D2)}\\ &= \mathrm{tr}(\rho_0 \ln \rho_0) - \mathrm{tr}\{\rho_0 \ln[e^{-\mathcal{X}}/\mathrm{tr}(e^{-\mathcal{X}})]\}\\ &\quad - \ln[\mathrm{tr}(e^{-\mathcal{X}})] + S(\rho_\tau) && \text{(D3)}\\ &= D[\rho_0 \| \omega] - \ln[\mathrm{tr}(e^{-\mathcal{X}})] + S(\rho_\tau), && \text{(D4)} \end{aligned}$$

where $\omega := e^{-\mathcal{X}}/\mathrm{tr}(e^{-\mathcal{X}})$.

For reliable reset processes, for which $\rho_\tau \approx r_\tau$ is very nearly independent of the input, it is clear that ω minimizes $f_{\rho_0}^{(\mathcal{X})}$, where it takes on the minimal value $f_\omega^{(\mathcal{X})} = S(r_\tau) - \ln[\mathrm{tr}(e^{-\mathcal{X}})]$.

In Ref. [5], it has been pointed out that entropy production can be written in this form. Here, we note that this type of relation also applies to a larger family of thermodynamic quantities.

APPENDIX E: GENERALIZED GELL-MANN MATRICES AS A STANDARD OPERATOR BASIS

For concreteness, we will describe an operator basis that can be used in any finite dimension d.

Let $\{|1\rangle, \ldots, |d\rangle\}$ be an orthonormal basis for the d-dimensional Hilbert space of our system under study. We can then construct the generalized Gell-Mann matrices to complete a matrix basis for density matrices acting on this Hilbert space.

Let $n \in \mathbb{R}$ be any convenient constant. $\vec{\Gamma}$ will contain $d-1$ diagonal matrices

$$n\left(\frac{2}{\ell^2+\ell}\right)^{1/2}\left[\left(\sum_{j=1}^{\ell} |j\rangle\langle j|\right) - \ell\,|\ell+1\rangle\langle \ell+1|\right], \quad \text{(E1)}$$

with $\ell \in \{1, \ldots, d-1\}$. It will contain $(d^2-d)/2$ distinct nondiagonal symmetric matrices

$$n\left(|k\rangle\langle j| + |j\rangle\langle k|\right), \quad \text{(E2)}$$

with $1 \leq j < k \leq d$. It will also contain $(d^2-d)/2$ distinct antisymmetric matrices

$$in\left(|k\rangle\langle j| - |j\rangle\langle k|\right), \quad \text{(E3)}$$

with $1 \leq j < k \leq d$. The ordering of these d^2-1 matrices is arbitrary.

Note that these operators are traceless and satisfy $\mathrm{tr}(\Gamma_m \Gamma_n) = 2n^2 \delta_{m,n}$. Hence, $\eta = 2n^2$. If we choose $n = \sqrt{(d-1)/2d}$, then $\eta = \frac{d-1}{d}$, pure states will always have a generalized Bloch vector of unit length [56] and the generalized Bloch vector will reduce to the standard Bloch vector in $d=2$.

APPENDIX F: COMPOSITE OPERATOR BASES

Suppose that we have normalized Hermitian operator bases, $(I_d/d, \vec{\Gamma})$ for a d-dimensional subsystem and $(I_{d'}/d', \vec{\Gamma}')$ for a d'-dimensional subsystem. The operator bases satisfy

$$\mathrm{tr}(\Gamma_n) = 0 \quad \text{and} \quad \mathrm{tr}(\Gamma_m \Gamma_n) = \eta \delta_{m,n} \quad \text{(F1)}$$

and

$$\mathrm{tr}(\Gamma'_n) = 0 \quad \text{and} \quad \mathrm{tr}(\Gamma'_m \Gamma'_n) = \eta' \delta_{m,n}. \quad \text{(F2)}$$

We can then construct a normalized Hermitian basis for operators acting on the composite d''-dimensional Hilbert space, where $d'' = dd'$: The new operator basis $(I_{d''}/d'', \vec{\Gamma}'')$ satisfies

$$\mathrm{tr}(\Gamma''_n) = 0 \quad \text{and} \quad \mathrm{tr}(\Gamma''_m \Gamma''_n) = \eta'' \delta_{m,n}, \quad \text{(F3)}$$

where the new composite operator basis contains

$$\vec{\Gamma}'' = \left(\sqrt{\frac{\eta''}{\eta' d}} I_d \otimes \vec{\Gamma}', \sqrt{\frac{\eta''}{\eta d'}} \vec{\Gamma} \otimes I_{d'}, \sqrt{\frac{\eta'}{\eta \eta'}} \vec{\Gamma} \otimes \vec{\Gamma}'\right). \quad \text{(F4)}$$

Equation (F3) can be verified via the identities $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ and $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B)$.

This allows us to, e.g., build up a normalized Hermitian operator basis for:

(a) many qubits, using tensor products of the local Pauli operators, or
(b) a qubit and qutrit, using tensor products of their local Pauli and Gell-Mann operators
(c) etc.

APPENDIX G: EXISTENCE OF, AND EXPRESSION FOR, THE ENTROPY-FLOW VECTOR

Leveraging the decomposition of the system state in terms of its generalized Bloch vector, $\rho_t = I/d + \vec{b}_t \cdot \vec{\Gamma}$, we can expand the general expression for entropy flow given in Ref. [4] and discussed in the main text to express entropy flow in terms of the entropy-flow vector and initial Bloch vector:

$$
\begin{aligned}
\langle \Phi \rangle_{\rho_0} &= -k_B \int_0^\tau \mathrm{tr}\left(\dot{\rho}_t^{\mathrm{env}} \ln \boldsymbol{\pi}_t^{\mathrm{env}}\right) dt && \text{(G1)}\\
&= -k_B \int_0^\tau \mathrm{tr}\left\{\left[\frac{d}{dt}\mathrm{tr}_{\mathrm{sys}}\left(U_{0:t}\rho_0 \otimes \rho_0^{\mathrm{env}} U_{0:t}^\dagger\right)\right] \ln \boldsymbol{\pi}_t^{\mathrm{env}}\right\} dt && \text{(G2)}\\
&= -k_B \int_0^\tau \mathrm{tr}\left\{\left[\frac{d}{dt}\mathrm{tr}_{\mathrm{sys}}\left(U_{0:t}(I/d + \vec{b}_0 \cdot \vec{\Gamma}) \otimes \rho_0^{\mathrm{env}} U_{0:t}^\dagger\right)\right] \ln \boldsymbol{\pi}_t^{\mathrm{env}}\right\} dt && \text{(G3)}\\
&= \underbrace{-k_B \int_0^\tau \mathrm{tr}\left\{\left[\frac{d}{dt}\mathrm{tr}_{\mathrm{sys}}\left(U_{0:t} I/d \otimes \rho_0^{\mathrm{env}} U_{0:t}^\dagger\right)\right] \ln \boldsymbol{\pi}_t^{\mathrm{env}}\right\} dt}_{=\langle \Phi \rangle_{I/d}} + \vec{b}_0 \\
&\quad \times \left(\underbrace{-k_B \int_0^\tau \mathrm{tr}\left\{\left[\frac{d}{dt}\mathrm{tr}_{\mathrm{sys}}\left(U_{0:t}\vec{\Gamma} \otimes \rho_0^{\mathrm{env}} U_{0:t}^\dagger\right)\right] \ln \boldsymbol{\pi}_t^{\mathrm{env}}\right\} dt}_{=:\vec{\varphi}}\right) && \text{(G4)}\\
&= \langle \Phi \rangle_{I/d} \\
&\quad + \vec{b}_0 \cdot \vec{\varphi}. && \text{(G5)}
\end{aligned}
$$

Other thermodynamic quantities can be decomposed similarly.

APPENDIX H: EXPRESSION FOR THE SECOND DERIVATIVE OF VON NEUMANN ENTROPY

The second derivative of the von Neumann entropy, with respect to the elements of the Bloch vector of the quantum state, can be calculated via the spectral decomposition of the quantum state $\tilde{\rho}(\vec{b}) = \sum_k \lambda_k |k\rangle\langle k|$ and its logarithm $\ln \tilde{\rho}(\vec{b}) = \sum_k \ln(\lambda_k) |k\rangle\langle k|$. Assuming nondegeneracy of the eigenvalues, we find

$$
\begin{aligned}
\partial_{b_m}\partial_{b_n} S[\tilde{\rho}(\vec{b})] &= -\partial_{b_m} \mathrm{tr}\left[\Gamma_n \ln \tilde{\rho}(\vec{b})\right] && \text{(H1)}\\
&= -\mathrm{tr}\left[\Gamma_n\, \partial_{b_m} \ln \tilde{\rho}(\vec{b})\right], && \text{(H2)}
\end{aligned}
$$

where

$$
\partial_{b_m} \ln \tilde{\rho}(\vec{b}) = \sum_k \left\{ \frac{1}{\lambda_k} \underbrace{(\partial_{b_m}\lambda_k)}_{\langle k|\Gamma_m|k\rangle} |k\rangle\langle k| + \ln(\lambda_k) \left[\underbrace{\left(\partial_{b_m}|k\rangle\right)}_{\sum_{\ell\neq k} \frac{\langle \ell|\Gamma_m|k\rangle}{\lambda_k - \lambda_\ell}|\ell\rangle} \langle k| + |k\rangle \underbrace{\left(\partial_{b_m}\langle k|\right)}_{\sum_{\ell\neq k} \frac{\langle k|\Gamma_m|\ell\rangle}{\lambda_k - \lambda_\ell}\langle \ell|} \right] \right\}. \tag{H3}
$$

Hence,

$$
\partial_{b_m}\partial_{b_n} S[\tilde{\rho}(\vec{b})] = \left(\sum_k \frac{1}{\lambda_k} \langle k|\Gamma_n|k\rangle \langle k|\Gamma_m|k\rangle \right) + \left(\sum_{\substack{k,\ell \\ \ell\neq k}} \frac{\ln \lambda_k}{\lambda_k - \lambda_\ell} \langle k|\Gamma_n|\ell\rangle \langle \ell|\Gamma_m|k\rangle \right) + \left(\sum_{\substack{k,\ell \\ \ell\neq k}} \frac{\ln \lambda_k}{\lambda_k - \lambda_\ell} \langle \ell|\Gamma_n|k\rangle \langle k|\Gamma_m|\ell\rangle \right). \tag{H4}
$$

Upon swapping the names of the indices in the last of the three terms on the right-hand side of Eq. (H4), we find

$$\sum_{\substack{k,\ell\\ \ell\neq k}} \frac{\ln \lambda_k}{\lambda_k - \lambda_\ell} \langle \ell|\Gamma_n|k\rangle \langle k|\Gamma_m|\ell\rangle = \sum_{\substack{\ell,k\\ k\neq \ell}} \frac{\ln \lambda_\ell}{\lambda_\ell - \lambda_k} \langle k|\Gamma_n|\ell\rangle \langle \ell|\Gamma_m|k\rangle \tag{H5}$$

$$= \sum_{\substack{k,\ell\\ \ell\neq k}} \frac{-\ln \lambda_\ell}{\lambda_k - \lambda_\ell} \langle k|\Gamma_n|\ell\rangle \langle \ell|\Gamma_m|k\rangle . \tag{H6}$$

Noting that $\lim_{b\to a} (\ln a - \ln b)/(a-b) = 1/a$, we can thus combine all three terms on the right-hand side of Eq. (H4) to find

$$\partial_{b_m}\partial_{b_n} S[\tilde{\rho}(\vec{b})] = -\sum_{k,\ell} \phi(\lambda_k, \lambda_\ell) \langle k|\Gamma_n|\ell\rangle \langle \ell|\Gamma_m|k\rangle , \tag{H7}$$

where we have defined $\phi(a,b) := (\ln a - \ln b)/(a-b)$ (with $\phi(a,a) = 1/a$ by continuity). Note that $\phi(a,b)$ is the reciprocal of the logarithmic mean of the eigenvalues a and b.

Similarly, using the spectral decomposition of $\tilde{\rho}'(\vec{b}) = \sum_k \lambda'_k |k'\rangle\langle k'|$, we find

$$\partial_{b_m}\partial_{b_n} S[\tilde{\rho}'(\vec{b})] = -\sum_{k,\ell} \phi(\lambda'_k, \lambda'_\ell) \langle k'|\Gamma'_n|\ell'\rangle \langle \ell'|\Gamma'_m|k'\rangle . \tag{H8}$$

APPENDIX I: TIME-DEPENDENT DOUBLE-WELL DYNAMICS

Here, we construct a simple model for the dynamics of a nonrelativistic charged particle in a time-dependent double well of potential energy across one spatial dimension while it is immersed in a bosonic bath at temperature T. The protocol is capable of extracting work from some states that are initially localized in the right well.

In particular, we simulate a work-extraction protocol via the time-dependent system Hamiltonian:

$$H_t = \frac{p_x^2}{2m_q} + V_t, \tag{I1}$$

where $p_x = -i\hbar\partial_x$ is the momentum operator and m_q is the mass of the system with charge q. The time-dependent potential-energy landscape is

$$V_t = 16h_0\left(\frac{x}{w_0}\right)^4 - 8h_0\left(\frac{x}{w_0}\right)^2 g_t - h_0\frac{x}{w_0}f_t, \tag{I2}$$

where g_t and f_t are non-negative scalar functions of time, with $g_0 = g_\tau = 1$ and $f_0 = f_\tau = 0$. To enter an interesting thermodynamic regime, we choose the initial energy barrier between the two wells to be $h_0 = 8k_BT$, where $k_BT = 1/\beta$ is the thermal energy of the environment. To enter an interesting quantum regime, we choose the initial separation between the bottom of the two wells to be $w_0 = 3\lambda_{\text{th}}$, where $\lambda_{\text{th}} = \hbar\sqrt{2\pi\beta/m_q}$ is the thermal de Broglie wavelength. To induce nonequilibrium quantum dynamics, we choose a sufficiently fast protocol.

g_t controls the height of the barrier while f_t controls the tilt of the energetic landscape throughout the protocol. During the protocol, the potential is tilted to the right, the barrier is lowered, the potential is untilted, and the barrier is then reintroduced, in that order. The exact form of the control protocol is given by

$$g_t = \begin{cases} 1, & \text{if } 0 \le t < \tau/4, \\ \sin^2(2\pi t/\tau), & \text{if } \tau/4 \le t < \tau/2, \\ 0, & \text{if } \tau/2 \le t < 3\tau/4, \\ \cos^2(2\pi t/\tau), & \text{if } 3\tau/4 \le t \le \tau, \end{cases} \tag{I3}$$

and

$$f_t/f^{\max} = \begin{cases} \sin^2(2\pi t/\tau), & \text{if } 0 \le t < \tau/4, \\ 1, & \text{if } \tau/4 \le t < \tau/2, \\ \cos^2(2\pi t/\tau), & \text{if } \tau/2 \le t < 3\tau/4, \\ 0, & \text{if } 3\tau/4 \le t \le \tau. \end{cases} \tag{I4}$$

We have chosen the maximal tilt to be $f^{\max} = 8/(3\sqrt{2\pi})$.

The system Hamiltonian $H_t = p_x^2/2m_q + V_t$ has a countably infinite orthonormal set of instantaneous energy eigenstates $\{|E_t^{(n)}\rangle\}_n$ with corresponding instantaneous energy eigenvalues $\{E_t^{(n)}\}_n$, ordered such that $E_t^{(n)} \ge E_t^{(m)}$ if $n > m$. Hence, $H_t|E_t^{(n)}\rangle = E_t^{(n)}|E_t^{(n)}\rangle$ and the Hamiltonian has the simple eigenrepresentation $H_t = \sum_{n=1}^\infty E_t^{(n)}|E_t^{(n)}\rangle\langle E_t^{(n)}|$. It is useful to represent differences among eigenenergies via the relevant angular frequencies $\omega_t^{(n,m)}$, such that $\hbar\omega_t^{(n,m)} = E_t^{(n)} - E_t^{(m)}$.

The system is immersed in a photon bath throughout the protocol. Photons in the environment can induce transitions among the instantaneous energy eigenstates of the system, with absorption and stimulated-emission transition rates proportional to the number of photons with the relevant transition energy $E_t^{(n)} - E_t^{(m)}$. In general, this expected number of bath photons, per mode with frequency $\omega/2\pi$, can be a function of time $\langle N^{(\omega)}\rangle$, which is related to the intensity.

Suppose for now that the energy eigenstates of the system are nondegenerate [57]. Let $L_t^{(m,n)} = |E_t^{(m)}\rangle\langle E_t^{(n)}|$. If $E_t^{(n)} > E_t^{(m)}$, this can be interpreted as the lowering operator between these two instantaneous energy eigenstates. It is useful to reflect on the physical implications of the mathematical dissipator $\mathcal{D}[L](\rho) = L\rho L^\dagger - \frac{1}{2}\{L^\dagger L, \rho\}$. Note that the dissipator

$$\mathcal{D}[L_t^{(m,n)}](\rho_t) = |E_t^{(m)}\rangle\langle E_t^{(n)}|\,\rho_t\,|E_t^{(n)}\rangle\langle E_t^{(m)}| - \frac{1}{2}|E_t^{(n)}\rangle\langle E_t^{(n)}|\,\rho_t - \frac{1}{2}\rho_t\,|E_t^{(n)}\rangle\langle E_t^{(n)}| \tag{I5}$$

fully removes the $|E_t^{(n)}\rangle$ population and shifts it down to $|E_t^{(m)}\rangle$. Analogously, $\mathcal{D}[(L_t^{(m,n)})^\dagger](\rho_t) = \mathcal{D}[L_t^{(n,m)}](\rho_t)$ fully removes the $|E_t^{(m)}\rangle$ population and shifts it up to $|E_t^{(n)}\rangle$. The *rates* at which these two processes happen can be denoted by $r_t^{(n\to m)}$ and $r_t^{(m\to n)}$, respectively. This can then be integrated into a quantum master equation that takes on the Lindblad form:

$$\dot{\rho}_t = \mathcal{L}_t(\rho_t) = \frac{i}{\hbar}[\rho_t, H_t] + \sum_n \sum_{m<n} r_t^{(n\to m)}\mathcal{D}[L_t^{(m,n)}](\rho_t) + r_t^{(m\to n)}\mathcal{D}[L_t^{(n,m)}](\rho_t). \tag{I6}$$

As already mentioned, the rate of excitation $r_t^{(m\to n)}$ will be proportional to the expected number of bath particles $\langle N^{(\omega)}\rangle$ with the relevant energy $\hbar\omega_t^{(n,m)}$. The rate of emission is, however, more subtle, since it involves both stimulated and spontaneous emission.

A standard quantum electrodynamic calculation (introduced a century ago by Dirac [58]) relates absorption and emission between any two energy eigenstates of a system, given the photon intensity of the environment. In particular, following Fermi's golden rule for transitions on the joint state space of the system and the photons [59, Sec. 1.3], for $E_t^{(m)} < E_t^{(n)}$, the rate of absorption is given by

$$r_t^{(m\to n)} = \frac{2\pi}{\hbar}|\mathcal{M}_t^{(m,n)}|^2\,\langle N^{(\omega_t^{(n,m)})}\rangle\,, \tag{I7}$$

while the net rate of both stimulated and spontaneous emission is given by

$$r_t^{(n\to m)} = \frac{2\pi}{\hbar}|\mathcal{M}_t^{(m,n)}|^2\left(\langle N^{(\omega_t^{(n,m)})}\rangle + 1\right). \tag{I8}$$

Here, $\mathcal{M}_t^{(m,n)}$ is the transition amplitude between $|E_t^{(m)}\rangle$ and $|E_t^{(n)}\rangle$ induced by the background radiation. We use the standard interaction Hamiltonian $H_I = (-q/m_q c)A_x \otimes p_x$ on the joint state space of photons and system, where c is the speed of light and A_x is the x component of polarization of the quantum field for the electromagnetic vector potential $\vec{A}(\vec{r}, t)$. For simplicity, we assume that the system is some distance z from the photon source, so that the wave vector of radiation is $\vec{k} = k\hat{z} = (\omega/c)\hat{z} \perp \hat{x}$. In the standard electric dipole approximation, the absorption cross section—i.e., (power absorbed by the $m \to n$ transition)/(incident power per area)—is given by $\sigma_{\text{abs}} = 4\pi^2(q/\text{e})^2\alpha\omega\left|\langle E_t^{(m)}|x|E_t^{(n)}\rangle\right|^2\delta(\omega - \omega_t^{(n,m)})$, where e is the charge of a single electron, and $\alpha \approx 1/137$ is the fine-structure constant [60]. Hence the transition rate of absorption is related to the spectral intensity (power per area per angular frequency) $\mathcal{I}(\omega)$ via

$$\begin{aligned} r_t^{(m\to n)} &= \int_\omega \frac{\mathcal{I}(\omega)\sigma_{\text{abs}}}{\hbar\omega}\,d\omega & \text{(I9)}\\ &= \frac{4\pi^2\alpha}{\hbar}\left(\frac{q}{\text{e}}\right)^2\left|\langle E_t^{(m)}|x|E_t^{(n)}\rangle\right|^2\mathcal{I}(\omega_t^{(n,m)}). & \text{(I10)}\end{aligned}$$

Arbitrarily far from equilibrium, the intensity and the expected occupation are related by $\mathcal{I}(\omega) = \hbar\omega^3/\pi^2c^2\,\langle N^{(\omega)}\rangle$. Incorporating this, we find

$$\begin{aligned} r_t^{(m\to n)} &= \frac{4\alpha}{c^2}(q/\text{e})^2(\omega_t^{(n,m)})^3\left|\langle E_t^{(m)}|x|E_t^{(n)}\rangle\right|^2\langle N^{(\omega_t^{(n,m)})}\rangle & \text{(I11)}\\ &= \frac{k_BT}{\hbar}8\pi\alpha(q/\text{e})^2\frac{k_BT}{m_qc^2}\left(\frac{\hbar\omega_t^{(n,m)}}{k_BT}\right)^3 \\ &\quad\times\left|\langle E_t^{(m)}|(x/\lambda_{\text{th}})|E_t^{(n)}\rangle\right|^2\langle N^{(\omega_t^{(n,m)})}\rangle & \text{(I12)}\end{aligned}$$

for $n > m$. Comparing with Eqs. (I7) and (I8), this also implies

$$r_t^{(n\to m)} = \frac{k_B T}{\hbar} 8\pi\alpha (q/\mathrm{e})^2 \frac{k_B T}{m_q c^2} \left(\frac{\hbar\omega_t^{(n,m)}}{k_B T} \right)^3 \times \left| \langle E_t^{(m)} | (x/\lambda_{\mathrm{th}}) | E_t^{(n)} \rangle \right|^2 \left(\langle N^{(\omega_t^{(n,m)})} \rangle + 1 \right) \tag{I13}$$

for $n > m$.

The dynamics can thus be expressed as

$$\dot{\rho}_t = \frac{i}{\hbar}[\rho_t, H_t] + \gamma \sum_n \sum_{m<n} \left(\frac{\hbar\omega_t^{(n,m)}}{k_B T} \right)^3 \left| \langle E_t^{(m)} | \frac{x}{\lambda_{\mathrm{th}}} | E_t^{(n)} \rangle \right|^2 \times \left\{ \left(\langle N^{(\omega_t^{(n,m)})} \rangle + 1 \right) \mathcal{D}[L_t^{(m,n)}](\rho_t) + \langle N^{(\omega_t^{(n,m)})} \rangle \times \mathcal{D}[L_t^{(n,m)}](\rho_t) \right\}, \tag{I14}$$

where $\gamma = (k_B T/\hbar) 8\pi\alpha (q/\mathrm{e})^2 (k_B T/m_q c^2)$.

In the main text, we assume that the photon bath is always in equilibrium at temperature T. In this case, the photons exhibit the standard Bose-Einstein statistics $\langle N^{(\omega)} \rangle = (e^{\beta\hbar\omega} - 1)^{-1}$ and the dynamics reduce to

$$\dot{\rho}_t = \frac{i}{\hbar}[\rho_t, H_t] + \gamma \sum_n \sum_{m<n} \left(\frac{\hbar\omega_t^{(n,m)}}{k_B T} \right)^3 \left| \langle E_t^{(m)} | \frac{x}{\lambda_{\mathrm{th}}} | E_t^{(n)} \rangle \right|^2 \times \left\{ \frac{e^{\beta\hbar\omega_t^{(n,m)}}}{e^{\beta\hbar\omega_t^{(n,m)}} - 1} \mathcal{D}[L_t^{(m,n)}](\rho_t) + \frac{1}{e^{\beta\hbar\omega_t^{(n,m)}} - 1} \mathcal{D}[L_t^{(n,m)}](\rho_t) \right\}. \tag{I15}$$

Note that in the presence of the equilibrium photon bath, the relative transition rates between the instantaneous energy eigenstates of the system satisfy detailed balance, such that

$$\frac{r_t^{(n\to m)}}{r_t^{(m\to n)}} = e^{\beta(E_t^{(n)} - E_t^{(m)})}. \tag{I16}$$

We have chosen a regime for the quantum nonequilibrium thermodynamics that is, at least plausibly, experimentally accessible. Note that coherent quantum dynamics are very fast, since $k_B/\hbar = 131$ GHz/K. Meanwhile, for a single electron, $q/\mathrm{e} = 1$, $k_B/(m_\mathrm{e} c^2) = 1.69 \times 10^{-10}$/K and $\lambda_{\mathrm{th}} = 74.6\ \mathrm{nm}/\sqrt{T/\mathrm{K}}$. At $T = 300$ K, the thermal energy is $k_B T = 25.9$ meV and the thermal wavelength of the electron is $\lambda_{\mathrm{th}} = 4.31$ nm, while the coherent dynamics occur on the very fast time scale of $\beta\hbar = 25.4$ fs and the relaxation dynamics occur at the much slower rate on the order of $\gamma = 9.3 \times 10^{-9}/(\beta\hbar) = 366$ kHz.

1. Quantum evolution with separation of time scales

The factor $k_B/(m_\mathrm{e} c^2) = 1.69 \times 10^{-10}$/K (in the dissipative γ term) induces a separation of time scales between the coherent dynamics and the consequently much slower relaxation dynamics. In particular, if, over the time scale $\beta\hbar$, the Hamiltonian is approximately constant and the approximately constant relaxation rates are very small (such that $\beta\hbar r^{(m\to n)} \ll 1$), then it is useful to derive and employ the following discrete-time dynamics.

It will be fruitful to write the full Lindblad superoperator of Eq. (I6) as

$$\mathcal{L} = C + D \tag{I17}$$

and explore the discrete-time evolution superoperator $e^{t\mathcal{L}}$ over a duration t where $\mathcal{L}$ is approximately unchanging, and in a regime where $(tD)^2$ is negligible. In our case, $C(\cdot) = (i/\hbar)[\cdot, H]$ is the typical coherent superoperator and D describes the dissipative dynamics. Employing the Lie product formula and noting that $e^{tD/N} = I + tD/N + \mathcal{O}\left((tD)^2\right)$, we find

$$e^{t\mathcal{L}} = e^{t(C+D)} = \lim_{N\to\infty} \left(e^{tC/N} e^{tD/N} \right)^N \tag{I18}$$

$$\approx e^{tC} + \lim_{N\to\infty} \sum_{n=0}^{N-1} e^{ntC/N} \frac{t}{N} D e^{(N-n-1)tC/N} \tag{I19}$$

$$= e^{tC} + \int_0^t dt'\, e^{t'C} D e^{(t-t')C}. \tag{I20}$$

Using Eq. (I20) together with Eq. (I6) explicitly, a long calculation yields

$$e^{t\mathcal{L}}(\rho) = U\rho U^\dagger - \frac{1}{2} t \sum_n [I - |n\rangle\langle n|\, \mathrm{tr}] \times \left(\left\{ \sum_{m\neq n} r^{(m\to n)}\ |m\rangle\langle m|\, , U\rho U^\dagger \right\} \right), \tag{I21}$$

where $U = e^{-iHt/\hbar}$ and $\{\cdot, \cdot\}$ is the anticommutator.

In our simulations, we have restricted the dynamics at each time step to the 20 lowest-lying instantaneous energy eigenstates, so that m and n in Eq. (I21) both range over these 20 states. To approximate the dynamics, we have compounded a sequence of many discrete steps—each of the form $e^{(\delta t)\mathcal{L}_t}$ of Eq. (I21) but with each step using the instantaneous Lindblad superoperator $\mathcal{L}_t$ induced by (i) the instantaneous energy eigenstates of the Hamiltonian H_t and (ii) the instantaneous transition rates given by Eqs. (I12) and (I13).

Recall that our simulations assume room temperature (300 K) and so the transition rates are on the order of $\gamma = 9.3 \times 10^{-9}/(\beta\hbar) = 366$ kHz. Accordingly, we have taken $\delta t = 4 \times 10^5 \beta\hbar \approx 10$ ns as the duration of each discrete time step [61]. Each step of the discrete dynamic thus allows for many coherent oscillations with only perturbative decoherence. We have chosen the total duration of the protocol to be $\tau = 4 \times 10^9 \beta\hbar \approx 100$ µs, which is long enough to allow for many transitions, yet short enough to keep the system away from equilibrium throughout the protocol.

2. Work with separation of time scales

Recall that the expectation value of work can be calculated as $\langle W\rangle_{\rho_0} = \int_0^\tau \mathrm{tr}\left(\rho_t \dot{H}_t\right) dt$. During a simulation, the integral is approximated numerically via breaking time into many bins, each much smaller than τ. In typical simulations, with N time steps of duration $\tau' = \tau/N$, the expected work can be approximated as [62]

$$\langle W\rangle_{\rho_0} \overset{?}{\approx} \sum_{n=0}^{N-1} \mathrm{tr}\left[\rho_{n\tau'}(H_{(n+1)\tau'} - H_{n\tau'})\right], \tag{I22}$$

which implicitly assumes that the density matrix does not change appreciably during the time step.

However, because of the separation of time scales in our simulations, the density matrix goes through many coherent oscillations (in the relative phase between energy eigenstates) during a single time step. Accordingly, treating the density matrix as a constant during each time step, as done in Eq. (I22), would not be appropriate in our situation.

Rather—with N time steps of duration $\tau' = \tau/N$, and in a regime in which the rate of change of the Hamiltonian is approximately constant throughout each time step—the expected work is well approximated by

$$\begin{aligned}\langle W\rangle_{\rho_0} &= \int_0^\tau \mathrm{tr}\left(\rho_t \dot{H}_t\right) dt \\ &= \sum_{n=0}^{N-1} \int_{n\tau'}^{(n+1)\tau'} \mathrm{tr}\left(\rho_t \dot{H}_t\right) dt \qquad \text{(I23)} \\ &\approx \sum_{n=0}^{N-1} \mathrm{tr}\left[\left(\int_{n\tau'}^{(n+1)\tau'} \rho_{n\tau'}\, dt\right) \frac{H_{(n+1)\tau'} - H_{n\tau'}}{\tau'}\right] \qquad \text{(I24)} \\ &= \sum_{n=0}^{N-1} \mathrm{tr}\left[\overline{\rho_{n\tau'}}\, \left(H_{(n+1)\tau'} - H_{n\tau'}\right)\right], \qquad \text{(I25)}\end{aligned}$$

where

$$\overline{\rho_{n\tau'}} := \frac{1}{\tau'} \int_{n\tau'}^{(n+1)\tau'} \rho_t\, dt \tag{I26}$$

is the time-averaged density matrix, averaged over the duration of the nth time step. (Recall that our simulations use a time step of $4 \times 10^5 \beta\hbar$.) It is easy to see how Eq. (I25) would reduce to Eq. (I22) in a regime in which ρ_t is approximately constant throughout the duration of a time step.

As a first-order approximation to the time-averaged density matrix within a single time step, we can assume that the system does not make any dissipative transition. In this case, $e^{t\mathcal{L}} = e^{t(C+D)} \approx e^{tC}$ and the time-averaged density matrix will be well approximated by

$$\begin{aligned}\overline{\rho_{n\tau'}} &\approx \frac{1}{\tau'} \int_{n\tau'}^{(n+1)\tau'} e^{(t-n\tau')C_{n\tau'}}(\rho_{n\tau'})\, dt \qquad \text{(I27)} \\ &= \frac{1}{\tau'} \int_0^{\tau'} e^{tC_{n\tau'}}(\rho_{n\tau'})\, dt \qquad \text{(I28)} \\ &= \frac{1}{\tau'} \sum_{E,E' \in V_{H_{n\tau'}}} |E\rangle\langle E|\, \rho_{n\tau'}\, |E'\rangle\langle E'| \int_0^{\tau'} e^{i(E'-E)t/\hbar}\, dt \qquad \text{(I29)} \\ &= \sum_{E \in V_{H_{n\tau'}}} |E\rangle\langle E|\, \rho_{n\tau'}\, |E\rangle\langle E| \\ &\quad + \sum_{E' \in V_{H_{n\tau'}} \setminus \{E\}} \frac{\sin(\omega\tau'/2)}{\omega\tau'/2}\, e^{i\omega\tau'/2}\, |E\rangle\langle E|\, \rho_{n\tau'}\, |E'\rangle\langle E'|, \qquad \text{(I30)}\end{aligned}$$

where we have used the shorthand $\omega = (E' - E)/\hbar$. Note that energetic coherences get averaged out via a decaying envelope with magnitude $2/(\omega\tau')$.

[1] A. Kolchinsky and D. H. Wolpert, Dependence of dissipation on the initial distribution over states, J. Stat. Mech.: Theory Exp. **2017**, 083202 (2017).

[2] A. Kolchinsky, I. Marvian, C. Gokler, Z. Liu, P. Shor, O. Shtanko, K. Thompson, D. Wolpert, and S. Lloyd, Maximizing free energy gain, ArXiv:1705.00041.

[3] P. M. Riechers and M. Gu, Initial-state dependence of thermodynamic dissipation for any quantum process, Phys. Rev. E **103**, 042145 (2021).

[4] P. M. Riechers and M. Gu, Impossibility of achieving Landauer's bound for almost every quantum state, Phys. Rev. A **104**, 012214 (2021).

[5] A. Kolchinsky and D. H. Wolpert, Dependence of integrated, instantaneous, and fluctuating entropy production on the initial state in quantum and classical processes, Phys. Rev. E **104**, 054107 (2021).

[6] E. T. Jaynes, The minimum entropy production principle, Annu. Rev. Phys. Chem. **31**, 579 (1980).

[7] G. W. Paltridge, Climate and thermodynamic systems of maximum dissipation, Nature **279**, 630 (1979).

[8] W. F. Stinespring, Positive functions on C*-algebras, Proc. Am. Math. Soc. **6**, 211 (1955).

[9] T. A. B. Pinto Silva and R. M. Angelo, Quantum mechanical work, Phys. Rev. A **104**, 042215 (2021).

[10] R. Clausius, On a modified form of the second fundamental theorem in the mechanical theory of heat, London Edinb. Dublin Philos. Mag. J. Sci. **12**, 81 (1856).

[11] These expressions for entropy flow follow from from $\boldsymbol{\pi}_t^{\mathrm{env}} = (1/Z_t^{(b)})e^{-H^{(b)}/k_B T_t^{(b)}}$ and $\boldsymbol{\pi}_t^{\mathrm{env}} = (1/Z^{(b)})e^{-H^{(b)}/k_B T^{(b)}} \otimes (1/Z^{(b')})e^{-H^{(b')}/k_B T^{(b')}}$, respectively, with $H^{(b)}$ as the Hamiltonian for bath b and corresponding partition function $Z^{(b)}$.

[12] M. Esposito, K. Lindenberg, and C. Van den Broeck, Entropy production as correlation between system and reservoir, New J. Phys. **12**, 013013 (2010).

[13] M. Esposito and C. Van den Broeck, Second law and Landauer principle far from equilibrium, Eur. Phys. Lett. **95**, 40004 (2011).

[14] G. T. Landi and M. Paternostro, Irreversible entropy production: From classical to quantum, Rev. Mod. Phys. **93**, 035008 (2021).

[15] P. Talkner, E. Lutz, and P. Hänggi, Fluctuation theorems: Work is not an observable, Phys. Rev. E **75**, 050102 (2007).

[16] S. Hassani, *Mathematical Physics* (Springer, New York, 1999).

[17] C. Elouard and C. Lombard Latune, Extending the laws of thermodynamics for arbitrary autonomous quantum systems, PRX Quantum **4**, 020309 (2023).

[18] P. Strasberg and A. Winter, First and second law of quantum thermodynamics: A consistent derivation based on a microscopic definition of entropy, PRX Quantum **2**, 030202 (2021).

[19] A. Hewgill, G. De Chiara, and A. Imparato, Quantum thermodynamically consistent local master equations, Phys. Rev. Res. **3**, 013165 (2021).

[20] Á. Rivas, Strong coupling thermodynamics of open quantum systems, Phys. Rev. Lett. **124**, 160601 (2020).

[21] M. Perarnau-Llobet, E. Bäumer, K. V. Hovhannisyan, M. Huber, and A. Acin, No-go theorem for the characterization of work fluctuations in coherent quantum systems, Phys. Rev. Lett. **118**, 070601 (2017).

[22] M. Esposito, U. Harbola, and S. Mukamel, Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems, Rev. Mod. Phys. **81**, 1665 (2009).

[23] S. Deffner, J. P. Paz, and W. H. Zurek, Quantum work and the thermodynamic cost of quantum measurements, Phys. Rev. E **94**, 010103 (2016).

[24] K. Beyer, K. Luoma, and W. T. Strunz, Work as an external quantum observable and an operational quantum work fluctuation theorem, Phys. Rev. Res. **2**, 033508 (2020).

[25] A. Sone, Y. Liu, and P. Cappellaro, Quantum Jarzynski equality in open quantum systems from the onetime measurement scheme, Phys. Rev. Lett. **125**, 060602 (2020).

[26] It is interesting to note that there appears to be a hierarchy of operators that range from "universal" ones that require no knowledge of the system (e.g., coordinate operators) to those that require very detailed knowledge of both the system and external influences (e.g., the type-II operators defined later).

[27] A. E. Allahverdyan, Nonequilibrium quantum fluctuations of work, Phys. Rev. E **90**, 032137 (2014).

[28] S. Gherardini and G. De Chiara, Quasiprobabilities in quantum thermodynamics and many-body systems: A tutorial, ArXiv:2403.17138.

[29] A. E. Allahverdyan and Th. M. Nieuwenhuizen, Fluctuations of work from quantum subensembles: The case against quantum work-fluctuation theorems, Phys. Rev. E **71**, 066102 (2005).

[30] L. Jakóbczyk and M. Siennicki, Geometry of Bloch vectors in two-qubit system, Phys. Lett. A **286**, 383 (2001).

[31] The choice of inputs is arbitrary, except that they should all be linearly independent. However, this is not much of a restriction: almost any d^2 or fewer inputs chosen at random will be linearly independent.

[32] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, J. Mod. Opt. **44**, 2455 (1997).

[33] This initial state $\rho_0 = |w_{\min}\rangle \langle w_{\min}|$ yields the minimal expected work and thus the maximal expected work-extraction value $-\langle w_{\min}|\mathcal{W}|w_{\min}\rangle = \max\{-\lambda : \lambda \in \Lambda_{\mathcal{W}}\}$.

[34] D. S. Gonçalves, M. A. Gomes-Ruggiero, and C. Lavor, A projected gradient method for optimization over density matrices, Optim. Methods Softw. **31**, 328 (2016).

[35] L. Banchi, J. Pereira, S. Lloyd, and S. Pirandola, Convex optimization of programmable quantum computers, npj Quantum Inf. **6**, 1 (2020).

[36] M. Jaggi, in *Proceedings of the 30th International Conference on Machine Learning*, edited by Sanjoy Dasgupta and David McAllester, Proceedings of Machine Learning Research (PMLR, Atlanta, Georgia, 2013), p. 427, Vol. 28.

[37] For comparison with the previous work, note that the entropy flow vector $\vec{\varphi}$ as defined here is scaled by a factor of 1/2 relative to the $\vec{\phi}$ of Ref. [4]: $\vec{\varphi} = \vec{\phi}/2$. This rescaling has a number of esthetic benefits and leads more naturally to our high-dimensional generalization.

[38] Hermiticity and trace-one conditions are automatically satisfied by any generalized Bloch vector, so we only need to be careful about ensuring positive semidefiniteness of the induced matrix.

[39] D. Petz and G. Toth, The Bogoliubov inner product in quantum statistics, Lett. Math. Phys. **27**, 205 (1993).

[40] P. M. Riechers and J. P. Crutchfield, Beyond the spectral theorem: Decomposing arbitrary functions of nondiagonalizable operators, AIP Adv. **8**, 065305 (2018).

[41] H. J. D. Miller, G. Guarnieri, M. T. Mitchison, and J. Goold, Quantum fluctuations hinder finite-time information erasure near the Landauer limit, Phys. Rev. Lett. **125**, 160602 (2020).

[42] G. Manzano, J. M. Horowitz, and J. M. R. Parrondo, Nonequilibrium potential and fluctuation theorems for quantum maps, Phys. Rev. E **92**, 032129 (2015).

[43] F. Ciccarello, S. Lorenzo, V. Giovannetti, and G. Massimo Palma, Quantum collision models: Open system dynamics from repeated interactions, Phys. Rep. **954**, 1 (2022).

[44] M. Ziman, P. Štelmachovič, V. Bužek, M. Hillery, V. Scarani, and N. Gisin, Diluting quantum information: An analysis of information transfer in system-reservoir interactions, Phys. Rev. A **65**, 042105 (2002).

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

[46] M. Blaauboer and R. L. de Visser, An analytical decomposition protocol for optimal implementation of two-qubit entangling gates, J. Phys. A: Math. Theor. **41**, 395307 (2008).

[47] N. H. Y. Ng and M. P. Woods, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (Springer, Cham, Switzerland, 2019), p. 625.

[48] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Quantum and information thermodynamics: A unifying framework based on repeated interactions, Phys. Rev. X **7**, 021003 (2017).

[49] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, Nat. Commun. **5**, 1 (2014).

[50] D. P. DiVincenzo, Double quantum dot as a quantum bit, Science **309**, 2173 (2005).

[51] This work-extraction protocol is the time-reversed control sequence of a reset protocol.

[52] Notably, this initial condition was not used to infer the expected-work operator.

[53] N. Yunger Halpern and D. T. Limmer, Fundamental limitations on photoisomerization from thermodynamic resource theories, Phys. Rev. A **101**, 042116 (2020).

[54] F. Anza and J. P. Crutchfield, Beyond density matrices: Geometric quantum states, Phys. Rev. A **103**, 062218 (2021).

[55] This generalizes the classical case, in which the most fine-grained version of the random variable X is uniquely fixed by the probability distribution in the classical basis. Classical subensemble decompositions include coarse grainings but also include collections of initial distributions with overlapping support. The quantum case furthermore allows for basis freedom. In each case, the physical relevance of subensemble decompositions correspond to possible physical preparations of the system.

[56] This can be verified through the expression $\mathrm{tr}(\rho^2) = \mathrm{tr}\left[(I/d + \vec{b}\cdot\vec{\Gamma})(I/d + \vec{b}\cdot\vec{\Gamma})\right]$, from which we find $b = \sqrt{(\mathrm{tr}(\rho^2)d - 1)/(d-1)}$.

[57] Since the energy eigenstates are nondegenerate almost everywhere along the continuous time interval from 0 to τ, we will not worry about the potential complications of choosing a preferred basis for transitions to or from an energy-degenerate subspace.

[58] P. A. M. Dirac, The quantum theory of the emission and absorption of radiation, Proc. R. Soc. London A **114**, 243 (1927).

[59] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, UK, 2014).

[60] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics* (Addison-Wesley, San Francisco, USA, 2014), 2nd ed.

[61] This regime of $(\gamma\delta t)^2 \ll 1$ indeed satisfies the assumption of negligible $(\delta t D)^2$ that justifies the use of Eq. (I21) in our simulations.

[62] G. E. Crooks, Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems, J. Stat. Phys. **90**, 1481 (1998).