

Information Processing and the Second Law of Thermodynamics: An Inclusive, Hamiltonian Approach

Sebastian Deffner* and Christopher Jarzynski†

Department of Chemistry and Biochemistry and Institute for Physical Science and Technology,
University of Maryland, College Park, Maryland 20742, USA

(Received 31 May 2013; published 17 October 2013)

We obtain generalizations of the Kelvin-Planck, Clausius, and Carnot statements of the second law of thermodynamics for situations involving information processing. To this end, we consider an information reservoir (representing, e.g., a memory device) alongside the heat and work reservoirs that appear in traditional thermodynamic analyses. We derive our results within an inclusive framework in which all participating elements—the system or device of interest, together with the heat, work, and information reservoirs—are modeled explicitly by a time-independent, classical Hamiltonian. We place particular emphasis on the limits and assumptions under which cyclic motion of the device of interest emerges from its interactions with work, heat, and information reservoirs.

DOI: 10.1103/PhysRevX.3.041003

Subject Areas: Chemical Physics, Interdisciplinary Physics,
Statistical Physics

I. INTRODUCTION

Three classic expressions of the second law of thermodynamics are formulated in terms of cyclic processes. The Kelvin-Planck statement asserts that [1,2]

no process is possible whose sole result is the extraction of energy from a heat bath, and the conversion of all that energy into work.

The Clausius statement reads [3],

no process is possible whose sole result is the transfer of heat from a body of lower temperature to a body of higher temperature.

Finally, the Carnot statement declares that [4]

no engine operating between two heat reservoirs can be more efficient than a Carnot engine operating between those same reservoirs.

These formulations refer to processes involving the exchange of energy among idealized subsystems: one or more heat reservoirs; a work source—for example, a mass that can be raised or lowered against gravity; and a device that operates in cycles and affects the transfer of energy among the other subsystems. All three statements follow from simple entropy-balance analyses and offer useful, logically transparent reference points as one navigates the application of the laws of thermodynamics to real systems.

*sebastian.deffner@gmail.com

†cjarzyns@umd.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 3.0 License](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

This paper concerns extensions of these classic statements to situations involving information processing. In addition to the above-mentioned elements, we will consider an information reservoir—a system that exchanges information but not energy with the device. As we will show, the Kelvin-Planck, Clausius, and Carnot statements are each generalized in a natural way in the presence of such a reservoir. Although these generalized statements can be derived *ad hoc*, simply by including the Shannon entropy of the information reservoir in the entropy-balance analysis, our aim is to obtain these results directly from microscopic, Hamiltonian dynamics, highlighting the assumptions and approximations that are made along common idealizations.

Among the various connections that exist between information theory and thermodynamics, two are relevant in the present context. The first involves the relationship between the thermodynamic entropy defined via the Clausius relation $\int dQ/T = \Delta S$ and the Shannon entropy of information theory [5]

$$\mathcal{H} = -\text{tr}\{\rho \ln \rho\} \equiv -\int \rho \ln \rho, \quad (1)$$

where \int denotes an integral over phase space. Since these definitions coincide for a system in canonical equilibrium with a heat reservoir [6], it is highly tempting to use Eq. (1) to define the entropy of a nonequilibrium state. Indeed, if a system in contact with one or more thermal reservoirs evolves from an initial statistical state ρ_i to a final state ρ_f —neither of which is assumed to correspond to thermal equilibrium—then the Clausius-like inequality

$$\int_i^f \frac{dQ}{T} \leq -\text{tr}\{\rho_f \ln \rho_f\} + \text{tr}\{\rho_i \ln \rho_i\} \equiv \Delta \mathcal{H} \quad (2)$$

can be established from microscopic principles, as shown in Ref. [7] under assumptions similar to those we will make in the present paper; see also Refs. [8–21] for related

results and alternative derivations. On the other hand, for an isolated classical system, the Shannon entropy \mathcal{H} remains constant with time (by Liouville's theorem), which conflicts with the observation that the entropy of an isolated physical system increases until equilibrium is attained. Such considerations show that, at the very least, one must be careful when identifying Shannon entropy with thermodynamic entropy, away from thermal equilibrium.

The second connection involves the question of whether information about molecular-scale motions, gained by external observation, can be used to subvert the second law, in the sense suggested by Maxwell's famous thought experiment [22]. In an illuminating refinement of the Maxwell demon framework, Szilárd described a hypothetical scenario in which an intelligent being takes advantage of microscopic observations to manipulate a single-particle gas, so as to extract energy systematically from a reservoir and convert it to work [23]. Szilárd explicitly raised the possibility that this humanlike intelligence could be replaced by a purely physical device, in apparent violation of the Kelvin-Planck statement. By current consensus, the resolution of this paradox resides in Landauer's principle [24], which assigns a minimal thermodynamic cost to the erasure of the information gathered by the device; see Refs. [25–27] for details, Ref. [28] for an experimental treatment, Refs. [19,29–35] for illustrative models, and Refs. [36–39] for dissenting perspectives.

These topics have gained recent prominence in the context of microscopic feedback control. Sagawa and Ueda have analyzed the amount of work that can be delivered by the measurement and manipulation of small, fluctuating systems [40–44]. Their predictions have been verified experimentally using trapped colloidal particles [45] and mathematically illustrated for a solvable system with a linear feedback protocol with a Kalman filter [46]. Their analyses are based on an approach considering an integral fluctuation theorem. The corresponding detailed fluctuation theorems are discussed in Refs. [47–50]. These results have been extended to systems prepared in initial nonequilibrium stationary states [51–54] and to quantum systems [55–57]. Alternative treatments of feedback control, which do not rely on fluctuation theorems, can be found in Refs. [58–62] with applications to theoretic models [63–66] and experimental systems [67,68].

In the feedback control paradigm, the microscopic state of the system of interest (or of a measurement device [69]) is observed, and on the basis of those observations, a protocol is adapted to manipulate the system. Implicit in this paradigm is an external agent or apparatus—the demon or feedback controller [19]—who makes these observations and implements the feedback. The results derived within this approach are thus expressed as relationships between thermodynamic quantities such as work and information-theoretic quantities that measure the quality of the observations.

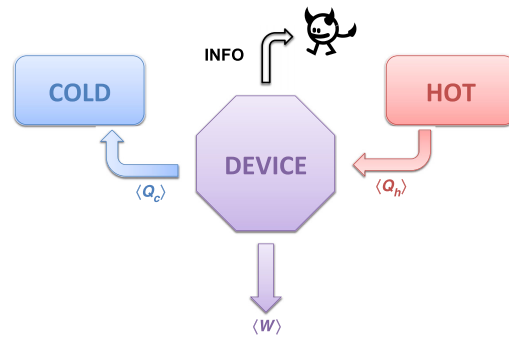


FIG. 1. Thermodynamic setup: A device exchanges heat with thermal reservoirs and work with a work source. The process is observed by a Maxwell demon.

In this paper, we aim at a treatment that does not involve an external agent. Instead, we consider a self-contained universe, a composite system containing the elements mentioned earlier: a device, one or more thermal reservoirs, a work source, and an information reservoir; cf. the illustration in Fig. 1. This composite system evolves autonomously under Hamilton's equations of motion, and any effective feedback control arises entirely from the interplay of the subsystems. Within this inclusive framework, we will obtain inequalities that generalize the Kelvin-Planck, Clausius, and Carnot statements to processes involving the exchange of information.

We will begin in Sec. II by specifying our theoretical framework and terminology. In Sec. III, we will obtain formal inequalities, which will then be combined in Sec. IV with physical interpretations, to obtain generalized statements of the second law for the cyclic process. In order to complete the analysis, we will derive a generalized maximum work theorem for the noncyclic process in Sec. V. Finally, we will conclude in Sec. VI. In obtaining these results, we will make a number of assumptions and approximations, reflecting idealizations that commonly arise in analyses of thermodynamic principles, and we will discuss the roles of these assumptions in our treatment.

II. THERMODYNAMICS WITHIN A HAMILTONIAN FRAMEWORK

In this section, we describe our framework, beginning with concepts and terminology. For our purposes, systems are categorized as devices, heat sources, work sources, and information sources. It is important to understand this categorization as an idealization of real physical systems with one dominant behavior.

A heat source (or sink) is a system that exchanges energy with other systems, in the form of heat but not work. Relaxation processes within a heat source are generally assumed to occur rapidly, implying that its temperature remains well defined throughout any process under consideration [70]. Moreover, if its heat capacity is sufficiently large, then that

temperature remains effectively constant and the heat source can be viewed as a heat reservoir. Because we wish to make contact with the usual formulations of the Kelvin-Planck, Clausius, and Carnot statements, we will use the term heat reservoir in the analysis that follows. In particular, we will see that the assumption of a large heat capacity is crucial for the emergence of cyclic motion of the device.

Analogously, a work source (or sink) can exchange energy in the form of work but not heat, and its internal relaxation processes are again assumed to be rapid [70]. As a result, the entropy of the work source remains constant and can be neglected. The assumption of rapid relaxation will appear implicitly in our treatment: We will model the work source as a single degree of freedom—effectively, a collective coordinate such as the center of mass of a macroscopic system—while ignoring its internal degrees of freedom; this assumption prevents the energy of the work source from being lost to internal dissipation. If the inertia of a work source is sufficiently large, then its motion is largely unaffected by interactions with other systems and it can be viewed as a work reservoir. In order to model cyclic processes, we will assume a work source with an effectively infinite inertia; we will return to this point in greater detail shortly.

Heat and work sources are convenient conceptual idealizations, familiar from classic thermodynamic treatments [2,71]. To this list, we add an information source, a system that exchanges information but not energy with other systems. The information source can exist in a number of physically distinct accessible states with identical free energies. A useful example is a memory register with N bits, hence 2^N energetically degenerate states. The capacity of the information source is measured by the natural logarithm of the number of accessible states, i.e., $N \ln 2$ in the case of the memory register. When this capacity is sufficiently large, the information source becomes an information reservoir.

Finally, we will consider a device, or (sub)system, of interest, which interacts with the above-mentioned elements. These interactions give rise to the exchange of energy with the heat and work sources, and they influence the dynamics of the information source among its degenerate states. As a result, data relating to the evolution of the device may become encoded in the information source. It is precisely this possibility that adds a new element to the standard analyses of the Kelvin-Planck, Clausius, and Carnot statements.

A. First law of thermodynamics

We proceed by formulating the first law of thermodynamics within our classical, Hamiltonian framework. To begin, we restrict the analysis to include only a device, a work source, and a single heat source (later, we will add multiple heat sources and an information source), and we model these elements with a Hamiltonian

$$H_{\text{uni}}(\gamma) = H_0(\mathbf{x}, \mathbf{p}; X) + h(\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}) + H_{\text{work}}(X, P) + H_{\text{heat}}(\boldsymbol{\xi}, \boldsymbol{\varphi}), \quad (3)$$

with $\gamma = (\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}; X, P)$, and where H_{uni} is our notation for the Hamiltonian of the universe. Here, (\mathbf{x}, \mathbf{p}) denotes the microstate of the device, and $(\boldsymbol{\xi}, \boldsymbol{\varphi})$ is that of the heat reservoir. The bold letters indicate vectors in the configuration and momentum spaces of these subsystems. We use (X, P) to specify the microstate of the work reservoir, which we model with a single degree of freedom. Finally, γ denotes a point in the full phase space, describing the combined microstate of all three subsystems.

We view the first term on the right side of Eq. (3) H_0 as the bare Hamiltonian for the device, parametrized by the configuration of the work source X . The second term h gives the interaction between the device and the heat source, and the third and fourth terms are the bare Hamiltonians for the work and heat sources. Defining

$$H_{\text{dev}}(\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}; X) \equiv H_0(\mathbf{x}, \mathbf{p}; X) + h(\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}), \quad (4)$$

we have

$$H_{\text{uni}} = H_{\text{dev}} + H_{\text{heat}} + H_{\text{work}}. \quad (5)$$

We interpret the three terms on the right side of Eq. (5) to be the instantaneous energies of the device, heat source, and work source, respectively. In our accounting, all interaction terms contribute to the energy of the device.

The microscopic evolution of our composite system is described by a Hamiltonian trajectory $\gamma(t)$, along which the value of H_{uni} remains constant:

$$\frac{d}{dt} H_{\text{uni}}[\gamma(t)] = \frac{dH_{\text{dev}}}{dt} + \frac{dH_{\text{work}}}{dt} + \frac{dH_{\text{heat}}}{dt} = 0. \quad (6)$$

The three subsystems exchange energy among themselves, with the total energy remaining fixed. The heat and work sources are not directly coupled to one another, but each is coupled to the device. Therefore, the rate at which the work source loses energy is interpreted as the rate at which work is performed on the device:

$$\frac{dW}{dt} \equiv -\frac{dH_{\text{work}}}{dt}. \quad (7)$$

Similarly, energy lost by the heat source is equated with heat absorbed by the device:

$$\frac{dQ}{dt} \equiv -\frac{dH_{\text{heat}}}{dt}. \quad (8)$$

Combining Eqs. (6)–(8), we arrive at

$$\dot{H}_{\text{dev}} = \dot{W} + \dot{Q}, \quad (9)$$

where the dots indicate derivatives with respect to time. Equation (9) constitutes the first law of thermodynamics in our framework.

By direct evaluation—using Hamilton's equation for the work source $\dot{X} = \partial H_{\text{uni}} / \partial P$, $\dot{P} = -\partial H_{\text{uni}} / \partial X$, together with Eq. (3)—we obtain

$$\frac{d}{dt}H_{\text{work}}(X, P) = \dot{X} \frac{\partial H_{\text{work}}}{\partial X} + \dot{P} \frac{\partial H_{\text{work}}}{\partial P} = -\dot{X} \frac{\partial H_0}{\partial X}. \quad (10)$$

Thus, with Eq. (7), the work performed on the device from time t_1 to t_2 is given by

$$W = \int_{t_1}^{t_2} dt \dot{X} \frac{\partial H_0}{\partial X}, \quad (11)$$

which is the familiar integral of displacement \times force used in thermodynamics [72].

B. Work reservoir

As mentioned earlier, there are two assumptions one might make about the properties of the work source: rapid self-equilibration and large inertia. We have built the first assumption into our framework by modeling the work source with a single degree of freedom X . We will now also make the second assumption, which will allow us to address cyclic processes.

To formalize the assumption of large inertia, let us consider a specific example, in which a massive piston (the work source) confines a rarefied gas (the device) within a cylinder. The bare Hamiltonian for the work source is

$$H_{\text{work}}(X, P) = \frac{P^2}{2M} + \frac{M\omega^2 X^2}{2}, \quad (12)$$

where the potential energy term models an ideal spring attached to the piston, as illustrated in Fig. 2. The piston begins in a microstate (X_0, P_0) , then evolves together with the gas and the surrounding thermal reservoir over a time interval $0 < t < t_f$, where t_f specifies the duration of the process in which we are interested.

In the limit $M \rightarrow \infty$, with initial conditions $(\mathbf{x}_0, \mathbf{p}_0; \boldsymbol{\xi}_0, \boldsymbol{\varphi}_0; X_0, P_0/M)$ held fixed, the motion of the massive piston becomes unaffected by the remaining degrees of freedom and is given by its free dynamics

$$\lim_{M \rightarrow \infty} X(t) = X_0 \cos(\omega t) - \frac{V_0}{\omega} \sin(\omega t), \quad (13a)$$

$$\lim_{M \rightarrow \infty} V(t) = V_0 \cos(\omega t) + X_0 \omega \sin(\omega t), \quad (13b)$$

where $V = P/M$ is the piston speed. (See Appendix A for details.) Thus, for sufficiently large M , we can treat the

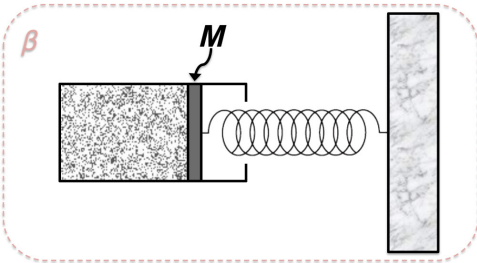


FIG. 2. Example of a work reservoir: The spring is attached to a piston with large mass M and confining a rarefied gas in a cylinder. The gas is in thermal contact with a heat bath of temperature β^{-1} .

motion of the piston as fully prescribed, given the initial conditions. This observation allows us to simplify the description of the total system. The piston now evolves independently, and the device and heat source evolve under a Hamiltonian with an externally imposed time dependence determined by $X(t)$. Introducing the notation

$$H_0(\mathbf{x}, \mathbf{p}; t) = H_0[\mathbf{x}, \mathbf{p}; X(t)], \quad (14)$$

$$H_{\text{dev}}(\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}; t) = H_{\text{dev}}[\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}; X(t)],$$

with $X(t)$ given by Eq. (13a), we now define

$$\begin{aligned} H_{\text{tot}}(\boldsymbol{\zeta}; t) &\equiv H_{\text{dev}}(\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}; t) + H_{\text{heat}}(\boldsymbol{\xi}, \boldsymbol{\varphi}) \\ &= H_0(\mathbf{x}, \mathbf{p}; t) + h(\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi}) + H_{\text{heat}}(\boldsymbol{\xi}, \boldsymbol{\varphi}), \end{aligned} \quad (15)$$

where $\boldsymbol{\zeta} = (\mathbf{x}, \mathbf{p}; \boldsymbol{\xi}, \boldsymbol{\varphi})$ specifies a point in the reduced phase space of the device and heat source. The Hamiltonian H_{tot} gives the combined energy of these two subsystems and generates their motion via Hamilton's equations. Because it is explicitly time dependent, its value is not preserved. Rather, the net change in H_{tot} along a trajectory $\boldsymbol{\zeta}(t)$ corresponds to the work performed on the device. This conclusion follows from energy conservation in the full phase space [$\dot{H}_{\text{tot}} = \dot{H}_{\text{dev}} + \dot{H}_{\text{heat}} = -\dot{H}_{\text{work}} = \dot{W}$; see Eqs. (6) and (7)] as well as directly from Eq. (11):

$$\begin{aligned} W &= \int_{t_1}^{t_2} dt \dot{X} \frac{\partial H_0}{\partial X} \xrightarrow{M \rightarrow \infty} \int_{t_1}^{t_2} dt \frac{\partial H_0}{\partial t} = \int_{t_1}^{t_2} dt \frac{\partial H_{\text{tot}}}{\partial t} \\ &= H_{\text{tot}}[\boldsymbol{\zeta}(t_2); t_2] - H_{\text{tot}}[\boldsymbol{\zeta}(t_1); t_1]. \end{aligned} \quad (16)$$

Here, we have made use of the Hamiltonian identity $dH_{\text{tot}}/dt = \partial H_{\text{tot}}/\partial t$ [73].

By construction, $H_{\text{tot}}(\boldsymbol{\zeta}; t)$ is a periodic function of time, with period $\tau = 2\pi/\omega$. The limit of large inertia [Eq. (13)] thus takes us from an inclusive description involving three subsystems (the device, the heat source, and the work source) to a reduced description in which the device, coupled to the heat source, is subjected to time-periodic external driving. We will continue our analysis within the reduced framework, making use of time-periodic Hamiltonians of the form given by Eq. (15). However, we emphasize that the explicit time dependence of H_{tot} is entirely induced by the dynamics of the massive work source.

We have used the piston and spring as an illustrative example, but the work source can equally well be modeled using a generic one-dimensional potential, provided the limit $M \rightarrow \infty$ is taken (as above) with P_0/M held fixed. Thus, the time dependence of the coordinate $X(t)$, while periodic, need not be sinusoidal. In the remainder of the paper, we will use τ to denote the period of the motion of the coordinate X (in the large-inertia limit), whether it is harmonic or not.

C. Heat reservoir

The limit of large work-source inertia gives us time-periodic driving, as we have just argued, but does not yet guarantee that the device itself relaxes to a time-periodic

steady state. For that, we will require two assumptions about the heat source, namely, that it is self-equilibrating and has a large heat capacity. In classical, macroscopic thermodynamics, these properties are among the defining properties of an idealized heat reservoir [70]. We now discuss these assumptions in the context of our explicitly microscopic setup and we formulate a plausibility argument for the emergence of a periodic steady state [Eq. (19)].

A large heat capacity implies that the number of degrees of freedom of our heat source N_{heat} far exceeds that of the device. This statement can be formalized by considering the thermodynamic limit $N_{\text{heat}} \rightarrow \infty$ while holding fixed the intensive properties of the heat source—its temperature, density, and chemical composition. In this limit, the characteristic energy exchanged between the device and the heat source, during the process in question, becomes a negligible fraction of the total energy of the heat source. Therefore, its intensive properties, and particularly its temperature, remain unchanged.

We take the assumption of self-equilibration to mean the following: From a generic initial microstate and in the absence of external influences, the heat source evolves to a microstate that—for the purpose of subsequent calculations—can be treated as a random sample from an equilibrium probability distribution [74]. A first-principles justification of this assumption involves issues that are well beyond the scope of this paper [75,76]. Empirically, however, macroscopic systems do relax to equilibrium when left undisturbed (leaving aside special cases such as glassy systems), and these equilibrium states are accurately modeled by the standard probability distributions of classical statistical mechanics. We will therefore assume that the heat source satisfies the property of self-equilibration, and we will investigate the consequences of this assumption.

Let us first consider the extreme limit, in which relaxation to equilibrium occurs on a time scale that is much faster than any other relevant time scale in our problem. In this case, even when the heat source interacts and exchanges energy with the device, its microstate at any instant can be treated as a random sample from an equilibrium ensemble. Effectively, then, the heat source evolves through a sequence of equilibrium states, as it absorbs or releases energy. Moreover, in the limit of infinite heat capacity, its temperature remains constant $\beta_\tau \xrightarrow{N_{\text{heat}} \rightarrow \infty} \beta_0$, as discussed in a previous paragraph.

Now, let $\mathbf{z}_n = [\mathbf{x}(n\tau), \mathbf{p}(n\tau)]$ denote the microstate of the device at the start of the n th period, and similarly define $\mathbf{Z}_n = [\boldsymbol{\xi}(n\tau), \boldsymbol{\varphi}(n\tau)]$ for the heat source. The evolution of the combined system from one period to the next is given by the iteration of a deterministic mapping:

$$\cdots \rightarrow (\mathbf{z}_{n-1}; \mathbf{Z}_{n-1}) \rightarrow (\mathbf{z}_n; \mathbf{Z}_n) \rightarrow (\mathbf{z}_{n+1}; \mathbf{Z}_{n+1}) \rightarrow \cdots \quad (17)$$

Each microstate ζ_n in this sequence is reached from the previous one, by evolving under Hamilton's equations for

one period of the time-dependent Hamiltonian $H_{\text{tot}}(\boldsymbol{\zeta}; t)$. In the limit of extremely rapid self-equilibration of the heat source, the \mathbf{Z}_n 's effectively become uncorrelated random samples from a fixed equilibrium distribution. Abstractly, we can view \mathbf{Z}_n as a set of freshly generated random numbers that collectively determine the value of \mathbf{z}_{n+1} , given \mathbf{z}_n ; in the next iteration, a new set of random numbers \mathbf{Z}_{n+1} determines the transition from \mathbf{z}_{n+1} to \mathbf{z}_{n+2} , and so forth. Adopting this perspective, the stroboscopic evolution of the device from one period to the next

$$\cdots \rightarrow \mathbf{z}_{n-1} \rightarrow \mathbf{z}_n \rightarrow \mathbf{z}_{n+1} \rightarrow \cdots \quad (18)$$

is given by the iteration of a stationary, stochastic, Markovian mapping. This Markov chain relaxes to a unique stationary state described by a fixed, generally nonequilibrium distribution $\bar{\rho}_{\text{dev}}(\mathbf{x}, \mathbf{p})$. (This result is a consequence of the Perron-Frobenius theorem [77], under standard assumptions.) Therefore, the time-dependent probability distribution for the device relaxes to a periodic steady state

$$\rho_{\text{dev}}(\mathbf{x}, \mathbf{p}, t + \tau) = \rho_{\text{dev}}(\mathbf{x}, \mathbf{p}, t), \quad (19)$$

where $\rho_{\text{dev}}(\mathbf{x}, \mathbf{p}, n\tau) = \bar{\rho}_{\text{dev}}(\mathbf{x}, \mathbf{p})$.

To reach Eq. (19), we have assumed that the self-equilibration of the heat source occurs, in effect, infinitely rapidly. Now, we loosen this assumption by allowing the relaxation time scale of the heat source to be comparable to other time scales in the problem. In this case, \mathbf{Z}_{n+1} in Eq. (17) may be statistically correlated with \mathbf{Z}_n and with \mathbf{z}_n . Nevertheless, it is reasonable to assume that there exists some integer $K > 0$, such that \mathbf{Z}_{n+K} is statistically uncorrelated with \mathbf{Z}_n and \mathbf{z}_n . In other words, a time interval of duration $K\tau$ is sufficient for the heat source to “forget” its microstate. Then, the stroboscopic evolution of the device in time increments $K\tau$

$$\cdots \rightarrow \mathbf{z}_{n-K} \rightarrow \mathbf{z}_n \rightarrow \mathbf{z}_{n+K} \rightarrow \cdots \quad (20)$$

is a Markov chain. The \mathbf{Z}_n 's are no longer necessarily sampled from equilibrium. However, if the heat source itself reaches a stationary state, in which the energy exchanged with the device is transported at a fixed rate to more distant regions of the heat source, then Eq. (20) becomes a stationary Markov chain, and the final arguments of the previous paragraph continue to apply: The device eventually relaxes to a periodic steady state.

As mentioned, the reasoning of the preceding paragraphs is intended as a plausibility argument for the emergence of cyclic motion of the device, under the conditions and limits we have discussed: The large inertia of the work source induces a time-periodic Hamiltonian for the device, and the large heat capacity and self-equilibration of the heat source cause the relaxation of the device into a time-periodic steady state. For the remainder of this paper, we will assume that these arguments apply—hence, the device reaches a periodic steady state—and we will explore their

consequences. As suggested at the beginning of Sec. II, we will henceforth use the terms work reservoir and heat reservoir.

We note that once the device has reached a periodic steady state, both its internal energy and its Shannon entropy become time periodic as well:

$$\langle H_{\text{dev}}(t + \tau) \rangle = \langle H_{\text{dev}}(t) \rangle, \quad \mathcal{H}_{\text{dev}}(t + \tau) = \mathcal{H}_{\text{dev}}(t). \quad (21)$$

We will make use of this observation in our later analysis.

D. Information reservoir

In the preceding subsections, work and heat reservoirs have been discussed within a classical, Hamiltonian framework. We now complete this framework by introducing the possibility of information processing. In effect, we aim to describe thermodynamic processes in the presence of a physical device capable of acting like Maxwell's demon, performing microscopic measurements and feedback on the other subsystems in our picture. The key feature that we wish to capture is the demon's memory, where it stores information that it has gathered. To this end, we introduce an idealized information reservoir, representing the demon's memory. All other components of the mechanical demon are implicitly treated as belonging to the device of interest.

To describe the complete system consisting of device, heat reservoir, and information reservoir, the total Hamiltonian (15) is extended to read

$$H_{\text{tot}}(\zeta; \Xi, \Phi; t) = H_{\text{dev}}(\mathbf{x}, \mathbf{p}; \xi, \varphi; t) + H_{\text{heat}}(\xi, \varphi) + H_{\text{info}}(\Xi, \Phi) + \mathfrak{h}(\zeta; \Xi, \Phi), \quad (22)$$

where (Ξ, Φ) is the microstate of the information reservoir and H_{info} is its bare Hamiltonian. The term $\mathfrak{h}(\zeta; \Xi, \Phi)$ describes the interaction between the information reservoir and the device and thermal reservoir. The assumption that the information and thermal reservoirs are coupled is important for the following discussion.

Let us first describe the information reservoir in the presence of a single thermal reservoir, at inverse temperature β , before discussing its interaction with the device.

For specificity, we will take the information reservoir to be a memory register consisting of N bits [78]. A single bit is physically implemented using a large collection of atoms or molecules, whose total magnetization (or some other collective observable) acts as a binary order parameter. We will distinguish between the microstate of the information reservoir $\psi \equiv (\Xi, \Phi)$ and its informational state σ . The microstate ψ is a point in the phase space of the entire collection of atoms and molecules comprising the memory register, whereas the informational state σ is a given sequence of bit values, e.g., 0110...10. We assume that each microstate ψ corresponds to a particular informational state σ , and we will use the function $\hat{\sigma}(\psi)$ to specify the informational state associated with the microstate ψ .

The variables ψ and σ thus represent fine-grained and coarse-grained descriptions of the state of the information reservoir.

The function $\hat{\sigma}(\psi)$ partitions the phase space of the information reservoir into 2^N distinct regions, each corresponding to one informational state. To guarantee a stable and reliable memory register, we assume these regions are separated by large free-energetic barriers, so that over the time scales that concern us, the probability of a spontaneous, thermally driven transition from one informational state to another is negligible. It then becomes useful to consider a constrained equilibrium state, described by a conditional probability distribution

$$p^{\text{eq}}(\psi|\sigma) = \delta_{\sigma, \hat{\sigma}(\psi)} \exp(-\beta[H_{\text{info}}(\psi) - F_{\text{info}}^{\sigma}]). \quad (23)$$

Here, the Kronecker δ function acts as an indicator variable; hence, $p^{\text{eq}}(\psi|\sigma)$ is simply a canonical probability distribution, restricted to the region of phase space corresponding to the information state σ . The free energy F_{info}^{σ} is determined by normalization $\int d\psi p^{\text{eq}}(\psi|\sigma) = 1$, and in the usual manner, we can define an equilibrium internal energy and entropy

$$\begin{aligned} \langle H_{\text{info}} \rangle^{\text{eq}, \sigma} &= \int d\psi p^{\text{eq}}(\psi|\sigma) H_{\text{info}}(\psi), \\ \mathcal{H}_{\text{info}}^{\text{eq}, \sigma} &= - \int d\psi p^{\text{eq}}(\psi|\sigma) \ln p^{\text{eq}}(\psi|\sigma). \end{aligned} \quad (24)$$

Equation (23) represents the statistical state of the information reservoir, when it has been left undisturbed in the informational state σ , in the presence of a thermal reservoir.

Following Bennett [79], we will refer to the N bits as information-bearing degrees of freedom, or IBD, and the remaining microscopic variables as non-information-bearing degrees of freedom, or NBD. Using this terminology, Eq. (23) represents an equilibrium state of the NBD $(\psi|\sigma)$ for a given state of the IBD (σ) .

Let us now consider the behavior of the information reservoir in the presence of the device of interest. We explicitly assume that interactions with the device can give rise to transitions among the informational states. In this manner, information about the evolution of the device of interest becomes encoded in the IBD. Let us further assume that (1) the 2^N informational states have the same equilibrium energies and entropies and (2) after a transition from one informational state to another, thermal equilibration of the NBD occurs rapidly. Under these assumptions, the energy of the information reservoir effectively remains constant, aside from equilibrium thermal fluctuations. In the presence of the device of interest and thermal reservoir, the evolution of the information reservoir is a sequence of transitions from one equilibrated informational state to another.

The total information encoded in the reservoir is quantified by its Shannon entropy, which can formally be

decomposed into contributions from the information-bearing and non-information-bearing degrees of freedom:

$$\mathcal{H}_{\text{info}}(t) = -\text{tr}\{\rho_{\text{info}}(t)\ln\rho_{\text{info}}(t)\} = \mathcal{H}_{\text{info}}^{\text{IBD}}(t) + \mathcal{H}_{\text{info}}^{\text{NBD}}(t), \quad (25)$$

as we show in Appendix B. Moreover, under the assumptions of the previous paragraph, $\mathcal{H}_{\text{info}}^{\text{NBD}}(t)$ does not vary with time and is given simply by the equilibrium entropy of the microscopic, non-information-bearing degrees of freedom (again, see Appendix B for details). As a result, any change in the Shannon entropy of the information reservoir, resulting from its interactions with the device over an interval of time, is entirely captured by the net change in the probability distribution of the mesoscopic, information-bearing degrees of freedom:

$$\Delta\mathcal{H}_{\text{info}} = \Delta\mathcal{H}_{\text{info}}^{\text{IBD}}. \quad (26)$$

In Secs. III, IV, and V below, we will use the notation $\Delta\mathcal{H}_{\text{info}}$ rather than $\Delta\mathcal{H}_{\text{info}}^{\text{IBD}}$, to avoid clutter, but it will be understood that the net change in the Shannon entropy of the information reservoir refers to the change in its information-bearing degrees of freedom.

III. NON-NEGATIVITY OF INFORMATION EXCHANGE

The rest of this paper is devoted to investigating specific thermodynamic processes within the framework introduced above. To this end, we begin by obtaining an inequality for the sum of changes of the Shannon entropy for the individual subsystems [Eq. (33)] from which we derive an inequality related to the behavior of our system in the periodic steady state [Eq. (43)]. The latter result and its generalization [Eq. (44)] will then be exploited in Sec. IV. As in the recent work of Hasegawa *et al.* [15,17] and Esposito *et al.* [16,18], our approach in this section will draw on properties of the canonical distribution, the Shannon entropy, and the Kullback-Leibler divergence [80], as well as assumptions about the initial state of the system.

We adopt an explicitly statistical perspective, in which we consider an ensemble representing different possible microscopic realizations of the process. The probability distribution in the full phase space at an initial time $t = 0$ reflects the preparation of the system prior to this time, and we now spell out the assumptions that we make regarding this preparation. As in Refs. [7,8,15–18], we assume that the total system begins in a product state

$$\rho_{\text{tot}}(\zeta; \Xi, \Phi; 0) = \rho_{\text{dev}}(\mathbf{x}, \mathbf{p}; 0) \times \rho_{\text{heat}}(\xi, \varphi; 0)\rho_{\text{info}}(\Xi, \Phi; 0). \quad (27)$$

This assumption does not substantially restrict the generality of the following discussion, as we expect the device to relax into a time-periodic steady state that is independent

of its initial preparation (see Sec. II). For the time being, we restrict ourselves to considering only a single thermal reservoir, but as discussed below, the results generalize easily to multiple reservoirs; see, e.g., Eq. (44).

We take the initial state of the heat reservoir to be given by the canonical distribution

$$\rho_{\text{heat}}(\xi, \varphi; 0) = \frac{1}{Z_{\text{heat}}} \exp[-\beta H_{\text{heat}}(\xi, \varphi)] \equiv \rho_{\text{heat}}^{\text{eq}}(\xi, \varphi), \quad (28)$$

with free energy $F_{\text{heat}} = -\beta^{-1} \ln Z_{\text{heat}}$. For the distribution $\rho_{\text{info}}(\Xi, \Phi; 0)$, we assume that the microscopic, non-information-bearing degrees of freedom are in equilibrium with the thermal reservoir (see Sec. IID), whereas the distribution of the mesoscopic, information-bearing degrees of freedom reflects the manner in which the information reservoir was prepared. For instance, the memory register might be initialized in a blank state $000\cdots 0$, in which case $\mathcal{H}_{\text{info}}^{\text{IBD}} = 0$. At the other extreme, it may be prepared so that every possible N -bit sequence (informational state) is equally likely; hence, $\mathcal{H}_{\text{info}}^{\text{IBD}} = N \ln 2$. We will not place any restrictions on the initial statistical state of the device $\rho_{\text{dev}}(\mathbf{x}, \mathbf{p}; 0)$.

After the initial preparation, the full system evolves under the time-periodic Hamiltonian given by Eq. (22). In general, the total density at time $t > 0$, $\rho_{\text{tot}}(\zeta; \Xi, \Phi; t)$, will not be a product state, and the reduced densities for device, heat reservoir, and memory are obtained by integrating out the other subsystems. Thus, for the device, we have

$$\rho_{\text{dev}}(\mathbf{x}, \mathbf{p}; t) = \int d\xi d\varphi \int d\Xi d\Phi \rho_{\text{tot}}(\zeta; \Xi, \Phi; t), \quad (29)$$

and similarly for the heat and information reservoir. We can use these reduced densities to define the Shannon entropy [81] of each subsystem, e.g.,

$$\begin{aligned} \mathcal{H}_{\text{dev}}(t) &= -\text{tr}\{\rho_{\text{dev}}(t)\ln\rho_{\text{dev}}(t)\} \\ &= -\int d\mathbf{x}d\mathbf{p} \rho_{\text{dev}}(\mathbf{x}, \mathbf{p}; t) \ln\rho_{\text{dev}}(\mathbf{x}, \mathbf{p}; t), \end{aligned} \quad (30)$$

and analogously for the heat and information reservoir.

By Liouville's theorem, the Shannon entropy of the total system remains constant under Hamiltonian dynamics: $\mathcal{H}_{\text{tot}}(t) = \mathcal{H}_{\text{tot}}(0)$. Moreover, since the system is prepared in a product state (27), we have

$$\mathcal{H}_{\text{tot}}(0) = \mathcal{H}_{\text{dev}}(0) + \mathcal{H}_{\text{heat}}(0) + \mathcal{H}_{\text{info}}(0). \quad (31)$$

At later times t , we have

$$\mathcal{H}_{\text{tot}}(t) \leq \mathcal{H}_{\text{dev}}(t) + \mathcal{H}_{\text{heat}}(t) + \mathcal{H}_{\text{info}}(t), \quad (32)$$

due to the subadditivity of the Shannon entropy [81]. Subtracting Eq. (31) from Eq. (32), we obtain

$$\Delta\mathcal{H}_{\text{dev}} + \Delta\mathcal{H}_{\text{heat}} + \Delta\mathcal{H}_{\text{info}} \geq \Delta\mathcal{H}_{\text{tot}} = 0, \quad (33)$$

where $\Delta\mathcal{H}_{\text{dev}}$ denotes the net change in the Shannon entropy of the device of interest, and similarly for the other subsystems.

For the heat reservoir, we can write

$$\begin{aligned}\mathcal{H}_{\text{heat}}(t) &= -\text{tr}\{\rho_{\text{heat}}(t) \ln \rho_{\text{heat}}(t)\} \\ &= \beta E_{\text{heat}}(t) - \beta F_{\text{heat}} - D(\rho_{\text{heat}}(t) \parallel \rho_{\text{heat}}^{\text{eq}}),\end{aligned}\quad (34)$$

where

$$E_{\text{heat}}(t) \equiv \text{tr}\{H_{\text{heat}}\rho_{\text{heat}}(t)\} \quad (35)$$

is the average energy of the reservoir, and $D(\cdot \parallel \cdot)$ denotes the Kullback-Leibler divergence [80]

$$D(\rho \parallel \sigma) = \text{tr}\{\rho \ln \rho\} - \text{tr}\{\rho \ln \sigma\} \geq 0. \quad (36)$$

Combining Eq. (34) with our assumption that the heat reservoir is prepared in equilibrium [Eq. (28)], we get

$$\begin{aligned}\Delta\mathcal{H}_{\text{heat}} &= \beta E_{\text{heat}}(t) - \beta E_{\text{heat}}(0) - D(\rho_{\text{heat}}(t) \parallel \rho_{\text{heat}}^{\text{eq}}) \\ &\leq \beta(E_{\text{heat}}(t) - E_{\text{heat}}(0)) = \beta\Delta E_{\text{heat}},\end{aligned}\quad (37)$$

using Eq. (36). This result in turn combines with Eq. (33) to give

$$\Delta\mathcal{H}_{\text{dev}} + \beta\Delta E_{\text{heat}} + \Delta\mathcal{H}_{\text{info}} \geq 0. \quad (38)$$

Note that we have taken two distinct steps to arrive at Eq. (38). First, we have obtained Eq. (33) from our assumption that the subsystems are statistically uncorrelated at the initial time [Eq. (27)]. In fact, the left side of Eq. (33) quantifies the degree to which correlations develop between the subsystems because of their mutual interactions; Esposito *et al.* [16] have explicitly interpreted this buildup of correlations as representing entropy production. Next, to get to Eq. (38), we have used the assumption that the reservoir is initialized in the canonical distribution [Eq. (28)], together with the non-negativity of the Kullback-Leibler divergence. Similar manipulations appear in Refs. [15–18]. We will now use Eq. (38) to arrive at inequalities that characterize the behavior of our system in the periodic steady state.

A natural time scale for our process is given by the driving period τ (Sec. II B). Let us set $t = n\tau > n_0\tau$, where n_0 is the number of periods needed for the device to relax into its periodic steady state. Then, the process in question can be divided into a transient interval ($0 \rightarrow n_0\tau$) followed by an interval of time-periodic behavior ($n_0\tau \rightarrow n\tau$). Expressing each term in Eq. (38) as a sum of contributions from these two intervals, we get

$$\begin{aligned}\Delta\mathcal{H}_{\text{dev}}^{0 \rightarrow n_0} + \beta\Delta E_{\text{heat}}^{0 \rightarrow n_0} + \Delta\mathcal{H}_{\text{info}}^{0 \rightarrow n_0} + \Delta\mathcal{H}_{\text{dev}}^{n_0 \rightarrow n} \\ + \beta\Delta E_{\text{heat}}^{n_0 \rightarrow n} + \Delta\mathcal{H}_{\text{info}}^{n_0 \rightarrow n} \geq 0,\end{aligned}\quad (39)$$

which can further be rewritten as

$$\begin{aligned}\Delta\mathcal{H}_{\text{dev}}^{0 \rightarrow n_0} + \beta\Delta E_{\text{heat}}^{0 \rightarrow n_0} + \Delta\mathcal{H}_{\text{info}}^{0 \rightarrow n_0} \\ + (n - n_0)(\beta\Delta E_{\text{heat}}^{\text{cyc}} + \Delta\mathcal{H}_{\text{info}}^{\text{cyc}}) \geq 0,\end{aligned}\quad (40)$$

where

$$\Delta E_{\text{heat}}^{\text{cyc}} = \frac{\Delta E_{\text{heat}}^{n_0 \rightarrow n}}{n - n_0} \quad (41)$$

is the average heat absorbed by the heat reservoir, per cycle, in the periodic steady state, and

$$\Delta\mathcal{H}_{\text{info}}^{\text{cyc}} = \frac{\Delta\mathcal{H}_{\text{info}}^{n_0 \rightarrow n}}{n - n_0} \quad (42)$$

is the average change in the Shannon entropy of the information reservoir, per cycle, in the periodic steady state. Note that the similarly defined quantity $\Delta\mathcal{H}_{\text{dev}}^{\text{cyc}}$ vanishes, by Eq. (21). Dividing both sides of the inequality in Eq. (40) by $(n - n_0)$, then taking the limit $n \rightarrow \infty$, we finally obtain

$$\beta\Delta E_{\text{heat}}^{\text{cyc}} + \Delta\mathcal{H}_{\text{info}}^{\text{cyc}} \geq 0. \quad (43)$$

In Sec. IV, we will exploit this result [or its generalization, Eq. (44)] to obtain generalized versions of the Kelvin-Planck, the Clausius, and the Carnot statements of the second law.

Equation (43) has a simple interpretation: The first term on the left represents the net change in the thermodynamic entropy of the heat reservoir, and the second term is the net change in the Shannon entropy of the information reservoir, specifically, its information-bearing degrees of freedom. Either term can be positive, negative, or zero, but their sum must be non-negative.

In the preceding analysis, for convenience, we have restricted ourselves to a single heat reservoir. The arguments are readily generalized to the case of multiple heat reservoirs by replacing the change of Shannon entropy for one reservoir by a sum over all reservoirs, in Eq. (33), and by assuming that each heat reservoir is independently prepared in a canonical distribution corresponding to a particular temperature. In particular, for one hot and one cold reservoir, Eq. (43) becomes

$$\beta_{\text{hot}}\Delta E_{\text{hot}}^{\text{cyc}} + \beta_{\text{cold}}\Delta E_{\text{cold}}^{\text{cyc}} + \Delta\mathcal{H}_{\text{info}}^{\text{cyc}} \geq 0. \quad (44)$$

In the case of multiple reservoirs, we will assume that the information reservoir is coupled only to a single thermal reservoir, and its microscopic degrees of freedom remain in equilibrium at the corresponding temperature. The results that we derive in the following section do not depend on which reservoir is selected for this role.

IV. THE SECOND LAW AND INFORMATION PROCESSING

In the periodic steady state, $\langle H_{\text{dev}}(t + \tau) \rangle = \langle H_{\text{dev}}(t) \rangle$ [Eq. (21)]. Therefore, by the first law of thermodynamics, integrating Eq. (9) over a single cycle, we have

$$\langle W^{\text{cyc}} \rangle + \langle Q^{\text{cyc}} \rangle = 0, \quad (45)$$

where W^{cyc} is the net work performed on the device, and Q^{cyc} is the net heat absorbed by the device (from one or more heat reservoirs) over one cycle in the periodic steady state, and angular brackets denote averages over many realizations. By Eq. (8), the heat absorbed by the system is defined as the net decrease in the bare energies of the heat reservoir(s); hence, Eq. (45) becomes

$$\langle W^{\text{cyc}} \rangle = \Delta E_{\text{heat}}^{\text{cyc}} \quad (46)$$

Kelvin-Planck statement.—The Kelvin-Planck statement [1] expresses the observation that in cyclic, isothermal processes, the average work is always non-negative $\langle W^{\text{cyc}} \rangle \geq 0$. To generalize this statement, we consider a device of interest, coupled to a single heat reservoir at inverse temperature β , a work reservoir, and an information reservoir. Equation (46) then combines with Eq. (43) to give

$$\beta \langle W^{\text{cyc}} \rangle \geq -\Delta \mathcal{H}_{\text{info}}^{\text{cyc}}, \quad (47)$$

which constitutes a generalized version of the Kelvin-Planck statement. For processes during which information is written to the information reservoir ($\Delta \mathcal{H}_{\text{info}}^{\text{cyc}} > 0$), the net work over one cycle can be negative. In other words, there can be a systematic transfer of energy from the heat reservoir to the work reservoir, provided the Shannon entropy of the information reservoir increases. This conclusion is consistent with the current consensus regarding the Maxwell demon paradox [25–27,31]. For processes during which information is erased ($\Delta \mathcal{H}_{\text{info}}^{\text{cyc}} < 0$), Eq. (47) becomes equivalent to Landauer's principle [8,15,17,18,24], placing a lower limit on the amount of work that must be expended in order to accomplish this erasure.

Clausius statement.—To generalize the Clausius statement, we consider a device interacting with two heat reservoirs, one hot and one cold, as well as a work reservoir and an information reservoir. As above, these interactions produce exchanges of both energy and information. In general, the net work performed on the device over one cycle can have either sign, and the device may be able to operate as either a heat engine or a refrigerator. For the Clausius statement, we restrict our attention to processes for which $\langle W^{\text{cyc}} \rangle = 0$. In this case, Eq. (46) becomes

$$0 = \Delta E_{\text{hot}}^{\text{cyc}} + \Delta E_{\text{cold}}^{\text{cyc}}. \quad (48)$$

Consequently, Eq. (44) can be written as

$$(\beta_{\text{cold}} - \beta_{\text{hot}}) \langle Q_{\text{hot}}^{\text{cyc}} \rangle \geq -\Delta \mathcal{H}_{\text{info}}^{\text{cyc}}, \quad (49)$$

which generalizes the Clausius statement. Note that the left side of Eq. (49) represents classical thermodynamic entropy, i.e., the heat exchanged over temperature, whereas the right side quantifies the internal information gain in the memory. Since $\beta_{\text{cold}} > \beta_{\text{hot}}$, Eq. (49) allows for processes during which heat flows systematically from cold to hot ($\langle Q_{\text{hot}} \rangle < 0 < \langle Q_{\text{cold}} \rangle$), provided information is written to the memory,

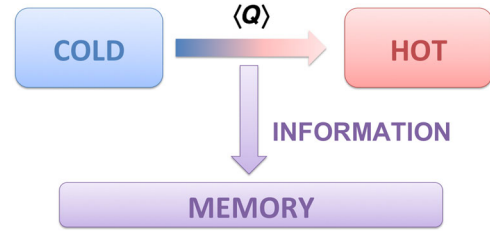


FIG. 3. Illustration of the generalized Clausius statement [Eq. (49)]. Heat can flow from the cold to the hot reservoir if information is written to memory.

as illustrated schematically in Fig. 3. Conversely, if information is to be erased, then the right side of the inequality is positive, and we get a lower bound on the amount of heat that must flow from the hot to the cold reservoir. For the erasure of one bit of information per cycle $\Delta \mathcal{H}_{\text{info}}^{\text{cyc}} = -\ln 2$, the average heat flow must satisfy

$$\langle Q_{\text{hot}}^{\text{cyc}} \rangle \geq (\beta_{\text{cold}} - \beta_{\text{hot}})^{-1} \ln 2, \quad (50)$$

which represents a modified version of Landauer's principle [35].

Carnot statement.—The Carnot statement asserts that the efficiency of a heat engine is always less than the Carnot efficiency $\eta \leq \eta_C \equiv 1 - \beta_{\text{hot}}/\beta_{\text{cold}}$ [4]. To generalize this result, we again consider a device interacting with two heat reservoirs, a work reservoir, and an information reservoir, but now we consider processes for which the work performed on the device over one cycle is negative (in other words, the device delivers work) and the heat absorbed from the hot reservoir is positive; hence, the device operates as a heat engine, with efficiency $\eta = -\langle W^{\text{cyc}} \rangle / \langle Q_{\text{hot}}^{\text{cyc}} \rangle > 0$. Equation (44) takes the form

$$-\beta_{\text{hot}} \langle Q_{\text{hot}}^{\text{cyc}} \rangle - \beta_{\text{cold}} \langle Q_{\text{cold}}^{\text{cyc}} \rangle + \Delta \mathcal{H}_{\text{info}}^{\text{cyc}} \geq 0, \quad (51)$$

and Eq. (45) can be written as

$$\langle Q_{\text{cold}}^{\text{cyc}} \rangle = -\langle W^{\text{cyc}} \rangle - \langle Q_{\text{hot}}^{\text{cyc}} \rangle. \quad (52)$$

Combining these equations, we obtain

$$-\beta_{\text{hot}} \langle Q_{\text{hot}}^{\text{cyc}} \rangle + \beta_{\text{cold}} \langle W^{\text{cyc}} \rangle + \beta_{\text{cold}} \langle Q_{\text{hot}}^{\text{cyc}} \rangle + \Delta \mathcal{H}_{\text{info}}^{\text{cyc}} \geq 0. \quad (53)$$

After rearrangement of terms, we find that the efficiency must satisfy

$$\eta \leq \left(1 - \frac{\beta_{\text{hot}}}{\beta_{\text{cold}}}\right) + \frac{\Delta \mathcal{H}_{\text{info}}^{\text{cyc}}}{\beta_{\text{cold}} \langle Q_{\text{hot}} \rangle} = \eta_C + \frac{\Delta \mathcal{H}_{\text{info}}^{\text{cyc}}}{\beta_{\text{cold}} \langle Q_{\text{hot}} \rangle}. \quad (54)$$

Thus, for cyclic processes in which information is systematically written to the memory, the efficiency can exceed the Carnot limit. Note that Eq. (54) does not depend on whether the information reservoir is coupled to the hot or the cold heat reservoir.

V. MAXIMUM WORK THEOREM

In Sec. IV, we considered only cyclic processes. Now, let us briefly consider what happens when we relax this restriction. For noncyclic processes in the presence of a single heat reservoir, the second law is formulated in terms of the Helmholtz free energy $F = E - \beta^{-1}S$, where $E = \langle H \rangle$ is the mean internal energy, F the free energy, and S the thermodynamic entropy of the system in question, in a state of thermal equilibrium. If the system begins in one equilibrium state and ends in another, then the average work performed on the system during the process satisfies $\langle W \rangle \geq \Delta F$, where the equality holds for reversible processes. Equivalently, the decrease in free energy gives the maximum usable, i.e., extractable, work during such a process.

To generalize this result, we consider a device of interest, coupled to a single heat reservoir, a work reservoir, and an information reservoir, without assuming cyclic motion. As before, we assume an initial product state [Eq. (27)] without imposing any restrictions on the initial state of the device and we imagine observing the entire system over some interval of time. Integrating Eq. (9) over this interval, we get

$$\Delta E_{\text{dev}} \equiv \langle \Delta H_{\text{dev}} \rangle = \langle W \rangle + \langle Q \rangle = \langle W \rangle - \Delta E_{\text{heat}}, \quad (55)$$

averaging over many realizations of the process. Combining this result with Eq. (38), which was derived without assuming cyclic processes, we obtain

$$\Delta \mathcal{H}_{\text{dev}} - \beta \Delta E_{\text{dev}} + \beta \langle W \rangle + \Delta \mathcal{H}_{\text{info}} \geq 0. \quad (56)$$

In order to further simplify Eq. (56), we introduce the information free energy

$$\mathcal{F} = E_{\text{dev}} - \beta^{-1} \mathcal{H}_{\text{dev}} = F + D(\rho \parallel \rho^{\text{eq}}), \quad (57)$$

which generalizes the equilibrium free energy to an arbitrary nonequilibrium state characterized by a probability distribution ρ . This nonequilibrium free energy has previously appeared in both Hamiltonian treatments [15–17], for instance, to derive Landauer’s principle [18], as well as stochastic treatments [19,82–85]. A generalized free energy of this form has also appeared in the thermodynamic description of open system dynamics [86]. More recently, it was shown that \mathcal{F} is a Lyapunov function for nonequilibrium stationary states [54].

In terms of this quantity, Eq. (56) becomes

$$\beta \langle W \rangle \geq \beta \Delta \mathcal{F} - \Delta \mathcal{H}_{\text{info}}, \quad (58)$$

which is a generalized version of the maximum work theorem. If the device begins and ends in equilibrium, $\Delta \mathcal{F}$ is replaced by the equilibrium free energy difference ΔF .

Equation (58) is similar to a version of the maximum work theorem applicable to systems with external feedback control [42,52,54,66,69]:

$$\beta \langle W \rangle \geq \beta \Delta F - \langle J \rangle. \quad (59)$$

Here, $\langle J \rangle$ denotes the mutual information that quantifies the quality of the measurements that are performed by an external agent. In Eq. (58), by contrast, $\Delta \mathcal{H}_{\text{info}}$ is the change of Shannon entropy of an explicitly modeled subsystem (our information reservoir), without reference to feedback control. See Ref. [19] for a treatment that combines both perspectives.

VI. CONCLUDING REMARKS

By categorizing thermodynamics systems as devices, thermal reservoirs, work reservoirs, and information reservoirs, we have developed an inclusive approach for investigating the thermodynamics of information processing, in which all participating subsystems are explicitly modeled. This approach is based on autonomous evolution under a time-independent Hamiltonian, supplemented by a number of limits, approximations, and assumptions, spelled out in Sec. II. Our main results in Sec. IV generalize the Kelvin-Planck, Clausius, and Carnot statements for cyclic thermodynamic processes and they support the consensus view [24–27] that the Shannon entropy in a random data set (as encoded by a memory register’s information-bearing degrees of freedom, for instance) should be placed on the same footing as the Clausius entropy when analyzing the second law of thermodynamics. Thus, for example, work can systematically be extracted from a single heat bath, heat can flow from cold to hot, and the Carnot efficiency can be exceeded, provided these entropy-decreasing consequences are compensated by the writing of information to a memory register. Section V extends these results to noncyclic processes in the form of a generalized maximum work principle.

As mentioned, our derivations have elements in common with previous treatments, particularly those of Refs. [15–19]. However, our focus on a fully autonomous, inclusive framework, on cyclic processes, and on the designation of an information reservoir as a separate element in thermodynamic analyses distinguishes our approach. In the spirit of a fully inclusive framework, Maes and Tasaki [21] have derived the maximum work statement of the second law of thermodynamics using a time-independent Hamiltonian. Their emphasis is on a mathematically rigorous treatment and does not focus on information processing.

Very recently, Tasaki [87] has analyzed a Hamiltonian model of Maxwell’s demon, involving an engine and a memory that interact by the exchange of information, and Barato and Seifert [88] have investigated feedback control with an explicit information reservoir, within the framework of stochastic thermodynamics.

Finally, it is worth mentioning that, at least formally, the present analysis can be extended to quantum-mechanical systems. In place of Hamiltonian dynamics, one would use the unitary dynamics of the “universe” under consideration, the Shannon entropy would be replaced by the von Neumann entropy, and classical ensemble averages

would be replaced by quantum expectation values taken with respect to density operators. Aside from these modifications, the mathematical steps in the derivation remain the same. However, because our analysis in this paper has relied heavily on classical reasoning and interpretation, it is not clear whether formally analogous quantal manipulations lead to physically meaningful results. These conceptual difficulties may perhaps be addressed by an appeal to decoherence by an external environment, so as to induce classicality [89], but such a treatment is beyond the scope of our approach, which is based on a self-contained, isolated universe. We leave these subtle interpretational issues to future work.

ACKNOWLEDGMENTS

S.D. acknowledges financial support by a fellowship from the German Academic Exchange Service (DAAD, Contract No. D/11/40955). Both authors acknowledge support from the National Science Foundation (U.S.A.), under Grant No. DMR-1206971, and stimulating conversations with Dibendu Mandal and Jordan M. Horowitz.

APPENDIX A: WORK RESERVOIR

In this Appendix, we analyze the illustrative example sketched in Fig. 2. A rarefied gas is confined by a cylinder, whose piston is coupled to a harmonic oscillator. The Hamiltonian of the work source is

$$H_{\text{work}}(X, P) = \frac{P^2}{2M} + \frac{M\omega^2 X^2}{2}, \quad (\text{A1})$$

where M is the mass of the piston and ω the angular frequency of the spring. From Hamilton's equations applied to the total Hamiltonian [Eq. (3)], we get for the piston

$$\dot{X} = V, \quad \dot{V} = -\omega^2 X - \frac{1}{M} \frac{\partial H_0}{\partial X}, \quad (\text{A2})$$

where $V = P/M$ is the piston velocity. In the limit $M \rightarrow \infty$, with the initial position X_0 and velocity V_0 held fixed, the last term above can be neglected. The work source then becomes a work reservoir whose time dependence is given by its free dynamics

$$\begin{aligned} \lim_{M \rightarrow \infty} X(t) &= X_0 \cos(\omega t) - \frac{V_0}{\omega} \sin(\omega t), \\ \lim_{M \rightarrow \infty} V(t) &= V_0 \cos(\omega t) + X_0 \omega \sin(\omega t), \end{aligned} \quad (\text{A3})$$

which are the periodic solutions used in Sec. II B.

APPENDIX B: INFORMATION RESERVOIR

Here, we provide justification for Eqs. (25) and (26), drawing on the assumptions that we have made about the information reservoir.

Given an arbitrary probability distribution $p(\psi)$ on the phase space of the information reservoir, we define marginal and conditional distributions

$$p_\sigma = \int d\psi \delta_{\sigma, \hat{\sigma}(\psi)} p(\psi), \quad p(\psi|\sigma) = \delta_{\sigma, \hat{\sigma}(\psi)} \frac{p(\psi)}{p_\sigma}. \quad (\text{B1})$$

p_σ is the probability distribution of informational states; $p(\psi|\sigma)$ is the conditional probability distribution of microstates, given an informational state σ ; and

$$p(\psi) = \sum_\sigma p_\sigma p(\psi|\sigma). \quad (\text{B2})$$

Introducing the shorthand notation $\int_\sigma d\psi \dots \equiv \int d\psi \delta_{\sigma, \hat{\sigma}(\psi)} \dots$, we obtain

$$\begin{aligned} \mathcal{H}_{\text{info}} &= - \int d\psi p(\psi) \ln p(\psi) \\ &= - \sum_\sigma p_\sigma \int d\psi p(\psi|\sigma) \ln \left[\sum_{\sigma'} p_{\sigma'} p(\psi|\sigma') \right] \\ &= - \sum_\sigma p_\sigma \int_\sigma d\psi p(\psi|\sigma) \ln [p_\sigma p(\psi|\sigma)] \\ &= - \sum_\sigma p_\sigma \ln p_\sigma - \sum_\sigma p_\sigma \int_\sigma d\psi p(\psi|\sigma) \ln p(\psi|\sigma) \\ &= \mathcal{H}_{\text{info}}^{\text{IBD}} + \sum_\sigma p_\sigma \mathcal{H}_{\text{info}}^{\text{NBD}}(\sigma), \end{aligned} \quad (\text{B3})$$

where $\mathcal{H}_{\text{info}}^{\text{NBD}}(\sigma)$ is the Shannon entropy associated with the conditional distribution $p(\psi|\sigma)$. Defining $\mathcal{H}_{\text{info}}^{\text{NBD}} \equiv \sum_\sigma p_\sigma \mathcal{H}_{\text{info}}^{\text{NBD}}(\sigma)$, we get the decomposition given by Eq. (25).

Making use of the assumption that the NBD equilibrate rapidly with the thermal reservoir, we replace $\mathcal{H}_{\text{info}}^{\text{NBD}}(\sigma)$ with its equilibrium value [see Eq. (24)]:

$$\mathcal{H}_{\text{info}} = \mathcal{H}_{\text{info}}^{\text{IBD}} + \sum_\sigma p_\sigma \mathcal{H}_{\text{info}}^{\text{eq}, \sigma}. \quad (\text{B4})$$

Combining this result with the assumption that the equilibrated informational states all have the same entropy ($\mathcal{H}_{\text{info}}^{\text{eq}, \sigma} = \mathcal{H}_{\text{info}}^{\text{eq}, \sigma'}$), we conclude that, of the two terms on the right of Eq. (B4), only the first one changes as the distribution p_σ evolves with time. This conclusion finally justifies Eq. (26).

-
- [1] Sir W. Thomson, *Mathematical and Physical Papers* (Cambridge University Press, Cambridge, England, 1882).
 - [2] M. Planck, *Thermodynamik* (Walter de Gruyter, Berlin, 1954).
 - [3] R. Clausius, *Abhandlungen über die mechanische Wärmetheorie* (Vieweg, Braunschweig, Germany, 1864).
 - [4] S. Carnot, *Réflexions sur la puissance motrice de feu et sur les machines propres à développer cette puissance* (Bachelier, Paris, 1824).

- [5] C.E. Shannon, *A Mathematical Theory of Communication*, *Bell Syst. Tech. J.* **27**, 379 (1948).
- [6] T.L. Hill, *Statistical Mechanics: Principles and Selected Applications* (McGraw-Hill, New York, 1956).
- [7] C. Jarzynski, *Microscopic Analysis of Clausius-Duhem Processes*, *J. Stat. Phys.* **96**, 415 (1999).
- [8] B. Piechocinska, *Information Erasure*, *Phys. Rev. A* **61**, 062314 (2000).
- [9] H. Qian, *Mesoscopic Nonequilibrium Thermodynamics of Single Macromolecules and Dynamic Entropy-Energy Compensation*, *Phys. Rev. E* **65**, 016102 (2001).
- [10] C. Maes and K. Netočný, *Time-Reversal and Entropy*, *J. Stat. Phys.* **110**, 269 (2003).
- [11] U. Seifert, *Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem*, *Phys. Rev. Lett.* **95**, 040602 (2005).
- [12] D. Andrieux and P. Gaspard, *Quantum Work Relations and Response Theory*, *Phys. Rev. Lett.* **100**, 230404 (2008).
- [13] D. Andrieux and P. Gaspard, *Molecular Information Processing in Nonequilibrium Copolymerizations*, *J. Chem. Phys.* **130**, 014901 (2009).
- [14] S. Vaikuntanathan and C. Jarzynski, *Dissipation and Lag in Irreversible Processes*, *Europhys. Lett.* **87**, 60005 (2009).
- [15] H.-H. Hasegawa, J. Ishikawa, K. Takara, and D. J. Driebe, *Generalization of the Second Law for a Nonequilibrium Initial State*, *Phys. Lett. A* **374**, 1001 (2010).
- [16] M. Esposito, K. Lindenberg, and C. van den Broeck, *Entropy Production as Correlation between System and Reservoir*, *New J. Phys.* **12**, 013013 (2010).
- [17] K. Takara, H.-H. Hasegawa, and D. J. Driebe, *Generalization of the Second Law for a Transition between Nonequilibrium States*, *Phys. Lett. A* **375**, 88 (2010).
- [18] M. Esposito and C. van den Broeck, *Second Law and Landauer Principle Far from Equilibrium*, *Europhys. Lett.* **95**, 40004 (2011).
- [19] T. Sagawa and M. Ueda, *Nonequilibrium Thermodynamics of Feedback Control*, *Phys. Rev. E* **85**, 021104 (2012).
- [20] H. Tasaki, *The Coefficient of Restitution Does Not Exceed Unity*, *J. Stat. Phys.* **123**, 1361 (2006).
- [21] C. Maes and H. Tasaki, *Second Law of Thermodynamics for Macroscopic Mechanics Coupled to Thermodynamic Degrees of Freedom*, *Lett. Math. Phys.* **79**, 251 (2007).
- [22] J.C. Maxwell, *Theory of Heat* (Longmans, London, 1871).
- [23] L. Szilárd, *Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen*, *Z. Phys.* **53**, 840 (1929).
- [24] R. Landauer, *Irreversibility and Heat Generation in the Computing Process*, *IBM J. Res. Dev.* **5**, 183 (1961).
- [25] O. Penrose, *Foundations of Statistical Mechanics: A Deductive Treatment* (Pergamon, Oxford, England, 1970).
- [26] C.H. Bennett, *The Thermodynamics of Computation—A Review*, *Int. J. Theor. Phys.* **21**, 905 (1982).
- [27] *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*, edited by H.S. Leff and A.F. Rex (Princeton University Press, Princeton, NJ, 2003).
- [28] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, *Experimental Verification of Landauer's Principle Linking Information and Thermodynamics*, *Nature (London)* **483**, 187 (2012).
- [29] H. T. Quan, Y. D. Wang, Y.-X. Liu, C. P. Sun, and F. Nori, *Maxwell's Demon Assisted Thermodynamic Cycle in Superconducting Quantum Circuits*, *Phys. Rev. Lett.* **97**, 180402 (2006).
- [30] M. Bier and F.J. Cao, *Szilard-Machine-like Features in a Processive Motor Protein*, *Acta Phys. Pol. B* **43**, 889 (2012).
- [31] D. Mandal and C. Jarzynski, *Work and Information Processing in a Solvable Model of Maxwell's Demon*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11641 (2012).
- [32] J.M. Horowitz, T. Sagawa, and J.M.R. Parrondo, *Imitating Chemical Motors with Optimal Information Motors*, *Phys. Rev. Lett.* **111**, 010602 (2013).
- [33] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, *Thermodynamics of a Physical Model Implementing a Maxwell Demon*, *Phys. Rev. Lett.* **110**, 040601 (2013).
- [34] A.C. Barato and U. Seifert, *An Autonomous and Reversible Maxwell's Demon*, *Europhys. Lett.* **101**, 60001 (2013).
- [35] D. Mandal, H. T. Quan, and C. Jarzynski, *Maxwell's Refrigerator: An Exactly Solvable Model*, *Phys. Rev. Lett.* **111**, 030602 (2013).
- [36] J. Earman and J.D. Norton, *Exorcist XIV: The Wrath of Maxwell's Demon. Part I. From Maxwell to Szilard*, *Stud. Hist. Phil. Mod. Phys.* **29**, 435 (1998).
- [37] J. Earman and J.D. Norton, *Exorcist XIV: The Wrath of Maxwell's Demon. Part II. From Szilard to Landauer and Beyond*, *Stud. Hist. Phil. Mod. Phys.* **30**, 1 (1999).
- [38] M. Hemmo and O.R. Shenker, *Maxwell's Demon*, *J. Philos.* **107**, 389 (2010) [<http://www.journalofphilosophy.org/articles/index.htm>].
- [39] J.D. Norton, *Waiting for Landauer*, *Stud. Hist. Phil. Mod. Phys.* **42**, 184 (2011).
- [40] T. Sagawa and M. Ueda, *Second Law of Thermodynamics with Discrete Quantum Feedback Control*, *Phys. Rev. Lett.* **100**, 080403 (2008).
- [41] T. Sagawa and M. Ueda, *Minimal Energy Cost for Thermodynamic Information Processing: Measurement and Information Erasure*, *Phys. Rev. Lett.* **102**, 250602 (2009).
- [42] T. Sagawa and M. Ueda, *Generalized Jarzynski Equality under Nonequilibrium Feedback Control*, *Phys. Rev. Lett.* **104**, 090602 (2010).
- [43] T. Sagawa, *Hamiltonian Derivations of the Generalized Jarzynski Equalities under Feedback Control*, *J. Phys. Conf. Ser.* **297**, 012015 (2011).
- [44] T. Sagawa, *Thermodynamics of Information Processing in Small Systems*, *Prog. Theor. Phys.* **127**, 1 (2012).
- [45] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, *Experimental Demonstration of Information-to-Energy Conversion and Validation of the Generalized Jarzynski Equality*, *Nat. Phys.* **6**, 988 (2010).
- [46] Y. Fujitani and H. Suzuki, *Jarzynski Equality Modified in the Linear Feedback System*, *J. Phys. Soc. Jpn.* **79**, 104003 (2010).
- [47] M. Ponnurugan, *Generalized Detailed Fluctuation Theorem under Nonequilibrium Feedback Control*, *Phys. Rev. E* **82**, 031129 (2010).
- [48] J.M. Horowitz and S. Vaikuntanathan, *Nonequilibrium Detailed Fluctuation Theorem for Repeated Discrete Feedback*, *Phys. Rev. E* **82**, 061120 (2010).

- [49] J.M. Horowitz and J.M.R. Parrondo, *Thermodynamic Reversibility in Feedback Processes*, *Europhys. Lett.* **95**, 10005 (2011).
- [50] J.M. Horowitz and J.M.R. Parrondo, *Designing Optimal Discrete-Feedback Thermodynamic Engines*, *New J. Phys.* **13**, 123019 (2011).
- [51] D. Abreu and U. Seifert, *Extracting Work from a Single Heat Bath through Feedback*, *Europhys. Lett.* **94**, 10001 (2011).
- [52] D. Abreu and U. Seifert, *Thermodynamics of Genuine Nonequilibrium States under Feedback Control*, *Phys. Rev. Lett.* **108**, 030601 (2012).
- [53] S. Lahiri, S. Rana, and A.M. Jayannavar, *Fluctuation Theorems in the Presence of Information Gain and Feedback*, *J. Phys. A* **45**, 065002 (2012).
- [54] S. Deffner and E. Lutz, *Information Free Energy for Nonequilibrium States*, [arXiv:1201.3888v1](https://arxiv.org/abs/1201.3888v1).
- [55] Y. Morikuni and H. Tasaki, *Quantum Jarzynski-Sagawa-Ueda Relations*, *J. Stat. Phys.* **143**, 1 (2011).
- [56] V. Vedral, *An Information-Theoretic Equality Implying the Jarzynski Relation*, *J. Phys. A* **45**, 272001 (2012).
- [57] D. Kafri and S. Deffner, *Holevo's Bound from a General Quantum Fluctuation Theorem*, *Phys. Rev. A* **86**, 044302 (2012).
- [58] K.H. Kim and H. Qian, *Entropy Production of Brownian Macromolecules with Inertia*, *Phys. Rev. Lett.* **93**, 120602 (2004).
- [59] K.H. Kim and H. Qian, *Fluctuation Theorems for a Molecular Refrigerator*, *Phys. Rev. E* **75**, 022102 (2007).
- [60] F.J. Cao and M. Feito, *Thermodynamics of Feedback Controlled Systems*, *Phys. Rev. E* **79**, 041118 (2009).
- [61] H. Suzuki and Y. Fujitani, *One-Dimensional Shift of a Brownian Particle under the Feedback Control*, *J. Phys. Soc. Jpn.* **78**, 074007 (2009).
- [62] K. Jacobs, *Second Law of Thermodynamics and Quantum Feedback Control: Maxwell's Demon with Weak Measurements*, *Phys. Rev. A* **80**, 012322 (2009).
- [63] F.J. Cao, L. Dinis, and J.M.R. Parrondo, *Feedback Control in a Collective Flashing Ratchet*, *Phys. Rev. Lett.* **93**, 040603 (2004).
- [64] M. Feito, J.P. Baltanás, and F.J. Cao, *Rocking Feedback-Controlled Ratchets*, *Phys. Rev. E* **80**, 031128 (2009).
- [65] F.J. Cao, M. Feito, and H. Touchette, *Information and Flux in a Feedback Controlled Brownian Ratchet*, *Physica (Amsterdam)* **388A**, 113 (2009).
- [66] S. Vaikuntanathan and C. Jarzynski, *Modeling Maxwell's Demon with a Microcanonical Szilard Engine*, *Phys. Rev. E* **83**, 061120 (2011).
- [67] B.J. Lopez, N.J. Kuwada, E.M. Craig, B.R. Long, and H. Linke, *Realization of a Feedback Controlled Flashing Ratchet*, *Phys. Rev. Lett.* **101**, 220601 (2008).
- [68] M. Bonaldi, L. Conti, P. De Gregorio, L. Rondoni, G. Vedovato, A. Vinante, M. Bignotto, M. Cerdonio, P. Falferi, N. Liguori, S. Longo, R. Mezzena, A. Ortolan, G.A. Prodi, F. Salemi, L. Taffarelli, S. Vitale, and J.-P. Zengri, *Nonequilibrium Steady-State Fluctuations in Actively Cooled Resonators*, *Phys. Rev. Lett.* **103**, 010601 (2009).
- [69] L. Granger and H. Kantz, *Thermodynamic Cost of Measurements*, *Phys. Rev. E* **84**, 061110 (2011).
- [70] H. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, New York, 1985).
- [71] E. Fermi, *Thermodynamics* (Blackie & Son, London, 1938).
- [72] L.D. Landau and M.E. Lifshitz, *Statistical Physics* (Pergamon, Oxford, England, 1980), 3rd ed.
- [73] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980), 2nd ed.
- [74] For the purpose of the present discussion, in the thermodynamic limit, it is not particularly relevant whether we view the equilibrium distribution to be microcanonical or canonical.
- [75] A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949).
- [76] J.R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics* (Cambridge University Press, Cambridge, England, 1999).
- [77] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier Science B.V., Amsterdam, 1992).
- [78] In general, of course, information can be stored by other means, for instance, using trits—trinary digits—rather than bits. Our choice of using bits is motivated by convenience and familiarity, and does not restrict the validity of our conclusions.
- [79] C.H. Bennett, *Notes on Landauer's Principle, Reversible Computation, and Maxwell's Demon*, *Stud. Hist. Phil. Mod. Phys.* **34**, 501 (2003).
- [80] S. Kullback, *Information Theory and Statistics* (Peter Smith, Gloucester, MA, 1978).
- [81] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 1991).
- [82] H. Qian, *Relative Entropy: Free Energy Associated with Equilibrium Fluctuations and Nonequilibrium Deviations*, *Phys. Rev. E* **63**, 042103 (2001).
- [83] G.E. Crooks, *Beyond Boltzmann-Gibbs Statistics: Maximum Entropy Hyperensembles Out of Equilibrium*, *Phys. Rev. E* **75**, 041119 (2007).
- [84] H. Ge and H. Qian, *Physical Origins of Entropy Production, Free Energy Dissipation, and Their Mathematical Representations*, *Phys. Rev. E* **81**, 051133 (2010).
- [85] S. Still, D.A. Sivak, A.J. Bell, and G.E. Crooks, *Thermodynamics of Prediction*, *Phys. Rev. Lett.* **109**, 120604 (2012).
- [86] Yu.L. Klimontovich, *Entropy and Information of Open Systems*, *Phys. Usp.* **42**, 375 (1999) [<http://iopscience.iop.org/1063-7869/42/4/A05/>].
- [87] H. Tasaki, *Unified Jarzynski and Sagawa-Ueda Relations for Maxwell's Demon*, [arXiv:1308.3776v1](https://arxiv.org/abs/1308.3776v1).
- [88] A.C. Barato and U. Seifert, *Unifying Three Perspectives on Information Processing in Stochastic Thermodynamics*, [arXiv:1308.4598v1](https://arxiv.org/abs/1308.4598v1).
- [89] W.H. Zurek, *Decoherence, Einselection, and the Quantum Origin of the Classical*, *Rev. Mod. Phys.* **75**, 715 (2003).