

**Microcanonical Szilárd engines beyond the quasistatic regime**

Thiago V. Acconcia\* and Marcus V. S. Bonança†

*Instituto de Física “Gleb Wataghin,” Universidade Estadual de Campinas, 777 Rua Sérgio Buarque de Holanda, Campinas, São Paulo 13083-859, Brazil*

(Received 2 June 2016; revised manuscript received 30 September 2017; published 14 December 2017)

We discuss the possibility of extracting energy from a single thermal bath using microcanonical Szilárd engines operating in finite time. This extends previous works on the topic which are restricted to the quasistatic regime. The feedback protocol is implemented based on linear response predictions of the excess work. It is claimed that the underlying mechanism leading to energy extraction does not violate Liouville’s theorem and preserves ergodicity throughout the cycle. We illustrate our results with several examples including an exactly solvable model.

DOI: [10.1103/PhysRevE.96.062117](https://doi.org/10.1103/PhysRevE.96.062117)**I. INTRODUCTION**

Over the last 20 years there has been an intense research activity on the interplay between information and thermodynamics. This field of investigation dates back to Maxwell, who introduced a *gedanken* experiment, where a being, the so-called Maxwell’s demon, is able to violate the second law based on the knowledge acquired about the microscopic states of the system of interest [1–3]. This paradox was later on reformulated by Szilárd [4], who devised an engine able to extract energy from a single thermal bath. In Szilárd’s setup, not only the information gathered by the demon is more easily quantified but also its role on the conversion of heat into work becomes more transparent. Hence, it has become crucial to understand how thermodynamic entropy and information are related.

Major efforts on this direction have been made, for instance, by Brillouin, Landauer, Bennett, and Penrose [5–8], who have discussed the energetic costs of measurement and measurement erasure. Along these lines, the demon operating the Szilárd engine needs a memory device to store the information gathered by him. To reset this memory at the end of the cyclic process, the demon has to dissipate an amount of energy that is generally greater than the extracted work. Hence, the second law is rescued within this scenario by attributing an entropic cost for the measurement erasure. Besides the conceptual debate, several experimental setups have recently analyzed these issues [9–12].

In the last decade, feedback-controlled processes such as those performed by a Maxwell’s demon have been incorporated into the frameworks of fluctuation theorems [13–19] and stochastic thermodynamics [20–24], leading to significant progress on the history of information in thermodynamics. More recently, the demon itself has been modeled as a physical system, yielding an antonymous formulation of the original paradox [25–31]. In this new point of view, a self-contained universe is composed of a device, the thermal reservoirs, a work source, and an information reservoir. This composite system evolves autonomously, and any effective feedback control arises from the interplay of the different subsystems.

In the present work, we focus on a particular kind of feedback control, namely, the microcanonical Szilárd engine [32,33]. In this setup, the demon performs an energy measurement on a system initially equilibrated with a heat bath. The demon then acts according to the outcome of energy measurement, keeping the system *isolated* from the reservoir. Effectively, one may think of a demon acting on a system initially prepared in a microcanonical ensemble. The examples of microcanonical Szilárd engines available in the literature so far deal with the quasistatic regime, which means zero power extracted by the demon. Hence, it is highly desirable to construct examples operating in finite time.

In the quasistatic regime, the energy extraction is a consequence of an effective symmetry breaking [12,34,35]. Nevertheless, we show that it is also possible to extract energy in finite time without splitting the phase space in two or more disconnected regions. Moreover, we show that for certain protocols the work performed in finite time can be equal to the quasistatic work. These are essentially the basic ingredients we use to construct microcanonical Szilárd engines producing finite power.

The outline of the paper is as follows. In Sec. II we discuss which are the essential features of our construction of microcanonical Szilárd engines. In Sec. III, these features are illustrated with an exactly solvable model; an interpretation of our results in terms of phase space is given in Sec. IV. Additional examples and final remarks are presented in Secs. V and VI, respectively.

**II. MICROCANONICAL SZILÁRD ENGINES IN FINITE TIME**

It will be shown in Secs. III and V that, for a certain class of systems, it is possible to design *finite-time* cyclic processes to extract energy from a single heat bath on average. Thus, any intermediate step in the cycle occurs in finite time. Besides, such cyclic processes require *feedback*; i.e., they depend crucially on the information gathered by the external observer or demon via an *energy* measurement. In particular, it will be shown that these cycles are two-stage processes such that each stage is a *linear* variation of a single external control parameter taking time intervals  $\tau_1$  and  $\tau_2$ . The energy extraction is possible only when these two switching times are carefully chosen based on the outcome of the energy

\*thiagova@if.unicamp.br

†mbonanca@if.unicamp.br

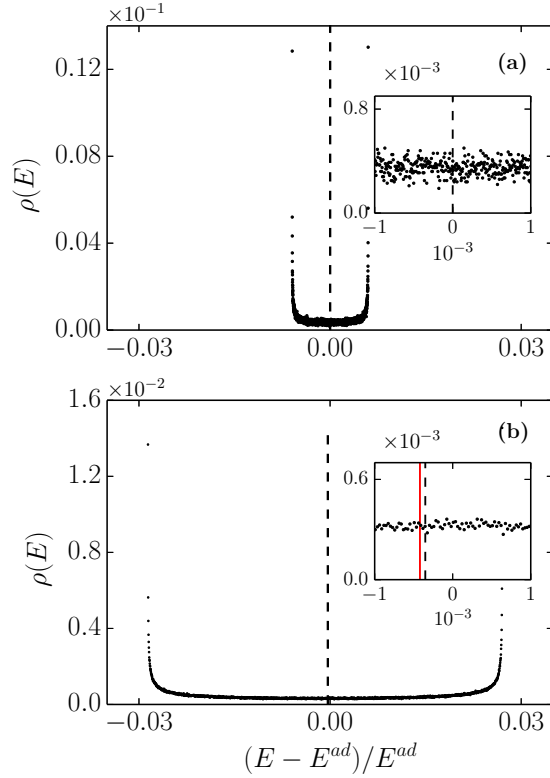


FIG. 1. Energy distributions after a finite-time driving of system (1) using the linear protocol  $\lambda(t) = \lambda_0 + \delta\lambda(t - t_0)/\tau$ , with  $\delta\lambda/\lambda_0 = 0.1$ . Vertical dashed lines represent the average energy for  $10^6$  initial conditions. The energy was rescaled in terms of the  $E^{ad}$ , the energy corresponding to the quasistatic evolution of the protocol (see Appendix A). (a) Energy distribution after a single linear switching such that  $W_{ex} = 0$  (with  $\omega\tau \approx 1.8$ ; see Fig. 2). (b) Energy distribution after a cyclic switching of  $\lambda$  in two linear steps, first with  $W_{ex} = 0$  ( $\omega\tau_1 \approx 1.8$ ) and after with  $W_{ex} < 0$  ( $\omega\tau_2 \approx 2.4$ ; see Fig. 2). The values of  $\omega\tau_{1,2}$  were calculated using Eq. (8). The continuous vertical red line corresponds to the linear response prediction for the average work along the cycle.

measurement. In Sec. III, we show how to predict these values of  $\tau$  for an exactly solvable model using a linear response approach for the excess work [see Eqs. (6) and (8)]. In Sec. V, we claim that  $\tau_{1,2}$  can also be obtained from numerical simulations of noncyclic process performed on anharmonic oscillators.

We shall first describe the general features of the cyclic processes we are going to perform. First, the system representing the engine is put in contact with a reservoir at temperature  $T$ . After its relaxation, the system is found in a Boltzmann-Gibbs distribution  $\rho_{eq}(\Gamma, \lambda) = \exp[-\beta\mathcal{H}(\Gamma; \lambda)]/\mathcal{Z}(\beta, \lambda)$ , where  $\Gamma$  is a point in phase space,  $\lambda$  is an external control parameter,  $\mathcal{H}$  is the system's Hamiltonian, and  $\beta = (k_B T)^{-1}$ , with  $k_B$  being the Boltzmann constant. The quantity expressed by  $\mathcal{Z}(\beta, \lambda)$  is the partition function of the system.

The engine is then decoupled from the heat bath, and the demon measures its energy, causing the collapse of the initial Boltzmann-Gibbs distribution into a microcanonical one. After that, the engine will be driven by the demon in a two-stage process while it is kept *isolated* from the reservoir. We denote

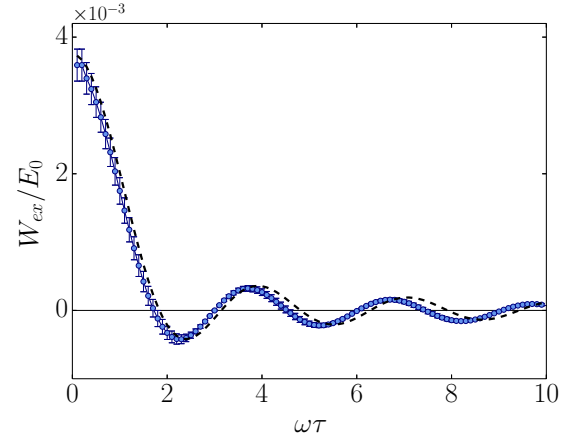


FIG. 2. Comparison between numerical results (blue circles) and linear response prediction [Eq. (8)] (dashed line) of  $W_{ex}$  as a function of the switching time  $\tau$  for system (1). The linear protocol  $g(t) = (t - t_0)/\tau$  and  $\delta\lambda/\lambda_0 = 0.1$  was chosen. We used  $10^6$  initial conditions to calculate numerically the excess work for each  $\tau$ .

by step 1 the part of the cyclic process in which the system starts with a well-defined energy  $E_1$  and the external parameter is driven in *finite time* from  $\lambda_1$  to  $\lambda_2$ . This process takes a time interval  $\tau_1$  that has been carefully chosen by the demon based on the value of  $E_1$ . To complete the cycle, the external parameter is driven in the opposite direction, i.e., from  $\lambda_2$  to  $\lambda_1$ . This step 2 takes a time interval  $\tau_2$  that is different from  $\tau_1$  but that also depends crucially on the energy measured by the demon.

In step 1, the demon chooses a particular protocol and finite switching time  $\tau_1$  so that the average work performed is the *quasistatic* work  $W_{qs}$ . In other words, after several realizations of the cycle, the work performed along this stage is, on average, equal to the value obtained after a quasistatic switching from  $\lambda_1$  to  $\lambda_2$ . This might sound wrong because we usually think that only quasistatic processes yield work equal to the quasistatic value. However, it has already been discussed elsewhere [36,37] that this is not always the case. Numerical evidence of this fact is presented in Fig. 1(a). There, the energy distribution after a finite-time process is shown to give an average value exactly equal to the energy predicted by the adiabatic invariant (see Appendix A) for a quasistatic process. We give more details in the next section.

In step 2, the demon chooses another particular protocol and finite switching time  $\tau_2$  to drive  $\lambda$  from  $\lambda_2$  to  $\lambda_1$ . This process is such that the average work  $W$  is equal to the negative of the quasistatic work  $W_{qs}$  obtained in step 1 plus the excess work  $W_{ex}$  (see, for instance, Ref. [36] for a discussion about  $W_{ex}$ ). The protocol and switching time  $\tau_2$  are such that  $W_{ex} < 0$  (see Fig. 2 and Sec. V for examples). Hence, the net work that has been performed in the cycle is negative, meaning that energy was extracted,  $W_{cycle} = W_{ex} < 0$  [see Fig. 1(b)]. It is noteworthy that, in contrast to the corresponding quasistatic engines [32,33], no ergodicity breaking is necessary; i.e., it is not necessary to split the phase space into disconnected regions. It will become clear in what follows how the information gathered by the demon

determines the values of  $\tau$  for which the above-mentioned features happen for each part of the cycle.

In summary, our finite-time microcanonical Szilárd engine is based on the existence of *finite*-time protocols leading to  $W = W_{qs}$  and  $W_{ex} < 0$  according to the value of the switching time. It remains to be explained how it is possible to find such protocols and why (or more precisely for which class of systems) they yield such features. Based on extensive numerical investigation, we claim that several anharmonic oscillators of one degree of freedom share these properties (see Sec. V). In the next section, we take an analytically solvable example to show how to find the protocols we are interested in. It will be shown that all we have to do is to study the behavior of the excess work,  $W_{ex} \equiv W - W_{qs}$ , as a function of switching time  $\tau$  for *non*-cyclic processes. Every time  $W_{ex}$  is zero or negative for a *finite*  $\tau$ , we have found protocols for steps 1 and 2, respectively.

### III. EXACTLY SOLVABLE MODEL

We will consider in this section the system

$$\mathcal{H}(t) = \epsilon^{-1} H_{HO}^2 = \frac{1}{\epsilon} \left[ \frac{p^2}{2m} + \lambda(t) \frac{x^2}{2} \right]^2, \quad (1)$$

where  $H_{HO}$  is the harmonic oscillator Hamiltonian,  $\epsilon$  is a constant, and  $\lambda(t)$  is the externally controlled parameter whose time dependence is expressed in general by  $\lambda(t) = \lambda_0 + \delta\lambda g(t)$ , where  $g(t)$  is such that  $g(t_0) = 0$  and  $g(t_f) = 1$ ,  $\tau \equiv t_f - t_0$ . The solutions of Hamilton's equations are

$$x(t) = x(t_0) \cos(\omega t) + \frac{p(t_0)}{m\omega_0} \sin(\omega t), \quad (2a)$$

$$p(t) = p(t_0) \cos(\omega t) - m\omega_0 x(t_0) \sin(\omega t), \quad (2b)$$

where  $\omega_0 = \sqrt{\lambda_0/m}$  and  $\omega = 2\omega_0\sqrt{E/\epsilon}$  is the natural frequency of oscillations for an energy  $E$ . They describe an oscillatory motion with an energy-dependent angular frequency and whose phase space is bounded as, for instance, a pendulum.

The thermodynamic work  $W$  performed when the external parameter  $\lambda$  is switched from  $\lambda_0$  to  $\lambda_f = \lambda_0 + \delta\lambda$  reads

$$W = \int_{t_0}^{t_f} dt \dot{\lambda} \overline{\frac{\partial \mathcal{H}}{\partial \lambda}}, \quad (3)$$

where  $\overline{A}$  is the out-of-equilibrium average of the quantity  $A$ .

Restricting ourselves to processes in which  $\delta\lambda/\lambda_0 \ll 1$  (but which are not necessarily slow), we can apply linear response theory to relate the out-of-equilibrium average  $\overline{\partial \mathcal{H}/\partial \lambda}$  to its corresponding relaxation function  $\Psi_0(t)$  [38,39]. For a system initially in a microcanonical equilibrium distribution whose energy is  $E_0 = \mathcal{H}(t_0)$ ,  $\Psi_0(t)$  is given by [40]

$$\Psi_0(t) = \frac{1}{Z(\lambda_0, E_0)} \frac{\partial}{\partial E_0} [Z(\lambda_0, E_0)(C(t) - C)], \quad (4)$$

where  $Z(\lambda_0, E_0) = \int dx dp \delta[E_0 - \mathcal{H}(x, p; \lambda_0)]$ ,  $C(t) = \langle \partial \mathcal{H}(0)/\partial \lambda \partial \mathcal{H}(t)/\partial \lambda \rangle_0$ , and  $C$  is the asymptotic value, for large  $t$ , of the correlation function  $C(t)$  (for simply oscillatory functions, this has to be properly defined; see, for instance,

Sec. 4.2.2 of [39]). We denote by  $\langle A \rangle_0$  the microcanonical average of the observable  $A$ .

After some algebra linking Eq. (3) to the relaxation function (4) (see Appendix A), we obtain an expression for  $W$  given by the sum of two contributions. The first one can be identified with the quasistatic work along a equivalent quasistatic process and reads

$$W_{qs} = \delta\lambda \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_0 - \frac{(\delta\lambda)^2}{2} \tilde{\Psi}_0(0), \quad (5)$$

where  $\tilde{\Psi}_0 \equiv \Psi_0(0) - \chi_0^\infty$ , with  $\chi_0^\infty = \langle \partial^2 \mathcal{H}/\partial \lambda^2 \rangle_0$ .

It is worth notice that both terms *do not* depend on the protocol  $g(t)$ . Indeed, it can be verified that they are the first terms of the series expansion of the quasistatic work for  $\delta\lambda/\lambda_0 \ll 1$ .

The second contribution for the total work is a term that vanishes in the quasistatic limit and clearly depends on  $g(t)$ . Therefore, it is called excess work and reads [36]

$$W_{ex} = \frac{(\delta\lambda)^2}{2} \int_{t_0}^{t_f} dt \int_{t_0}^{t_f} dt' \Psi_0(t-t') \dot{g}(t) \dot{g}(t'). \quad (6)$$

In other words,  $W_{ex} \equiv W - W_{qs}$  is the extra amount of energy the external agent has to pump into the system during a finite-time process.

The previous analysis about the linear response expression for  $W$  is very general and can now be applied to system (1). The correlation function  $C(t)$  can be calculated analytically from the previous expressions for  $x(t)$  and  $p(t)$  [Eq. (2)], yielding the following relaxation function:

$$\Psi_0(t) = \frac{E}{4\lambda_0^2} [3 \cos(2\omega t) - 2\omega t \sin(2\omega t)]. \quad (7)$$

Finally, we can calculate  $W_{ex}$  analytically for a given protocol  $g(t)$  using Eq. (6). In particular, for the linear protocol  $g(t) = (t - t_0)/\tau$  the result reads

$$W_{ex}(\tau) = \frac{E_0}{8} \left( \frac{\delta\lambda}{\lambda_0} \right)^2 \frac{\sin(\omega\tau) [(2\omega\tau) \cos(\omega\tau) + \sin(\omega\tau)]}{(\omega\tau)^2}. \quad (8)$$

The previous expression has several important features. As mentioned before, it goes to zero when we take the limit  $\omega\tau \rightarrow \infty$ . For large values of  $\omega\tau$ , we find that  $W_{ex} \propto 1/\tau$ , which can be related to the Sekimoto-Sasa relation presented in Refs. [41,42]. Equation (8) also reveals that  $W_{ex}$  can be null for *finite* values of  $\tau$ . This happens either (i) when  $\sin(\omega\tau)$  is zero, i.e.,  $\omega\tau = 2\pi k$ , with  $k$  an integer, or (ii) when  $\tan(\omega\tau)/2 = -\omega\tau$ . It also predicts *negative* values of  $W_{ex}$  for specific ranges of  $\tau$ . Figure 2 shows the comparison between Eq. (8) and numerical simulations using the linear protocol. Numerics were implemented using symplectic integrators [43].

The agreement between theoretical prediction for  $W_{ex}$  and the numerical calculation is very good. Although the theoretical prediction is restricted to weak driving, i.e.,  $\delta\lambda/\lambda_0 \ll 1$ , it includes arbitrarily fast processes. Besides, Eq. (8) predicts the special switching times for which the excess work is either zero or negative. Additionally, we have investigated numerically the behavior of the excess work up to  $\delta\lambda/\lambda_0 = 1.0$ . Figure 3 shows

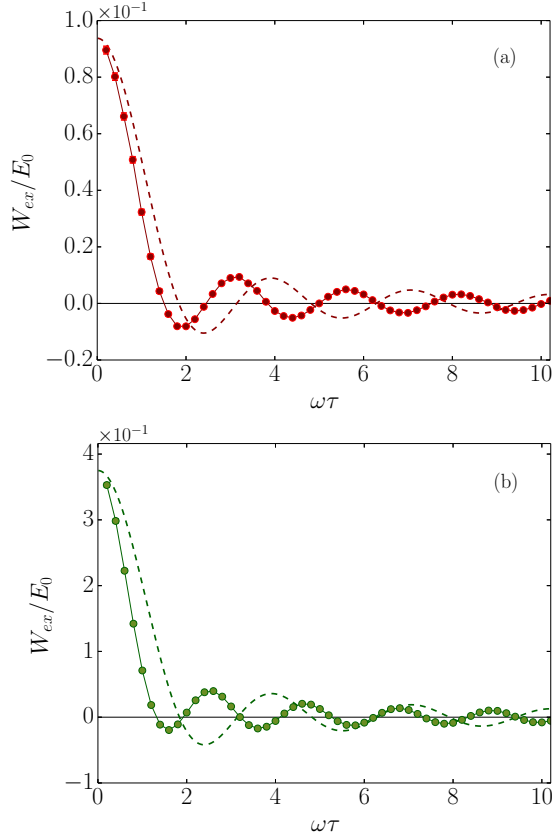


FIG. 3. Comparison between numerical calculations (circles) and linear response predictions [Eq. (8)] (dashed line) of  $W_{ex}$  as a function of the switching time  $\tau$  for system (1). The driving was performed using the linear protocol  $g(t) = (t - t_0)/\tau$ , with (a)  $\delta\lambda/\lambda_0 = 0.5$  and (b)  $\delta\lambda/\lambda_0 = 1.0$ . We used  $10^6$  initial conditions to calculate numerically the excess work for each  $\tau$ .

that the linear response expression clearly deviates from the numerical calculations as  $\delta\lambda/\lambda_0$  increases. However, it also shows that negative values of excess work are not restricted to our linear response regime. This suggests that our construction of a microcanonical Szilárd engine might be valid beyond the weak driving regime. We must be careful about this because, as will be explained next, our proposal depends crucially on how small energy fluctuations are after the first step of the cycle and we have not studied them in detail.

In summary, it is possible to implement a microcanonical Szilárd engine driving system (1) with linear protocols as follows. In step 1, the control parameter is driven from  $\lambda_0 + \delta\lambda$  to  $\lambda_0$  in a time interval  $\tau_1$  yielding the first zero of  $W_{ex}$ . Since the energy distribution after this step is a very narrow one (see Fig. 1), we assume that the initial energy distribution for step 2 is again a microcanonical one. Hence, the linear response results for  $W_{ex}$  can be used to predict what is going to happen in both steps of the cycle. This is the reason why we first perform a process with  $W_{ex} = 0$ . Although the inset in Fig. 1(b) shows that this is indeed a good approximation, this might not be the case beyond the linear response regime. In step 2, another linear protocol drives  $\lambda$  back to  $\lambda_0$ , taking a time interval  $\tau_2$  such that  $W_{ex} < 0$ . Thus, the net work performed in the cycle is  $W_{ex}(\tau_2)$ . In Fig. 1(b), we show the numerical result

for the average energy after this cycle. Since  $\omega$  is a function of the energy, the duration of each protocol is chosen based on the information gathered by the demon about the initial energy.

In contrast to the quasistatic regime, the finite-time driving of the system leads to different values of work after each single realization of the protocol. Hence, we want to stress that our results are valid on average. Indeed, every time the demon starts a new cycle, the outcome of the energy measurement differs from the previous one. To obtain the curve in Fig. 2, the demon has then to sort his ensemble of trajectories by their initial energy values and take an average using only trajectories with essentially the same initial energy. Instead, if all trajectories are considered, the work performed in a cycle is, on average, equal to

$$\begin{aligned} \langle W_{\text{cycle}} \rangle &= \frac{1}{Z(\beta)} \int_0^\infty dE e^{-\beta E} Z(E) W_{ex}[E, \tau_2(E)] \\ &= -\frac{1}{16\beta} \left( \frac{\delta\lambda}{\lambda_0} \right)^2 |f(\omega\tau_2)|, \end{aligned} \quad (9)$$

where  $W_{ex}(\tau_2)$  is the value of the first minimum of Eq. (8),  $f(\omega\tau) = \sin(\omega\tau)[2\omega\tau \cos(\omega\tau) + \sin(\omega\tau)]/(\omega\tau)^2$  and  $Z(E)$  is the density of states. Since  $\omega\tau_2 \approx 2.4$  does not depend on the energy,  $f(\omega\tau_2) \approx -0.34$  can be taken outside the integral. It is worth mentioning that the demon can always find a suitable protocol yielding  $W_{ex} < 0$  no matter the value of the energy measured.

#### IV. PHASE SPACE DYNAMICS

After presenting the linear response description of finite-time microcanonical Szilárd engines, we shall discuss the physical mechanism behind the energy extraction in our setup. The protocols we have discussed previously never split the phase space of the system in two or more disconnected parts [12,34,35]. Nevertheless, Fig. 1 shows that, after a finite-time process, the energy distribution is essentially zero except for two particular values of energy; i.e., the system basically assumes either one or the other value of energy.

It is necessary to demand a bit more from the energy distribution in order to extract energy. The balance between the two peaks of the distribution has to be such that the average energy is *smaller* than the energy obtained after an equivalent quasistatic process; otherwise,  $W_{ex}$  never becomes negative. This seems to be impossible for a harmonic oscillator, as shown below. Its energy distribution after a finite-time driving also presents two pronounced peaks (see Fig. 4), but  $W_{ex}$  is always nonnegative [see Eq. (12) and Fig. 5]. Our claim about why the anharmonic oscillators yield  $W_{ex} < 0$  is based on the phase-space plots in Fig. 6 for the quartic oscillator of Eq. (13).

The initial energy shell is deformed along the momentum and coordinate directions when the system is driven quasistatically. Nevertheless, our numerical simulations indicate that the stretching and contraction directions rotate when an anharmonic oscillator is driven in finite time. This happens in such way that the portions of the deformed curve with energy below the energy after the quasistatic process are slightly larger than those with energy higher than that.



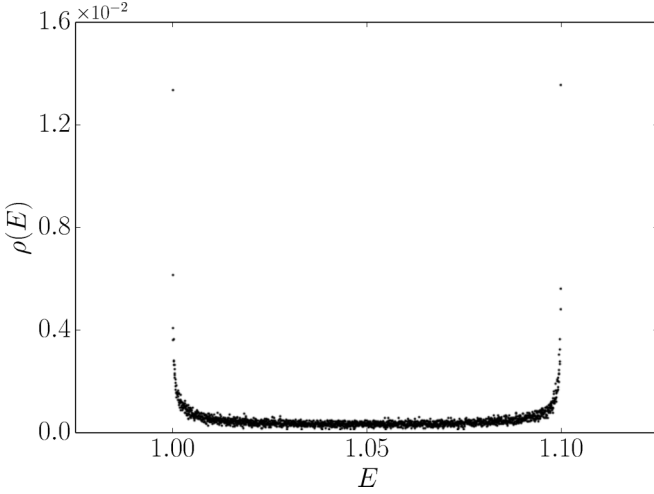


FIG. 4. Energy distribution of the harmonic oscillator (10) after a finite-time linear switching of  $\lambda$ . We set  $\delta\lambda/\lambda_0 = 0.1$  and  $\tau = 0.1\omega_0^{-1}$ . We used  $5 \times 10^4$  initial conditions.

### Harmonic oscillator

We briefly discuss now why the harmonic oscillator does not reproduce the behavior we observe in Fig. 2. For the Hamiltonian

$$H_{HO} = \frac{p^2}{2m} + \lambda(t)\frac{x^2}{2}, \quad (10)$$

the relaxation function can be obtained as described previously, and it reads

$$\Psi_0(t) = \frac{E}{4\lambda_0^2} \cos(2\omega_0 t). \quad (11)$$

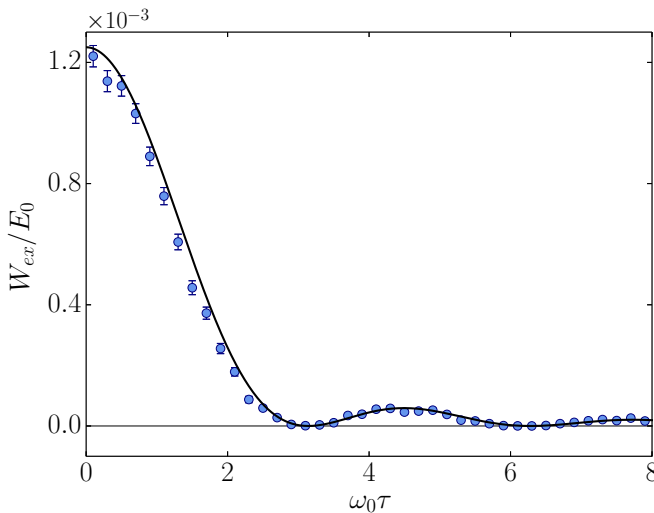


FIG. 5. Excess work as a function of the switching time for the harmonic oscillator [Eq. (10)]. We compare numerical calculations (blue circles) and linear response prediction (solid line) when  $g(t) = (t - t_0)/\tau$ . We used  $10^6$  initial conditions for each switching time  $\tau$  and fixed  $\delta\lambda/\lambda_0 = 0.1$ .

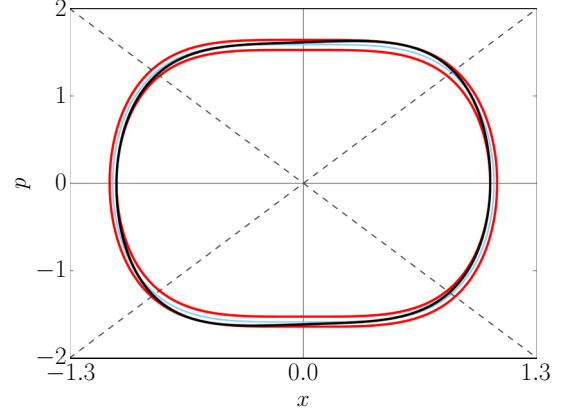


FIG. 6. Phase space distribution for the quartic oscillator [Eq. (13)]. The black curve represents the deformed energy shell after the finite-time protocol yielding  $W_{ex} < 0$  (see Fig. 11). The outer and inner red (dark gray) curves are the energy shells corresponding, respectively, to the maximum and minimum values of the energy distribution (see Fig. 11) after the following finite-time protocol:  $g(t) = (t - t_0)/\tau$ ,  $\delta\lambda/\lambda_0 = 1.0$ ,  $\lambda_0 = 1.0$ , and  $W_{ex}(\tau) < 0$ . The blue (light gray) curve corresponds to the energy shell obtained after the corresponding quasistatic driving.

Equations (6) and (11) then yield

$$\begin{aligned} W_{ex} &\propto \int_0^1 ds \int_0^1 du \cos[2\omega_0\tau(s-u)]\dot{g}(s)\dot{g}(u) \\ &= \left[ \int_0^1 ds \cos(2\omega_0\tau s)\dot{g}(s) \right]^2 \\ &\quad + \left[ \int_0^1 du \sin(2\omega_0\tau u)\dot{g}(u) \right]^2. \end{aligned} \quad (12)$$

The excess work is therefore always positive for the harmonic oscillator. In other words, we could use this system to obtain zero excess work in finite time, but it would be impossible to model a microcanonical Szilárd engine with it along the same lines presented previously. Figure 5 shows the comparison between numerical calculations and the linear response expression for the linear protocol.

Figure 4 shows the energy distribution of the harmonic oscillator after a linear switching of  $\lambda$ . Although it presents the same structure we find in anharmonic oscillators, it does not favor the lower value of energy, and, consequently, it does not give rise to negative values of  $W_{ex}$ .

Figure 7 shows the phase-space representation of the states leading to the energy distribution in Fig. 4. Since the switching time was chosen such that  $W_{ex} > 0$ , the initial set of phase-space points evolves to a curve that, although very close, is not an energy shell. The initial energy shell deforms along the momentum and coordinate axis even in finite time. This contrasts with what happens in anharmonic oscillators where the stretching and contraction directions rotate and coincide with the momentum and coordinate axis only in the quasistatic regime.

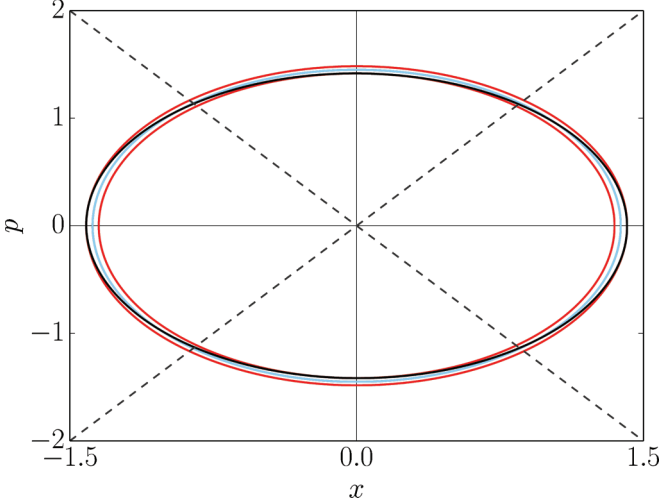


FIG. 7. Deformation of the initial energy shell of system (10) after a finite-time driving of  $\lambda$ . The outer and inner red (dark gray) curves are energy shells corresponding, respectively, to the maximum and minimum values of the energy distribution in Fig. 4. The black curve represents the finite-time deformation of the initial energy shell after the linear protocol  $g(t) = (t - t_0)/\tau$ , with  $\delta\lambda/\lambda_0 = 0.1$  and  $\omega_0\tau = 0.1$ . The blue (light gray) curve corresponds to the energy shell obtained after a quasistatic driving. We used  $5 \times 10^4$  initial conditions for each curve.

## V. OTHER EXAMPLES

The behavior of  $W_{ex}$  for the system (1) is typically found in other anharmonic oscillators. In what follows, we show numerical calculations of  $W_{ex}$  for a few examples: the quartic oscillator,

$$\mathcal{H}[\lambda(t)] = \frac{p^2}{2m} + \lambda(t)\frac{x^4}{2}, \quad (13)$$

the pendulum,

$$\mathcal{H}[\lambda(t)] = \frac{p^2}{2m} + 2\lambda(t)\sin^2\left(\frac{x}{2}\right), \quad (14)$$

and the logarithmic oscillator,

$$\mathcal{H}[\lambda(t)] = \frac{p^2}{2m} + \lambda(t)\ln\left(\frac{x^2 + b^2}{b^2}\right), \quad (15)$$

where  $b$  is a fixed parameter.

The corresponding excess work as a function of the switching time is presented in Figs. 8 to 10 for a linear switching of  $\lambda$ . We would like to highlight that the system (15) was used as an ideal Hamiltonian thermostat in [44]. We also present the energy distribution of the system (13) in Fig. 11 to stress the average negative excess work after the finite-time driving for the particular switching time  $\omega\tau \approx 2.6$ .

## VI. CONCLUDING REMARKS

In conclusion, we have shown how to construct finite-time microcanonical Szilárd engines for systems whose excess work presents finite-time zeros and negative values in non-cyclic processes. The exact values of switching times allowing for such values of excess work have to be carefully chosen

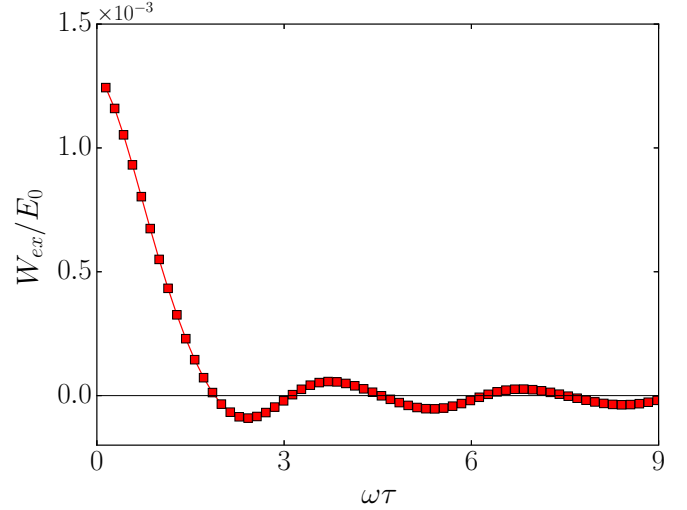


FIG. 8. Excess work as a function of the switching time for the system (13) using the linear protocol  $g(t) = (t - t_0)/\tau$ ,  $\delta\lambda/\lambda_0 = 0.1$ , and  $10^6$  initial conditions for each value of  $\tau$ .

based on the information gathered by the demon. In contrast to previous examples available in the literature, these engines produce finite power on average. The cyclic process was obtained using a linear response approach which can be further explored to furnish the optimal protocols yielding the maximum energy extraction. Despite of the unexpected absence of ergodicity breaking, we have shown that it is still possible to extract energy due to special features of the energy distribution. This question deserves further investigation, especially if one wants to extend the present results to systems with a larger number of degrees of freedom.

The energy extraction in Szilárd engines is usually attributed to the sudden reduction of phase space volume due to a symmetry breaking [12,34,35]. In the present case, although there is no such mechanism, we may have a reduction of phase space volume on average. Since the volume  $\Omega$  enclosed by an

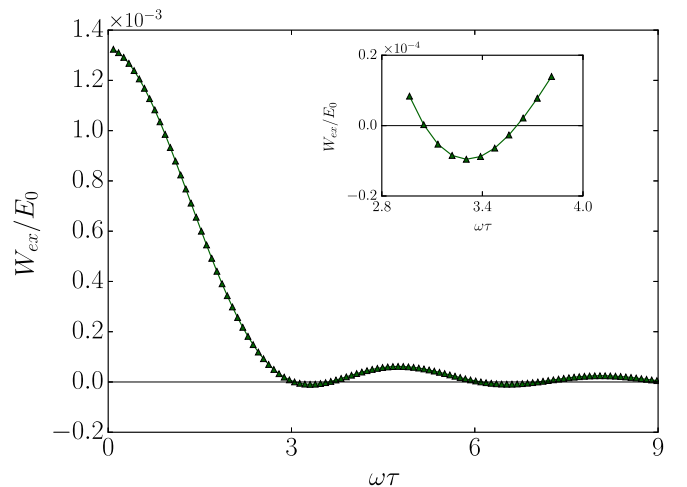


FIG. 9. Excess work as a function of the switching time for the system (14) using the linear protocol  $g(t) = (t - t_0)/\tau$ ,  $\delta\lambda/\lambda_0 = 0.1$ , and  $10^6$  initial conditions for each value of  $\tau$ .

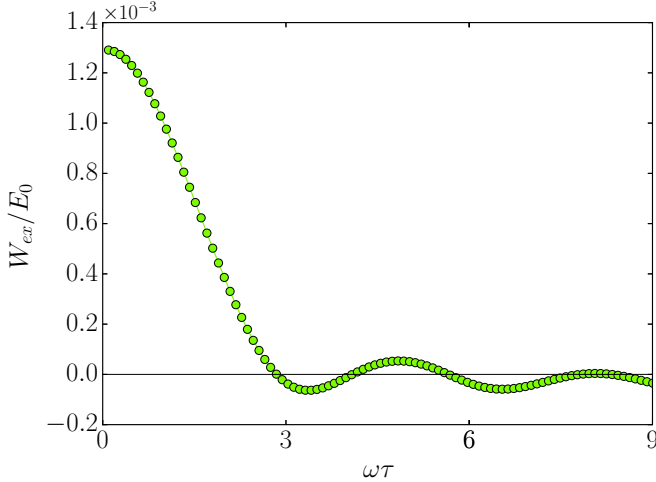


FIG. 10. Excess work as a function of the switching time for the system (15) using the linear protocol  $g(t) = (t - t_0)/\tau$ ,  $\delta\lambda/\lambda_0 = 0.1$ , and  $10^6$  initial conditions for each value of  $\tau$ .

energy shell is a monotonic increasing function of the energy  $E$ , there is a one-to-one mapping between  $\Omega$  and  $E$ . Thus, from an energy distribution as Fig. 1, it is possible to obtain the corresponding distribution of  $\Omega$ , whose average value is going to be smaller than the initial one every time  $W_{ex} < 0$ . In this sense, we might connect the energy extraction to a reduction of phase space volume on average.

As a final remark, we should mention that the quantification of the information gathered by the demon in our examples does not seem as straightforward as in Ref. [33]. This analysis will be left for a future work. Nevertheless, it is clear that we have shown examples of a feedback process in which the outcome of an energy measurement allows for energy extraction.

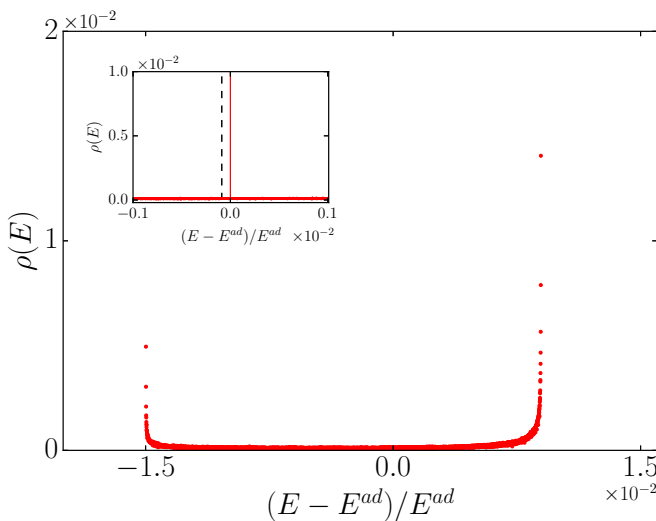


FIG. 11. Energy distribution (see Appendix A for a definition of  $E^{ad}$ ) after a finite-time driving of system (13) using a linear protocol  $g(t) = (t - t_0)/\tau$  such that  $\delta\lambda/\lambda_0 = 1.0$  and  $W_{ex}(\tau) < 0$ . We used  $10^6$  initial conditions. The values of energy were rescaled by the corresponding value  $E^{ad}$  after a quasistatic driving (see Appendix A).

## ACKNOWLEDGMENTS

The authors acknowledge J. P. Pekola, P. Talkner, M. Campisi, S. Deffner, and W. H. Zurek for enriching discussions. We especially thank É. Roldán for the careful analysis of the manuscript and inspiring discussions. T.A. acknowledges support from “Gleb Wataghin” Physics Institute (Brazil) and CAPESP (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) (Brazil) (Grant No. 1504869). M.B. acknowledges financial support from FAEPEX (Fundo de Apoio ao Ensino, à Pesquisa e Extensão, Universidade Estadual de Campinas) (Brazil) (Grant No. 0031/15) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) (Brazil) (Grant No. 2016/01660-2).

## APPENDIX A: ADIABATIC INVARIANT

We define here what we mean by  $E^{ad}$ , the energy we use to rescale the energy distributions in Figs. 1 and 11 and to calculate the work  $W_{qs}$  along a quasistatic process. First, it is necessary to define the “volume”  $\Omega(E, \lambda)$  enclosed by the surface of constant energy  $H(\mathbf{q}, \mathbf{p}; \lambda) = E$ ,

$$\Omega(E, \lambda) = \int d\mathbf{q}d\mathbf{p} \Theta[E - H(\mathbf{q}, \mathbf{p}; \lambda)], \quad (\text{A1})$$

where  $H$  is the system Hamiltonian and  $\Theta(x)$  is the step function. For systems with one degree of freedom,  $\Omega(E, \lambda)$  is an adiabatic invariant since it is the action [45]. Thus, after a quasistatic change of  $\lambda$ , an initial energy shell is mapped into a final energy shell such as  $\Omega(E_i, \lambda_i) = \Omega(E_f, \lambda_f)$ . This relation can be seen as an expression for the final energy,  $E_f^{ad}$ , as function of  $E_i$ ,  $\lambda_i$ , and  $\lambda_f$  and hence gives the energy at the end of the quasistatic process.

## APPENDIX B: EXCESS WORK WITHIN LINEAR RESPONSE THEORY

In this Appendix we derive the linear response expression for excess and quasistatic work. In the inclusive picture, the thermodynamic work produced during a finite-time driving of a control parameter  $\lambda$  is given by (3). Assuming that  $|\delta\lambda g(t)/\lambda_0| \ll 1$  for  $t_0 \leq t \leq t_f$ , linear response theory provides the following relation between the out-of-equilibrium average and its corresponding response function  $\phi_0(t)$  [38],

$$\overline{\frac{\partial \mathcal{H}}{\partial \lambda}}(t) = \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_0 + \chi_0^\infty \delta\lambda g(t) - \delta\lambda \int_{t_0}^t ds \phi_0(t-s) g(s), \quad (\text{B1})$$

where  $\langle \cdot \rangle_0$  denotes the average on the initial microcanonical ensemble and the subscript refers to the value  $\lambda_0$ . The second term in the right-hand side of Eq. (B1) describes the *instantaneous* response, which is due to  $\partial \mathcal{H}/\partial \lambda$  being a function of the external control  $\lambda$ . In particular, we have

$$\chi_0^\infty = \left\langle \frac{\partial^2 \mathcal{H}}{\partial \lambda^2} \right\rangle_0. \quad (\text{B2})$$

The second term describes the *delayed* response. It is convenient to express it in terms of the *relaxation* function as  $\phi_0(t) = -d\Psi_0(t)/dt$ . Thus, integrating by parts Eq. (B1),

we find

$$\begin{aligned} \frac{\overline{\partial \mathcal{H}}}{\partial \lambda}(t) &= \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_0 - \delta \lambda \tilde{\Psi}_0(0) g(t) \\ &+ \delta \lambda \int_0^{t-t_0} du \Psi_0(u) \frac{dg}{dt'} \Big|_{t'=t-u}, \end{aligned} \quad (\text{B3})$$

where  $\tilde{\Psi}_0 \equiv \Psi_0(0) - \chi_0^\infty$ . Finally, substituting Eq. (B3) in expression (3), we obtain

$$W = \delta \lambda \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_0 - \frac{(\delta \lambda)^2}{2} \tilde{\Psi}_0(0)$$

$$+ (\delta \lambda)^2 \int_{t_0}^{t_f} dt \frac{dg}{dt} \int_{t_0}^t dt' \Psi_0(t-t') \frac{dg}{dt'}, \quad (\text{B4})$$

where the following boundary conditions,  $g(t_0) = 0$  and  $g(t_f) = 1$ , were used. The first two terms of the previous expression *do not* depend on the protocol  $g(t)$ . Indeed, it can be verified that they are the first terms of the series expansion of the quasistatic work for  $\delta \lambda / \lambda_0 \ll 1$  given in (5). The last term clearly depends on  $g(t)$  and therefore represents the *excess* work given by (6).

- 
- [1] J. Maxwell, *Theory of Heat*, Text-books of Science (Longmans, London, 1871).
- [2] H. Leff and A. Rex, *Maxwell's Demon: Entropy, Information, Computing*, Princeton Legacy Library (Princeton University Press, Princeton, NJ, 2014).
- [3] K. Maruyama, F. Nori, and V. Vedral, *Rev. Mod. Phys.* **81**, 1 (2009).
- [4] L. Szilárd, *Z. Phys.* **53**, 840 (1929).
- [5] L. Brillouin, *J. Appl. Phys.* **22**, 334 (1951).
- [6] R. Landauer, *Phys. Today* **44**(5), 23 (1991).
- [7] C. Bennett, *Int. J. Theor. Phys.* **21**, 905 (1982).
- [8] O. Penrose, *Foundations of Statistical Mechanics: A Deductive Treatment* (Dover Publications, Mineola, NY, 2014).
- [9] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, *Nat. Phys.* **6**, 988 (2010).
- [10] A. Berut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, *Nature (London)* **483**, 187 (2011).
- [11] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola, *Phys. Rev. Lett.* **113**, 030601 (2014).
- [12] E. Roldán, I. A. Martínez, J. M. R. Parrondo, and D. Petrov, *Nat. Phys.* **10**, 457 (2014).
- [13] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **100**, 080403 (2008).
- [14] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **102**, 250602 (2009).
- [15] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **104**, 090602 (2010).
- [16] J. M. Horowitz and S. Vaikuntanathan, *Phys. Rev. E* **82**, 061120 (2010).
- [17] M. Ponnuragan, *Phys. Rev. E* **82**, 031129 (2010).
- [18] T. Sagawa and M. Ueda, *Phys. Rev. E* **85**, 021104 (2012).
- [19] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **109**, 180602 (2012).
- [20] M. Esposito and C. V. den Broeck, *Europhys. Lett.* **95**, 40004 (2011).
- [21] D. Abreu and U. Seifert, *Europhys. Lett.* **94**, 10001 (2011).
- [22] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, *Phys. Rev. Lett.* **110**, 040601 (2013).
- [23] A. C. Barato and U. Seifert, *Phys. Rev. Lett.* **112**, 090601 (2014).
- [24] A. C. Barato and U. Seifert, *Phys. Rev. E* **90**, 042150 (2014).
- [25] D. Mandal and C. Jarzynski, *Proc. Natl. Acad. Sci. USA* **109**, 11641 (2012).
- [26] A. Barato and U. Seifert, *Europhys. Lett.* **101**, 60001 (2013).
- [27] S. Deffner and C. Jarzynski, *Phys. Rev. X* **3**, 041003 (2013).
- [28] S. Deffner, *Phys. Rev. E* **88**, 062128 (2013).
- [29] Z. Lu, D. Mandal, and C. Jarzynski, *Phys. Today* **67**(8), 60 (2014).
- [30] J. V. Koski and J. P. Pekola, *C. R. Phys.* **17**, 1130 (2016).
- [31] A. Kutvonen, J. Koski, and T. Ala-Nissila, *Sci. Rep.* **6**, 21126 (2016).
- [32] R. Marathe and J. M. R. Parrondo, *Phys. Rev. Lett.* **104**, 245704 (2010).
- [33] S. Vaikuntanathan and C. Jarzynski, *Phys. Rev. E* **83**, 061120 (2011).
- [34] J. M. R. Parrondo, *Chaos* **11**, 725 (2001).
- [35] J. M. R. Parrondo and L. Granger, *Eur. Phys. J. Spec. Top.* **224**, 865 (2015).
- [36] T. V. Acconcia and M. V. S. Bonança, *Phys. Rev. E* **91**, 042141 (2015).
- [37] T. V. Acconcia, M. V. S. Bonança, and S. Deffner, *Phys. Rev. E* **92**, 042148 (2015).
- [38] R. Kubo, *J. Phys. Soc. Jpn.* **12**, 570 (1957).
- [39] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II-Nonequilibrium Statistical Mechanics* (Springer-Verlag, Berlin, 1985).
- [40] M. V. S. Bonança, *Phys. Rev. E* **85**, 031122 (2012).
- [41] K. Sekimoto and S. I. Sasa, *J. Phys. Soc. Jpn.* **66**, 3326 (1997).
- [42] I. A. Martínez, E. Roldán, L. Dinis, and R. A. Rica, *Soft Matter* **13**, 22 (2017).
- [43] P. J. Channel and C. Scovel, *Nonlinearity* **3**, 231 (1990).
- [44] M. Campisi, F. Zhan, P. Talkner, and P. Hänggi, *Phys. Rev. Lett.* **108**, 250601 (2012).
- [45] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (Addison-Wesley, Boston, 2002).