Second Law for Active Heat Engines

Arya Datta,[1] Patrick Pietzonka,[2] and Andre C. Barato[1]
[1]*Department of Physics, University of Houston, Houston, Texas 77204, USA*
[2]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, 01187 Dresden, Germany*

(Received 30 November 2021; revised 27 May 2022; accepted 7 July 2022; published 6 September 2022)

Macroscopic cyclic heat engines have been a major motivation for the emergence of thermodynamics. In the past decade, cyclic heat engines that have large fluctuations and operate at finite time were studied within the more modern framework of stochastic thermodynamics. The second law for such heat engines states that the efficiency cannot be larger than the Carnot efficiency. The concept of cyclic active heat engines for a system in the presence of hidden dissipative degrees of freedom, also known as a nonequilibrium or active reservoir, has also been studied in theory and experiment. Such active engines show rather interesting behavior such as an "efficiency" larger than the Carnot bound. They are also likely to play an important role in future developments, given the ubiquitous presence of active media. However, a general second law for cyclic active heat engines has been lacking so far. Here, by using a known inequality in stochastic thermodynamics for the excess entropy, we obtain a general second law for active heat engines, which does not involve the energy dissipation of the hidden degrees of freedom and is expressed in terms of quantities that can be measured directly from the observable degrees of freedom. Besides heat and work, our second law contains an information-theoretic term, which allows an active heat engine to extract work beyond the limits valid for a passive heat engine. To obtain a second law expressed in terms of observable variables in the presence of hidden degrees of freedom, we introduce a coarse-grained excess entropy and prove a fluctuation theorem for this quantity.

DOI: 10.1103/PhysRevX.12.031034 Subject Areas: Soft Matter, Statistical Physics

I. INTRODUCTION

Thermodynamics [1] is a major theory in physics that deals with the transfer of heat into other forms of energy. This theory provides universal bounds through its second law on processes that involve heat transfer. Cyclic engines that transform part of the heat taken from a hot reservoir into useful work are central objects in thermodynamics. One of the most prominent statements of the second law of thermodynamics is the Carnot bound on the efficiency of heat engines. Yet standard thermodynamics has considerable limitations: only heat engines with negligible fluctuations and that operate in the quasistatic limit can be analyzed quantitatively.

Real engines, however, operate at finite time. Furthermore, it is now possible to realize in experiments heat engines made of a small number of constituents, for which fluctuations cannot be neglected. The modern theory of stochastic thermodynamics [2,3] does not suffer from the same limitations as standard thermodynamics and is able to deal with such heat engines. Within stochastic thermodynamics a model for a cyclic heat engine with a colloidal particle in a harmonic potential has been proposed in Ref. [4]. This model has been realized in an experiment in Ref. [5]. Much theoretical [6–12] and experimental [13–17] work has been done on such engines. For such heat engines, the Carnot bound on efficiency remains valid.

A novel perspective was put forward in an experiment of a so-called active heat engine [18]. This cyclic heat engine is likewise composed of a colloidal particle in a harmonic trap. However, in contrast to a passive heat engine the colloidal particle is not immersed in a solution that simply acts as an equilibrium heat bath. Rather, the solution has bacteria that collide with the particle. Such a bacterial bath may be called a nonequilibrium or active reservoir. Key observations made in this work were that an active heat engine can extract more work than its passive counterpart and that the pseudo-efficiency (work divided by heat taken from hot reservoir) of the engine can be larger than the Carnot efficiency, which might be called a "violation" of the second law.

Theoretical work on cyclic active heat engines has been prompted by this experiment [19–27]. Furthermore, a general theoretical framework for steady state or autonomous active engines has been proposed in Ref. [28]. Overall, active matter [29–32] is a novel key concept in

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

nonequilibrium statistical mechanics and there has been considerable interest in the stochastic thermodynamics of systems immersed in active external media [33–40].

A standard explanation that can be given to the observation of a pseudo-efficiency larger than the Carnot efficiency is that if we properly account for the energy dissipated by the bacteria, then there is no violation of the second law. This explanation is in principle correct: if we also consider the dissipation of the bacteria, then the total entropy production must be positive. However, such statement of the second law is rather limited. First, the energy dissipation of the bacteria is probably much larger than the extracted work. Hence, a second law that contains this energy dissipation would not represent a meaningful bound on the amount of extracted work. Second, this energy dissipation is related to degrees of freedom that are not accessible in the experiment since only the position of the colloidal particle is measured.

Beyond this particular experiment, an active heat engine is an interesting concept. We can imagine a cyclic heat engine with a system composed of a few degrees of freedom. This system interacts with, possibly hidden, dissipative degrees freedom that make the external medium active. A second law that contains the energy dissipation due to this activity has two main issues. First, the energy dissipated per cycle period due to the activity increases indefinitely if we increase the cycle period, since a nonequilibrium system dissipates energy even after reaching a steady state. However, the extracted work per period saturates and becomes a constant for large period. Second, this energy dissipation might be related to degrees of freedom that cannot be observed in an experiment and their measurement is not necessary to calculate extracted work.

What is the correct statement of the second law for active heat engines? Such a statement must fulfill the following requirements. First, the bound on extracted work can only contain terms that have the same scaling with the period of the engine. Second, it must accommodate the observations made in the experiment from Ref. [18], such as a pseudo-efficiency larger than the Carnot efficiency. Third, for a passive heat engine, it must become the known standard second law. Fourth, it must contain terms that depend only on the observable degrees of freedom.

In this paper we show that the appropriate statement of the second law for generic cyclic active heat engines, which fulfills all the aforementioned requirements, is obtained from a well-known quantity in stochastic thermodynamics called excess entropy [3,41–46] (or the related nonadiabatic entropy [47–49]). Crucially, we obtain this result with a suitable decomposition of the excess entropy. Beyond the known excess entropy, in order to fulfill the fourth requirement, we introduce a coarse-grained excess entropy and prove a fluctuation theorem for it.

This decomposition of the excess entropy for cyclic active heat engines leads to three terms, the heat taken from the hot reservoir, the extracted work, and a novel term that has an information-theoretical interpretation. Interestingly, this term relates the nonequilibrium steady state distribution associated with the active heat engine to the equilibrium Boltzmann distribution associated with its passive counterpart. Our results are illustrated with three models that allow for analytical calculations and one numerical model inspired by the experiment in Ref. [18].

The paper is organized as follows. In Sec. II we use a two-state model to illustrate our main result. The general formalism of our second law for active heat engines is introduced in Sec. III. In Sec. IV we introduce the coarse-grained second law, which is expressed in terms of quantities that only depend on an observable degree of freedom. In Sec. V we discuss three simple models for active heat engines that allow for analytical calculations. Section VI contains a more complex model that is analyzed with numerics. We conclude in Sec. VII. The fluctuation theorem for the coarse-grained excess entropy is proved in the Appendix.

II. SECOND LAW FOR TWO-STATE HEAT ENGINE

Consider the two-state heat engine shown in Fig. 1. It has a down state with energy zero and an up state with energy larger than zero. The protocol of this engine with period τ is as follows. It has four parts, each with duration τ_l, with $\tau_1+\tau_2+\tau_3+\tau_4=\tau$. There are two energy changes: at time τ_1 the energy changes from E to $E+\delta E$ and at time $\tau_1+\tau_2+\tau_3$ the energy changes back from $E+\delta E$ to E. Moreover, there are two temperature changes: at time $\tau_1+\tau_2$ the temperature changes from β_c to β_h and at time τ the temperature changes back from β_h to β_c.

Work can be extracted in this cycle in the following way. If the system finishes the third part at the up state, then an amount δE of work is extracted. Moreover, if the system finishes the first part at the up state, then an amount δE of work is spent on the system ($-\delta E$ of extracted work). Hence, the maximum amount of work that can be extracted in a cycle is δE.

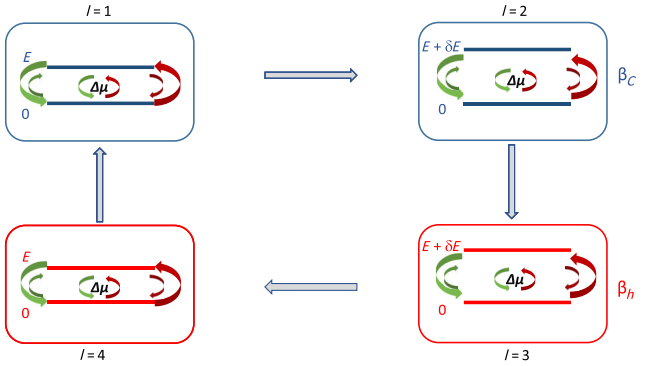
A similar passive two-state heat engine has been studied in Ref. [6]. The known second law for this passive heat engine reads

$$(\beta_c-\beta_h)Q_h-\beta_c W\geq 0, \tag{1}$$

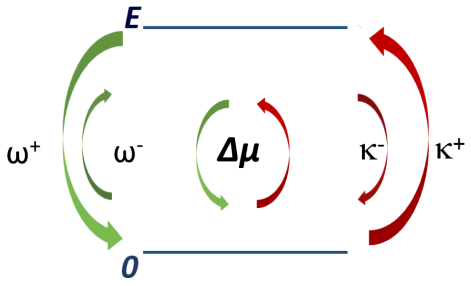
where Q_h is the average heat taken from the hot reservoir during a period and W is the average extracted work during a period. This second law implies the well-known bound on the efficiency of a passive engine,

$$W/Q_h\leq \eta_C\equiv 1-\beta_h/\beta_c, \tag{2}$$

where η_C is the Carnot efficiency.



(a)



(b)

FIG. 1. (a) Depiction of the protocol for the two-state heat engine with a period divided into four parts labeled by l. The four changes are as follows: from $l = 1$ to $l = 2$ the energy changes from E to $E + \delta E$, from $l = 2$ to $l = 3$ the inverse temperature changes from β_c to β_h, from $l = 3$ to $l = 4$ the energy changes from $E + \delta E$ to E, and from $l = 4$ to $l = 1$ the inverse temperature changes from β_h to β_c. (b) Possible reactions for the two-state system. The transition rate for the forward (backward) phosphorylation reaction in Eq. (3) is κ_+ (κ_-) and the transition rate for the forward (backward) dephosphorylation reaction in Eq. (4) is ω_+ (ω_-). We parametrize the transition rates as $\kappa^+(t) = \omega^-(t) = k$, $\kappa^-(t) = ke^{\beta(t)[E(t)+\Delta\mu]}$, and $\omega^+(t) = ke^{\beta(t)E(t)}$.

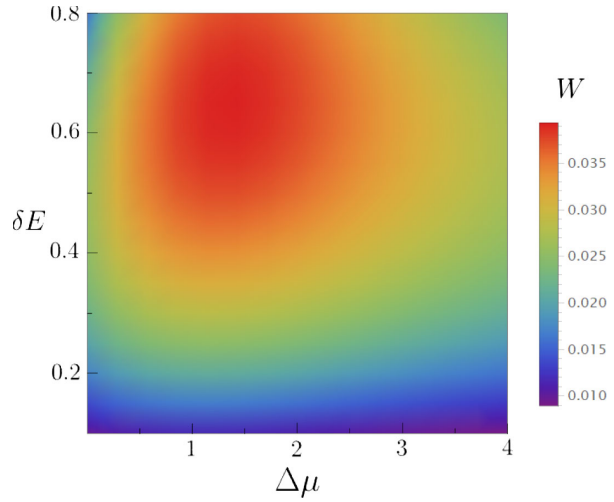
Now we consider an active two-state heat engine. The simplest way to make the engine active is to include two different transition paths between the two states. A possible physical interpretation shown in Fig. 1(b) would be an enzyme that can be in two states, X and XP, where XP is the phosphorylated form of the enzyme. The state X has energy 0 and the state XP has nonzero energy, which can be either E or $E + \delta E$. One transition path from X to XP is a phosphorylation reaction

$$X + \mathrm{ATP} \rightleftharpoons XP + \mathrm{ADP}. \tag{3}$$

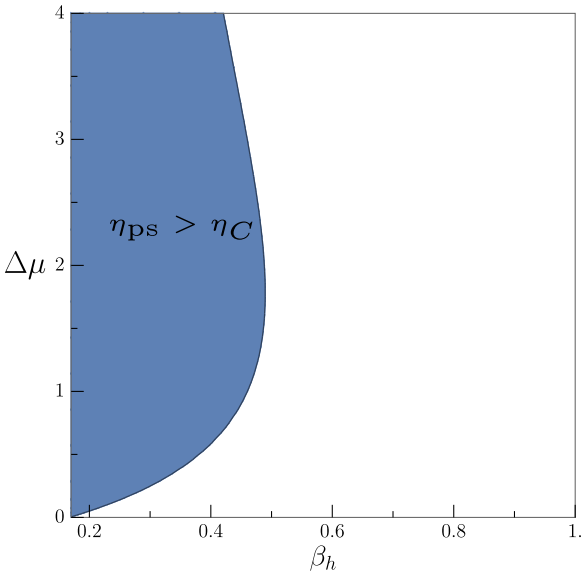
The other transition path is a dephosphorylation reaction

$$XP \rightleftharpoons X + P. \tag{4}$$

During any of the four steps of the cycle the enzyme burns ATP. The free energy of one ATP hydrolysis is written as



(a)



(b)

FIG. 2. Work and efficiency for the two-state heat engine depicted in Fig. 1. (a) Work as a function of $\Delta\mu$ and δE for $E = 1$ and $\beta_h = 0.5$. Work for an active heat engine can be larger as compared to the work of its passive counterpart for $\Delta\mu = 0$. (b) The pseudo-efficiency η_{ps} is larger than the Carnot efficiency η_C in the blue filled region, for $E = 0.1$ and $\delta E = 0.5$. For both figures $\beta_c = 1$, $k = 1$, and $\tau = 1$. These parameters correspond to an arbitrary cold temperature. The hot temperature is the cold temperature times β_h^{-1}. The parmaeters E, δE, and $\Delta\mu$ are all in units of β_c^{-1}.

$\Delta\mu = \mu_{\mathrm{ATP}} - \mu_{\mathrm{ADP}} - \mu_P$, where μ is the chemical potential. This exactly solvable model for an active heat engine is fully defined in Sec. V B. In this section we use the results obtained with this model for illustrative purposes.

If $\Delta\mu = 0$, the heat engine is a standard passive heat engine and the second law in Eq. (1) applies. For $\Delta\mu \neq 0$, the heat engine is an active heat engine. As shown in Fig. 2,

this simple active heat engine produces remarkable behavior. First, making the engine active with a nonzero $\Delta\mu$ can increase the amount of extracted work. Second, the pseudo-efficiency,

$$\eta_{\mathrm{ps}} \equiv W/Q_h, \tag{5}$$

can be larger than the Carnot efficiency η_C. The heat Q_h requires a word of caution. This quantity is not exactly the heat taken from the hot reservoir for an active heat engine. This passive heat Q_h only accounts for the contributions coming from the energy changes of the system associated with transitions between the two states during the part of the period for which the inverse temperature is β_h. This Q_h does not account for dissipated heat associated with ATP consumption.

The first solution for the issue $\eta_{\mathrm{ps}} > \eta_C$ is to account for the average entropy production associated with the consumption of ATP during a cycle ΔS_{act}. The standard second law from stochastic thermodynamics for this model reads

$$\Delta S_{\mathrm{act}} + (\beta_c - \beta_h)Q_h - \beta_c W \geq 0. \tag{6}$$

The condition $\eta_{\mathrm{ps}} > \eta_C$ is then fully consistent with this second law since η_{ps} can be larger than η_C if $\Delta S_{\mathrm{act}} > 0$.

However, it is often the case that this second law does not produce a relevant bound on the rate of extracted work W. Let us consider the period of the engine τ. For large τ, the amount of entropy produced due to the ATP consumption ΔS_{act} increases linearly with τ. This scaling with the period τ is different from the scaling of Q_h and W. For instance, the stochastic work that can be extracted in a cycle is at most δE. Hence, for large τ the average work extracted per cycle cannot be more than the constant δE. This difference in scaling is illustrated in Fig. 3. We conclude that in the regime of large τ the standard second law in Eq. (6) does not produce a relevant bound, ΔS_{act} dominates the inequality, which simply states that ΔS_{act} is positive.

What is the appropriate statement of the second law that does not suffer from the limitation discussed above? Before we answer this question let us consider the following question. How can an active heat engine deliver more work than a passive heat engine? If we introduce a nonzero $\Delta\mu$ in our model the probability of a state is shifted. If this shift favors work extraction, then the active heat engine will deliver more work than its passive counterpart.

Our main finding is that the appropriate statement of the second law for active heat engines comes from a quantity known as excess entropy. A main interest in this quantity in stochastic thermodynamics comes from the fact that it fulfills a fluctuation theorem [3]. Here we show that it has a central role for cyclic active heat engines. For the present model this second law derived from the excess entropy reads

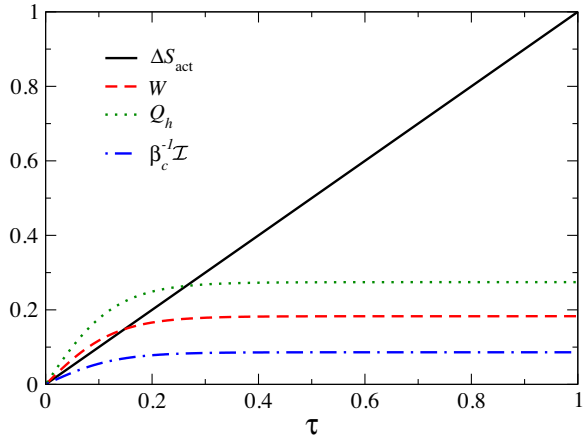


FIG. 3. The different scaling with τ of ΔS_{act} in comparison to W, Q_h, and $\mathcal{I}$ for an active heat engine with fixed $\Delta\mu$. The inequality $\Delta S_{\mathrm{act}} \geq \mathcal{I}$ holds true for any value of τ. The parameters are set to $\beta_c = 1$, $\beta_h = 0.1$, $E = 0.5$, $\delta E = 1$, $k = 10$, and $\Delta\mu = 2$.

$$\mathcal{I} + (\beta_c - \beta_h)Q_h - \beta_c W \geq 0. \tag{7}$$

The term $\mathcal{I}$ is the novel term compared to Eq. (6). It has an information-theoretic expression and it quantifies the shift in the probability of a state due to a nonzero $\Delta\mu$. A positive $\mathcal{I}$ allows an active heat engine to operate beyond the limits that are valid for a passive heat engine, as illustrated in Fig. 2.

Concerning the scaling with the period τ, $\mathcal{I}$ scales in the same way as W and Q_h. Furthermore, we show that $\Delta S_{\mathrm{act}} \geq \mathcal{I}$ independently of the value of τ, as illustrated in Fig. 3. The second law above is consistent with $\eta_{\mathrm{ps}} > \eta_C$. For a passive heat engine $\mathcal{I}$ is zero and Eq. (7) becomes the known second law for passive heat engines in Eq. (1).

This minimal two-state model does not allow for the discussion of the issue of coarse graining. For the case with possible hidden dissipative degrees of freedom, we obtain a second law that has the exact same structure of Eq. (7), as shown in Sec. IV.

III. GENERAL FRAMEWORK

A. Invariant distribution and accompanying distribution

We consider Markov processes with discrete states. A transition rate from state i to state j is denoted $k_{ij}(t)$. Since we want to model cyclic engines, the transition rate is periodic in time with a period τ, i.e., $k_{ij}(t) = k_{ij}(t+\tau)$. The master equation for the probability to be in state i at time t is

$$\frac{d}{dt}P_i(t) = \sum_j [P_j(t)k_{ji}(t) - P_i(t)k_{ij}(t)]. \tag{8}$$

The following condition is assumed to hold: if $k_{ij}(t) \neq 0$, then $k_{ji}(t) \neq 0$. In the longtime limit, $P_i(t)$ becomes an

invariant time-periodic distribution with period τ. We denote this invariant distribution simply by $P_i(t)$.

The accompanying distribution [50,51] $P_i^S(t)$ is generally different from $P_i(t)$. It is defined as the stationary distribution the system would reach if the transition rates were fixed and given by $k_{ij}(t)$. Hence, $P_i^S(t)$ is also time periodic with period τ. Specifically, $P_i^S(t)$ is the solution of the equation

$$\sum_j [P_j^S(t)k_{ji}(t) - P_i^S(t)k_{ij}(t)] = 0, \tag{9}$$

which is the stationary form of Eq. (8) for fixed t. If each term in the sum in j is zero, i.e., if detailed balance is fulfilled, then this accompanying distribution $P_i^S(t)$ is an equilibrium distribution denoted by $P_i^{\mathrm{eq}}(t)$. Otherwise, the accompanying distribution corresponds to a nonequilibrium stationary distribution denoted by $P_i^S(t)$.

We reiterate the difference between three probability distributions we consider here. First, $P_i(t)$ is the time-periodic distribution of the system, corresponding to the longtime solution of Eq. (8). Second, $P_i^S(t)$ is the accompanying distribution corresponding to a nonequilibrium steady state. Third, $P_i^{\mathrm{eq}}(t)$ is the accompanying distribution corresponding to an equilibrium steady state.

B. Generalized detailed balance

The periodicity of the transition rates comes from the fact that energies and temperature are periodic. The energy of state i at time t is denoted $E_i(t) = E_i(t+\tau)$. The inverse temperature remains between β_h and β_c and is written as

$$\beta(t) = \beta_c[1 - \eta_C h(t)], \tag{10}$$

where $h(t)$ is a dimensionless time-periodic function bounded by the inequalities $0 \le h(t) \le 1$. If $h(t) = 0$, then $\beta(t) = \beta_c$, and if $h(t) = 1$, then $\beta(t) = \beta_h$.

An active heat engine must contain internal thermodynamic forces or affinities. An internal affinity is denoted by $\mathcal{A}_\alpha$. We use the index α since the system can have more than one internal affinity. The generalized distance associated with this affinity for a transition from i to j is $d_{ij}^\alpha(t) = d_{ij}^\alpha(t+\tau)$. This distance is antisymmetric, i.e., $d_{ij}^\alpha(t) = -d_{ji}^\alpha(t)$. For example, if $\mathcal{A}_\alpha$ is a force applied to a colloid, then $d_{ij}^\alpha(t)$ is the spatial distance between i and j.

The generalized detailed balance relation [3] is a postulate of stochastic thermodynamics that relates transition rates to physical parameters such as temperature, free energy, and affinities. This relation is written in the form

$$\ln \frac{k_{ij}(t)}{k_{ji}(t)} = \beta(t)\left[\Delta E_{ji}(t) + (\beta_c)^{-1}\sum_\alpha \mathcal{A}_\alpha d_{ij}^\alpha(t)\right], \tag{11}$$

where $\Delta E_{ji}(t) \equiv E_i(t) - E_j(t)$ and the product $\mathcal{A}_\alpha d_{ij}^\alpha(t)$ has dimension of energy in units of β_c^{-1}. There is no loss of generality in assuming a constant $\mathcal{A}_\alpha$ since a time dependence of $\mathcal{A}_\alpha$ can be absorbed by the time dependence of the distance $d_{ij}^\alpha(t)$.

We can now give a clear mathematical definition for active and passive heat engines. If $\mathcal{A}_\alpha = 0$ for all α, then the heat engine is a passive heat engine. In this case the stationary distribution the system would have if the protocol was frozen, i.e., the accompanying distribution, which corresponds to the solution of Eq. (9), is the Boltzmann distribution $P_i^{\mathrm{eq}}(t) = e^{-\beta(t)E_i(t)}/Z(t)$, where $Z(t) = \sum_i e^{-\beta(t)E_i(t)}$ is the partition function. Note that a passive heat engine is a nonequilibrium system due to the periodic variation of temperature and energies. For nonzero $\mathcal{A}_\alpha$ the accompanying distribution $P_i^S(t)$ is not the equilibrium distribution and the heat engine is active.

C. Entropy production

The elementary (probability) current from state i to state j is defined as

$$J_{ij}(t) \equiv P_i(t)k_{ij}(t) - P_j(t)k_{ji}(t). \tag{12}$$

The mathematical form of the average entropy production during a period is [3]

$$\Delta S \equiv \int_0^\tau dt \sum_{i<j} J_{ij}(t) \ln \frac{k_{ij}(t)}{k_{ji}(t)} \ge 0, \tag{13}$$

where the sum $\sum_{i<j}$ is over all pairs of states. Using the generalized detailed balance relation in Eq. (11) we obtain

$$\Delta S = \sum_\alpha \mathcal{A}_\alpha J_\alpha + \int_0^\tau dt \sum_{i<j} J_{ij}(t)\beta(t)\Delta E_{ji}(t), \tag{14}$$

where

$$J_\alpha \equiv \int_0^\tau dt \sum_{i<j} (\beta_c)^{-1}\beta(t) J_{ij}(t) d_{ij}^\alpha(t). \tag{15}$$

The second term in the expression for ΔS in Eq. (14) can be written as

$$\int_0^\tau dt J_{ij}(t)\beta(t)\Delta E_{ji}(t) = \int_0^\tau dt \beta_c J_{ij}(t)\Delta E_{ji}(t) + \eta_C J_q, \tag{16}$$

where

$$J_q \equiv \int_0^\tau dt \sum_{i<j} h(t)\beta_c J_{ij}(t)\Delta E_{ij}(t). \tag{17}$$

The physical interpretation for this term is as follows. The energy change of the system in a jump from i to j is

$\Delta E_{ij}(t)$. The probability current $J_{ij}(t)$ is the average number of jumps from i to j minus the average number of jumps from j to i per unit of time. By summing over all pairs of states we have the average energy gain of the system at time t. The multiplication by β_c means that energies are measured in units of β_c^{-1}. If the engine is passive the term J_q is a generalized heat flux during a period, since the energy change in a jump is the energy taken from the thermal reservoir.

The generalized heat flux J_q for an active heat engine is not the sole contribution to heat flux since the internal affinities $\mathcal{A}_\alpha$ can also contribute to heat dissipation. However, as we show next, there is a refined first law that also holds true for an active heat engine and does not involve heat dissipation due to the internal affinities $\mathcal{A}_\alpha$. We refer to J_q as passive generalized heat flux.

For the models analyzed here, the inverse temperature $\beta(t)$ only takes the two values β_c and β_h during the period τ. The function $h(t)$ only takes the values $h(t) = 0$ for $\beta(t) = \beta_c$ and $h(t) = 1$ for $\beta(t) = \beta_h$. For this case, $J_q = \beta_c Q_h$, where Q_h is the heat taken from the hot reservoir during a period.

We point out that even though J_q is not the only contribution to heat, this quantity can be measured in an experiment. One way to measure it would be to observe the trajectory of the system and sum up $h(t)\Delta E_{ij}(t)$ whenever there is a jump from i to j. An alternative way to measure it would be to evaluate the total heat flux and the heat flux due to the activity. Hypothetically, the first quantity could be measured by monitoring small temperature differences in a large but finite reservoir and the second one by monitoring changes in the concentration of a chemical fuel such as ATP. The difference between both gives the passive contribution.

D. Refined first law and standard second law

The entropy production in Eq. (14) can be written as

$$\Delta S = \sum_\alpha \mathcal{A}_\alpha J_\alpha + \eta_C J_q - \int_0^\tau dt \sum_{i<j} J_{ij}(t)\beta_c \Delta E_{ij}(t), \quad (18)$$

where we used Eq. (16). The time derivative of the average energy $\bar{E}(t) = \sum_i P_i(t)E_i(t)$ reads

$$\frac{d}{dt}\bar{E}(t) = \sum_{i<j} J_{ij}(t)\Delta E_{ij}(t) + \sum_i P_i(t)\frac{d}{dt}E_i(t). \quad (19)$$

Since $\bar{E}(t) = \bar{E}(t+\tau)$, integration over a period leads to

$$\int_0^\tau dt \sum_{i<j} J_{ij}(t)\Delta E_{ij}(t) = -\int_0^\tau dt \sum_i P_i(t)\frac{d}{dt}E_i(t) \equiv W. \quad (20)$$

The term on the right-hand side is the extracted work W from stochastic thermodynamics [3]. This equation constitutes the refined first law for active heat engines. If the inverse temperature only takes the values β_c and β_h, i.e., $h(t)$ is either 0 or 1, this equation becomes

$$Q_h + (-Q_c) = W, \quad (21)$$

where

$$-Q_c = -\int_0^\tau dt[1 - h(t)]\sum_{i<j} J_{ij}(t)\Delta E_{ij}(t) \quad (22)$$

is the average passive heat delivered to the cold reservoir during a period. We reiterate that this refined first law for active heat engines does not involve energy dissipation associated with the affinities $\mathcal{A}_\alpha$.

The entropy production in Eq. (18) then becomes

$$\Delta S = \Delta S_{\mathrm{act}} + \eta_C J_q - \beta_c W \geq 0, \quad (23)$$

where $\Delta S_{\mathrm{act}} \equiv \sum_\alpha \mathcal{A}_\alpha J_\alpha$ is the contribution due to the internal affinities. This contribution to the entropy production is the one that makes the heat engine active. This standard form of the second law has been obtained in Ref. [11].

A key issue is the scaling with the period τ of the different terms in this entropy production. A cycle in the network of states is a sequence of jumps that finishes in the state it started. Consider a cycle that happens while the external protocol is fixed. The total change of energy associated with this cycle, i.e., the net exchange of passive heat, is zero. The entropy change during the cycle associated with $\mathcal{A}_\alpha$ is not necessarily zero since the distance $d_{ij}^\alpha(t)$ cannot be written as a difference such as $\Delta E_{ij}(t) = E_j(t) - E_i(t)$. The larger period τ, the larger the number of times cycles are performed. Therefore, the active part of the entropy change increases indefinitely with τ, unlike passive heat and work.

This difference in scaling is an important limitation of the second law in Eq. (23). For large enough τ this second law does not provide a relevant bound on W since the term ΔS_{act} dominates the inequality. A relevant bound on extracted work can only contain terms that scale with τ in the same way that W does. We point out that it is possible to make ΔS_{act} scale in the same way as heat and work with τ by using infinite energy barriers that block the system from performing cycles in the space of states during the period. Models for molecular pumps do have such energy barriers [11,52].

E. Excess entropy and second law for active heat engines

The entropy change ΔS can be written as

$$\Delta S = \Delta S_{\rm ex} + \Delta S_{\rm hk}, \tag{24}$$

where $\Delta S_{\rm ex}$ is the excess entropy and $\Delta S_{\rm hk}$ is the housekeeping entropy. A main interest in these quantities within stochastic thermodynamics is related to the fact that each of these terms is positive and their fluctuating versions each fulfill a separate fluctuation theorem [3]. Here we are interested in the mathematical expression of the excess entropy for periodically driven systems. As shown in the Appendix, the average excess entropy change during a period is given by

$$\Delta S_{\rm ex} = \int_0^\tau dt \sum_{i<j} J_{ij}(t) \ln \frac{P_j^S(t)}{P_i^S(t)} \geq 0. \tag{25}$$

We now show that this inequality provides the appropriate statement of the second law for active heat engines.

The main step is to define

$$\mathcal{I} \equiv \int_0^\tau dt \sum_{i<j} J_{ij}(t) \left[\ln \frac{P_j^S(t)}{P_j^{\rm eq}(t)} - \ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)} \right]. \tag{26}$$

The excess entropy in Eq. (25) can be written as

$$\Delta S_{\rm ex} = \mathcal{I} + \int_0^\tau dt \sum_{i<j} J_{ij}(t) \ln \frac{P_j^{\rm eq}(t)}{P_i^{\rm eq}(t)}. \tag{27}$$

The relation $\ln[P_j^{\rm eq}(t)/P_i^{\rm eq}(t)] = \beta(t)\Delta E_{ji}(t)$, together with Eqs. (17) and (20), leads to our second law for active heat engines:

$$\Delta S_{\rm ex} = \mathcal{I} + \eta_C J_q - \beta_c W \geq 0. \tag{28}$$

This inequality involves the extracted work W, the generalized heat flux J_q, and the new term $\mathcal{I}$.

Our second law allows for the following insight about active heat engines. For an active heat engine there is a shift from $P^{\rm eq}$ to P^S; this shift can allow for a larger amount of work extraction, beyond the bound given by the second law for a passive heat engine. The amount of energy dissipated by the active process is not directly relevant; rather, the way the probability P^S is shifted in relation to $P^{\rm eq}$, as quantified by $\mathcal{I}$, is the relevant limiting factor on W.

For an active heat engine the pseudo-efficiency $\eta_{\rm ps}$ can go beyond the Carnot bound η_C. If we consider the efficiency,

$$\eta \equiv \frac{W}{\beta_c^{-1} J_q + \mathcal{I}/(\beta_c - \beta_h)}, \tag{29}$$

which accounts for the term $\mathcal{I}$, then the second law for active heat engines in Eq. (28) becomes the traditional statement of the second law $\eta \leq \eta_C$. We note that an active heat engine is able to extract work even if the inverse temperature is fixed and equal to β_c, as we show with specific examples in Sec. V. In this case, $W \leq \beta_c^{-1}\mathcal{I}$.

Our second law in Eq. (28) has the correct scaling with the period τ and becomes the second law for passive heat engines for the case $P^S = P^{\rm eq}$ during the whole period. First, for a cycle in the space of states the change of $\ln[P_i^S(t)/P_i^{\rm eq}(t)] - \ln[P_j^S(t)/P_j^{\rm eq}(t)]$ is zero. Hence, for large τ, the scaling with τ of the term $\mathcal{I}$ is the same as the scaling of W and J_q. Second, for a passive heat engine $\mathcal{I} = 0$ and Eq. (28) becomes the standard second law for passive heat engines.

From the second law for the housekeeping entropy [3], we obtain

$$\Delta S_{\rm hk} = \Delta S - \Delta S_{\rm ex} \geq 0. \tag{30}$$

This inequality, together with Eqs. (23) and (28), leads to

$$\Delta S_{\rm act} \geq \mathcal{I}. \tag{31}$$

Hence, the second law in Eq. (28) provides a tighter bound on the extracted work W than the one in Eq. (23), for any period τ. The term $\mathcal{I}$ is the information-theoretic contribution to the second law for an active heat engine and $\Delta S_{\rm act}$ can be viewed as the total thermodynamic cost to "generate" this term. This total thermodynamic cost to generate an information-theoretic term shows up in many examples in the relation between information and thermodynamics [53–56].

We here assumed that there is only one transition path between a pair of states i, j. However, it is possible to have more than one transition path between a pair of states, which is the case for the two-state model in Sec. V B. The generalization of the expressions obtained here for the case that includes this possibility of more than one link between a pair of states is straightforward. If we use the label ν to represent the different links, the transition rate from i to j through link ν is k_{ij}^ν. For the generalized detailed balance relation in Eq. (11), the distances $d_{ij}^\alpha(t)$ depend on the index ν. The expression for quantities that show up in the second law inequalities remain similar, with the sum $\sum_{i<j}$ substituted by $\sum_\nu \sum_{i<j}$.

The second law obtained here is relevant for generic periodically driven active heat engines. Another important class of heat engines are those driven by constant time-independent forces that operate in a steady state. Since the average excess entropy in Eq. (25) is zero in a steady state, it provides no relevant information about the performance of steady state heat engines, which can be quantified with the standard entropy production [28].

F. Information-theoretic expression for $\mathcal{I}$

It turns out that $\mathcal{I}$ can be expressed as part of a time derivative of a difference of Kullback-Leibler distances. Consider the average,

$$\left\langle \ln \frac{P^S}{P^{\rm eq}} \right\rangle(t) \equiv \sum_i P_i(t) \ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)}. \tag{32}$$

This term can be written as

$$\left\langle \ln \frac{P^S}{P^{\rm eq}} \right\rangle(t) = D_{\rm KL}[P||P^{\rm eq}](t) - D_{\rm KL}[P||P^S](t), \tag{33}$$

where $D_{\rm KL}[p||q] \equiv \sum_i p_i \ln(p_i/q_i)$ is the Kullback-Leibler distance.

Taking a a derivative of this term we obtain

$$\frac{d}{dt}\left\langle \ln \frac{P^S}{P^{\rm eq}} \right\rangle = \sum_{i<j} J_{ij}(t) \left[\ln \frac{P_j^S(t)}{P_j^{\rm eq}(t)} - \ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)} \right] + \sum_i P_i(t) \frac{d}{dt}\left(\ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)} \right). \tag{34}$$

Since $\langle \ln(P^S/P^{\rm eq})\rangle(t) = \langle \ln(P^S/P^{\rm eq})\rangle(t+\tau)$, integration of the equation above yields

$$\int_0^\tau dt \sum_{i<j} J_{ij}(t) \left[\ln \frac{P_j^S(t)}{P_j^{\rm eq}(t)} - \ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)} \right] = -\int_0^\tau dt \sum_i P_i(t) \frac{d}{dt}\left(\ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)} \right). \tag{35}$$

This equality provides a different expression for $\mathcal{I}$ in Eq. (26), which is

$$\mathcal{I} = -\int_0^\tau dt \sum_i P_i(t) \frac{d}{dt}\left(\ln \frac{P_i^S(t)}{P_i^{\rm eq}(t)} \right). \tag{36}$$

Hence, the term $\mathcal{I}$ is part of the time derivative of the difference of Kullback-Leibler distances in Eq. (33). Furthermore, the expression in Eq. (36) can be more convenient to calculate $\mathcal{I}$.

IV. COARSE-GRAINED SECOND LAW

A. Observable variable

We now address the issue of coarse graining. The first step is to divide the state i into two variables, $i = (x, a)$. The variable x can be observed and the variable a is hidden. If we consider the experiment from Ref. [18], x would be the position of the colloidal particle and a would be related to the degrees of freedom that determine the state of the bacteria. If we consider the variable a as part of the reservoir, then we can say that the system labeled by the variable x is in contact with a nonequilibrium or active reservoir.

The master equation in Eq. (8) implies an equation for the time evolution of $P_x(t) \equiv \sum_a P_{x,a}(t)$. Writing k_{ij} as $k_{a,x;a',x'}$ and with the definition $K_{xx'}(t) = \sum_{a,a'} P_{a|x}(t) k_{x,a;x',a'}(t)$, where $P_{a|x}(t) \equiv P_{x,a}(t)/P_x(t)$ is a conditional probability, this equation becomes

$$\frac{d}{dt} P_x(t) = \sum_{x'} [P_x(t) K_{xx'}(t) - P_{x'}(t) K_{x'x}(t)]. \tag{37}$$

We point out that the dynamics of the variable x is non-Markovian since $K_{xx'}$ depends on the dynamics of the hidden degree of freedom represented by a.

The transition rates $k_{x,a;x',a'}(t)$ fulfill the generalized detailed balance relation from Eq. (11). The affinities $\mathcal{A}_\alpha$ are related to the variable a and they are responsible for the active nature of the engine. The accompanying distribution is still the solution of Eq. (9). For the coarse-grained variable x we have $P_x^S = \sum_a P_{x,a}^S$.

B. Passive counterpart

We assume that the time dependence of the energy $E_i(t) = E_{x,a}(t)$ is contained only in the observable variable x. If this assumption does not hold we would also need access to the variable a to evaluate the work. This energy can then be written as

$$E_{x,a}(t) = E_x(t) + E_{x,a}^{\rm int} + E_a, \tag{38}$$

where $E_{x,a}^{\rm int}$ represents the interaction energy between x and a.

The passive counterpart for our active heat engine is defined as follows. The equilibrium accompanying distribution associated with this passive counterpart is

$$P_x^{\rm eq}(t) = e^{-\beta E_x(t)}/Z(t), \tag{39}$$

where $Z(t) = \sum_x e^{-\beta E_x(t)}$. Such a passive heat engine corresponds to the complete absence of the degrees of freedom represented by the variable a. Physically, if we consider the experiment in Ref. [18], this passive counterpart corresponds to an experiment where all bacteria are removed from the solution.

A different passive counterpart would be to keep the variable a but set all affinities $\mathcal{A}_\alpha$ to zero. In this case, the accompanying equilibrium distribution for the passive counterpart would be

$$p_x^{\rm eq}(t) = \sum_a P_{x,a}^{\rm eq} \propto e^{-\beta(t) E_x^{\rm eff}(t)}, \tag{40}$$

where $E_x^{\rm eff}(t) = E_x(t) - \beta(t)^{-1} \ln[\sum_a e^{-\beta(t)(E_{x,a}^{\rm int} + E_a)}]$. The difference is that this effective energy also contains the

interaction energy with the degrees of freedom represented by a. If we consider the experiment in Ref. [18], this would correspond to an experiment with all bacteria present but with their internal driving forces set to zero.

We proceed with the elimination of the variable a as the passive counterpart of the engine. However, this choice does not represent a limitation in our framework. As we show below, our formalism also applies if we choose this second type of passive counterpart of the engine, with modified expressions for the terms in the second law for active heat engines.

C. Work, heat, and entropy

The work in Eq. (20) can be written as

$$W = -\int_0^\tau dt \sum_x P_x(t) \frac{d}{dt} E_x(t). \tag{41}$$

This expression follows from the fact that the time dependence of $E_{x,a}(t)$ is fully contained in $E_x(t)$. The coarse-grained generalized heat flux is defined as

$$J_q^{\rm cg} \equiv \int_0^\tau dt \sum_{x<x'} h(t)\beta_c J_{xx'}(t)\Delta E_{xx'}(t), \tag{42}$$

where $J_{xx'}(t) \equiv P_x(t)K_{xx'}(t) - P_{x'}(t)K_{x'x}(t)$ and $\Delta E_{xx'}(t) \equiv E_{x'}(t) - E_x(t)$. This expression is in general different from Eq. (17).

The coarse-grained variable x also has a refined first law, which is obtained in the following way. The time derivative of the average of the part of the energy that depends on x is

$$\frac{d}{dt}\sum_x P_x(t)E_x(t) = \sum_{x<x'} J_{xx'}(t)\Delta E_{xx'}(t) + \sum_x P_x(t)\frac{d}{dt}E_x(t). \tag{43}$$

From the fact that this average energy is periodic, we obtain

$$\int_0^\tau dt \sum_{x<x'} J_{xx'}(t)\Delta E_{xx'}(t) = W, \tag{44}$$

which is the refined first law for the variable x.

For the particular case where the inverse temperature only takes two values β_c and β_h, i.e., $h(t)$ is either 0 or 1, the coarse-grained generalized heat on Eq. (42) becomes the coarse-grained heat taken from the hot reservoir $Q_h^{\rm cg}$. The first law in Eq. (44) becomes

$$Q_h^{\rm cg} + (-Q_c^{\rm cg}) = W, \tag{45}$$

where $-Q_c^{\rm cg} \equiv -\int_0^\tau dt \sum_{x<x'}[1-h(t)]\beta_c J_{xx'}(t)\Delta E_{xx'}(t)$ is the coarse-grained heat released to the cold reservoir.

The coarse-grained excess entropy, as shown in the Appendix, follows the second law inequality:

$$\Delta S_{\rm ex}^{\rm cg} \equiv \int_0^\tau dt \sum_{x<x'} J_{xx'}(t) \ln \frac{P_{x'}^S(t)}{P_x^S(t)} \geq 0. \tag{46}$$

Analogously to the procedure used for $\Delta S_{\rm ex}$, the coarse-grained excess entropy can be written as

$$\Delta S_{\rm ex}^{\rm cg} = \mathcal{I}^{\rm cg} + \eta_C J_q^{\rm cg} - \beta_c W \geq 0, \tag{47}$$

where

$$\mathcal{I}^{\rm cg} \equiv \frac{1}{\tau}\int_0^\tau dt \sum_{x<x'} J_{xx'}(t)\left[\ln \frac{P_{x'}^S(t)}{P_{x'}^{\rm eq}(t)} - \ln \frac{P_x^S(t)}{P_x^{\rm eq}(t)}\right] \geq 0. \tag{48}$$

Similar to $\mathcal{I}$, the coarse-grained term $\mathcal{I}^{\rm cg}$ has the same information-theoretical expression as a part of the time derivative of $\sum_x P_x(t)\ln[P_x^S(t)/P_x^{\rm eq}(t)]$, which leads to the alternative formula for $\mathcal{I}^{\rm cg}$:

$$\mathcal{I}^{\rm cg} = -\int_0^\tau dt \sum_x P_x(t)\frac{d}{dt}\left(\ln \frac{P_x^S(t)}{P_x^{\rm eq}(t)}\right). \tag{49}$$

Remarkably, the coarse-grained second law in Eq. (47) has the exact same form as compared to the full second law in Eq. (28), in spite of the different values of the coarse-grained quantities. The physical interpretation of $\mathcal{I}^{\rm cg}$ as an extra resource that allows for more extracted work in an active heat engine also applies to this coarse-grained second law. The standard entropy production of stochastic thermodynamics also has this feature of invariance under coarse graining that we found here for the excess entropy [57].

The structure of the second law for active heat engines is independent of whether one has full access to the state i or only access to x. For the first kind of active heat engine the system represented by i is in contact with an equilibrium heat bath, and for the second kind the system represented by x is in contact with a nonequilibrium heat bath.

As proved in the Appendix, the excess entropy is larger than the coarse-grained excess entropy; i.e.,

$$\Delta S_{\rm ex} \geq \Delta S_{\rm ex}^{\rm cg}. \tag{50}$$

Hence, even if we do have access to the hidden variable a, the coarse-grained excess entropy provides a tighter bound on extracted work.

The passive counterpart of the engine was so far chosen as the absence of the variable a. For a passive counterpart that corresponds to setting all affinities $\mathcal{A}_\alpha$ to zero, the expressions above change as follows. The work W remains the same. The coarse-grained generalized heat $J_q^{\rm cg}$ has the expression in Eq. (42) with $E_x^{\rm eff}(t)$ instead of $E_x(t)$. The term $\mathcal{I}^{\rm cg}$ has the expression in Eq. (48) [or the alternative expression in Eq. (49)] with $p_x^{\rm eq}(t)$ in Eq. (40) instead of $P_x^{\rm eq}(t)$ in Eq. (39). The coarse-grained excess entropy $\Delta S_{\rm ex}^{\rm cg}$ is independent of the choice of the passive counterpart.

For all the models we consider here, either there are no interactions between x and a, i.e., $E^{\text{int}}_{x,a} = 0$, or the condition that $\sum_a e^{-\beta(t)(E^{\text{int}}_{x,a}+E_a)}$ is independent of x holds. For these cases, $p^{\text{eq}}_x(t) = P^{\text{eq}}_x(t)$ and both passive counterparts correspond to the same expression for the terms that show up in $\Delta S^{\text{cg}}_{\text{ex}}$.

In our framework we have considered a master equation with discrete states. Hence, our results are also valid for a continuous state space that follows overdamped Langevin dynamics, since an overdamped Langevin equation can be obtained as a limit of a master equation. In fact, we do consider two models with continuous variables in the next two sections. The generalization of our framework to underdamped Langevin dynamics is left for future work.

V. ANALYTICAL CASE STUDIES

A. Expressions for W, Q_h, $\mathcal{I}$, and $\mathcal{I}^{\text{cg}}$

For the models analyzed in this section the period has four parts illustrated in Fig. 4. Each part has duration τ_l, with $l = 1, 2, 3, 4$ and $\sum_{l=1}^{4} \tau_l = \tau$. There are two energy changes and two temperature changes. The time-dependent temperature is $\beta(t) = \beta_c$ for $0 \le t \le \tau_1 + \tau_2$ and $\beta(t) = \beta_h$ for $\tau_1 + \tau_2 < t \le \tau$. The time-dependent energy is E_x for $0 \le t \le \tau_1$ and for $\tau_1 + \tau_2 + \tau_3 \le t \le \tau$, and $E_x + \delta E_x$ for $\tau_1 \le t \le \tau_1 + \tau_2 + \tau_3$. Since the accompanying distributions $P^S_x(t)$ and $P^{\text{eq}}_x(t)$ are also divided into four parts, we write them as $P^{S,l}_x$ and $P^{\text{eq},l}_x$, respectively. The affinities $\mathcal{A}_\alpha$ can also change for different parts of the period.

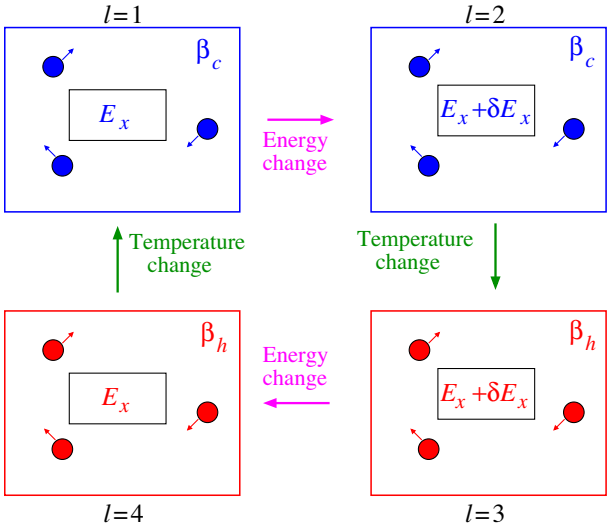


FIG. 4. Depiction of the protocol for the case studies. The inner box represents the observable degree of freedom x and the respective energy E_x. The particles with arrows represent hidden active degrees of freedom labeled by a. The four steps are as follows: from $l = 1$ to $l = 2$ energy changes from E_x to $E_x + \delta E_x$, from $l = 2$ to $l = 3$ the inverse temperature changes from β_c to β_h, from $l = 3$ to $l = 4$ the energy changes from $E_x + \delta E_x$ to E_x, and from $l = 4$ to $l = 1$ the inverse temperature changes from β_h to β_c.

The times are chosen such that $\tau_1 = \tau_3$ and $\tau_2 = \tau_4$. Furthermore, the temperature changes are assumed to be instantaneous, which corresponds to the limit $\tau_2 \to 0$. Physically, the temperature changes have to be much shorter than the typical time for a transition between states. In this limit, $\tau_1 = \tau/2$. For the two-state model we do obtain analytical expressions for finite time τ. For all other models we consider the limit of large τ, which means that at the end of the first and third parts the system has reached the respective stationary state. This assumption simplifies calculations since the distribution $P_x(t)$ at the end of these two parts is the same as the accompanying distribution $P^S_x(t)$. We point out that this case does not correspond to a quasistatic limit since the temperature changes are instantaneous. Within this particular limit the expressions for W, $Q_h = (\beta_c)^{-1} J_q$, $\mathcal{I}$, and $\mathcal{I}^{\text{cg}}$ acquire the following forms.

First, the work in Eq. (41) is

$$W = \sum_x (P^{S,3}_x - P^{S,1}_x)\delta E_x. \tag{51}$$

This expression can be directly understood from Fig. 4 as follows. From the first part to the second part the energy increases by δE_x, which gives the contribution $-P^{S,1}_x \delta E_x$. From the third part to the fourth part the energy decreases by δE_x, which gives the contribution $P^{S,3}_x \delta E_x$.

The heat taken from the hot reservoir in Eq. (17) is

$$Q_h = \sum_{x,x'} P^{S,1}_x P^{S,3}_{x'} (E_{x'} + \delta E_{x'} - E_x - + \delta E_x). \tag{52}$$

For the models in this section, the heat flux and the coarse-grained heat flux in Eq. (42) are equal. The term $\mathcal{I}$ in Eq. (36) becomes

$$\mathcal{I} = \sum_{x,a} (P^{S,3}_{x,a} - P^{S,1}_{x,a}) \left(\ln \frac{P^{S,3}_{x,a}}{P^{\text{eq},3}_{x,a}} - \ln \frac{P^{S,1}_{x,a}}{P^{\text{eq},1}_{x,a}} \right). \tag{53}$$

The coarse-grained term $\mathcal{I}^{\text{cg}}$ in Eq. (49) is given by

$$\mathcal{I}^{\text{cg}} = \sum_{x} (P^{S,3}_{x} - P^{S,1}_{x}) \left(\ln \frac{P^{S,3}_{x}}{P^{\text{eq},3}_{x}} - \ln \frac{P^{S,1}_{x}}{P^{\text{eq},1}_{x}} \right). \tag{54}$$

Since $J_q = J^{\text{cg}}_q$ for the models in this section, the inequality in Eq. (50) implies $\mathcal{I} \ge \mathcal{I}^{\text{cg}}$.

B. Two-state model

The two-state model is defined as follows. The transition rates of the model are given by $\kappa^+(t) = \omega^-(t) = k$,

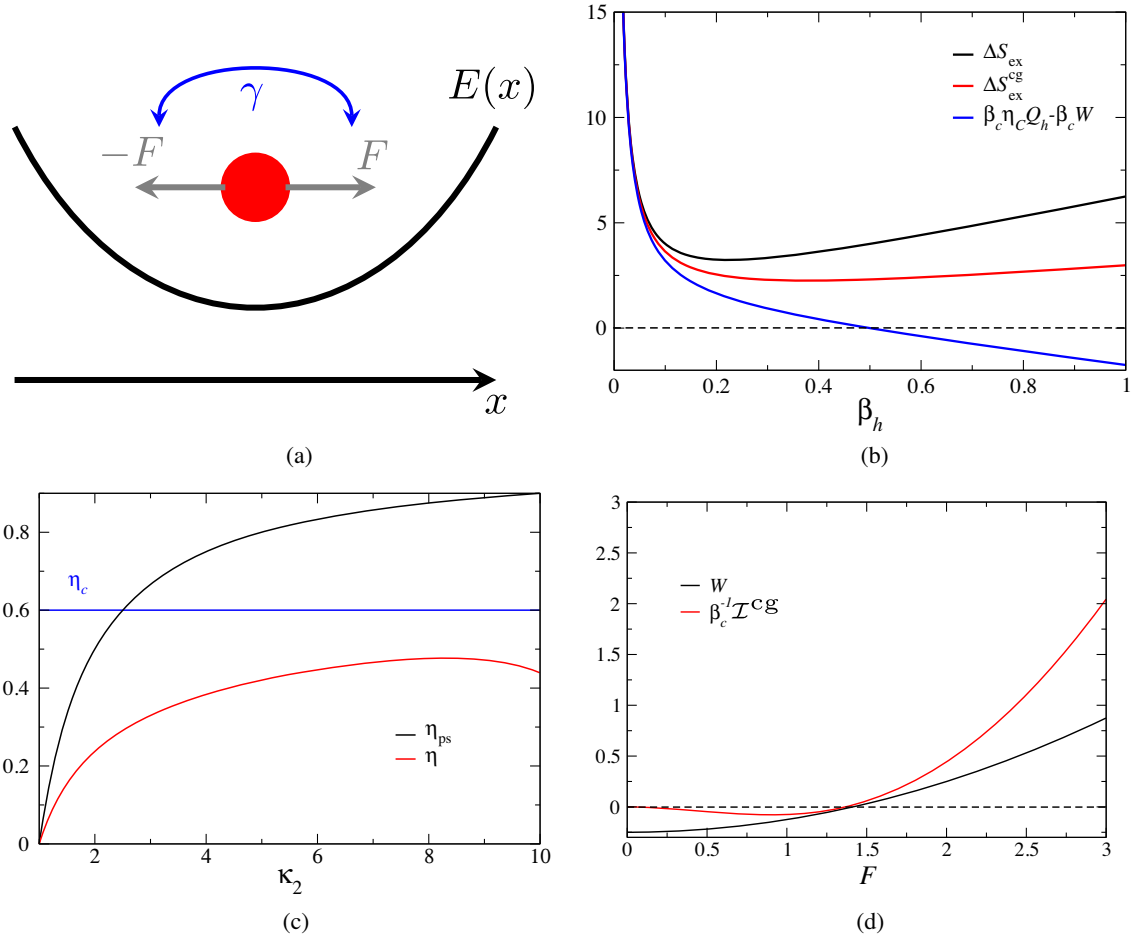


FIG. 5. (a) Depiction of an active particle with a force F that changes sign with a rate γ and in a harmonic potential $E(x)$. (b) The quantities $\beta_c\eta_C Q_h - \beta_c W$, which can be negative for an active heat engine, $\Delta S_{\rm ex}^{\rm cg}$, and $\Delta S_{\rm ex}$, as a function of the inverse hot temperature β_h for $\kappa_2 = 2$ and $F = 4$. (c) Efficiency η and pseudo-efficiency $\eta^{\rm ps}$ as functions of κ_2 and compared to the Carnot bound for $\beta_h = 0.4$, $F = 10$. (d) Work W and its upper bound $(\beta_c)^{-1}\mathcal{I}^{\rm cg}$ as functions of F for constant temperature $\beta_h = \beta_c$ and $\kappa_2 = 2$. For the three plots $\beta_c = 1$ and $\kappa_1 = 1$.

$\kappa^-(t) = ke^{\beta(t)[E(t)+\Delta\mu]}$, and $\omega^+(t) = ke^{\beta(t)E(t)}$, where the rates are represented in Fig. 1(b). The time-dependent energy $E(t)$ is E during the first and fourth parts of the period and $E + \delta E$ during the second and third parts of the period. The affinity $\mathcal{A}_\alpha$ is the free energy of one ATP hydrolysis $\Delta\mu$.

Using the methods from Ref. [58], which contains expressions for generic currents in two-state systems, we can evaluate Q_h, W, $\mathcal{I}$, and $\Delta S_{\rm act}$ analytically for finite τ. The expressions are too long to show here. The results for this model are shown in Sec. II.

C. Active particle under the action of a harmonic potential

We now consider a model of an active Brownian particle in one dimension under the action of a harmonic potential [59–61]. Similar models for cyclic active heat engines have been analyzed in Refs. [19,20,22,24]. Let us first consider the model for the case of fixed temperature and energy illustrated in Fig. 5(a). The position of the particle is labeled x. For this model x is a continuous variable; we then write the energy E_x as $E(x)$. The harmonic potential is written as $E(x) = \kappa x^2/2$. This model can be described by an overdamped Langevin equation that can be obtained as the continuous limit of the master equation for discrete x. A force of magnitude F acts on the active Brownian particle. The sign of the force is a Poisson process with a rate γ. The variable a then takes two values, one $a = +1$ corresponding to a force $+F$ and the other $a = -1$ corresponding to the force $-F$. The force F multiplied by distance is the affinity $\mathcal{A}_\alpha$ for this model.

We consider the limit for which γ is much smaller than the inverse relaxation time of the particle. Therefore, the particle reaches a stationary state with a given sign for the force F before the sign of the force changes. The stationary probability density for the particle to be at position x with $a = \pm 1$ is then

$$P^S_{\pm}(x) = \frac{\sqrt{\beta\kappa}e^{-\beta F^2/2\kappa}e^{-(1/2)\beta\kappa x^2 \mp \beta F x}}{2\sqrt{2\pi}}. \quad (55)$$

Furthermore, the probability density to be at position x is given by

$$P^S(x) = P^S_+(x) + P^S_-(x) = \frac{\sqrt{\beta\kappa}e^{-\beta F^2/2\kappa}e^{-(1/2)\beta\kappa x^2}}{\sqrt{2\pi}}\cosh(\beta F x). \quad (56)$$

The equilibrium distribution that corresponds to $F = 0$ is given by

$$P^{\mathrm{eq}}(x) = \frac{\sqrt{\beta\kappa}e^{-(1/2)\beta\kappa x^2}}{\sqrt{2\pi}}. \quad (57)$$

During the first part of the period the stiffness of the harmonic potential is κ_1 and the force is F_1. During the third part of the period the stiffness is κ_2 and the force is F_2. The second and fourth parts of the period are instantaneous. In terms of the notation in Sec. V A, $E_x = \kappa_1 x^2/2$ and $E_x + \delta E_x = \kappa_2 x^2/2$.

The work W in Eq. (51) is obtained with Eq. (56), written as an integral over x instead of a sum over x, which leads to the expression

$$W = \frac{(\kappa_2 - \kappa_1)[\beta_c\kappa_1^2(\beta_h F_2^2 + \kappa_2) - \beta_h\kappa_2^2(\beta_c F_1^2 + \kappa_1)]}{2\beta_c\beta_h\kappa_1^2\kappa_2^2}. \quad (58)$$

The heat taken from the hot reservoir in Eq. (52) for the present model is

$$Q_h = \frac{\beta_h F_2^2 + \kappa_2}{2\beta_h\kappa_2} - \frac{\kappa_2(\beta_c F_1^2 + \kappa_1)}{2\beta_c\kappa_1^2}. \quad (59)$$

The term $\mathcal{I}$ in Eq. (53) is obtained with Eqs. (57) and (55), which lead to

$$\mathcal{I} = \frac{(\beta_c F_1 - \beta_h F_2)(F_1\kappa_2 - F_2\kappa_1)}{\kappa_1\kappa_2}. \quad (60)$$

The coarse-grained term $\mathcal{I}^{\mathrm{cg}}$ in Eq. (54) follows from Eqs. (56) and (57):

$$\mathcal{I}^{\mathrm{cg}} = \int_{-\infty}^{\infty} dx\{P^{S,3}(x)\ln[\cosh(F_2\beta_h x)] - P^{S,1}(x)\ln[\cosh(F_1\beta_c x)]\}. \quad (61)$$

For this quantity the expression is in terms of an integral that has to be performed numerically.

We now restrict to the case $F_1 = 0$ and $F_2 = F$. In Fig. 5(b) we show that $\beta_c\eta_C Q_h - \beta_c W$ can be negative and that the inclusion of the term $\mathcal{I}^{\mathrm{cg}}$ recovers the appropriate statement of the second law for active heat engines. We also illustrate the inequality $\mathcal{I} \geq \mathcal{I}^{\mathrm{cg}}$ valid for the models analyzed here. In Fig. 5(c) we illustrate the second law for active heat engines from a different perspective. We compare the pseudo-efficiency $\eta^{\mathrm{ps}} = w/Q_h$ and the efficiency $\eta \equiv [W/Q_h + \mathcal{I}^{\mathrm{cg}}/(\beta_c - \beta_h)]$ with the Carnot efficiency η_C. Our second law does become the standard statement $\eta \leq \eta_C$ if we also account for the term $\mathcal{I}^{\mathrm{cg}}$, while the pseudo-efficiency does cross the Carnot bound.

This active heat engine can extract work even if the temperature is constant, i.e., $\beta_h = \beta_c$, as shown in Fig. 5(d), for which the extracted work becomes positive for $F > 1.4$ approximately. In this case, the second law becomes $W \leq (\beta_c)^{-1}\mathcal{I}^{\mathrm{cg}}$.

D. Two interacting particles

To illustrate the applicability of our second law to the case with interactions between x and a, we now consider a simple model for two interacting particles illustrated in Fig. 6(a). The particles can be in three different positions. The observable particle has its position labeled by $x = -1$, 0, 1. The energy for this particle is $E(t)$ for $x = \pm 1$ and zero for $x = 0$. The time-dependent energy is E for the first part of the period and $E + \delta E$ for the third part of the period. The position of the hidden particle is labeled by $a = -1, 0, 1$. The energy E_a is zero for all a. This particle is driven in the clockwise direction by an affinity $\mathcal{A}$. Finally, the particles interact with each other through a hardcore potential.

There are six possible states $i = a, x$. We label these states in the following way. State $i = 1$ corresponds to $x = -1$ and $a = 0$. State $i = 2$ corresponds to $x = -1$ and $a = 1$. State $i = 3$ corresponds to $x = 0$ and $a = -1$. State $i = 4$ corresponds to $x = 0$ and $a = 1$. State $i = 5$ corresponds to $x = 1$ and $a = -1$. State $i = 6$ corresponds to $x = 1$ and $a = 0$. Using this labeling of the states, the transpose of the stochastic matrix for this model reads

$$\begin{pmatrix} -r_1 & \gamma e^{\beta\mathcal{A}/6} & 0 & 0 & 0 & k \\ \gamma e^{-\beta\mathcal{A}/6} & -r_2 & 0 & k e^{\beta E/2} & 0 & 0 \\ 0 & 0 & -r_3 & \gamma e^{-\beta\mathcal{A}/6} & k e^{-\beta E/2} & 0 \\ 0 & k e^{-\beta E/2} & \gamma e^{\beta\mathcal{A}/6} & -r_4 & 0 & 0 \\ 0 & 0 & k e^{\beta E/2} & 0 & -r_5 & \gamma e^{\beta\mathcal{A}/6} \\ k & 0 & 0 & 0 & \gamma e^{-\beta\mathcal{A}/6} & -r_6 \end{pmatrix}, \quad (62)$$

where k (γ) sets the timescale for transitions of the observable (hidden) particle. The escape rates r_i are the sum of nondiagonal elements of a row. To shorten the matrix we have omitted the time dependence of $\mathcal{A}$, E, and β. For the first part of the period they are $\mathcal{A}_1$, E, and β_c, and for the third part of the period they are $\mathcal{A}_2$, $E + \delta_E$, and β_h. The stationary

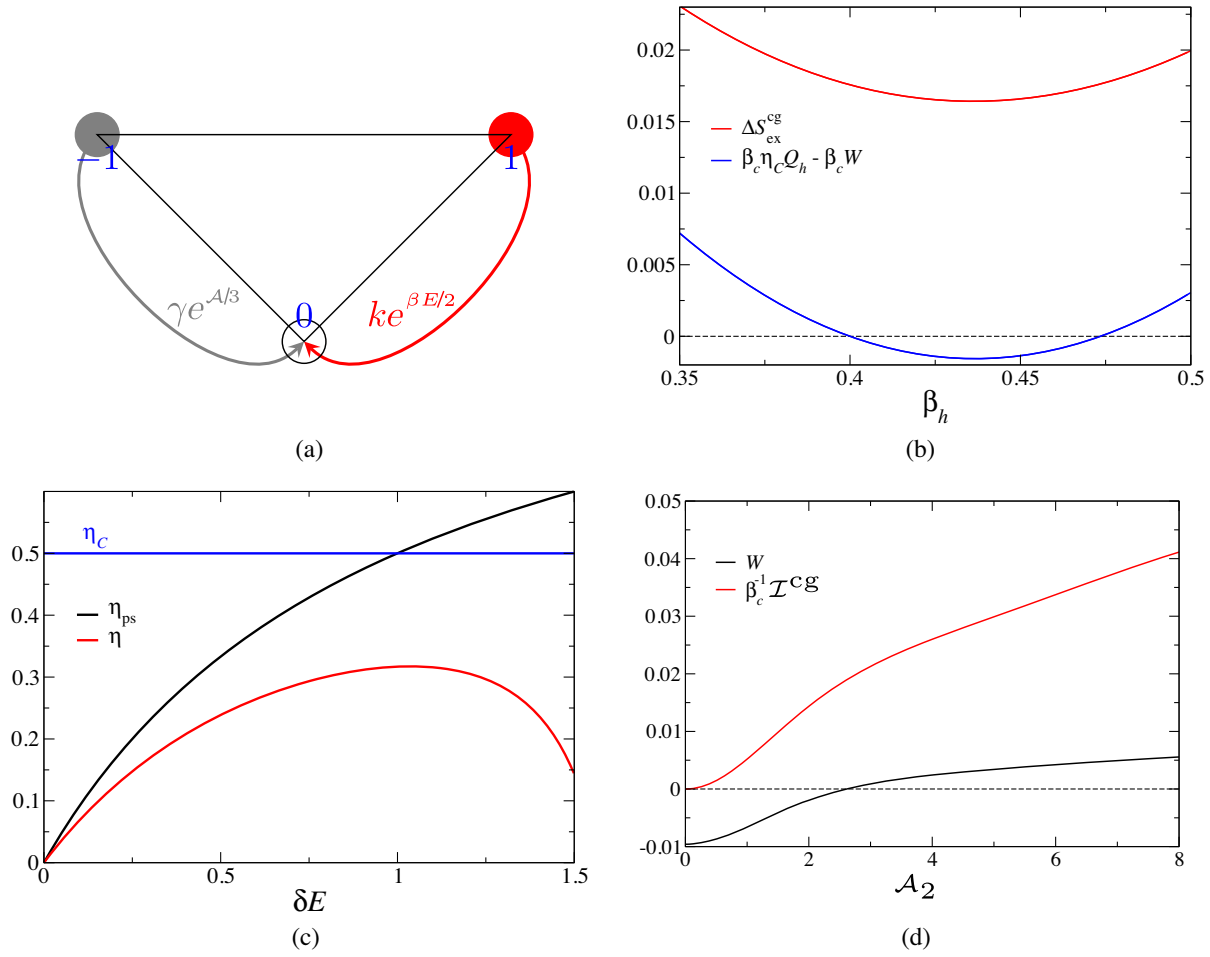


FIG. 6. (a) Depiction of the model for two interacting particles for the case the observable particle (in red) is at position $x = 1$ and the hidden particle (in gray) is at position $a = -1$. The particles can jump to position 0 with the rates given herein. (b) The quantities $\beta_c \eta_C Q_h - \beta_c W$, which can be negative for an active heat engine, and ΔS^{cg}_{ex} as a function of the inverse hot temperature β_h for $\delta E = 0.2$ and $\mathcal{A}_2 = 10/3$. (c) Efficiency η and pseudo-efficiency η^{ps} as functions of δE and compared to the Carnot bound for $\beta_h = 0.5$, $\mathcal{A}_2 = 10/3$. (d) Work W and its upper bound $(\beta_c)^{-1}\mathcal{I}^{cg}$ as functions of $\mathcal{A}_2$ for constant temperature $\beta_h = \beta_c$ and $\delta E = 1.5$. For the three plots $\beta_c = 1$, $\mathcal{A}_1 = 0$, $E = 1$, $k = 10$, and $\gamma = 1$.

distribution P^S is the eigenvactor associated with the eigenvalue 0 of this matrix. With this stationary distribution we obtain analytical expressions for all desired quantities defined in Sec. V A. The analytical expressions are too long to be displayed here.

This model displays all the same features from the previous model, as illustrated in Fig. 6, which reiterates the fact that the second law for active heat engines is independent of the particular model. The efficiency in Fig. 6(c) is $\eta \equiv [W/Q_h + \mathcal{I}^{cg}/(\beta_c - \beta_h)]$.

VI. NUMERICAL CASE STUDY

A. Relation to experiment

The models of the active particle in a harmonic potential and of the two interacting particles have some similarities with the experiment in Ref. [18]. There is an observable particle labeled by x. Furthermore, heat and work can be evaluated by an expression that only depends on the observation of the variable x. The difference is in the hidden degrees of freedom. For the first model, this hidden degree of freedom is the sign of the force F. For the second model, the hidden degree of freedom is the position of the driven particle. For the experiment involving a colloidal particle in a bacterial bath, the hidden degrees of freedom are more complex. They correspond to the positions and other variables that determine the state of the bacteria.

However, the fact that the hidden degrees of freedom in the experiment are more complex does not present a challenge for our second law. The new term $\mathcal{I}^{cg}$ can also be calculated by the sole observation of the variable x, no matter how complex the hidden degrees of freedom are. Our second law should be directly applicable to the experiment in Ref. [18] and to future experiments with cyclic active heat engines. To illustrate the generality of our second law we consider a more complex model than our

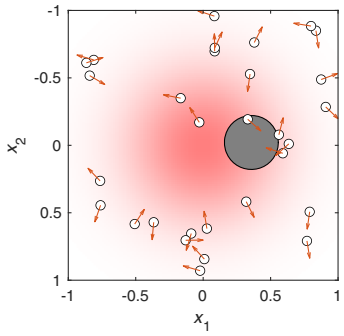


FIG. 7. Setup for our numerical case study. A passive particle (gray) trapped in a harmonic potential (indicated in red) interacts with many active Brownian particles (white) that are self-propelled in randomly varying directions (orange arrows).

analytical case studies. This model has several hidden degrees of freedom and is inspired by the experiment in Ref. [18].

B. Model definition

The two-dimensional model depicted in Fig. 7 has a passive particle and many active particles that make the reservoir active. These active particles correspond to the bacteria in the experiment. The passive particle has its position labeled by $\boldsymbol{x} = (x_1, x_2)$ and is subjected to a harmonic potential,

$$E(x) = \kappa x^2/2, \tag{63}$$

where κ is the stiffness of the potential and $x \equiv |\boldsymbol{x}|$. The active particles are not influenced by this harmonic potential and they do not interact with each other. They all interact with the passive particle via the repulsive potential $V(|\boldsymbol{r}|)$, where $\boldsymbol{r}$ is the distance to the passive particle. This potential is given by

$$V(|\boldsymbol{r}|) = \begin{cases} 0 & \text{if } |\boldsymbol{r}| > R \\ \kappa_{\text{rep}}(R - |\boldsymbol{r}|)^2/2 & \text{if } |\boldsymbol{r}| < R, \end{cases} \tag{64}$$

where R is the radius of the passive particle, and $\kappa_{\text{rep}} > 0$ such that the interaction is repulsive.

The passive particle obeys the Langevin equation,

$$\dot{\boldsymbol{x}} = -\mu_p \nabla_{\boldsymbol{x}} E(x) - \mu_p \sum_m \nabla_{\boldsymbol{x}} V(|\boldsymbol{a}_m - \boldsymbol{x}|) + \boldsymbol{\xi}(t), \tag{65}$$

where μ_p is the mobility of the passive particle. The term $\boldsymbol{\xi}(t)$ corresponds to Gaussian white noise with intensity μ_p/β. The sum over m is over all active particles and $\boldsymbol{a}_m$ is the position of the active particle indexed by m.

The active particle labeled by m obeys the Langevin equation,

$$\dot{\boldsymbol{a}}_m = u\boldsymbol{n}(\phi_m) - \mu_a \nabla_{\boldsymbol{a}_m} V(|\boldsymbol{a}_m - \boldsymbol{x}|) + \boldsymbol{\zeta}_m(t), \tag{66}$$

where μ_a is the mobility of the active particles and $\boldsymbol{\zeta}_m(t)$ is a Gaussian white noise with intensity μ_a/β. For the first term on the left-hand side, u is the active speed, $\boldsymbol{n}(\phi_m)$ is a unit vector pointing in the direction of angle ϕ_m. The angle ϕ_m undergoes Brownian diffusion with diffusion coefficient D_r [62,63].

For the active particles the vector $\boldsymbol{a}_m$ is constrained to the box shown in Fig. 7 with periodic boundary conditions. The passive particle is unconstrained. However, due to the harmonic potential $E(x)$ the passive particle rarely leaves this box.

The protocol has the same four parts as the generic one shown in Fig. 4. Each part of the protocol takes the time $\tau/4$. The temperature changes are the same as in all models analyzed here. The stiffness of the potential $E(x)$ switches from κ_1 to κ_2 from the first to the second part and changes back from κ_2 to κ_1 from the third to the fourth part. The velocity u of the active particles also changes with the protocol. During the parts of the protocol where the inverse temperature is β_c (β_h) this velocity is u_c (u_h).

We here have set the parameters to the following values: the period is $\tau = 20$; there are 30 active particles; the stiffness takes the values $\kappa_1 = 10$, $\kappa_2 = 15$; the inverse temperatures are $\beta_c^{-1} = 0.1$, $\beta_h^{-1} = 0.12$; the radius of the passive particle is $R = 0.2$; the mobilities are $\mu_p = 0.05$ and $\mu_a = 0.1$; the parameter of the repulsive potential is $\kappa_{\text{rep}} = 1000$; the diffusion coefficient associated with the direction of the active speed is $D_r = 1$. The active speed is set as $u_h = u_0$ and $u_c = u_0/100$, where u_0 is parameter that we vary. For $u_0 = 0$ the heat engine is passive.

C. Simulations and observables

Simulations are performed by integrating the Langevin equations (65) and (66) using a simple stochastic Euler scheme with a sufficiently small time step. We have used a time step of 10^{-4}. For each trajectory, we write stochastic work and heat as w and q_h, respectively. The heat Q_h and work W are related to the average of these quantities over stochastic trajectories.

From Eq. (41), the increments to the stochastic work w happen at the two times of the period that the stiffness changes. For the change in the stiffness from the first to the second part, the increment in the work is $(\kappa_1 - \kappa_2)\boldsymbol{x}^2/2$, where $\boldsymbol{x}$ is the position of the passive particle at the moment of the change. For the second change in the stiffness from the third to the fourth part, the increment in the work is $(\kappa_2 - \kappa_1)\boldsymbol{x}^2/2$.

From Eq. (42), the increments to the stochastic heat q_h happen when the passive particle changes position during the third and fourth parts of the period, for which the inverse temperature is β_h. This increment can be formally

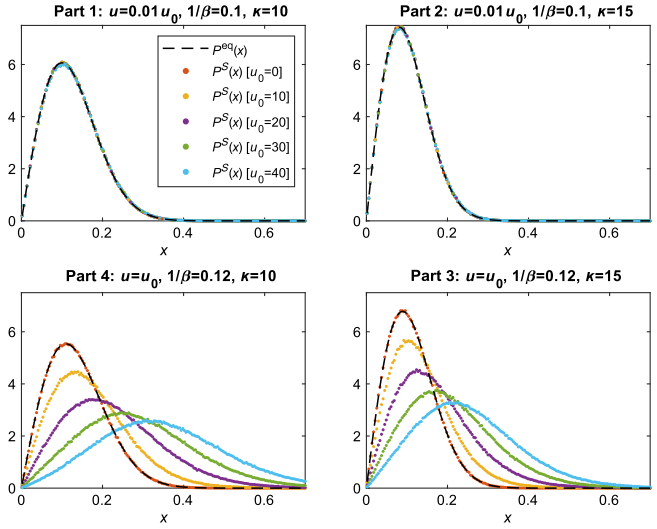


FIG. 8. Stationary distributions sampled with fixed parameters according to each of the four steps of the protocol. The reference active speed u_0 corresponds to different curves indicated by color. The equilibrium distribution is shown as a dashed black curve. For parts 1 and 2, the stationary distribution are closer to the equilibrium distribution due to the small driving speed u.

written as $dq_h = h(t)k(t)d(\boldsymbol{x}^2)/2$, where the function $h(t)$ in Eq. (10) is 1 if $\beta(t) = \beta_h$ and 0 if $\beta(t) = \beta_c$.

Estimates for the average heat Q_h and work W per cycle then follow from the final value of q_h and w divided by the number of cycles performed in a long trajectory (we average over 5000 cycles).

The calculation of $\mathcal{I}^{cg}$ is more involved. From Eq. (49), to calculate the increment associated with this quantity, we need the equilibrium distribution,

$$P^{eq}(x) = \beta k x \exp(-\beta k x^2/2), \tag{67}$$

and the accompanying distribution P^S. Since the protocol has four parts, there are four stationary distributions $P^{S,l}(x)$. To sample these distributions we run four separate simulations where κ, β, and u remain constant. The results for these distributions are shown in Fig. 8. This method of sampling $P^{S,l}(x)$ from realizations with fixed parameters that reach a steady state can also be used in experiments for which this probability is not known.

The stochastic quantity corresponding to $\mathcal{I}^{cg}$ can now be calculated in similar way to the stochastic work w. Whenever there is a change from part l to part $l+1$ of a cycle and the passive particle is in position $\boldsymbol{x}$, we add up the increment $\ln[P^{S,l}(x)P^{eq,l+1}(x)/P^{eq,l}(x)P^{S,l+1}(x)]$. To get an estimate of the average $\mathcal{I}^{cg}$ we divide the total sum of all increments during the trajectory by the number of cycles.

We reiterate that our estimates of W, Q_h, and $\mathcal{I}^{cg}$ only depend on the monitoring of the position of the passive particle $\boldsymbol{x}$. This scheme can also be applied to an experiment where one does not have access to the hidden active degrees of freedom, such as the experiment in Ref. [18].

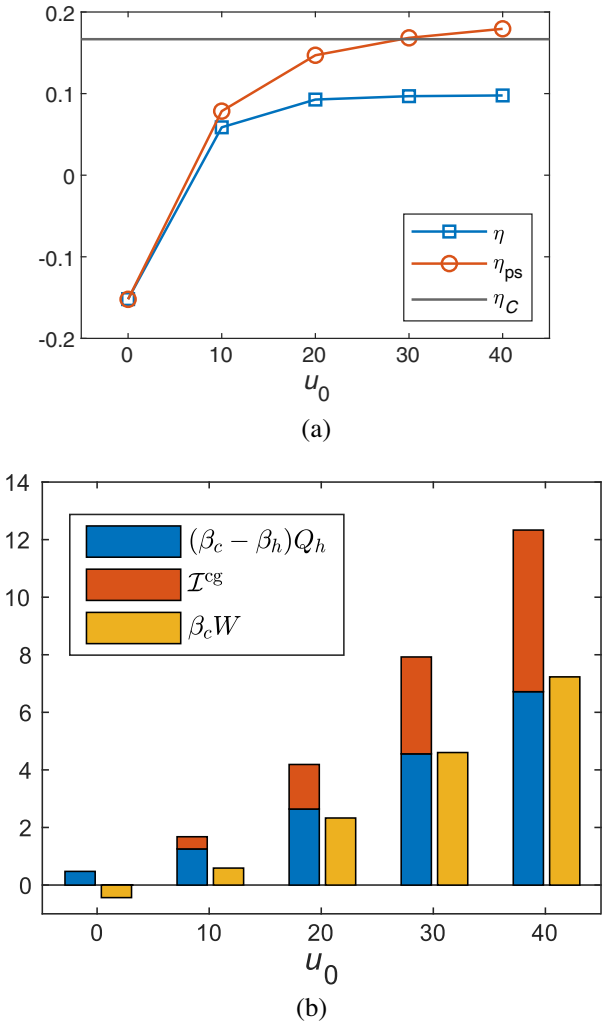


FIG. 9. Results for numerical model. (a) Efficiency (blue squares) and pseudo-efficiency (red circles), as a function of the reference active speed $u_0 = u_h$. Error bars calculated from the standard deviation with our data are not larger than the size of the points in the figure. (b) The quantities $\beta_c W$, $\mathcal{I}^{cg}$, and $(\beta_c - \beta_h)Q_c$.

D. Results

The results for the numerical model are shown in Fig. 9. They are plotted as a function of u_0, which parametrize the active speed. For $u_0 = 0$ the heat engine is passive. We observe that for increasing u_0 the active heat engine becomes more efficient. In fact, for our chosen parameters, activity allows the system to operate as a heat engine since for $u_0 = 0$ the extracted work W is negative. The pseudo-efficiency $\eta_{ps} \equiv W/Q_h$ does go beyond the Carnot bound while the efficiency $\eta \equiv [W/Q_h + \mathcal{I}^{cg}/(\beta_c - \beta_h)]$ stays below the bound, as predicted by our second law.

As shown in Fig. 8, with increasing active speed u_0 the stationary distribution departs further from the equilibrium distribution. In Fig. 9 we demonstrate that the term $\mathcal{I}^{\mathrm{cg}}$, which quantifies the difference between these distributions, increases, from about 30% of the value of $(\beta_c - \beta_h)Q_h$ for $u_0 = 10$ to 80% for $u_0 = 40$. When the engine is highly active, $\mathcal{I}^{\mathrm{cg}}$ acts as an important resource for the extraction of work.

Related models have been analyzed in Refs. [19,22], using Langevin equations for the passive particle with a noise that is not the standard Gaussian white noise. This noise effectively captures the influence of the active particles that interact with the passive particle. Both these studies use the concept of an effective temperature. An advantage of our approach is that we have no need to invoke an effective temperature; our second law inequality contains bona fide temperatures.

VII. CONCLUSION

The second law of thermodynamics determines which processes are possible. For well-known passive heat engines, this statement is expressed as the Carnot bound on the efficiency. We here derived the appropriate statement of the second law for cyclic active heat engines. Unlike a passive heat engine, the second law for an active heat engine includes an information-theoretic term, which allows active heat engines to perform tasks beyond the ones that can be performed by passive heat engines, such as work extraction at constant temperature.

Our second law for cyclic active heat engines has three main merits. First, it does not include the dissipation due to the hidden degrees of freedom. Second, it is expressed in terms of quantities that can be evaluated by the sole monitoring of the observable degrees of freedom. Third, it is general, as it only requires the mild assumption that at some level of description the observable degrees of freedom together with the hidden degrees of freedom follow Markovian dynamics. Furthermore, the system corresponding to the observable degrees of freedom does not have to be a single colloidal particle, which is the case for most models analyzed in the literature, but it can be any arbitrary collection of possibly interacting degrees of freedom.

From the perspective of stochastic thermodynamics, we have unveiled a central application of excess entropy in periodically driven systems. While this quantity has been known to fulfill a fluctuation theorem, we here show that it provides the correct statement of the second law for cyclic active heat engines. Beyond the known excess entropy we introduced a coarse-grained excess entropy that was key for the derivation of the second law in terms of quantities that only depend on the observable degrees of freedom. Interestingly, the second law for active heat engines retains its structure under coarse graining. This feature highlights the power of stochastic thermodynamics, which also applies to an active reservoir with hidden dissipative degrees of freedom that leads to non-Markovian dynamics.

Active matter is an important modern concept in physics. Quite generally, if we drive a system immersed in such active matter with an external periodic protocol with the objective of completing some task such as work extraction, the second law derived here constitutes a general bound on whether the completion of this task is possible. Therefore, our results are relevant for future experiments with active heat engines. From the theoretical side, the second law derived here opens up the possibility to develop a linear response theory for active heat engines, it is central for problems such as the optimization of work and efficiency, and it should be useful for future models of, possibly interacting, active heat engines.

APPENDIX: FLUCTUATION THEOREM FOR COARSE-GRAINED EXCESS ENTROPY

To prove the second law for coarse-grained excess entropy we must define quantities as functionals of stochastic trajectories from time $t = 0$ up to time $t = T$. The time t is discretized with time intervals δt to simplify the expression for the probability of a trajectory; the continuous time limit is recovered for $\delta t \to 0$. For discrete time, the transition probability from i to j is $M_{ij}(t) = \delta t k_{ij}(t)$ for $i \neq j$ and $M_{ii}(t) = 1 - \sum_{j\neq i} \delta t k_{ij}(t)$. For the proof in this appendix the time dependence of the transition rates is arbitrary; there is no need to assume a periodic time dependence. For the calculations below, a key quantity is the dual transition rate $k^\dagger_{ij}(t) = k_{ji}(t)P^S_j(t)/P^S_i(t)$ [the same equation applies for transition probabilities $M^\dagger_{ij}(t)$], where $P^S_i(t)$ is the stationary probability for fixed rates at time t; i.e., they are the solution of Eq. (9).

The probability of a trajectory $\Gamma = i_0 \to i_1 \ldots i_{N-1} \to i_N$, where $N = T/\delta t$, is given by

$$\mathrm{Prob}(\Gamma) = P_{i_0}(0) \prod_{n=0}^{N-1} M_{i_n i_{n+1}}(n\delta t). \tag{A1}$$

The probability of the reversed trajectory with dual dynamics (without reversing the time dependence of the transition rates) is

$$\mathrm{Prob}^\dagger(\Gamma) = P_{i_N}(N\delta t) \prod_{n=0}^{N-1} M^\dagger_{i_{n+1} i_n}(n\delta t). \tag{A2}$$

The trajectory-dependent excess entropy change is defined as

$$\Delta S_{\mathrm{ex}}(\Gamma) \equiv \sum_{n=0}^{N-1} \ln \frac{M_{i_n i_{n+1}}(n\delta t)}{M^\dagger_{i_{n+1} i_n}(n\delta t)} = \sum_{n=0}^{N-1} \ln \frac{P^S_{i_{n+1}}(n\delta t)}{P^S_{i_n}(n\delta t)}. \tag{A3}$$

Furthermore, the system entropy change is

$$\Delta S_{\rm syst}(\Gamma) \equiv \ln \frac{P_{i_0}(0)}{P_{i_N}(N\delta t)}. \tag{A4}$$

Therefore, the sum of these two terms gives

$$\Delta S_{\rm ex}(\Gamma) + \Delta S_{\rm syst}(\Gamma) = \ln \frac{\mathrm{Prob}(\Gamma)}{\mathrm{Prob}^*(\Gamma)}. \tag{A5}$$

From this formula it is easy to show that the excess entropy fulfills a fluctuation theorem [3]. A useful equation is

$$\langle \Delta S_{\rm ex} + \Delta S_{\rm syst} \rangle = D_{\rm KL}(\mathrm{Prob}||\mathrm{Prob}^\dagger) \geq 0, \tag{A6}$$

where the brackets indicate an average over stochastic trajectories and the Kullback-Leibler distance here is over the space of trajectories Γ. If the transition rates are time periodic with period τ, the final time is $T = \tau$, and at the starting time $t = 0$ the system has already reached the asymptotic time-invariant probability; this inequality becomes the second law used in Eq. (25). Note that the average change of system entropy is zero if the system entropy is time periodic.

To define the coarse-grained excess entropy we write $i = (x, a)$, where x is the observable variable and a is hidden. The dynamics for the coarse-grained variable x is not Markovian. The probability of a trajectory in the space of trajectories of x is defined as

$$\mathrm{Prob}(\Gamma^{\rm cg}) = \sum_{a_0,a_1,\ldots,a_N} \mathrm{Prob}(\Gamma), \tag{A7}$$

where $\Gamma^{\rm cg} = x_0 \to x_1 \ldots x_{N-1} \to x_N$. In order to express this probability in a form similar to Eq. (A1), we use the quantity $\mathcal{M}_{xx'}(t) = \sum_{a,a'} P_{a|x}(t) M_{x,a;x',a'}(t)$, which leads to the expression

$$\mathrm{Prob}(\Gamma^{\rm cg}) = P_{x_0}(0) \prod_{n=0}^{N-1} \mathcal{M}_{x_n x_{n+1}}(n\delta t). \tag{A8}$$

For this expression we used $P_{x_0,a_0}(0) = P_{x_0}(0) P_{a_0|x_0}(0)$ and

$$\sum_{a_n} P_{a_n|x_n}(n\delta t) M_{x_n,a_n;x_{n+1},a_{n+1}}(n\delta t) = \mathcal{M}_{x_n x_{n+1}}(n\delta t) P_{a_{n+1}|x_{n+1}}[(n+1)\delta t]. \tag{A9}$$

The same formula is valid for the reversed trajectory with dual rates, i.e.,

$$\mathrm{Prob}^\dagger(\Gamma^{\rm cg}) = P_{x_N}(N\delta t) \prod_{n=0}^{N-1} \mathcal{M}^\dagger_{x_n x_{n+1}}(n\delta t), \tag{A10}$$

where $\mathcal{M}^\dagger_{xx'}(t) = \sum_{a,a'} P_{a|x}(t) M^\dagger_{x,a;x',a'}(t)$.

The coarse-grained excess entropy is defined as

$$\Delta S^{\rm cg}_{\rm ex}(\Gamma) \equiv \sum_{n=0}^{N-1} \ln \frac{\mathcal{M}_{x_n x_{n+1}}(n\delta t)}{\mathcal{M}^\dagger_{x_{n+1} x_n}(n\delta t)} = \sum_{n=0}^{N-1} \ln \frac{P^S_{x_{n+1}}(n\delta t)}{P^S_{x_n}(n\delta t)}. \tag{A11}$$

The coarse-grained system entropy change is

$$\Delta S^{\rm cg}_{\rm syst}(\Gamma) \equiv \ln \frac{P_{x_0}(0)}{P_{x_N}(N\delta t)}. \tag{A12}$$

The sum of both terms gives

$$\Delta S^{\rm cg}_{\rm ex}(\Gamma) + \Delta S^{\rm cg}_{\rm syst}(\Gamma) = \ln \frac{\mathrm{Prob}(\Gamma^{\rm cg})}{\mathrm{Prob}^\dagger(\Gamma^{\rm cg})}. \tag{A13}$$

Hence,

$$\langle \Delta S^{\rm cg}_{\rm ex} + \Delta S^{\rm cg}_{\rm syst} \rangle = D^{\rm cg}_{\rm KL}(\mathrm{Prob}||\mathrm{Prob}^\dagger) \geq 0, \tag{A14}$$

where we have used the superscript cg in the Kullback-Leibler distance to indicate that we are considering probabilities in the space of trajectories $\Gamma^{\rm cg}$.

For transition rates that are time periodic with period τ, if the time interval is $T = \tau$ and if at the starting time $t = 0$ the system has already reached the asymptotic time-invariant probability, the inequality in Eq. (A14) becomes the coarse-grained second law used in Eq. (47). The average system entropy is periodic, and, therefore, the average system entropy change in a period is zero.

The average excess entropy is larger than the average coarse-grained excess entropy, since

$$D_{\rm KL}(\mathrm{Prob}||\mathrm{Prob}^\dagger) \geq D^{\rm cg}_{\rm KL}(\mathrm{Prob}||\mathrm{Prob}^\dagger), \tag{A15}$$

which implies the inequality in Eq. (50). A similar inequality for the standard entropy production has been used in Ref. [64].

Finally, we can prove a fluctuation theorem for the coarse-grained excess entropy using relation (A13):

$$\langle \exp(-\Delta S^{\rm cg}_{\rm ex} - \Delta S^{\rm cg}_{\rm syst}) \rangle^{\rm cg} = \sum_{\Gamma^{\rm cg}} \frac{\mathrm{Prob}^\dagger(\Gamma^{\rm cg})}{\mathrm{Prob}(\Gamma^{\rm cg})} \mathrm{Prob}(\Gamma^{\rm cg}) = 1. \tag{A16}$$

[1] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985).

[2] Y. Oono and M. Paniconi, *Steady State Thermodynamics*, Prog. Theor. Phys. Suppl. **130**, 29 (1998).

[3] U. Seifert, *Stochastic Thermodynamics, Fluctuation Theorems, and Molecular Machines*, Rep. Prog. Phys. **75**, 126001 (2012).

[4] T. Schmiedl and U. Seifert, *Efficiency at Maximum Power: An Analytically Solvable Model for Stochastic Heat Engines*, Europhys. Lett. **81**, 20003 (2008).

[5] V. Blickle and C. Bechinger, *Realization of a Micrometre-Sized Stochastic Heat Engine*, Nat. Phys. **8**, 143 (2012).

[6] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, *Quantum-Dot Carnot Engine at Maximum Power*, Phys. Rev. E **81**, 041106 (2010).

[7] Y. Izumida and K. Okuda, *Efficiency at Maximal Power of Minimal Nonlinear Irreversible Heat Engines*, Europhys. Lett. **97**, 10004 (2012).

[8] Z. C. Tu, *Stochastic Heat Engine with the Consideration of Inertial Effects and Shortcuts to Adiabaticity*, Phys. Rev. E **89**, 052148 (2014).

[9] K. Brandner, K. Saito, and U. Seifert, *Thermodynamics of Micro- and Nano-Systems Driven by Periodic Temperature Variations*, Phys. Rev. X **5**, 031019 (2015).

[10] O. Raz, Y. Subaşı, and C. Jarzynski, *Mimicking Nonequilibrium Steady States with Time-Periodic Driving*, Phys. Rev. X **6**, 021022 (2016).

[11] S. Ray and A. C. Barato, *Stochastic Thermodynamics of Periodically Driven Systems: Fluctuation Theorem for Currents and Unification of Two Classes*, Phys. Rev. E **96**, 052120 (2017).

[12] T. Koyuk and U. Seifert, *Operationally Accessible Bounds on Fluctuations and Entropy Production in Periodically Driven Systems*, Phys. Rev. Lett. **122**, 230601 (2019).

[13] P. Steeneken, K. Le Phan, M. Goossens, G. Koops, G. Brom, C. Van der Avoort, and J. Van Beek, *Piezoresistive Heat Engine and Refrigerator*, Nat. Phys. **7**, 354 (2011).

[14] I. A. Martínez, E. Roldán, L. Dinis, D. Petrov, and R. A. Rica, *Adiabatic Processes Realized with a Trapped Brownian Particle*, Phys. Rev. Lett. **114**, 120601 (2015).

[15] I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. Parrondo, and R. A. Rica, *Brownian Carnot Engine*, Nat. Phys.**12**, 67 (2016).

[16] J. Roßnagel, S. T. Dawkins, K. N. Tolazzi, O. Abah, E. Lutz, F. Schmidt-Kaler, and K. Singer, *A Single-Atom Heat Engine*, Science **352**, 325 (2016).

[17] I. A. Martinez, E. Roldan, L. Dinis, and R. A. Rica, *Colloidal Heat Engines: A Review*, Soft Matter **13**, 22 (2017).

[18] S. Krishnamurthy, S. Ghosh, D. Chatterji, R. Ganapathy, and A. Sood, *A Micrometre-Sized Heat Engine Operating between Bacterial Reservoirs*, Nat. Phys. **12**, 1134 (2016).

[19] R. Zakine, A. Solon, T. Gingrich, and F. Van Wijland, *Stochastic Stirling Engine Operating in Contact with Active Baths*, Entropy **19**, 193 (2017).

[20] A. Saha and R. Marathe, *Stochastic Work Extraction in a Colloidal Heat Engine in the Presence of Colored Noise*, J. Stat. Mech. (2019) 094012.

[21] T. Ekeh, M. E. Cates, and E. Fodor, *Thermodynamic Cycles with Active Matter*, Phys. Rev. E **102**, 010101(R) (2020).

[22] V. Holubec, S. Steffenoni, G. Falasco, and K. Kroy, *Active Brownian Heat Engines*, Phys. Rev. Research **2**, 043262 (2020).

[23] V. Holubec and R. Marathe, *Underdamped Active Brownian Heat Engine*, Phys. Rev. E **102**, 060101(R) (2020).

[24] A. Kumari, P. S. Pal, A. Saha, and S. Lahiri, *Stochastic Heat Engine Using an Active Particle*, Phys. Rev. E **101**, 032109 (2020).

[25] J. S. Lee, J.-M. Park, and H. Park, *Brownian Heat Engine with Active Reservoirs*, Phys. Rev. E **102**, 032116 (2020).

[26] É. Fodor and M. E. Cates, *Active Engines: Thermodynamics Moves Forward*, Europhys. Lett. **134**, 10003 (2021).

[27] G. Gronchi and A. Puglisi, *Optimization of an Active Heat Engine*, Phys. Rev. E **103**, 052134 (2021).

[28] P. Pietzonka, E. Fodor, C. Lohrmann, M. E. Cates, and U. Seifert, *Autonomous Engines Driven by Active Matter: Energetics and Design Principles*, Phys. Rev. X **9**, 041032 (2019).

[29] S. Ramaswamy, *The Mechanics and Statistics of Active Matter*, Annu. Rev. Condens. Matter Phys. **1**, 323 (2010).

[30] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Hydrodynamics of Soft Active Matter*, Rev. Mod. Phys. **85**, 1143 (2013).

[31] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, *Active Particles in Complex and Crowded Environments*, Rev. Mod. Phys. **88**, 045006 (2016).

[32] F. Jülicher, S. W. Grill, and G. Salbreux, *Hydrodynamic Theory of Active Matter*, Rep. Prog. Phys. **81**, 076601 (2018).

[33] A. Argun, A.-R. Moradi, E. Pinçe, G. B. Bagci, A. Imparato, and G. Volpe, *Non-Boltzmann Stationary Distributions and Nonequilibrium Relations in Active Baths*, Phys. Rev. E **94**, 062150 (2016).

[34] P. Pietzonka and U. Seifert, *Entropy Production of Active Particles and for Particles in Active Baths*, J. Phys. A **51**, 01LT01 (2018).

[35] D. Mandal, K. Klymko, and M. R. DeWeese, *Entropy Production and Fluctuation Theorems for Active Matter*, Phys. Rev. Lett. **119**, 258001 (2017).

[36] T. Speck, *Active Brownian Particles Driven by Constant Affinity*, Europhys. Lett. **123**, 20007 (2018).

[37] S. Shankar and M. C. Marchetti, *Hidden Entropy Production and Work Fluctuations in an Ideal Active Gas*, Phys. Rev. E **98**, 020604(R) (2018).

[38] L. Dabelow, S. Bo, and R. Eichhorn, *Irreversibility in Active Matter Systems: Fluctuation Theorem and Mutual Information*, Phys. Rev. X **9**, 021009 (2019).

[39] A. Gopal, É. Roldán, and S. Ruffo, *Energetics of Critical Oscillators in Active Bacterial Baths*, J. Phys. A **54**, 164001 (2021).

[40] T. Markovich, E. Fodor, E. Tjhung, and M. E. Cates, *Thermodynamics of Active Field Theories: Energetic Cost of Coupling to Reservoirs*, Phys. Rev. X **11**, 021057 (2021).

[41] T. Hatano and S. Sasa, *Steady-State Thermodynamics of Langevin Systems*, Phys. Rev. Lett. **86**, 3463 (2001).

[42] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt, *Experimental Test of Hatano and Sasa's Nonequilibrium Steady-State Equality*, Proc. Natl. Acad. Sci. U.S.A. **101**, 15038 (2004).

[43] T. Speck and U. Seifert, *Integral Fluctuation Theorem for the Housekeeping Heat*, J. Phys. A **38**, L581 (2005).

[44] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, *Path-Integral Analysis of Fluctuation Theorems for General Langevin Processes*, J. Stat. Mech. (2006) P08001.
[45] C. Perez-Espigares, A. B. Kolton, and J. Kurchan, *An Infinite Family of Second Law-like Inequalities*, Phys. Rev. E **85**, 031135 (2012).
[46] T. Sagawa and H. Hayaka, *Geometrical Expression of Excess Entropy Production*, Phys. Rev. E **84**, 051110 (2011).
[47] M. Esposito and C. Van den Broeck, *Three Detailed Fluctuation Theorems*, Phys. Rev. Lett. **104**, 090601 (2010).
[48] M. Esposito and C. Van den Broeck, *Three Faces of the Second Law. I. Master Equation Formulation*, Phys. Rev. E **82**, 011143 (2010).
[49] C. Van den Broeck and M. Esposito, *Three Faces of the Second Law. II. Fokker-Planck Formulation*, Phys. Rev. E **82**, 011144 (2010).
[50] P. Hänggi and H. Thomas, *Stochastic Processes: Time Evolution, Symmetries and Linear Response*, Phys. Rep. **88**, 207 (1982).
[51] R. Chetrite and S. Gupta, *Two Refreshing Views of Fluctuation Theorems through Kinematics Elements and Exponential Martingale*, J. Stat. Phys. **143**, 543 (2011).
[52] S. Rahav, *Extracting Work from Stochastic Pumps*, J. Stat. Mech. (2011) P09020.
[53] C. H. Bennett, *The Thermodynamics of Computation—A Review*, Int. J. Theor. Phys. **21**, 905 (1982).
[54] S. Deffner and C. Jarzynski, *Information Processing and the Second Law of Thermodynamics: An Inclusive, Hamiltonian Approach*, Phys. Rev. X **3**, 041003 (2013).
[55] D. Hartich, A. C. Barato, and U. Seifert, *Stochastic Thermodynamics of Bipartite Systems: Transfer Entropy Inequalities and a Maxwell's Demon Interpretation*, J. Stat. Mech. (2014) P02016.
[56] J. M. Horowitz and M. Esposito, *Thermodynamics with Continuous Information Flow*, Phys. Rev. X **4**, 031015 (2014).
[57] M. Esposito, *Stochastic Thermodynamics Under Coarse-Graining*, Phys. Rev. E **85**, 041125 (2012).
[58] A. C. Barato and R. Chetrite, *Current Fluctuations in Periodically Driven Systems*, J. Stat. Mech. (2018) 053207.
[59] A. Dhar, A. Kundu, S. N. Majumdar, S. Sabhapandit, and G. Schehr, *Run-and-Tumble Particle in One-Dimensional Confining Potentials: Steady-State, Relaxation, and First-Passage Properties*, Phys. Rev. E **99**, 032132 (2019).
[60] U. Basu, S. N. Majumdar, A. Rosso, S. Sabhapandit, and G. Schehr, *Exact Stationary State of a Run-and-Tumble Particle with Three Internal States in a Harmonic Trap*, J. Phys. A **53**, 09LT01 (2020).
[61] R. Garcia-Millan and G. Pruessner, *Run-and-Tumble Motion in a Harmonic Potential: Field Theory and Entropy Production*, J. Stat. Mech. (2021) 063203.
[62] Y. Fily and M. C. Marchetti, *Athermal Phase Separation of Self-Propelled Particles with No Alignment*, Phys. Rev. Lett. **108**, 235702 (2012).
[63] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, *Active Brownian Particles*, Eur. Phys. J. Special Topics **202**, 1 (2012).
[64] E. Roldan and J. M. R. Parrondo, *Entropy Production and Kullback-Leibler Divergence between Stationary Trajectories of Discrete Systems*, Phys. Rev. E **85**, 031129 (2012).